Study of transformation of defect states in GaN- and SiC-based materials and devices

Lorenzo Rigutti

Dottorato di Ricerca in Fisica
XVIII ciclo

Alma Mater Studiorum
Università degli Studi di Bologna

SETTORE DISCIPLINARE: FIS/03

TESI PER IL CONSEGUIMENTO DEL
TITOLO DI DOTTORE DI RICERCA

STUDY OF TRANSFORMATION OF DEFECT STATES IN GaN- AND SiC-BASED MATERIALS AND DEVICES

Candidato: Lorenzo Rigutti

Supervisore: Chiar.mo Prof. Anna Cavallini

_____

Coordinatore: Chiar.mo Prof. Roberto Soldati

_____

Table of Contents

0 Introduction

The study of defects in semiconductor Physics has to be considered one of the most important fields of the research. The semiconductors, indeed, can only approximately be described by the idealized model of the crystal lattice, although some fundamental properties, such as the energy bandgap, are predicted by this model. Defects are invariably present. Atoms from foreign species (*impurity atoms*) can be found either due to a specific choice or due to a lacking control over the conditions at which the semiconductor was produced. Intrinsic defects, i.e. crystal imperfections, have an equilibrium concentration depending on the temperature. A concentration of intrinsic defects exceeding the equilibrium can be created by specific treatments, or result from unwished interaction between material and environment during or after the growth. Furthermore, there are defects extending on scales going beyond those of atomic dimensions or lattice constants. Such defects, also called *extended defects*, may have various causes and effects. Dislocations are the most studied extended defect in the semiconductor science: these are line defects with intriguing electronic and structural properties, some of which will be illustrated in this thesis.
The study of defects is fundamental: their presence is required for the operation of a device, but on the other hand, their presence can significantly hinder the device working. As an example, some defects, such as shallow impurity dopants, are required if they fulfil certain design properties of the device, while other defects have parasitic effects, degrading the properties of the device by introducing in the gap electronic levels interfering with charge carriers in the valence and conduction band. However, *the kind of device and application determines which kind of defect is desirable or not*.
When a new material is developed, through a process beginning with the growth of good crystal quality and leading to the production of commercially available devices, it is necessary that this process is supported by an intensive characterization of the defects, their nature, the ways to control their presence and their influence on device performance.
For this reason also in the framework of wide bandgap (WBG) semiconductors, whose technology was developed from the 1980's and continues its progress in these days, many researchers focused on defect investigation in these materials.
Wide bandgap semiconductors, so called because their energy gap (E_g>2.5eV) exceeds significantly that of the most common Si (E_g=1.1eV) and GaAs (E_g=1.4 eV), is a category to which belong, among others, gallium nitride (GaN, E_g=3.39eV) and silicon carbide (SiC, E_g=2.3-3.4eV). These materials have made applications possible, which were previously unattainable by devices based on Si and GaAs. These are mostly niche applications, such as illumination or operation in high-temperature or high-radiation environments, for which only small market portions will be available. However, these applications are relevant for modern life: ultrabright light-emitting diodes (LEDs) are a more efficient and energy-saving form of illumination, the blue laser diode allows an increase in the density of data writable on or readable from DVDs, and SiC-based power devices are able to operate at high temperature without need of cooling.

The present thesis is a study of the evolution of defect states in devices based on wide bandgap semiconductors under various treatments. The attention has been focused on light-emitting diodes based on GaN and Schottky diodes based on SiC, these latter a

basic structure for the fabrication of high-power rectifiers and ionising particle detectors. In both cases, we studied the defects and their electronic properties by means of the following experimental techniques: current-voltage (I-V) measurements, in order to investigate the effect of imperfections on the transport properties of the material/device; capacitance-voltage (C-V) measurements, yielding the profile of concentration of charge carriers, and giving information on the influence of defects on this concentration; deep level transient spectroscopy (DLTS), a technique allowing for the identification and characterization of defect-originated electron levels in the gap. We also employed techniques, such as photocurrent spectroscopy (PC), allowing for the characterization of light absorption by the material and/or device versus varying photon energy.

Apparently, the research is divided into two main sub-fields, one regarding GaN, the other regarding SiC, as the defects found in the two materials are generally not comparable. However, the link between these two sub-fields is strong: first, both materials are often present in the same device (as in GaN-based LEDs having a SiC substrate); secondly, this study is not a mere catalogue of the defects introduced and their effects on device performance, but rather an analysis on how the defects, once introduced in the materials, can evolve according to the treatments which the material/device undergoes. In the case of LEDs, we studied the evolution of defect states with the increase of the time spent by the device at a moderate forward current level (DC current stress at 20 mA), while in the case of SiC-based diodes the evolution of the defects, formerly introduced by irradiation, was obtained by annealing treatments at low temperature (<500K).

In both cases of SiC and GaN, the defect characterization was always interpreted in the framework of its influence on device operation. In the analysed LEDs the defect evolution was connected to the evolution of quantum efficiency, and in the SiC diodes we studied the effects of defect introduction on the charge collection efficiency (CCE) and on the leakage current of the device. Furthermore, for the interpretation of photocurrent spectra, we developed a model describing the generation of photocurrent considering the dispersion relations for the absorption coefficient and refractive index in the various device layers, as well as the internal reflection, transmission and interference phenomena involving the optical field within the device.

The research yielded various interesting results: we detected many deep levels introduced by proton- and electron-irradiation in SiC. The concentration of some of these levels varied with low-temperature annealing. From the study of their annealing behaviour and from comparisons with literature data we concluded, for instance, that one of these levels is related to a particular lattice defect, the carbon interstitial. By means of the analysis of the introduction rates of the levels and comparisons between proton and electron irradiation, we were able to distinguish between deep levels related to simple intrinsic defects and to defect complexes.

In the case of the GaN LED, we found that the evolution of several independent properties are strongly correlated, meaning that a single degradation mechanism is responsible for the observed changes. In particular, we concluded that the degradation of the light emission intensity is due to a migration/generation of defects into/in the active region of the device. The characteristics of the main deep level detected showed that it is related to an extended defect, most likely a dislocation.

The thesis is organized as follows. The first part is divided into two chapters. In chapter 1 the wide bandgap semiconductors and their peculiar properties are introduced. In chapter 2 we illustrate some important examples of devices fabricated based on wide bandgap semiconductors. The second part, divided into three chapters, deals with the experimental techniques and with their role in the characterization of devices and defects. In chapter 3 we illustrate the electrical characterization techniques, such as the current-voltage and the capacitance-voltage measurements. Chapter 4 regards the deep level transient spectroscopy (DLTS), technique allowing for an accurate analysis of defect states giving rise to electron levels in the gap. Chapter 5 deals with optical spectroscopy techniques, such as photocurrent (PC) and electroluminescence (EL) spectroscopy. Results and discussion are illustrated in the third part, divided into two chapters. The evolution phenomena of defect states in GaN-based LEDs are analysed in chapter 6, and those in SiC-based diodes in chapter 7.

1 Wide bandgap semiconductors

Silicon is the material dominating the industry of electronics: no other semiconducting material has reached the level of silicon technology as far as low defect density, long minority carrier lifetimes, interface and doping control concerns. Other semiconducting materials, e.g. gallium arsenide (GaAs) have a very mature technology, but they only fill niche applications, such as light emission. Light emission is a field in which silicon is disadvantaged because of its indirect bandgap which makes the emission of phonons upon electron-hole pair recombination significantly unlikely. Many efforts are currently under way to create photonic devices based on silicon, and very important results have been reached so far, such as the demonstration of optical gain in silicon nanocrystals [Pavesi00]. However, optoelectronics is mostly based on III-V semiconductors with direct bandgap or doped with isoelectronic impurities, which make light emission highly efficient even in indirect bandgap materials. Furthermore, the gap of Si (1.1 eV) is in the infrared interval of the electromagnetic spectrum, thus making visible-light applications of this material impossible. Optoelectronic devices emitting in the visible range (1.8-3 eV) need materials with higher bandgap. Light-emitting diodes (LEDs) emitting light up to 2.5 eV (green) have been built based on gallium arsenide-phosphide (GaAsP), a III-V alloy, doped with N [Schubert]. Higher photon energy requires a wider bandgap. Emission of blue light was first reported by Pankove in 1971 [Pankove71] in a gallium nitride (GaN) based MIS diode structure. The difficulty to dope GaN p-type delayed the construction of bright blue LEDs until the beginning of the 1990s [NakamuraFasol]. In the last decade of the XX century UV, violet, blue and green LEDs, as well as blue laser diodes (LDs) were built based on GaN and on its alloys with In and Al. The full visible spectrum is therefore available with light emission from solid-state devices. Today, the market for GaN-based optoelectronics is growing due to the applications of these devices: among others, signals, displays, illumination and fibre-optics communication, as well as higher density DVD laser readers.

Another niche field where Si is not the viable material is high temperature operation. Temperature is a crucial factor in device operation, as it influences many important material parameters. Many negative effects of high temperature which are found in Si may be minimized by using wide bandgap materials (WBG), such as GaN or silicon carbide (SiC). As an example, due to the wide bandgap the intrinsic carrier density is many orders of magnitude lower in the two latter materials than in Si. At high temperature, the intrinsic carrier density can negatively influence the electrical properties of the device based on Si, while the properties of the device based on a WBG material remain almost unaffected [Neudeck02]. Moreover, the thermal conductivity of SiC is much higher than that of Si, leading to a more efficient heat dissipation.

WBG materials are also advantaged in high-frequency operation: in these materials higher Schottky barrier make the replacement of Si bipolar structures with SiC unipolar structures possible, in which minority carriers are not injected. This represents a significant enhancement of the performances of switching devices operating in the GHz range [Zolper05].

As it was previously mentioned, Si can rely on a very mature technology, which also means the possibility of a tight control on the defect density present in the material. WBG technology is not as mature as that of Si, so that the control on defect density and crystal structure is still not optimal. Therefore, the study of defects in WBG semiconductors is of paramount importance, as defect characterization and comprehension represents a useful feedback for growth and technology.

In the following sections, the properties of the analyzed WBG materials will be shortly revised, and a brief introduction to the problems related with defects in these materials will be given. Some concrete examples of electronic devices based on WBG semiconductors will be considered in the next chapter.

Table 1.1. Selected physical properties of Si, GaAs, InN, GaN, AlN and 4H-SiC. The quantities shown are intended at T=300K [IOFFE]

Material	Si	GaAs	InN	GaN	AlN	4H-SiC
Density (g/cm^3)	2.33	5.32	6.81	6.15	3.25	3.21
Lattice constant (Å)	5.43	5.65	3.53	3.16	3.11	3.07
Energy Gap (eV)	1.1	1.43	0.77	3.4	6.1	3.2
Gap type	indirect	direct	direct	direct	direct	indirect
Dielectric constant / ε_0	11.7	12.9	15.3	8.9	8.5	9.7
Thermal conductivity (W/cm s)	1.5	0.5	0.45	1.3	2.8	3.7
Peak electron mobility (cm^2/V s)	1400	8000	3200	900	300	800
Breakdown electric field (MV/cm)	0.6	0.65		3.5		3.5
Saturated electron velocity ($\times 10^7$ cm/s)	1	1	2.5	1.5		2
Electron DOS effective mass / m_0	0.36	0.85	0.11	0.2	0.4	0.77
Hole DOS effective mass / m_0	0.81	0.53	1.65	1.5	7.3	~1

1.1. Gallium Nitride (GaN)

1.1.1. Physical properties

The first analyzed WBG semiconductor is gallium nitride (GaN). It belongs to the family of III-V semiconductors, and, together with AlN and InN, to the group III nitrides. Differently from phosphides, arsenides and antimonides of group III elements, coming exclusively in the cubic zincblend structure, group III nitrides crystallize preferably in the hexagonal wurtzite structure, the zincblend structure being metastable. GaN is a direct bandgap semiconductor (Fig. [1.1]), and its gap at 300 K is 3.39 eV [Chow96].

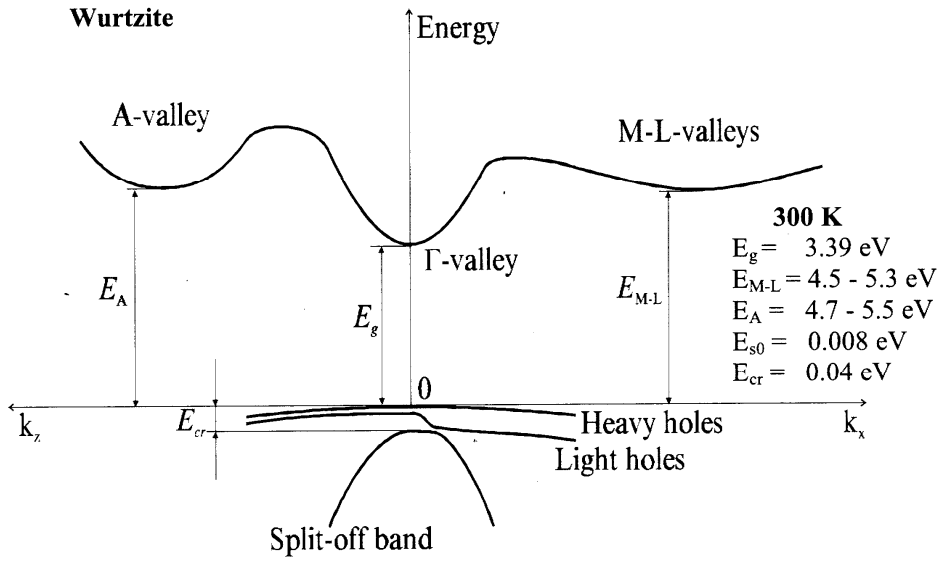


Fig. [1.1] The band structure of wurtzite GaN (from [Suzuki95])

The wide gap of GaN makes it a very attractive material for optoelectronic applications spanning the whole visible range of electromagnetic radiation. Yet, the historical development of GaN-based

devices was fairly slow, because of some material problems which will be addressed in the following. The first blue light emission was achieved in 1971 by Pankove et al. [Pankove71] in a metal-insulator-semiconductor (MIS) diode. Several examples of dim *p*-*n* junction LEDs are attested in the late 80s [Akasaki91], before the breakthrough of bright violet and blue LEDs and blue laser diodes (LDs) developed by Nakamura at Nichia Corporation [NakamuraFasol]. The main issue that slowed down the research on GaN devices was the extreme difficulty in obtaining p-type GaN, whereas nominally undoped GaN is *n*-type due to defects introduced during the growth. The *p*-type doping is achieved with Mg impurities. The presence of H atoms in the material tends to passivate the Mg acceptors by forming Mg-H complexes. The role of Mg as acceptor is restored either by low energy electron irradiation (LEEBI) [Amano89] or by thermal annealing [Nakamura92], which cause the dissociation of H from Mg. However, Mg is not a very shallow acceptor, its ionization energy being quite high, about 230 meV, so that only about 1-2% of Mg is ionized at room temperature. Other acceptor impurities have even higher ionization energies [DenBaars97]. The *n*-type doping, on the other hand, is quite easily achieved by the introduction of Si impurities, which substitute Ga atoms and acquire a donor configuration. Selected physical properties of GaN are listed in table 1.1.

1.1.2. Epitaxial growth

Another drawback of GaN, which significantly slowed down the development of GaN-based devices is the realization of bulk crystals. Bulk growth, in fact, is possible under disadvantageous conditions, such as 15 kbar pressure and 2000°C temperature, as reported from the Unipress group [Litwin99]. Therefore, GaN is usually grown heteroepitaxially, i.e. on substrates from other materials, which have a certain degree of lattice mismatch. The most common substrates are sapphire and SiC. Sapphire has a higher lattice mismatch (16%) than SiC (3.5%), but very good quality GaN films can be grown on it by using an AlN nucleation buffer layer [Hangleiter03]. A nucleation layer consists in a three-dimensional deposition of AlN on the Al_2O_3 substrate. In this way, islands forms which are then coalesced when GaN growth begins. This reduces the defectivity of the GaN film, but does not eliminate grain boundaries and dislocations: these latter can be divided into misfit dislocations, which are mostly confined close to the GaN-substrate interface, and threading dislocations, which extend in the crystal growth direction. Figure [1.2] shows the comparison between a secondary electron and an EBIC micrograph performed by means of a Schottky barrier deposited on the cross section of an n-type GaN film grown on sapphire and subsequently detached from the substrate. The secondary electron micrograph evidences only some linear defects on the lateral surface of the film. In the EBIC image, darker regions correspond to an enhanced recombination activity of the electron-hole pairs generated by the electron beam. It can be seen that the darkest regions are confined on the substrate side of the sample, where misfit dislocations are placed. However, strongly recombining regions extend throughout the layer.

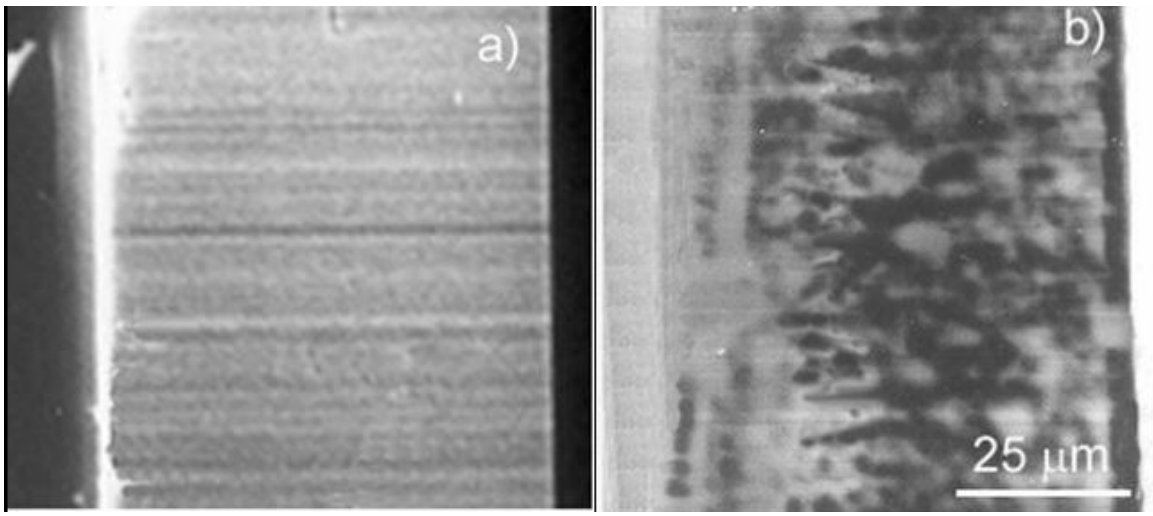


Fig. [1.2] Extended defects in GaN grown on sapphire substrate, the substrate being on the right hand side: a) Secondary electron and b) EBIC micrograph of the cross section of a GaN film.

SiC can be also used as a substrate: with respect to sapphire, its advantages are the lower lattice mismatch, the higher thermal conductivity and its adjustable doping (whereas sapphire is invariably insulating). Its main drawback is the higher cost, which is however expected to drop with the increase of SiC production connected to the growth of WBG semiconductors market. SiC is still lattice mismatched, which translates into the presence of misfit and threading dislocations in GaN. Typical threading dislocation densities in GaN are in the range 10^6-10^8 cm^{-2}.

The most used epitaxial growth techniques are either metalorganic vapour phase epitaxy (MOVPE) (also called "metalorganic chemical vapour deposition" – MOCVD) or molecular beam epitaxy (MBE). MOVPE is the leading growth technique because of its large-scale manufacturing potential. It consists in the transport of vapour species in a reactor, where they undergo a reaction yielding the deposition of GaN on the film surface. The basic reaction describing the GaN deposition process is the following:

$$Ga(CH_3)_{3(v)} + NH_{3(v)} \rightarrow GaN_{(s)} + 3CH_{4(v)} \qquad (1.1)$$

where (v,s)=vapour, solid. This expression ignores the specific reaction path and the reactive species. The growth technique is typically implemented by empirically studying the effect of external parameters (such as temperature, III/V ratio, substrate tilt, mass flow rates) on the crystal quality [DenBaars97].

MBE is an Ultra-High-Vacuum (UHV)-based technique for producing high quality epitaxial structures with monolayer control. MBE growth relies essentially on atoms, clusters of atoms, or molecules, which are produced by heating up a source. They then migrate in an UHV environment and impinge on a hot substrate surface, where they can diffuse and eventually incorporate into the growing film. Despite the conceptual simplicity, a great technological effort is required to produce systems that yield the desired crystal quality, material purity, composition, thickness, doping uniformity, and interface control, so that this technique is less viable from a large-scale point of view [Davies]. An advantage of MBE nitride growth is the low temperature that can be achieved due to the atomic nitrogen source, resulting in lower thermal stress upon cooling and less diffusion [DenBaars97]. Both epitaxial techniques are used to produce nitride-based heterostructure by alloying GaN with InN and AlN. Devices such as LEDs, detectors, lasers, modulators and devices containing distributed Bragg reflectors (DBRs) are fabricated with this technique [Moustakas01] This will be discussed in a later section.

1.2. GaN-related alloys

Typical applications of III-V semiconductors rely on heterostructures, which utilize various combinations of ternary alloys of the binary materials, which span the whole range of bandgaps illustrated in fig. [1.3]. In principle, the whole area delimited by the two solid lines and by the dashed line is exploitable to form an alloy with desired bandgap and lattice constant, but in practice it is common to grow ternary alloys made up of mixtures of binary materials, such as GaN/InN or GaN/AlN. The obtained alloys are then briefly indicated with the notation $In_xGa_{1-x}N$ and $Al_yGa_{1-y}N$, where x and y indicate the molar fraction of InN and AlN, respectively. Nitride based alloys span an extremely large range of energy gaps, from 0.77 eV to 6.1 eV, thereby covering the near-infrared, the visible and the UV parts of the electromagnetic spectrum. Due to the large bond strength of the nitrogen bonds, the lattice constants in the basal plane are much smaller than in the case of GaAs and other III-V materials [Hangleiter03].

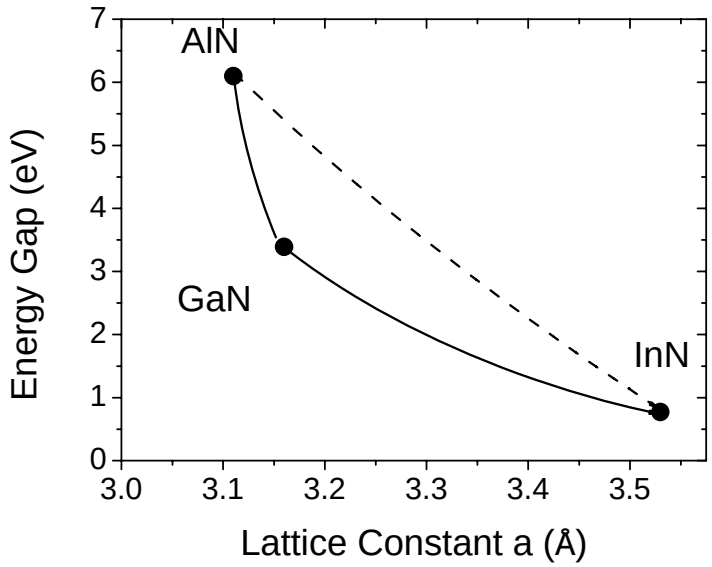


Fig. [1.3]. Fundamental bandgap versus basal-plane lattice constant of III-N materials at T=300K. The solid curves represent the ternary mixtures AlN/GaN and GaN/InN. The dotted curve qualitatively indicates the AlN/InN alloy.

1.2.1. Indium Gallium Nitride (InGaN)

$In_xGa_{1-x}N$ is the way to commonly indicate the ternary alloy obtained by mixing together InN and GaN, where x is the molar fraction of InN. $In_xGa_{1-x}N$ can be epitaxially grown by different techniques, such as the aforementioned MOVPE and MBE, provided the species participating to the reaction are correspondingly modified, for instance with the introduction of a certain amount of trimethilindium ($In(CH_3)_3$) vapour in the MOVPE reactor. The bandgap of this ternary alloy spans all the visible range and some part of the near infrared and near UV. It is therefore widely used in the production of optoelectronic devices operating in the visible range: as we will later describe, InGaN is the material found in the active regions of most nitride-based LEDs based on the Quantum Well emission mechanism. Selected physical properties of InN are listed in table 1.1. The properties of the $In_xGa_{1-x}N$ alloy are usually determined by interpolation of the properties of the binary materials, in the following way

$$A_x = A_{GaN}(1-x) + A_{InN}x + Bx(1-x) \qquad (1.2)$$

where A_x is any physical property of the alloy characterized by a x fraction of InN, A_{GaN} is the property for GaN, A_{InN} is the property for InN, and B is the so-called bowing parameter, describing the deviations from the linearity of the interpolation scheme up to the second order in the InN fraction x. The most important interpolation formula regards the bandgap of the alloy:

$$E_G(x) = 3.42(1-x) + 0.77x - 1.43x(1-x), \qquad (1.3)$$

as given by Wu et al [Wu02b]. It is interesting to notice that the bandgap of InN has been determined as 0.77 eV very recently through the work of two groups, one based at the Ioffe Institute of St. Petersburg, Russia [Davydov02], and the other based at the Lawrence Berkeley National Laboratory, U.S.A. [Wu02a]. The previously accepted bandgap of InN was about 1.9 eV [Teisseyre94], what also led to larger bowing parameters [ODonnell99a]. The previously accepted value was higher because in the early 90s the quality of the InN crystals was rather poor, and the intrinsic defects present, being of donor nature, made the material of *degenerate* *n*-type, i.e. with the Fermi level lying inside the conduction band. As the lowest energy levels of the conduction band were occupied, they could not be involved in the absorption of photons, what made the observation of the real band gap impossible (Burstein-Moss shift). With the improvements in the growth technique, better crystal quality was achieved, enabling thus to the observation of the fundamental band gap in non-degenerate *n*-type material. InN is, like GaN, *n*-type even without donor impurities because of the presence of intrinsic donor levels. So far, no successful *p*-type doping of InN has been attained.

This is not the only difficulty related to the $In_xGa_{1-x}N$ ternary alloy. The large ionic radius of In atoms makes them difficult to incorporate into GaN layers. It is generally accepted that $In_xGa_{1-x}N$ may be subjected to phase separation if not grown under optimum growth condition. For example, in MOVPE growth of $In_xGa_{1-x}N$ layers, the typical growth temperature is about 300 °C below the optimum growth temperature of GaN in order to incorporate enough In [Hangleiter03]. In many cases light emission from $In_xGa_{1-x}N$ layers has been explained in terms of phase separation with the formation of In rich quantum dots inside a matrix of In poor material [ODonnell99b], [Shapiro00].

If $In_xGa_{1-x}N$ is grown on GaN, the lattice mismatch is significant and the alloy layer is subjected to considerable pseudomorphic strain, i.e. the lattice constant in the basal plane of the alloy is compressed in order to accommodate to the lattice constant of GaN. This situation holds on up to a so-called critical thickness of the pseudomorphic layer. When the layer exceeds the critical thickness, relaxation phenomena occur, leading to the formation of a misfit dislocation network at the interface between the two materials and to the re-acquisition of the proper lattice constant in the previously strained layer. Surprisingly, the critical thickness in these nitride-based systems is rather large, much larger than in other III-V systems. For $In_xGa_{1-x}N$ with a fraction x=0.1 of InN grown on GaN, the elastic strain is about 1%, and pseudomorphic strained layers thicker than 100 nm can be grown before relaxation occurs [Hangleiter03], [Reed00].

1.2.2. Aluminium Gallium Nitride (AlGaN)

Mixing of AlN with GaN is not as difficult as in the previously discussed case of InN. The system AlN/GaN is relatively simple to control, because the ionic radii of Ga and Al are similar. Growth of $Al_yGa_{1-y}N$ layers is achieved either by MOVPE or by MBE, provided the fraction of the species participating to the reaction yielding the solid film is correspondingly varied. For instance, trimethilaluminum vapour ($Al(CH_3)_3$) is added to the gas flow in the MOVPE reactor. $Al_yGa_{1-y}N$ covers a wide gap interval which is completely contained in the UV part of the spectrum. This makes it an interesting material for applications in the UV optoelectronics, such as UV emitters or solar blind detectors. However, it can be used also in visible optoelectronics as non-active building block of the device [Hangleiter03]. As previously mentioned, AlN is also used as nucleation layer in the heteroepitaxial growth of GaN on sapphire substrates. Selected physical properties of AlN are listed in table 1.1.

1.2.3. Nitride-based heterostructures

Most devices based on III-nitride materials are not made of a single material layer, but of a certain number of layers of different materials. These are also called heterostructures. Heterostructure-based devices had a huge impact in the field of optoelectronics and electronics, and also represent systems allowing to learn more about the physics of electrons in the nanometre scale. The idea of heterostructure was so important for physics, that two of its pioneers were awarded the Nobel Prize [Kroemer], [Alferov]. The heterostructure technique has been developed since the 70s, and reached a mature level in the field of III-arsenides, -phosphides and –antimonides. Double-Heterostructure (DH) lasers and high electron mobility transistors (HEMTs) are some of the most important achievements in the field. With the development of the nitride technology in the 90s and with the successful p-type doping of GaN, the heterostructure scheme was considered also in the framework of III-nitrides research. Important examples of nitride-based heterostructure devices are the light-emitting-diode (LED), where light emission takes place in a very narrow plane of InGaN sandwiched between two GaN layers, and the aforementioned HEMT, where AlGaN/GaN heterostructures yield the formation of a so-called two dimensional electron gas (2DEG) at the interface, characterized by a very high mobility, which has its natural application in high frequency electronics. The growth of heterostructures can be achieved by any of the previously introduced epitaxial techniques, where the successful growth occurs upon a careful control of the growth parameters of each layer (temperature, pressure, gas flow, composition, etc.)

Heterostructures are grown because they offer to manipulate the behaviour of electrons and holes through band engineering. In fig. [1.4] the band diagrams of two different materials (A and B) are shown, where the energy gaps are $E_q^A < E_q^B$.

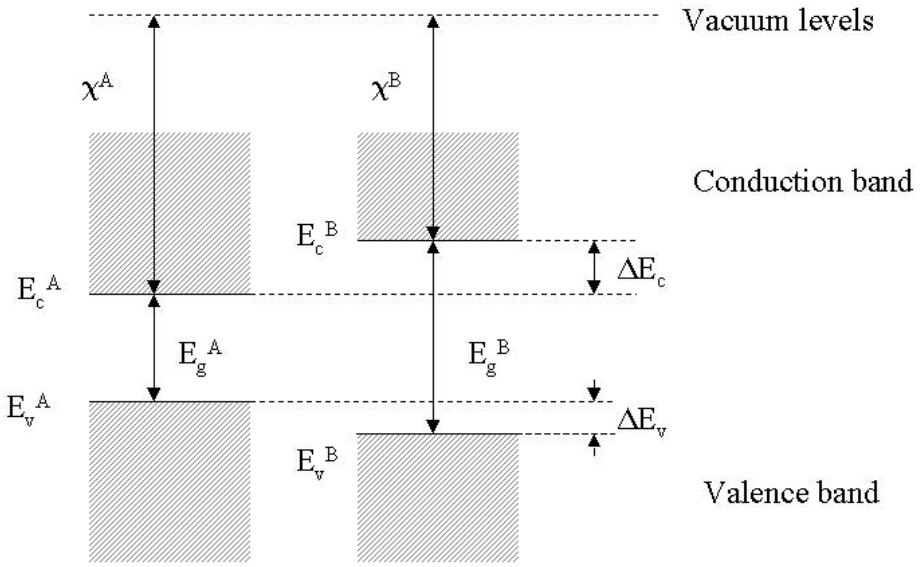


Fig. [1.4]. The alignment of the bands at a heterojunction between materials A and B following Anderson's rule, based on aligning vacuum levels.

The bands can be aligned following Anderson's rule, which is based on the electron affinity χ of the materials: this rule states that the vacuum levels of the two materials of a heterojunction should be aligned as in figure [1.4]. In the case illustrated in fig. [1.4], the narrower gap is enclosed within the wider bandgap: this is called a *type I* heterojunction, or, equivalently, a *straddling* alignment. In this case, both electrons and holes from material A must overcome a barrier in order to reach material B. There are other two possibilities, i.e. the *type II* or *staggered* alignment, where electrons (holes) from material A experience a barrier to material B but holes (electrons) from material A experience an energy drop to material B, and the *type III* or *broken gap* alignment, where the energy

gaps of A and B do not overlap [Davies]. In the present thesis, we will only consider type I heterojunctions.

1.2.4. Quantum wells

When a layer of material A with narrower gap is grown on the top of a material B with wider gap, and subsequently another layer of material B is grown, a so-called quantum well (QW) is formed, provided the alignment is of the straddling type. The situation is depicted in fig. [1.5].

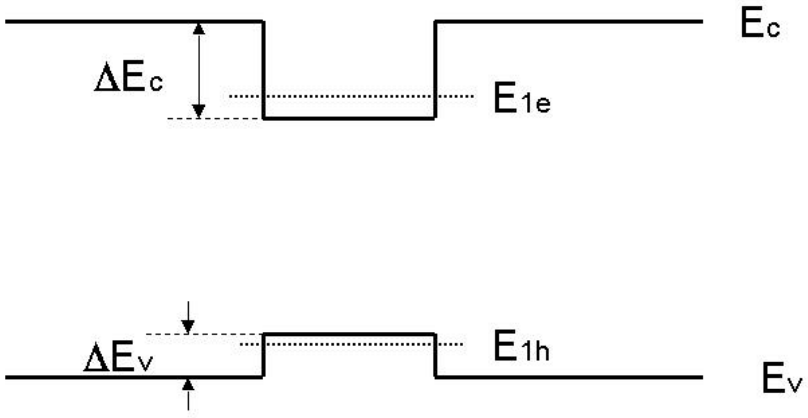


Fig. [1.5]. Formation of a quantum well (QW) from two type I heterojunctions.

In this case the well traps both electrons and holes, which are arranged in confined subbands. The energy levels of electrons and holes (in figure [1.5] only the ground state levels for electrons and holes are indicated) can be calculated in the framework of the effective mass approximation theory [Davies]. According to this theory, the total electron wave function is the product of the Bloch wave function solution of the local extremum of the host crystal energy band $\psi_{n0}(z)$ (virtual crystal approximation) and a slowly varying envelope function $\chi(z)$. The Bloch functions in the two materials must be similar for the effective-mass approximation to be valid, which yields the condition that they must belong to the same point in the $E_n(\boldsymbol{K})$ relation. This is equivalent with the statement that the conduction band bottom and the valence band top must be at the same point of the Brillouin zone (typically Γ) in both materials. Then, there are particular conditions for the matching of the wave function at the interfaces. If we consider a simple case, i.e. a junction at z=0 between neutral regions of materials A and B , we have the following Schrödinger equation for the envelope function (considering only one dimension, i.e. the growth direction):

$$\left(E_C^A - \frac{\hbar^2}{2m_A m_0}\frac{d^2}{dz^2}\right)\chi(z) = E\chi(z), \qquad (1.4)$$

$$\left(E_C^B - \frac{\hbar^2}{2m_B m_0}\frac{d^2}{dz^2}\right)\chi(z) = E\chi(z), \qquad (1.5)$$

where $m_{A,B}$ are the effective masses of electrons in the conduction band of material A and B, respectively. The difference in the bottoms of the conduction band must be treated as a step potential. If materials were the same the matching conditions would be the usual:

$$\chi(0_A) = \chi(0_B), \qquad \left.\frac{d\chi(z)}{dz}\right|_{z=0_A} = \left.\frac{d\chi(z)}{dz}\right|_{z=0_B}, \qquad (1.6)$$

where $0_{A,B}$ mean the side of the interface in material A and B, respectively. However, the materials are different and the matching conditions must be modified as

$$\chi(0_A) = \chi(0_B), \qquad \frac{1}{m_A}\frac{d\chi(z)}{dz}\bigg|_{z=0_A} = \frac{1}{m_B}\frac{d\chi(z)}{dz}\bigg|_{z=0_B} \tag{1.7}$$

in order to satisfy the conservation of current and the orthogonality of eigenfunctions relative to different energy levels [Davies].

Once the potential and the band offsets are known, effective mass theory can be used to find the subband energy levels in the QW. There are various numerical methods for this calculation, with different degree of complexity. Often, the calculation must be performed by means of a self-consistent algorithm, because the trapped charges influence the potential. Therefore, one usually speaks of self-consistent Schrödinger-Poisson algorithms.

Quantum wells are a systems extending in three dimensions, but the arrangement of the electrons can be described as though it is two-dimensional. Starting from the three-dimensional Schrödinger equation

$$\left[-\frac{\hbar^2}{2m}\nabla^2 + V(\vec{R})\right]\psi(\vec{R}) = E\psi(\vec{R}), \tag{1.8}$$

one can write, taking into account that the potential depends only on coordinate z, the wave function as a product of x, y and z components:

$$\psi(\vec{R}) = \exp[ik_x x]\exp[ik_y y]\chi(z); \tag{1.9}$$

and the energy as

$$\varepsilon = E - \frac{\hbar^2}{2m}\left(k_x^2 + k_y^2\right), \tag{1.10}$$

thus obtaining a 1-dimensional equation for the z-component:

$$\left[-\frac{\hbar^2}{2m}\frac{\partial^2}{\partial z^2} + V(z)\right]\chi(z) = \varepsilon\chi(z). \tag{1.11}$$

When the eigenvalues ε_n for the z- component are found by solving the Schrödinger equation, the electron states are described through

$$\psi_{k_x,k_y,n}(x,y,z) = \exp[ik_x x]\exp[ik_y y]\chi_n(z) \tag{1.12}$$

$$E_n(k_x,k_y) = \varepsilon_n + \frac{\hbar^2}{2m}\left(k_x^2 + k_y^2\right). \tag{1.13}$$

For a fixed value of n (*subband index*) the dispersion relation between energy and wave vector is that of a 2-dimensional electron with the bottom of the band shifted by ε_n. The relation is sketched in fig. [1.6] for a QW with the bottom of the conduction band slightly tilted by an electric field.

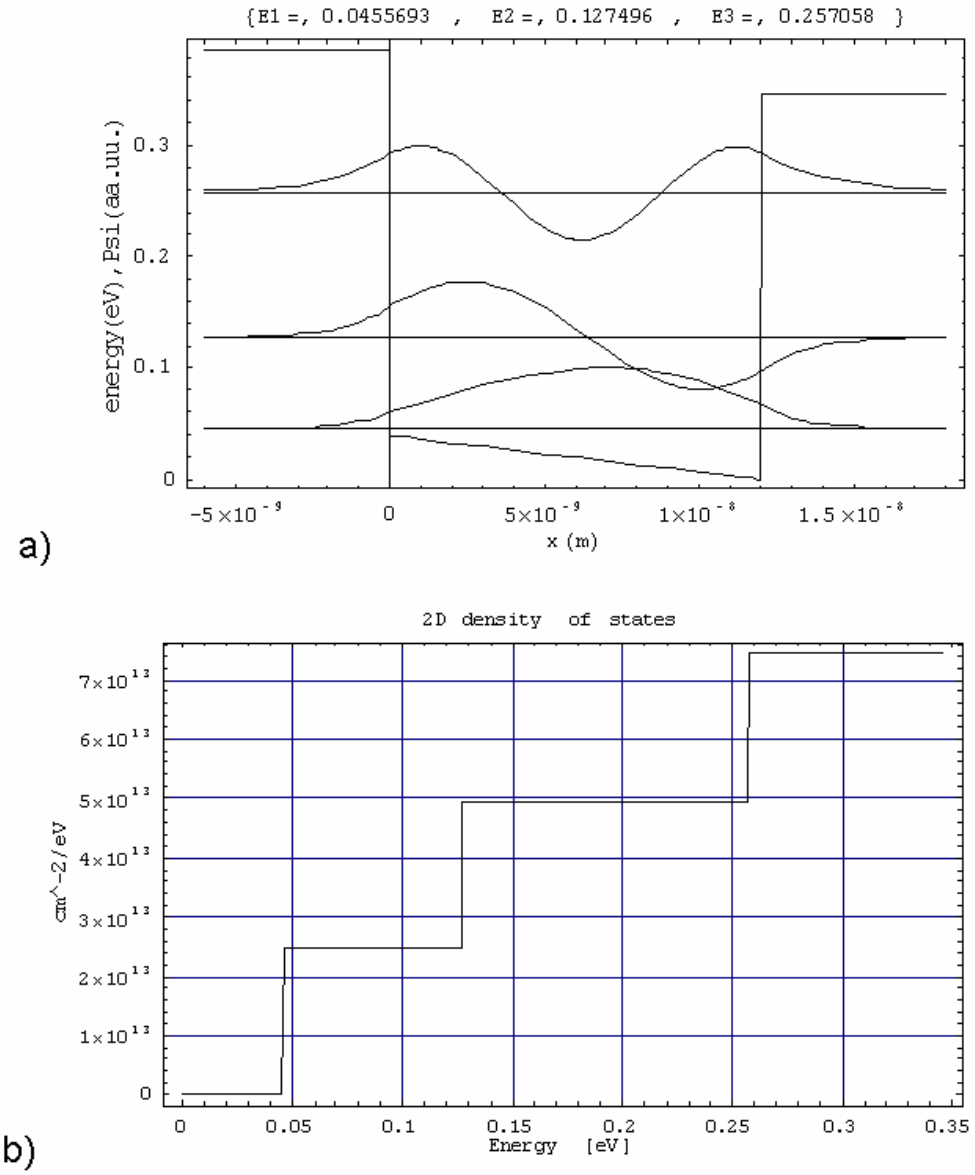


Fig. [1.6]. The solution of the Schrödinger equation for a particular QW configuration: a) potential, energy levels and envelope wavefunctions on (z-component); b) the 2-dimensional density of states.

The subbands determine the shape of the density of states *n(E)*, which is that of a 2-dimensional electron gas, a step function of height $m/\pi\hbar^2$, starting at ε_n. Each subband represents a step. The number of occupied levels can be found by integrating the product of the density of states and the Fermi-Dirac occupation function $f(E,E_F)$. One thus obtains the 2-dimensional electron density n_{2D} as the sum of the contributions from each subband n_j:

$$n_{2D} = \sum_j n_j \,, \tag{1.14}$$

$$n_j = \frac{m}{\pi\hbar^2}\int_{\varepsilon_j}^{\infty} f(E,E_F)dE = \frac{mk_BT}{\pi\hbar^2}\ln\left[1+\exp\left(\frac{E_F-\varepsilon_j}{k_BT}\right)\right]. \tag{1.15}$$

The same considerations apply to holes confined in the valence band QW [Davies]. The density of states is a key parameter for the calculation of absorption and emission properties of 2-dimensional electron gases, as it will be explained in the next chapters.

1.2.5. Piezoelectric fields in nitride-based heterostructures

Heterostructures based on III-N materials present peculiar characteristics, which are not found in other III-V systems, and that are due to the strain-induced polarization fields arising from the charge density originating at the interface between two III-N materials. Both AlN/GaN and GaN/InN systems are subject to considerable pseudomorphic strain, which is more pronounced in the latter case. This large strain has important consequences. III-nitrides have strongly polar bonds, and are highly piezoelectric materials. In $In_xGa_{1-x}N$/GaN heterostructures, piezoelectric fields of the order of 2 MV/cm for x=0.2. Moreover, the wurtzite structure allows not only piezoelectric polarization but also spontaneous polarization, because the bond lengths are different for different directions [Bernardini97]. Macroscopic fields can be screened by surface charges, but microscopic fields in heterostructures have a strong influence on the optoelectronic characteristics of the system.

One of the important consequences of the piezoelectric fields is its influence on the so-called quantum confined Stark effect (QCSE). This effect, first introduced to explain the modification in the optical absorption characteristics of InGaAs/GaAs quantum well systems under the application of an external bias across the QW [Miller84], [Miller85], consist in the decrease of the energy edge for absorption of photons by a QW system and in the diminution of the absorption strength, because confined electron and hole levels come closer in energy but their envelope functions overlap less and less when an external bias is applied. III-N system exhibit a somewhat different behaviour, because an electric field across the QW is already present even when no external bias is applied. The application of an external bias, depending on the sign of the field, can enhance the band tilt or restore the flat-band condition [Franssen04], [Renner02].

1.3. Silicon Carbide (SiC)

Silicon Carbide (SiC) is a material in which C and Si atoms form bonds which have 10% ionic character and 90% covalent character, where it is C to acquire a slightly negative charge. The lattice constant *a* in the basal plane amounts to about 3.1 Å. The crystal structure is similar to that of crystalline Si or diamond, where each atom forms bonds with the nearest neighbours by means of sp^3 orbitals.

1.3.1. SiC polytypes

While the crystal structure is unique in Si and diamond, where only one atomic species is present, the difference between the two atomic species Si and C makes possible the phenomenon called *polytypism*. Polytypism is due to the different sequences into which the lattice planes are arranged. The polytypes are indicated through the so-called Ramsdell notation, consisting in the integer number corresponding to the number of planes in a sequence, and in a letter (C, R or H) corresponding to the associated Bravais lattice, which can be cubic, rhombohedral or hexagonal. More than 150 SiC polytypes are known, but few are interesting for electronics applications. The sequences of lattice planes characterizing the most important SiC polytypes are depicted in the scheme of fig. [1.7].

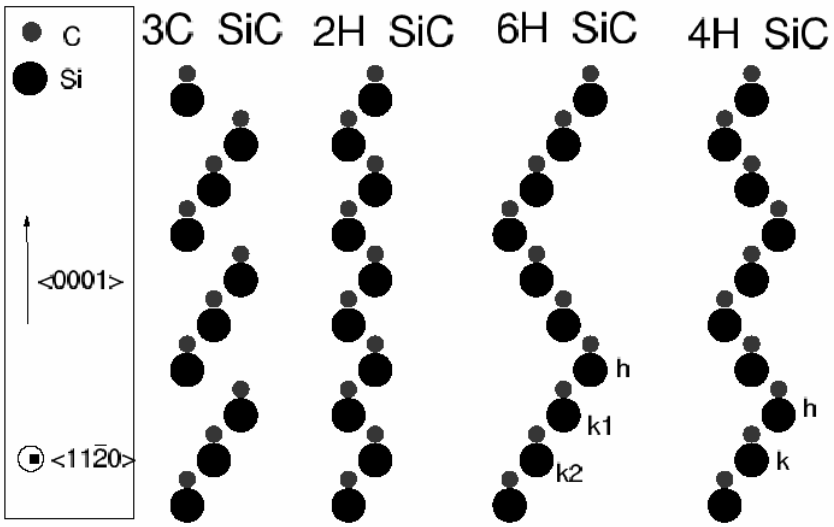


Fig. [1.7]. The sequences of lattice planes characterizing the four most important SiC polytypes.

In 3C-SiC the ABC sequence gives rise to a cubic zincblend structure; in 2H, 4H and 6H the AB, ABAC and ABCACB sequences give rise to the hexagonal wurtzite structure. One important consequence of polytypism is that the number of atoms per unit cell is different according to the number of lattice planes of the sequence. Another consequence is the presence of the so-called *inequivalent lattice sites*. In fig. [1.7] several inequivalent sites are indicated. The points labelled by h represent lattice points surrounded by a quasi-hexagonal environment, and the points labelled by k lattice points where the surrounding environment is quasi-cubic. A substitutional impurity, say N, is *feels* the different environments in terms of binding and ionization energy. It is a matter of fact that polytypes with more than one inequivalent site have more than one donor/acceptor level for the same doping species.

Polytypes are also distinguished by other important physical properties, such as the bandgap (ranging from 2.4 eV in 3C to 3.3 eV in 2H) and the electron mobility (highest in 4H, about 800 cm^2/Vs). Another factor distinguishing the polytypes is the difficulty with which their growth is controlled: inclusions of other polytypes in a host polytype are generally unwanted. So far, the most successful polytype for electronics applications has been 4H-SiC, because of the good quality of crystal achievable and of the relatively high electron mobility. In the following, we will shortly revise the properties of this polytype, of which consist the devices that will be studied in this thesis.

1.3.2. 4H-SiC – physical properties

Selected physical properties of 4H-SiC are listed in table 1.1. 4H-SiC is, like other SiC polytypes, an indirect-bandgap material, which makes it unsuitable for optoelectronics applications as active material (see fig. [1.8]). However, as already mentioned, it is often used as a substrate for III-nitride optoelectronic devices. The main applications of 4H-SiC are in the framework of high-temperature, high-frequency and high-radiation level operation. This is due to some of its properties, which we will briefly introduce.

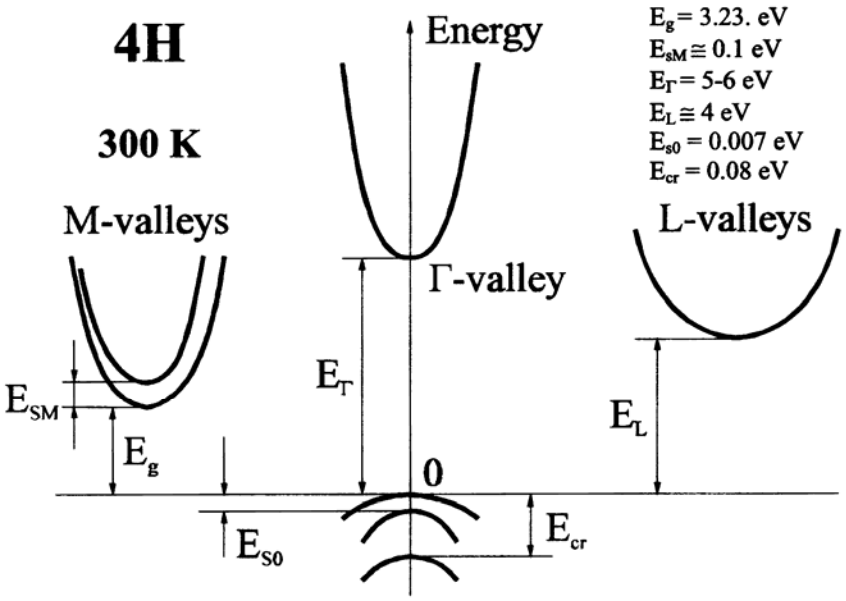


Fig. [1.8] Important minima of the conduction band and maxima of the valence band at T=300K (from [Persson97]).

As a wide bandgap semiconductor, 4H-SiC has a very low intrinsic carrier density, which is given by:

$$n_i = \sqrt{N_C N_V}\,\exp[-E_g / 2k_B T], \qquad (1.16)$$

where T is the temperature, k_B is the Boltzmann constant, E_g is the energy bandgap and $N_{C,V}$ are the equivalent densities of states for conduction and valence band, respectively. At 300 K 4H SiC has a vanishing intrinsic carrier density, of the order of about 10^{-7} cm^{-3}. Si, in comparison, has an intrinsic carrier density of about 10^{10} cm^{-3} at room temperature. When temperature increases, the intrinsic carrier density of Si becomes comparable with the doping density, while the intrinsic carrier density of SiC remains very low. This means that a Si-base device can undergo a failure due to high temperature in the same conditions in which a SiC-based device still works properly. As an example, at 600°C the intrinsic density of Si is over 10^{15}cm^{-3}, while that of SiC is under 10^{7} cm^{-3} [Neudeck02].

Another peculiarity making SiC a good choice for high temperature application is the thermal conductivity, which is very high (3.7 W/cm s) compared with that of Si and GaN (see table 1.1). High power devices and high frequency devices, as well as devices operating at high temperature, need to dissipate very efficiently the heat generated by the recombination and scattering phenomena involving charge carriers. This contributes to maintaining the temperature of the active device region as low as possible, therefore preventing thermally induced failure mechanisms. This is also the reason why in III-nitride technology SiC is preferred as a substrate material, if compared with sapphire, which has a lower thermal conductivity [Neudeck02].

The high saturated velocity of electrons in 4H-SiC is another property which should be mentioned. This parameter, corresponding to the velocity of electrons travelling in high electric fields, is a key parameter for field-effect transistors (FETs) based on this material, as it basically controls the transistor gain [Zolper05].

1.3.3. Substrate growth

The production of SiC substrates is a key issue in the technology related to this material and, more in general, to WBG technology. Like in the case of GaN, substrate growth is more difficult than in the case of materials like Si or GaAs; however, unlike GaN, substrate growth can be

achieved on large-scale facilities, and the crystal quality which has been reached is fairly good. Single crystal bulk SiC presents, in fact, a much lower dislocation density than GaN.

Large diameter SiC crystals are generally grown by sublimation rather than melt growth, as depicted in fig. [1.9]. Sublimation growth of SiC has been developed over the past 30 years, and typically employs a furnace in which graphite components constitute the hot zone. The SiC source sublimes in an inert ambient, the subliming species, including Si, SiC_2 and Si_2C, migrate down a temperature gradient and eventually deposits on a monocrystalline seed of desired orientation. Presently, six companies[1] offer single-crystal SiC substrates. The largest wafer diameter available is 3 in., and work is progressing towards 100 mm diameter. Meanwhile, the price of substrates has dropped from \$650/cm^2 in 1994 to \$16/cm^2 in 2004. Each step increasing the diameter of the bulk wafer is critical: the density of grain boundaries at the wafer periphery initially increases because of a lack of understanding of the new growth conditions. Then, new efforts are made for the optimization of all parameters governing the growth of the larger wafer [Sumakeris05].

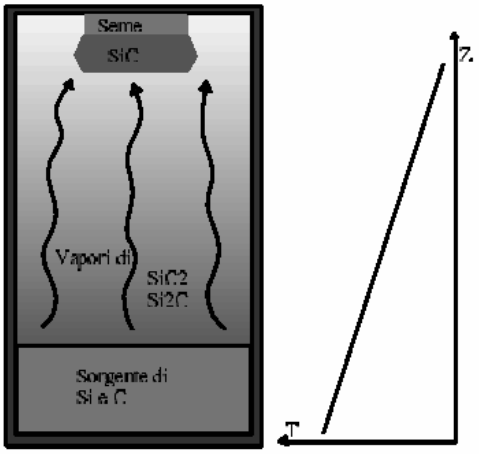


Fig. [1.9]. Schematic illustration of a silicon carbide vapour transport crucible.

1.3.4. Epitaxial growth

The usual procedure for the production of SiC-based devices is to exploit the superior quality of epitaxial layers for the active regions of the devices, although it is possible to produce some devices directly from the substrates. Each device type may have specific requirements, and a single epitaxial process may not be optimal for all applications. Usually, one can divide the epitaxial chemical vapour deposition (CVD) processes into two main categories, i.e. a) the warm-wall radial flow growth platform and b) the horizontal hot-wall epitaxial platform.

The warm-wall platform has the advantage of being able to process a large number of wafers, and it is scalable for larger wafer diameters: therefore, it is well suited to process large volumes of wafers with tight tolerances. The hot-wall platform, with planetary wafer rotation during growth, permits superior thickness and doping uniformity and minimal variation between different wafers.

One of the main concerns of epitaxial growers is the reduction of the defectivity of the material, which must be significantly lowered in the epitaxial layer with respect to the substrate. Examples of defect reduction during epitaxial growth are, for instance, the dissociation of highly harmful micropipes coming from the substrate into several less electrically active threading screw dislocations (TSDs), or the conversion of basal plane dislocations (BPDs) into more benign threading edge dislocations (TEDs) [Sumakeris05].

[1] The companies producing SiC substrates are Cree Inc., II-VI Inc., Intrinsic Semiconductors, Dow Corning Compound Semiconductor Solutions, SiCrystal AG, Sixon and Norstel AB.

1.4. Defects in wide bandgap semiconductors

A defect in a semiconductor crystal is, in general, a lack of crystal perfection in a certain region of the material. This region can be very small, for instance when an atom of one species is missing from the lattice of a compound semiconductor, or when a crystal site is occupied by in impurity atom. In other cases, the region can be large, sometimes extending throughout the whole thickness of the material. Moreover, the crystal structure usually considered for the deduction of many electrical and structural properties of the material is infinite, whereas every real semiconductor has surfaces. Surfaces can be viewed therefore as a huge defect. Some properties of the semiconducting material can be explained with the perfect crystal despite the presence of defects in the real material, but some other properties are significantly altered by the presence of defects even in very minute amounts. This represented a problem at the beginning of the study of semiconducting materials, because the amount of defects could not be well controlled, thus leading to a puzzling scattering in the experimentally observed properties [Queisser98].

Important steps forward were made when better growth methods led to the preparation of highly pure and structurally perfect crystals, and when processes able to control the amount of certain defects were discovered. This led to the invention of the transistor in 1947 and to its implementation by doping its regions with properly chosen impurity atoms conferring a specific electronic behaviour.

Doping of semiconductors is indeed a deliberate introduction of defects (impurity atoms) in the material. These defects determine the electrical properties of the region where they reside: donor impurities create a surplus of free negative charge in the conduction band (*n*-type material), acceptor impurities a surplus of free positive charge in the valence bands (*p*-type material). Thus, they make possible the formation of *p*-*n* junctions, which are a fundamental building block of many electronics applications.

The aforementioned is a case in which the presence of the defects is explicitly wanted, in order to have the semiconducting material to work as a device, but in most cases the presence of defects is unwanted, because many of them are electrically active and perturb the operation of the device: dislocations, for instance, or impurity levels lying deep inside the gap provide a non-radiative recombination path for electrons and holes: this can be harmful in optoelectronic devices, because these work on the principle that electron and holes recombine radiatively, i.e. emitting a photon. Recombination centres, however, could be desirable in certain devices, i.e. switching devices where the enhanced recombination of electron-holes pairs leads to a diminution of the dissipated power and to the possibility of operation at higher frequency.

These are only few examples indicating how various and how important for the technology is the study of defects in semiconductors. According to E. R. Weber [Weber03], the understanding of defects in semiconductors should be articulated in four steps:

- *Defect observation and characterization*: this is the first step, usually involving different kinds of spectroscopy, such as deep level transient spectroscopy (DLTS), which basically assess the presence of defects.

- *Defect identification*: once the presence of the defect has been ascertained, combination of experimental and theoretical methods is required to assess the microscopic nature of the defect.

- *Defect control*: this step involves the study of the thermal stability of the defect and the study of its introduction characteristics.

- *Influence of defects on device performance*: once it is known how to control the amount of defects present in the material, an optimum defect level can be identified according to the applications of the device under analysis.

The work done in this thesis is also articulated over these four points, as we will assess the presence of defects, draw (in some cases) conclusions about their microscopic nature, study their introduction and stability, and report on their influence on device performance in the framework of wide bandgap semiconductor electronics. Meanwhile, in the following subsections we will revise the properties of some important categories of defects.

1.4.1. Point defects

Point defects extend on spatial regions of the order of several times the lattice constant of the semiconducting material where they reside. These defects can be either intrinsic or impurity-related. *Intrinsic* (or *native*) defects consist in an erroneous arrangement of lattice atoms of the semiconducting material, while *impurity-related* defects originate from the presence of atom species which are not constituents of the semiconducting material. Intrinsic and impurity-related defects can also combine giving rise to defect complexes. In the following we will deal only with binary compound semiconductors, composed by atoms of species A and B; the materials will be shortly denoted as AB. Some of the most important intrinsic defects are

- The *vacancy* $V_{A,B}$, consisting in a missing atom of the atomic species A or B from its lattice site
- The *interstitial* $I_{A,B}$, consisting in an excess atom acquiring a position between lattice sites
- The *antisite* A_B, occurring when the lattice site of atom B is occupied by an atom of the A species

Vacancies and interstitials can constitute a so-called Frenkel pair V_A+I_A. In this case, if the defects composing the pair are not far from each other, they can recombine to yield the perfect crystal configuration. Intrinsic defects can also combine among themselves to form intrinsic complexes.

Impurities in a crystal can acquire different configurations. Denoting the impurity atom with X, one has:

- A *substitutional impurity* $X_{A,B}$, i.e. an impurity atom occupying the site reserved for species A or B in the crystal lattice
- An *interstitial impurity* I_X, occupying some other region in a lattice already occupied by A and B atoms.

In a binary compound semiconductor, it is important to distinguish between the impurity substituting an atom of A species or an atom of B species. In GaN, for instance, Si can substitute Ga or N. If it substitutes Ga the bonds formed have an extra electron which can easily leave the Si atom, thus becoming a free electron in the conduction band. Therefore, Si in a Ga site acts as a donor impurity. On the other hand, if Si occupies a N site, it easily attracts electrons to complete its Lewis octet, therefore generating free holes in the valence band, i.e., acting as an acceptor. In other terms, one says that Si is an *amphoteric* impurity, as it can act both as donor and as acceptor. Amphoteric behaviour is not only found for impurities, but also for charged intrinsic defects: as an

example in GaAs, the positively charged defect V_{As} can evolve into the negatively charged defect $Ga_{As}+V_{Ga}$ [Walukiewicz89], [Baraff85].

1.4.2. Deep and shallow defects

In order to better introduce the issues related to defects in semiconductors, we will here shortly revise what is the electronic fingerprint of the defects, i.e., what is their influence on the electronic properties of the material. The bandgap of a semiconductor (and of an insulator) is an energy region free of electron states. Impurities and intrinsic defects can significantly modify the electrical properties of the material as they can introduce electron states in the gap. The electrical characteristics of defects are mainly given by the position of their electron states in the energy gap. Shallow defects are called those whose levels lie few tens of millielectronvolts from the respective band edges, while deep defects have electron states lying further from the bands. As will be explained more in depth in chapter 4, shallow states resemble hydrogenic atoms, with energy levels and Bohr radii correspondingly scaled due to the relative dielectric constant of the material and to the effective electron mass. Shallow levels are usually related to substitutional impurities, but this role is sometimes played even by intrinsic defects, as in the case of the nitrogen vacancy V_N in GaN, likely acting as hydrogenic donor level [Yang02].

The deep defects, which will be dealt with extensively in chapter 4, are given by impurity atoms or intrinsic defects which badly fit in the lattice and distort it significantly. Their electron levels lie deep in the bandgap, the wavefunction associated to the electron states in the gap is highly localized and the lattice is distorted over many lattice constants. Deep defects yield a strong reduction of the lifetime of minority carriers, as they represent efficient recombination paths with majority carriers, and they can trap majority carrier from the relative band. In chapter 4 we will also describe, dealing with the DLTS technique, how deep levels can be characterized through their emission properties of trapped carriers.

1.4.3. Extended defects: dislocations

Extended defects, as the expression suggests, have a larger extension in the crystal than point defects. Extended defects can be well described by 1, 2 or 3-dimensional models, according to the geometry of the defect. Impurity precipitates and clusters, for instance, can be considered a 3-dimensional extended defect, whereas a stacking fault or an oxide platelet are 2-dimensional defects. 1-dimensional defects, or line defects, play a very important role in the semiconductor physics. The most important line defects are the so-called dislocations. The most simple dislocation types are the edge dislocation and the screw dislocation. The edge dislocation is schematically represented in fig. [1.10]: it can be fictitiously constructed by cutting a crystal halfway and inserting an extra half-plane of atoms. A dislocation can be described by the so-called Burgers vector. The Burgers vector is calculated as follows (see fig. [1.10]). First, one considers a closed path in a dislocation-free crystal region. This path (the Burgers' path) is characterized by a succession of steps over lattice sites. If the same sequence of steps is performed in a region surrounding a dislocation, the path is no longer closed. The further steps needed to close the path define thus the Burgers vector $\boldsymbol{b}$. The edge dislocation has $\boldsymbol{b}$ orthogonal to the dislocation line, whereas the screw dislocation has $\boldsymbol{b}$ parallel to the dislocation line. Dislocations play an important role in the determination of elastic properties of crystals and can be electrically active.

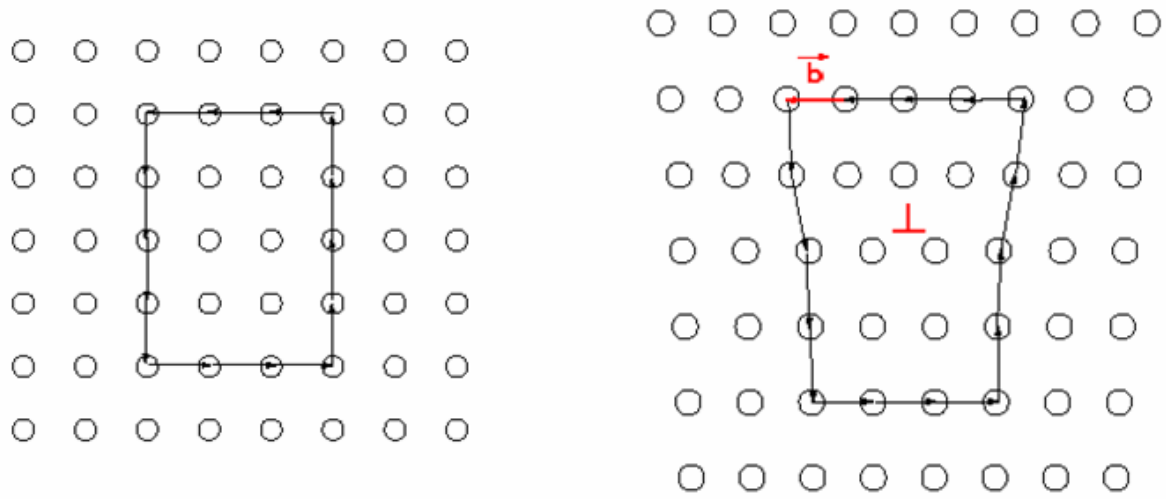


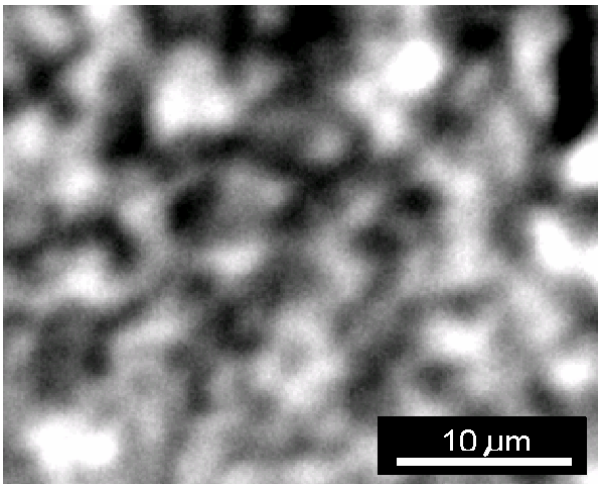
Fig. [1.10]. Schematic representation of the edge dislocation, with the definition of the Burgers path and Burgers vector.

The electrical activity of dislocations is an important parameter: similarly to point defect, dislocations may introduce electron states in the bandgap. However, there is an important difference between electron states related to point defects and to dislocations. When dislocations, for instance, trap electrons from the conduction band, the charge they acquire is not randomly distributed over the crystal, but it is organized along a line. This leads to the formation of space charge regions around dislocations (also called *Read cylinders*) which are very efficient in carrier scattering.

Scattering at dislocations can affect the electron mobility. In GaN, it is found that, due to the high dislocation density, vertical transport is favoured with respect to lateral transport; moreover, the scattering efficiency of dislocations depends on their charge, which, in turn, depends on the free carrier concentration [Ng98]. Many efforts have been made to reduce the dislocation densities both in GaN and SiC. The techniques of lateral epitaxial overgrowth (LEO) [DenBaars97] for GaN and repeated *a*-face [Nakamura04] for SiC exploit the propagation of dislocations mostly on the growth direction. Changing the growth direction, i.e. by growing epitaxially on the lateral surface of the slice of material, considerably helps in the reduction of this unwanted line defect.

Similarly to point defects, dislocations provide non-radiative recombination paths which can be harmful in optoelectronic devices, and represented one of the major degradation and failure sources in GaAs based LEDs and lasers. Unexpectedly, their role does not appear so dramatic in the case of III-nitride based optoelectronic devices. Despite huge dislocation densities, GaN-based LEDs show very efficient luminescence properties.

The recombining activity of dislocations can be characterized by different means. Their electron gap states can be characterized by means of spectroscopy techniques like DLTS, and their spatial arrangement can be evidenced by means of electron microscopy. Electron beam induced current (EBIC) is a powerful technique in the characterization of dislocations, because its micrographs are a spatially resolved image of their recombination strength [Donolato79]. In fig. [1.11] two examples of EBIC performed on the surface of a) GaN and b) SiC epitaxial layers are shown. The dark spots, very dense in GaN and more sparse in SiC are the surface terminations of threading dislocation lines coming from the bulk of the crystal. These images are also indicative of the generally higher defectivity of GaN with respect to SiC. In the present case, the threading dislocation density is about 10^8 cm^{-2} in GaN and about 10^5 cm^{-2} in SiC. Much lower levels are reached in more mature materials like GaAs and Si.

a) 

b) 

Fig. [1.11] EBIC micrographs of a) GaN and b) SiC epitaxial layers, showing two different threading dislocation densities in the two materials.

In many cases the electron states associated with a dislocation are not due to the presence of dangling bonds along the dislocation line (core), but to impurities residing in the neighbourhood of this line defect. It is possible, in fact, that the strain field surrounding a dislocation energetically favours the presence of interstitial or substitutional impurities. These impurities may then give rise to deep levels with a specific behaviour. Very detailed studies of this phenomena have been carried out in Ge, Si and GaAs materials [Figielski64], [Kveder82], [Omling85], [Schroter95] [Wosinski89], but this knowledge in the case of WBG semiconductors is not as advanced.

One of the most harmful manifestations of the dislocation defect is the so-called micropipe, a defect causing much concern among SiC growers. This defect is a hollow-core threading dislocation which can extend over the whole thickness of SiC substrates, and eventually penetrate also in the epitaxial layers. Any device built on one of such defects would invariably fail, because the micropipe yields a short-circuit between the device top and bottom. The reduction or elimination of this defect represented one of the most important efforts in SiC growth over the last decade [Sumakeris05].

1.4.4. Thermal properties of defects

Defects are entities subjected to specific thermodynamic laws. The concentration of a native defect, for instance, depends on temperature; the diffusivity of impurity atoms shows an important dependence on the concentration of intrinsic defects, as the interaction between impurity and intrinsic defect provides a path through which the impurity can move inside the crystal. The concentration of charged defects, moreover, depends on the Fermi energy, which, in turn, depends mostly on the concentration of shallow dopants in the material. As one can see, the interactions between defects are somewhat complex, and very various. In the following, we will introduce some thermal properties of defects which will be considered in the subsequent parts of the work.

Concentration of intrinsic defects.

The concentration of an intrinsic defect can be described through the expression:

$$C = C_0 \exp[-H_f / k_B T], \qquad (1.17)$$

where H_f is the formation enthalpy of the defect. If the defect is charged, as in the case of a donor-like defect which can donate *s* electrons, the formation enthalpy depends on the Fermi energy:

$$H_f = H_{f0} + \sum_{n=0}^{s-1} (E_F - E^{n+}) , \qquad (1.18)$$

where E^{n+} is the energy of the n+1/n transition state, E_F is the Fermi energy and H_{f0} is the formation enthalpy of the neutral defect. The concentration of the defect increases with increasing temperature and with decreasing Fermi energy. In terms of doping, *p*-type doping causes the shift of the Fermi energy towards the valence band top or below, therefore enhancing the formation of donor-like defects. This behaviour has far-reaching consequences: first, the donor-like defects compensate the acceptors conferring the *p* character to the material (self-compensation); secondly, the increased concentration of intrinsic defects enhances the diffusivity of impurity species. Correspondingly, acceptor-like defects are favoured in *n*-type material when the shallow donor density increases. These effects are very important in heavily doped semiconductors and devices, as the control of the doping of the active regions is crucial for proper device working [Walukiewicz94].

Diffusivity of defects.

The diffusion of defects is described through a diffusion coefficient *D*. The mechanism of acceptor impurity diffusion in a binary semiconductor is described by the so-called kick-out mechanism [Gösele81]:

$$X_i^{r+} + A \leftrightarrow I_A^{p+} + X_A^{p+} + (r - p + 1)h , \qquad (1.19)$$

evidencing the participation of interstitials. An increase in the concentration of interstitials can therefore lead to higher diffusivity of impurities. The example can be generalized to donor impurities.

Metastability.

A defect can change its geometric configuration, the transformation being accompanied by a certain energy exchange with the surrounding crystal. An example is the transformation of a substitutional impurity into an interstitial impurity. The defect tends of curse to acquire the configuration minimizing the total energy, but it can occur that the minimum-energy configuration is separated from higher-energy configurations by a potential barrier E_b. The rate of reconfiguration from one configuration to the other is therefore thermally activated:

$$R = R_0 \exp[-E_b / k_B T] . \qquad (1.20)$$

In certain cases the transformation between the two configurations is reversible, for instance by means of an electric field [Chantre86] or by optical excitation [Queisser98].

Annealing.

Annealing consists in the thermally induced modification of the concentration of defects in the material. When a material is heated, the concentration of defects tend to evolve toward the equilibrium concentration for the specific annealing temperature. If, for instance, a material at room temperature has a defect concentration exceeding the equilibrium concentration, it can be taken to higher temperature, so that the equilibrium concentration is restored. Then it can be cooled; depending on the velocity of the cooling process, the final defect concentration at room temperature will vary from the RT equilibrium concentration (for slow cooling) to the equilibrium concentration at annealing temperature (for quick cooling). Annealing is commonly used to characterize the evolution of the concentration of point defects as a function of temperature. There are two main types of annealing: isothermal and isochronal annealing.

Isothermal annealing consists in keeping the material at a certain fixed temperature and observing the temporal evolution of the parameters related to the defects. The concentration of defects follows generally the law:

$$\frac{dC}{dt} = -K \exp[-E_a / k_B T] C^n , \qquad (1.21)$$

where E_a is the activation energy for the process and n is the order of the process.

Isochronal annealing consists in keeping the material at a certain temperature for a fixed time t_{ann}. The material is then analyzed and subsequently taken to a higher temperature T_2 for the same fixed time t_{ann}. The procedure is then repeated for higher temperatures.

1.4.5. Introduction of point defects

The introduction of defects is important both in device technology and material characterization. Shallow impurities, for instance, need to be introduced in the material in order to confer to it specific electrical properties. From another point of view, defects can be introduced for the purpose of studying their (detrimental or benign) effect on the properties of the material. The ability to control the introduction of defects is therefore crucial for the study of their properties. The most usual ways by which defects are introduced are the following.

Incorporation during growth (impurities).

This technique is used during the growth of substrates or epitaxial layers of semiconducting material. the incorporated defects are usually shallow dopant impurities, which determine the type of conductivity of the material. In some cases impurities giving rise to deep levels can also be incorporated, when semi-insulating materials are needed: the presence of a significant amount of deep levels tends to pin the Fermi energy deep in the gap, thus depleting the bands from free carrier. An example of this procedure is V incorporation in SiC. The incorporation is achieved by adding appropriate amounts of the impurities species to the molecular flow in MOCVD epitaxy or to the molecular beams in MBE. This technique is not spatially resolved, due to the layer-by-layer growth mechanism.

Diffusion from the surface (impurities).

If the diffusivity of the semiconducting material is high enough, impurity species can be diffused to the bulk from the surface. This technique has the advantage of spatial resolution, as it is possible to choose a priori the surface regions where the impurities will diffuse into the bulk. Applications of this technique are in the fabrication of field-effect transistors (FETs) in Si and GaAs, where diffusion of acceptor impurity occurs in previously n-doped regions, thus creating p^+-n junctions. This process is not viable in WBG semiconductors like SiC or GaN. These have much stronger molecular bonds than Si and GaAs, resulting in a very low diffusivity for impurities. Spatially resolved doping of WBG semiconductors needs a somewhat more invasive technique.

Implantation (impurities and intrinsic defects).

Implantation is used for the spatially resolved doping of semiconductors. The spatial resolution of doping by implantation is superior to diffusion, but the technique is much more invasive. Making use of ions with energy of the order of kiloelectronvolts, it not only introduces impurity species in the material, but also creates intrinsic defects due to the interaction between impinging ions and lattice atoms. Displaced atoms can also give rise to reactions displacing other lattice atoms, what is also called an *atom cascade*. An estimate of atomic displacements N_{disp} by impinging ions can be made by means of the following formula:

$$N_{disp} = \frac{(1 - f_i)E_R}{2E_D}, \qquad (1.22)$$

where f_i is the fraction of energy lost by ionization, E_R is the energy of the impinging ion, and E_D is the mean energy for atomic displacement, which is characteristic of the material. The displaced atoms are not equally distributed along the path of the ion. The cross section for atom displacement is indeed higher for lower ion energy. A very energetic ion (some MeV) will produce a low number of intrinsic defects when just entering the material. When the ionizing and displacement interactions slow down the ion, it begins to interact more and more with the surrounding lattice, until it eventually stops. The ion ranges, i.e. the average distance travelled by the ion inside the crystal, and the distribution of introduced defects can be estimated by dedicated simulation programs, such as Srim2003 [SRIM]. In fig. [1.12] two examples of ion implantation are shown: a) 1 MeV protons and b) 15 keV phosphorus ions in SiC. The differences between the two implantations are dramatic. Protons penetrate for about 11 um in the crystal, and their distribution is peaked around their average range. The vacancies introduced are also concentrated around the region close to 11 um depth, whereas their distribution is almost uniform in the region close to the surface. Phosphorus penetrates much less due to the lower kinetic energy, and it yields much broader distributions of both ion ranges and damage events.

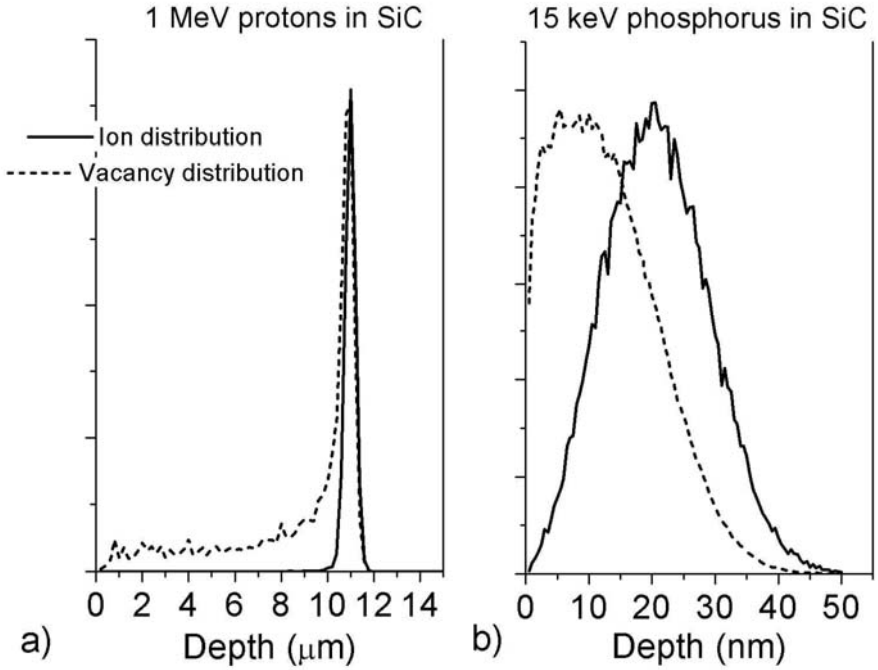


Figure [1.12]. Simulated distributions of ion ranges and introduced vacancies in two cases: a) 1 MeV protons in SiC and b) 15 keV P ions in SiC.

Irradiation (intrinsic defects)

Irradiation has the same principles of implantation. The difference is that irradiation does not stress the role played by the particle once at rest in the semiconducting material. Irradiation is performed with light ions, such as H^+ or $He^{+,2+}$, with electrons, with neutrons or with X and γ rays. Each particle has a specific interaction with the lattice atoms and, therefore, yields a specific damage. High energy protons or electrons are used to create significant amount of intrinsic defects, while neutrons are able to react with the nuclei of the lattice atoms, yielding their transmutation to other elements. Irradiation is used both as a way to introduce intrinsic defects in a controlled way, and as a test for the reliability of devices working in high-radiation environments.

REFERENCES

[Akasaki91] I. Akasaki, H. Amano, M. Kito and K. Iramatsu, *J. Lumin*. **48-49**, 666 (1991)

[Alferov] Zh. I. Alferov, Nobel Lecture, December 8, 2000.

[Amano89] H. Amano, M. Kito, K. Iramatsu, and I. Akasaki, *Jpn. J. Appl. Phys.* **28**, L2112 (1989)

[Baraff85] G.A. Baraff and M. Schlüter, *Phys. Rev. Lett*. 55, 1237 (1985)

[Chantre86] A. Chantre and L.C. Kimerling, *Appl. Phys. Lett.* **48**, 1000 (1986)

[Chow96] T.P. Chow, M. Ghezzo, *MRS Symposium Proceedings,* Pittsburgh, PA. **423** (1996), 69-73.

[Davies] J. H. Davies, *The physics of low-dimensional semiconductors*, (Cambridge University Press, 1998).

[Davydov02] V.Yu. Davydov, A.A. Klochikhin, V.V. Emtsev, S.V. Ivanov, V.V. Vekshin, F. Bechstedt, J. Furthmuller, H. Harima, A.V. Mudryi, A. Hashimoto, A. Yamamoto, J. Aderhold, J. Graul, E. E. Haller, *Phys. Stat. Solidi (b)* **230** (2002), R4.

[DenBaars97] S.P. DenBaars, *Proceedings of the IEEE*, **85**, 1740 (1997)

[Donolato79] C. Donolato, *Scann. Elec. Microscopy* **I**, 257 (1979)

[Figielski64] T. Figielski, *Phys. Stat. Sol.* **6** p.429 (1964).

[Franssen04] G. Franssen et al., *Phys. Rev. B* **69**, 045310 (2004)

[Gösele81] U. Gösele, F. Morehead, W. Frank and A. Seeger, *Appl. Phys. Lett*. **38**, 157 (1981)

[Hangleiter03] A. Hangleiter, *MRS bulletin*, p. 350 (May 2003)

[IOFFE] *http://www.ioffe.ru/SVA/NSM* archive of physical properties of semiconductors

[Kroemer] H. Kroemer, Nobel Lecture, December 8, 2000.

[Kveder82] V. Kveder, Y.A. Osypian, W. Schröter, G. Zoth, *Phys. Stat. Sol* **36**, 539 (1982)

[Litwin99] E. Litwin-Staszewska, T. Suski, I. Grzegory, S. Porowski, J. Robert, D. Wasik, A Witowski, D. Cote and B. Clerjund, *Proc. Intl. Conf. Nitride Semicond.*, Montpellier, p. 124 (1999)

[Miller84] D.A.B. Miller, D.S. Chemla, T.C. Damen, A.C. Gossard, W Wiegmann, T.H. Wood and C.A. Burrus, *Phys. Rev. Lett*. **53**, 2173 (1984)

[Miller85] D.A.B. Miller, D.S. Chemla, T.C. Damen, A.C. Gossard, W Wiegmann, T.H. Wood and C.A. Burrus, *Phys. Rev. B*. **32**, 1043 (1985)

[Moustakas01] T.D. Moustakas, E. Iliopoulos, A.V. Sampath, H.M. Ng, D. Doppalapudi, M. Misra, D. Korakakis and R. Singh, *J. Cryst. Growth* **227-228**, 13 (2001)

[Nakamura92] S. Nakamura, N. Iwasa, M. Senoh, and T. Mukai, *Jpn.J. Appl. Phys*. 31, 1258 (1992)

[Nakamura04] D. Nakamura, I. Gunjishima, S. Yamaguchi, T. Ito, A. Okamoto, H. Kondo, S. Onda and K. Takatori, *Nature* **430**, 1009 (2004)

[NakamuraFasol] S. Nakamura, G. Fasol, *The blue laser diode*, (Springer, Berlin, 1997)

[Neudeck02] P.G. Neudeck, R.S. Okojie and L-Y Chen, *Proceedings of the IEEE*, **90**, 1065 (2002)

[Ng98] H.M. Ng, D. Doppalapudi, T.D. Moustakas, N.G. Weimann and L.F. Eastman, *Appl. Phys. Lett.* **71**, 821 (1998)

[ODonnell99a] K.P. O'Donnell, R.W. Martin, S. Pereira, A. Bangura, M.E. White, W. van der Stricht, and K. Jacobs, *Phys. Stat. Sol. (b)* **216** , 141 (1999).

[ODonnell99b] K.P. O'Donnell, R.W. Martin and P.G. Middleton, *Phys. Rev. Lett*. **82,** 237 (1999)

[Omling85] P. Omling, L. Samuelson and H.G. Grimmeiss, *J. Appl. Phys*. **54**, p. 5117 (1983)

[Pankove71] J.I. Pankove, E.A. Miller and J.E. Berkeyheiser, *J. Lumin.* **4**, 63 (1971)

[Pavesi02] L. Pavesi, L. Dal Negro, C. Mazzoleni, G. Franzò and F. Priolo, *Nature* **408**, 440 (2000)

[Persson97] C. Persson, U. Lindefelt, *J. Appl. Phys.* **82,** 5496-5508 (19977).

[Queisser98] H.J. Queisser and E.E. Haller, *Science*, **281**, p.945 (1998)

[Reed00] M.J. Reed, N.A. El Masry, C.A. Parker, J.C. Roberts and S.M. Bedair, *Appl. Phys. Lett.* **77**, 4121 (2000).

[Renner02] F. Renner, P. Kiesel, G.H. Doehler, M. Kneissl, C.G. Van de Walle and N.M. Johnson, *Appl. Phys. Lett.* **81**, 490 (2002).

[Shapiro00] N.A. Shapiro, P. Perlin, C. Kisielowski, L.S. Mattos, J.W. Yang and E.R. Weber, *MRS internet J. Nitr. Semicond. Res.* **5**, 1 (2000)

[Schröter95] W. Schröter, J. Kronewitz, U. Gnauert, F. Riedel and M. Seibt, *Phys. Rev. B* **52**, p. 13726 (1995)

[Schubert] E.F. Schubert, *Light-Emiting Diodes*, (Cambridge University Press, 2003)

[SRIM] J.F. Ziegler and J.P. Biersack, SRIM 2000, http://www.srim.org

[Sumakeris05] J.J. Sumakeris, J.R. Jenny and A. R. Powell, *MRS bulletin* **30**, 280 (2005)

[Suzuki95] M. Suzuki, T. Uenoyama, A. Yanase, *Phys. Rev. B* **52,** 8132 (1995).

[Teisseyre94] H. Teisseyre , P. Perlin, T. Suski, I. Grzegory, S. Porowski, J. Jun, A. Pietraszko, T.D. Moustakas, *J. Appl. Phys.* **76, 4,** 2429-2434 (1994).

[Walukiewicz89] W. Walukiewicz, *Appl. Phys. Lett*. **54**, 2094 (1989)

[Weber03] E. R. Weber, *Physica B* **340-342**, p. 1 (2003)

[Wosinski89] T. Wosinski, *J. Appl. Phys.* **54**, p.1566 (1989)

[Wu02a] J. Wu, W. Walukiewicz, K.M. Yu, J.W. Ager III, E.E. Haller, H. Lu, W.J. Schaff, Y. Saito and Y. Nanishi, Appl. Phys. Lett. 80, 3967 (2002)

[Wu02b] J. Wu, W. Walukiewicz, K.M. Yu, J.W. Ager III, E.E. Haller, H. Lu, W.J. Schaff, Appl. Phys. Lett. 80, 4741 (2002)

[Yang02] Q.Yang, H. Feick nd E. R. Weber, *Appl. Phys. Lett.* **82**, 3002 (2002)

[Zolper05] J. C. Zolper and M. Skowronski, *MRS bulletin* **30**, 273 (2005)

2 Wide bandgap electronics

As already mentioned, WBG semiconductors are suitable for niche applications, for which silicon or GaAs can not be employed. Here we revise some of the properties making WBG semiconductors superior in specific electronic applications.

- The *wide optical bandgap* allows the design of devices emitting in the whole range of visible wavelengths and in the UV interval of the electromagnetic spectrum. The direct bandgap favours GaN rather than SiC for active optoelectronic devices.

- The *low intrinsic carrier density*, many orders of magnitude lower than that of Si, allowing the design of devices operating at high temperature, high frequency and high power applications.

- The *low leakage current* in both p-n and Schottky junction devices. This is important in high temperature and high frequency application, where it yields a reduction in power dissipation. A low leakage (dark) current is also important for the operation of radiation detectors, as it reduces noise and improves the energy resolution of the detector.

- The possibility of creating *Schottky contacts with higher barrier* than Si-based devices. This has the consequence that WBG based rectifiers can be built as Schottky diodes, with the important advantage of eliminating minority carrier transport from device operation. The presence of minority carriers, in fact, is a factor increasing the power dissipation and decreasing the maximum frequency of high frequency switching devices.

- The high *breakdown electric field*, sustaining an eight-fold larger voltage gradient, enabling thinner active regions, lower on-resistances and higher-voltage operation.

- The *high thermal conductivity*, highest in SiC (3.7 W/cm s), is a key parameter for heat dissipation, with benign effects on device operation and lifetime. It becomes crucial at high temperature operation. For this reason, SiC is often used as a substrate for GaN based devices, as in the case of transistors and LEDs.

- The *high electron mobility*, particularly in the 2-DEG forming at the interface of the heterostructure AlGaN/GaN. This makes possible the design of high speed transistors, such as the high electron mobility transistor (HEMT) or the heterostructure field-effect transistor (HFET). These devices can operate at very higher frequencies, in some cases in the range of hundreds of GHz.

- The *high saturated drift velocity*, which for SiC is about twice that of silicon. This is another factor useful for fast operating speed.

- The relatively *high binding energy*, giving high mechanical strength, chemical inertness and resistance to atomic displacement by impinging particles.

- The above mentioned qualities concur to the *high radiation hardness*, commonly understood as the stability of the parameters of a semiconductor device under nuclear irradiation [Lebedev02]. It is usually verified that WBG semiconductors can withstand radiation conditions at which Si and GaAs-based detectors quickly degrade.

Although the practical applications of WBG semiconductor-based devices will occupy a marginal market volume, estimated as about a tenth percent of the whole semiconductor market in 2008,and Si will remain the semiconductor dominating the market. However, the market will also be growing: new applications, for instance, will be found and implemented for high temperature operation of SiC power devices; though niche applications, they will be fairly important to many products and systems affecting modern human life [Neudeck02].

The device fabricated from WBG semiconductors are already a multitude. In the following, we will illustrate the properties of some selected examples: GaN LEDs and SiC Schottky diodes are tightly related to the development of this thesis; however, we will briefly mention some other devices which are also important in this overall look into the world of wide bandgap electronics.

2.1. GaN-based (opto)electronic devices

2.1.1. Light-Emitting Diodes (LEDs)

As far as 2003, the external quantum efficiency of GaN-based LEDs could reach up to 30% at blue-violet wavelengths [Hangleiter03]. Today, improvements have led to efficiencies as high as 60% [HärlePD]. At the current level of 20 mA the emitted power is up to 21 mW in the blue and up to 7 mW in the green. The reliability of these devices is high, as typical lifetimes can exceed 100,000 h. The applications made possible by the fabrication of blue and green LEDs are various, and a potentially large market has been opened. Some of the applications are: full colour displays with millions of LEDs, traffic signals, automotive applications such as tail- and front-lights, backlighting for LCD displays, etc. These applications make use of the single-colour efficiency of high-brightness LEDs, which is up to 50% for blue and up to 30% for green LEDs. Another important application, still under development, is white LEDs for general room-lighting. White LEDs are realized by using blue or UV-emitting LEDs and a luminescence conversion medium, such as inorganic and organic phosphors [Hangleiter03]. Figure [2.1] is an optical micrograph of one of the LEDs studied in this thesis, representing the cross section of the device with the active region in evidence.

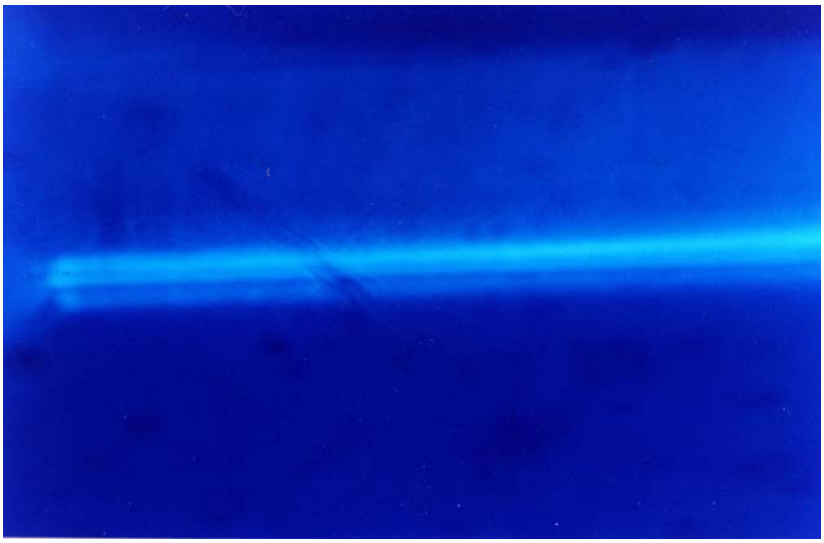


Fig. [2.1] Optical micrograph of the cross section of a GaN/InGaN quantum well light-emitting diode during operation. The bright stripe corresponds to the active region.

These devices are based on a *p*-*n* junction. When the junction is forward-biased, both electrons and holes populate the region in correspondence of the junction. As both carrier species are here

present, there is the possibility for them to recombine. If the recombination is radiative, i.e. band-to band or assisted by radiative recombination centres (usually isoelectronic impurities), the energy set free is transferred to a photon [Sze] [Schubert].

Typical LED structures are based on an active region made of an InGaN/GaN multi-quantum well (MQW). The MQW scheme is particularly efficient, for the following regions:

- The carriers are confined in the QWs, and dwell longer in the active region
- The spatial confinement increases the overlap of the electron and hole wavefunction, thus increasing the recombination probability
- The x mole fraction of InN in the QWs tunes the emission wavelength
- The active region is surrounded by material of higher bandgap, which is transparent to photon of the emitted wavelength.

Usually, LEDs are fabricated starting with SiC or Sapphire substrates. The SiC substrate, more expensive, has the advantages of being both electrically and thermally conductive, thus giving the possibility of creating back contacts and yielding a more efficient heat dissipation, which, in turn, has benign effects on the internal quantum efficiency and on the device lifetime. The structures of LEDs with SiC and sapphire substrates are schematically illustrated in fig. [2.2].

The peculiarities of junction band structure, device working, emission and absorption of photons by the device will be addressed to later, in the sections on the experimental techniques.

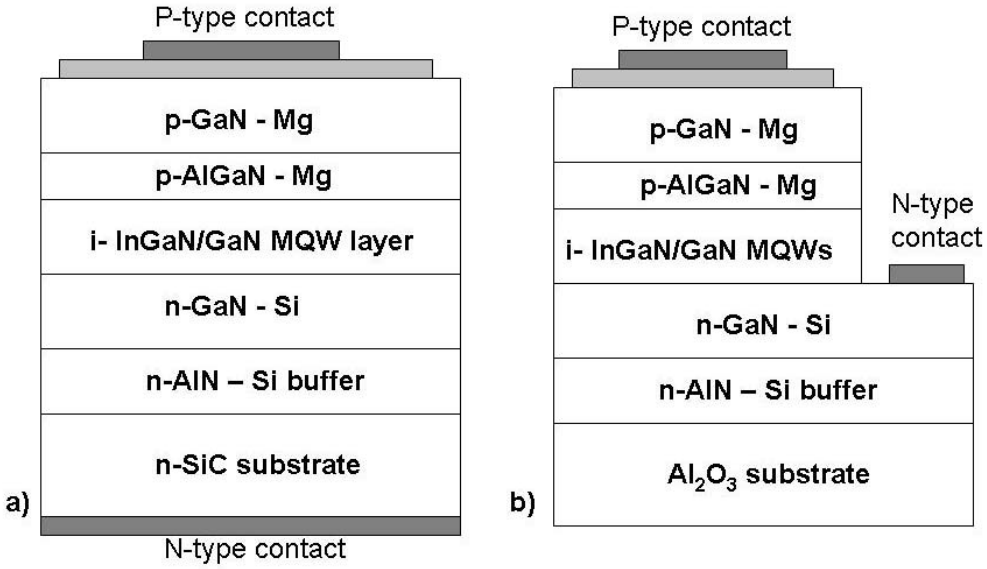


Fig. [2.2] Schematic representation of the cross section of an InGaN/GaN multi quantum well light-emitting diode a) with conducting SiC substrate and b) with insulating sapphire substrate. The relative layer thicknesses are not in scale.

2.1.2. Laser diodes

Laser diodes (LDs) represent a step further in the emission of electromagnetic radiation. Unlike LEDs, emitting a relatively broad spectrum of incoherent radiation, LDs yield emission with a very narrow lineshape of coherent radiation. This is due to the feedback mechanism, which concurs to the stimulation of the emission of photons from an extremely narrow energy interval. The feedback

is usually achieved by means of a resonant cavity, which can be either internal (i.e. the diode dice itself) or external (mirrors) [Saleh&Teich].

The first pulsed operation of a GaN-based laser diode was in 1996, followed by the continuous-wave operation one year later [NakamuraFasol]. These laser diodes with lifetimes greater than 10,000 h were also based on the InGaN/GaN MQW system. Today's nitride lasers can work for over 15,000 h at continuous-wave operation , with an output power of about 30 mW and a threshold current density as low as 1.2 kA/cm^2. The range of wavelengths at which laser emission has been achieved in these devices is 370-450 nm. The most promising application of nitride-based laser diodes will be ultrahigh-density optical storage in next generation DVDs [Hangleiter03].

2.1.3. AlGaN/GaN transistors

Another recent development of III-nitrides is the use of heterostructures for electronics application, exploiting the high electron mobility of the 2-DEG forming at the interface of the AlGaN/GaN system or in an AlGaN/GaN quantum well. In this way, transistors for high power, high frequency and high temperature applications have been fabricated and are approaching commercial distribution, with power densities outperforming those of silicon or GaAs-based devices. The most common design schemes are the high electron mobility transistor, coming in the simple (HEMT) and in the pseudomorphic (p-HEMT) version. In the simple version, the 2DEG resides at the interface between an AlGaN and a GaN layer, while in the pseudomorphic version the 2DEG is confined in a strained (pseudomorphic) GaN quantum well sandwiched between two AlGaN layers. The high mobility is achieved by modulation doping, i.e., by the spatial separation of the Si donors from the 2DEG system, preventing impurity scattering in the transistor channel. A layer scheme of a HEMT is depicted in fig. [2.3]. Nitride-based HEMTs provide good linearity, operating frequencies as high as several tenths of GHz (in some cases even hundredths), and high power output. SiC substrate are preferred because of their heat-sinking properties, although, unlike in the case of LEDs, are not meant to provide the possibility of a back-contact fabrication [Mills04]. The AlGaN/GaN heterostructure can also offer potential application in the field of sensor devices [Stutzmann02].

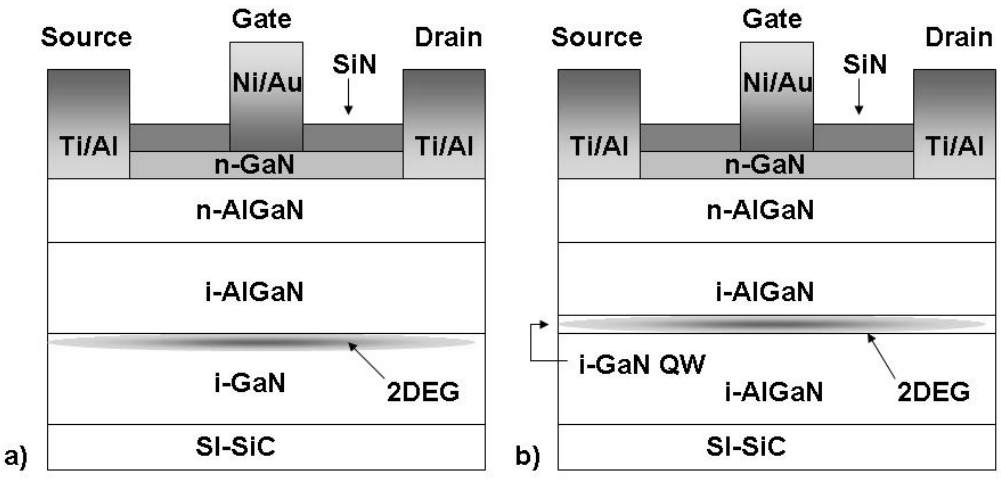


Fig. [2.3]. Schematic section of a) a HEMT b) a p-HEMT based on the AlGaN/GaN heterostructure.

2.2. SiC-based electronic devices

2.2.1. SiC radiation detectors

Radiation damage can affect various properties of a detector. Phenomena connected with the alteration of these properties are the increase of the leakage current, the decrease of charge collection efficiency (CCE), the removal of free carriers from the conductive regions of the device, and sometimes the inversion of conduction type due to the introduction of electrically active defects. Radiation hardness is the inertness of these device parameters to high doses of radiation [Lebedev02].

Due to its physical properties SiC, due to the wide gap and to the strength of its chemical bonds, has been seriously considered as a valid alternative to silicon for the production of radiation hard ionizing particle detectors. An important driving force for the development of SiC-based detectors has come from the construction of the large hadron collider (LHC) at CERN in Geneva. The collaboration CERN RD-50 aims to the development of detectors able to withstand the higher irradiation fluences that will affect them: a large amount of research is made for the improvement of silicon-based detectors, but also the possibility of using SiC is investigated [Moll02] [Sellin06]. Our group at the University of Bologna is currently involved in this collaboration, and our contributions span from the characterization of the radiation hardness of several SiC detector schemes to the study of the defects introduced by irradiation by means of DLTS and other experimental techniques.

The usual design of a detector includes a p^+-n, n^+-p or a p^+-i-n^+ diode structure, operating under reverse bias, as illustrated in fig. [2.4]. A space charge region is formed, and the p^+ and n^+ regions act as electrodes. Ionizing particles produce ionization in a semiconductor when they are slowed down or absorbed. Thus, electron-hole pairs are formed, which are then separated by the electric field and collected at the electrodes, yielding a current pulse in the detection circuit. The current generated is well correlated with the impinging particle energy.

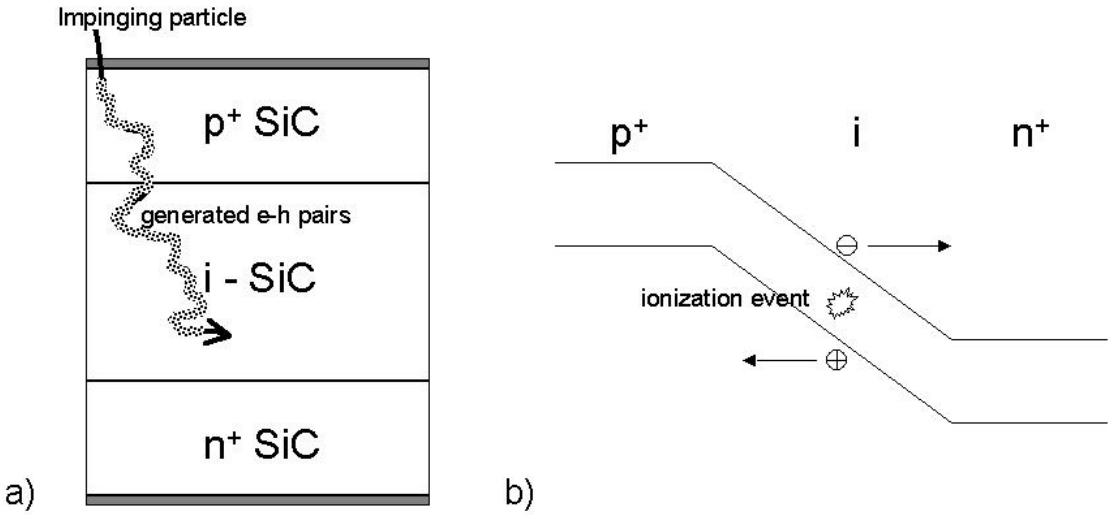


Fig. [2.4]. a) Scheme of cross section and b) band diagram of a radiation detector based on SiC operating under reverse bias.

A detector should have a low concentration of impurities and defects, because these cause a diminution in the current pulse amplitude due to recombination of electron-hole pairs and scattering of charge carriers. Moreover, a low concentration of dopant impurities extends the thickness of the space charge region.

A wide bandgap is useful in a detector, as it reduces significantly the rate of thermally generated charge carriers raising the noise level. On the other hand, it also represents a disadvantage: a

particle with a certain energy, ideally transforming all its energy for the generation of electron-hole pairs, generates 3 times more particles in Si than in SiC. Detectors based on SiC, therefore, have lower pulse amplitudes. However, for low signals, the reduction of the noise level is more important than the reduction of the signal level, so that the overall signal-to-noise ratio (SNR) is improved for SiC-based detectors. Furthermore, SiC-based detectors still have a high SNR at temperatures which are unattainable for Si-based devices, needing external cooling to keep the intrinsic carrier level sufficiently low [Lebedev04].

2.2.2. SiC power devices

The main driving force for SiC market is represented by the application of SiC-based power devices to high temperature, high frequency and high power applications. The factors favouring SiC for such applications have already been discussed: high breakdown field, low leakage current, high Schottky barrier height, high thermal conductivity, etc.

Particularly for switching applications, there is a peculiarity definitely favouring SiC with respect to Si. The high Schottky barrier makes possible the fabrication of SiC Schottky barrier Diodes (SBDs) which outperform the Si p-i-n rectifiers in the 600-1200 V range. Silicon p-i-n diodes, in fact, tend to store large amounts of minority carriers in the forward-biased state, whereas this does not happen in SiC SBDs, being unipolar devices. The stored minority charge has to be removed by minority carrier recombination before the diode enters the proper off-state, This causes long storage and turn-off times. The benefit of using SiC SBD lies in its ability to switch faster with almost zero recovery charge. Furthermore, the recovery time grows in silicon p-i-n diodes with the temperature, due to the increase in majority carrier concentration, while such dependence is not found for SiC SBDs, which can easily operate at 200°C [AgarwalCREE].

SiC is especially considered for high-temperature power electronics. High temperature at the junction can be either a consequence of heat production at high power operation, or due to the ambient temperature at which the device is required to operate. In both cases, it is crucial that device operation is not hindered by thermally activated mechanisms, and that an efficient heat dissipation is provided by a high thermal conductivity. SiC has all these properties, and rectifiers based on it can properly operate even at T=600°C [Neudeck02].

Our group has been recently involved in the study of defects in SiC-based power diodes, in a collaboration with Politecnico di Torino and International Rectifier Corporation Italia (IRCI).

2.3. Integrated GaN-SiC electronics

At the end of this part about wide-bandgap electronics, it is worth mentioning some examples of integrated electronics between SiC and GaN, going beyond the usual “epitaxial layer – substrate” relationship.

2.3.1. GaN-SiC heterojunctions

Growing GaN on SiC substrates has become a common procedure in the last few years. SiC substrates offer a low lattice mismatch and both electrical and thermal conductivity. However, there are some examples of devices where the functional relationship between the GaN and the SiC layers are of more complex nature. Although these devices remain somewhat exotic and do not represent a large impact for the market, they are nevertheless interesting from a more fundamental point of view. Some of these devices have been developed as early as 1994, like the GaN/SiC heterojunction bipolar transistor [Pankove94], when the growth techniques did not offer very good crystal qualities

yet. This device, consisting in the sequence of an *n*-type GaN emitter, a *p*-type SiC base and an *n*-type SiC collector/substrate, was among the first examples of transistor able to operate at T=260°C with very little degradation of its properties. The fabrication of GaN/SiC n-p heterojunction diodes made possible the determination of the band offsets between GaN and 6H-SiC. These heterojunction diodes were electrically characterized, showing also ideality factors (see section 3.1.3) very close to one. The band offsets were determined as E_C^{SiC}-E_C^{GaN}=0.11eV and E_V^{SiC}-E_V^{GaN}=0.48eV, thus describing a type II alignment [Torvik98].

2.3.2. GaN-SiC integrated devices

Radiation detectors, such as those illustrated in section 2.2.1 based on SiC need a front end electronics as an implementation. The front end electronics usually consists in a transistor acting asa a signal amplifier. Si-based detectors are driven by Si electronics, but they can not operate for harsh environment conditions, at which SiC is the only feasible choice. Currently, a project founded by the Italian Ministry of University and Research (MIUR) in cooperation with former Alenia Marconi Systems is under way, with the participation of our group at the University of Bologna, aiming to the fabrication of SiC radiation detectors with GaN-based front end electronics (consisting in an AlGaN/GaN heterostructure transistor). The potential advantage of such GaN/SiC integration are a high energy and spatial resolution, the possibility to operate at high temperatures without the need of expensive cooling systems, the large area and volume and the radiation hardness.

REFERENCES

[AgarwalCREE] A. Agarwal, R. Singh, S-H Ryu, J. Richmond, C. Capell, S. Schwab, B. Moore and J. Palmour, *technical paper* at *www.cree.com*.

[Friel05] I. Friel, C. Thomidis and T.D. Moustakas, *J. Appl. Phys.* **97**, 123515 (2005)

[Hangleiter03] A. Hangleiter, *MRS bulletin*, p. 350 (May 2003)

[HärlePD] V. Härle, talk held at the University of Padova, September 2005

[Lebedev02] A.A. Lebedev, V.V. Kozlovski, N.B. Strokan, D.V. Davydov, A.M. Ivanov, A.M. Strel'chuk and R. Yakimowa, *Fiz. Tekh. Poluprovodnikov* **36**, 1354 (2002)

[Lebedev04] A.A. Lebedev, A.M. Ivanov, N.B. Strokan, *Fiz. Tekh. Poluprovodnikov* **38**, 129 (2004)

[Mills04] A. Mills, *III-V review* **18**, 5, p. 38 (2004)

[Moll02] M. Moll, *RD50 internal note* – RD50/2003/001

[NakamuraFasol] S. Nakamura, G. Fasol, *The blue laser diode*, (Springer, Berlin, 1997)

[Neudeck02] P.G. Neudeck, R.S. Okojie and L-Y Chen, *Proceedings of the IEEE*, **90**, 1065 (2002)

[Pankove94] J. Pankove, S.S. Chang, H.C. Lee, R.J. Molnar, T.D. Moustakas and B. Van Zeghbroeck, *Proceedings of IEDM* 15.6.1 (1994)

[Saleh&Teich] B. Saleh, M.C Teich, *Fundamentals of Photonics*, (1992)

[Sellin06] P.J. Sellin, J. Vaitkus, *Nucl. Instr. and Meth. in Phys. Res. A* **557**, 479 (2006)

[Stutzmann02] M. Stutzmann, G. Steinhoff, M. Eickhoff, O. Ambacher, C.E. Nebel, J. Schalwig, R. Neuberger and G. Muller, *Diamond and Rel. Mater.* **11**, 886 (2002)

[Sze] S.M. Sze, *Physics of semiconductor devices* (Wiley Interscience, 1981)

[Torvik98] J.T. Torvik, M. Leksono, J.I. Pankove, B. Van Zeghbroeck, H.M. Ng and T.D. Moustakas, *Appl. Phys. Lett.* **72**, 1371 (1998)

3 Electrical Characterization

3.1 Current-Voltage (I-V) characterization

Current-Voltage (I-V) characterization is among the most common experimental techniques in semiconductor device physics, for various reasons. First, the I-V characteristics gives information about the operation of the sample as a device: the I-V of a diode, evidencing its rectifying properties, is completely different from the I-V of an ohmic resistor; the I-V of a transistor evidences what the amplification or logical capabilities of the device are, and many other examples could be mentioned. Secondly, the I-V characteristics give a significant amount of physical information. From the analysis of I-V data it is possible to draw conclusions about the presence of defects, about the goodness of the contacts, the potential barriers governing the transport of charge, the presence of hetero-interfaces, etc.

In the following the main applications of this characterization will be shortly revised. For a more in-depth analysis, many semiconductor materials and devices handbooks offer the details [Sze], [McKelvey], [Neamen].

3.1.1. I-V characterization of a Schottky junction

In a Schottky-n-type junction the current transport happens mostly by thermionic emission of majority carriers over the metal-semiconductor barrier. Other transport mechanisms involve diffusion, recombination in the junction region, or tunnelling through the barrier.

Thermionic model

The dominant process for current transport in the analysed samples is thermionic emission over the barrier. According to this model, the current depends on the voltage according to the following expression [Sze]:

$$J = A^* T^2 \exp\left[-\frac{e\phi_b}{kT}\right] \times \left\{\exp\left[\frac{eV}{kT}\right] - 1\right\} \tag{3.1}$$

with

$$A^* = \frac{4\pi e m^* k^2}{h^3} \tag{3.2}$$

also called the theoretical *Richardson constant* of the material. The value

$$J_s = A^* T^2 \exp\left[-\frac{e\phi_b}{kT}\right] \tag{3.3}$$

is the saturation current of the diode, which is strongly temperature dependent.

Diffusion model

According to this model the current-voltage characteristics follows the law [Sze]

$$J = J_{SD}\left\{\exp\left[\frac{eV}{kT}\right]-1\right\} \tag{3.4}$$

with

$$J_{SD} = \frac{e^2 D_n N_C}{kT}\left(\frac{e(V_{bi}-V)2N_D}{\varepsilon_s}\right)^{1/2}\exp\left[-\frac{e\phi_b}{kT}\right], \tag{3.5}$$

where D_n is the diffusion coefficient for electrons, N_C is the equivalent density of states in the conduction band, V_{bi} is the built-in potential. The expression for the diffusion current is similar to that for the thermionic current; however, the saturation current has a weak dependence on the temperature and a strong dependence on the applied voltage in the diffusive model, while it depends mostly on the temperature in the thermionic model. This is useful in order to discriminate the dominant current transport mechanism. In many cases, both mechanisms coexist at the same time, though thermionic emission dominates.

Tunnelling model. Ideality factor

The tunnelling mechanism for current transport is usually found in heavily doped semiconductors or at low temperature operation [Sze]. The current due to tunnelling can be expressed through

$$J \propto \exp[-e\phi_b / E_{00}] \tag{3.6}$$

with

$$E_{00} = \frac{e\hbar}{2}\sqrt{\frac{N_D}{\varepsilon_s m^*}}. \tag{3.7}$$

The resulting forward current density, given by the superposition of tunnelling current and thermionic current, can be re-expressed as

$$J \cong J_S \exp[eV / nkT], \tag{3.8}$$

where n is called the *ideality factor*, and is extracted from the experimental characteristics as

$$n \equiv \frac{e}{kT}\frac{\partial V}{\partial(\ln J)}. \tag{3.9}$$

The ideality factor is close to one when the diode forward characteristics resemble closely that of the thermionic model, and tends to grow when other factors play a significant role in current transport.

Other charge transport mechanisms in Schottky junctions can be the injection of minority carriers at large forward bias, and the generation current due to deep levels in the gap in reverse bias. This latter mechanism is also found in p-n junctions, so it will be treated in the following sections.

Soft breakdown. Dislocation-assisted tunnelling

Under certain circumstances, for instance when a high dislocation density is present in the device, tunnelling mechanisms can be enhanced under reverse bias, which is also called *soft breakdown*. This case has been reported in InGaN blue LEDs grown on sapphire [Fang00]. When this dislocation-assisted tunnelling occurs, the reverse I-V characteristics tend to have the form $I=V^{m}$.

3.1.2. I-V characterization of a p-n homojunction

The current-voltage characteristics of a heterostructure-based LED are essentially governed by the same mechanisms of the abrupt *p-n* homojunction, with some differences due to the presence of bandgap discontinuities. In any case, the heterostructure-based LED is a *p-n* junction, therefore it is a bipolar device, where charge is transported by both electrons and holes. The equations describing the current-voltage characteristics of *p-n* junctions have a form which is very similar to those for the Schottky diode, yet the physical parameters appearing are different.

Diffusion model

The simplest model for charge transport in abrupt *p-n* junctions is the diffusive model, which predicts current-voltage characteristics described by the Shockley equation [Shockley50], [Sze]:

$$J = J_s\left\{\exp\left[\frac{eV}{kT}\right]-1\right\} \qquad (3.10)$$

with the saturation current density given by the sum of the saturation current densities of minority electrons J_n and holes J_p

$$J_S = J_p + J_n = e\left(\frac{D_p p_{n0}}{L_p}+\frac{D_n n_{p0}}{L_n}\right), \qquad (3.11)$$

where $D_{p,n}$ are the hole and electron diffusion coefficients, respectively, $L_{p,n}$ are the hole and electron diffusion lengths, respectively, p_{n0} and n_{p0} are the equilibrium concentrations of holes on the n-side and of electrons on the p-side of the device, respectively. This expression can be

simplified in the case of one-sided p^+-n junctions like those we analysed. In this case $p_{n0}>>n_{p0}$, so that the contribution to the current comes mainly from the holes on the n-side of the junction:

$$J_S \cong \frac{eD_p p_{n0}}{L_p} \cong e\sqrt{\frac{D_p}{\tau_p}}\frac{n_i^2}{N_D} \propto T^{(3+\gamma/2)} \exp\left[-\frac{E_g}{kT}\right] \qquad (3.12)$$

where it is supposed that $D_p/\tau_p \propto T^{\gamma}$. The most significant temperature dependence is provided by the exponential term. Plotting the J_s values in an Arrhenius Plot allows thus to extract the bandgap value.

3.1.3. Generation-Recombination currents

An important deviation from the ideal characteristics of current transport is represented by the so-called generation-recombination currents. These currents can be explained by the theory of recombination of excess carriers by Shockley, Read and Hall (SRH) [McKelvey]. This theory describes basically the recombination of electrons from the conduction band with holes from the valence bands assisted by electron states in the bandgap, also called *deep levels* or *deep states* (see section 4 for more detail). According to the SRH theory, the recombination rate of excess electrons and holes is

$$R = \frac{C_n C_p N_t (np - n_i^2)}{C_n(n+n') + C_p(p+p')} \qquad (3.13)$$

where $C_{p,n}$ is the probability with which the deep level captures holes and electrons per unit time, respectively, $n=n_0+\delta n$, $p = p_0+\delta p$ are the actual electron and hole concentrations, n_i is the intrinsic concentration and n',p' are defined as

$$n' = N_C \exp\left[-\frac{E_C - E_t}{kT}\right] \qquad (3.14)$$

$$p' = N_V \exp\left[-\frac{E_t - E_V}{kT}\right] \qquad (3.15)$$

where E_t is the energy level of the deep state. Under extrinsic doping and low injection conditions in an n-type semiconductor, the following relationships are satisfied

$$n_0 >> p_0, \quad n_0 >> \delta p, \quad n_0 >> n', \quad n_0 >> p'$$

and eq. (3.13) reduces to

$$R = C_p N_t \delta p = \frac{\delta p}{\tau_{p0}} \qquad (3.16)$$

with

$$\tau_{p0} = \frac{1}{C_p N_t}. \qquad (3.17)$$

A similar expression is valid for the case of a p-type semiconductor under low-injection conditions:

$$R = C_n N_t \delta n = \frac{\delta n}{\tau_{n0}} \qquad (3.18)$$

with

$$\tau_{n0} = \frac{1}{C_n N_t}. \qquad (3.19)$$

The quantities τ_{n0} and τ_{p0} are the mean excess minority carrier lifetimes. Equation (3.13) can therefore be applied in the two different situations of generation of carriers in reverse bias, and of recombination in forward bias.

Forward bias recombination current. Ideality factor

Under forward bias, an excess concentration of electrons and holes is injected in the junction region. If deep levels are present, they can assist the carriers in the recombination process. Using the definition of the minority carrier lifetimes and the approximation that $\tau_{p0} = \tau_{n0} = \tau_0$, and supposing that the trap level is at the intrinsic Fermi level, one obtains the maximum recombination rate at the junction:

$$R_{\max} = \frac{n_i}{2\tau_0} \exp\left[\frac{eV}{2kT}\right], \qquad (3.20)$$

from which the forward recombination current can be expressed as

$$J_{rec} = \frac{eWn_i}{2\tau_0} \exp\left[\frac{eV}{2kT}\right], \qquad (3.21)$$

where W is the bias-dependent space charge width. As the forward recombination current requires the presence of both electrons and holes in the space charge region, it will not be found in unipolar devices, such as Schottky diodes. The total forward current is then the sum of diffusion and recombination components (and possibly of other components, such as tunnelling)

$$J_{fw} = J_D + J_{rec}. \qquad (3.22)$$

This will yield forward characteristics like that illustrated in fig. [3.1], and described through the expression

$$J = J_S \exp[eV / nkT] \qquad (3.23)$$

where the n parameter is the already introduced ideality factor (see eq. (3.9)). Usually, for a large forward bias n tends to unity, as diffusion dominates. For lower forward bias, n is close to

two, as recombination is more significant [Neamen]. Higher n values can be found in presence of tunnelling or in heterojunctions.

Reverse bias generation current

The generation current is a phenomenon regarding both p-n and Schottky junction. To describe the generation of carriers in the space charge region under reverse bias, one makes use of eq. (3.13), considering the fact that the space charge region is depleted of free carriers, and therefore $n = p \cong 0$. Thus the recombination rate becomes

$$R = -\frac{C_n C_p N_t n_i^2}{C_n n' + C_p p'} \cong \frac{-n_i}{\frac{1}{N_t C_p} + \frac{1}{N_t C_n}} = -\frac{n_i}{\tau_0}. \qquad (3.24)$$

Here the negative sign implies a negative recombination rate, that is a generation rate $G = -R$. This leads to the expression of the generation current

$$J_{gen} = \int_0^w eG\,dx = \frac{e n_i w}{2\tau_0}. \qquad (3.25)$$

The most important feature of the generation current is that, unlike the reverse diffusion current, it does not saturate, but it increases regularly until the bias is so high that it provokes the junction breakdown [Neamen]. The total reverse bias current is given by the sum of diffusion and generation components, and its behaviour is reported in fig. [3.1].

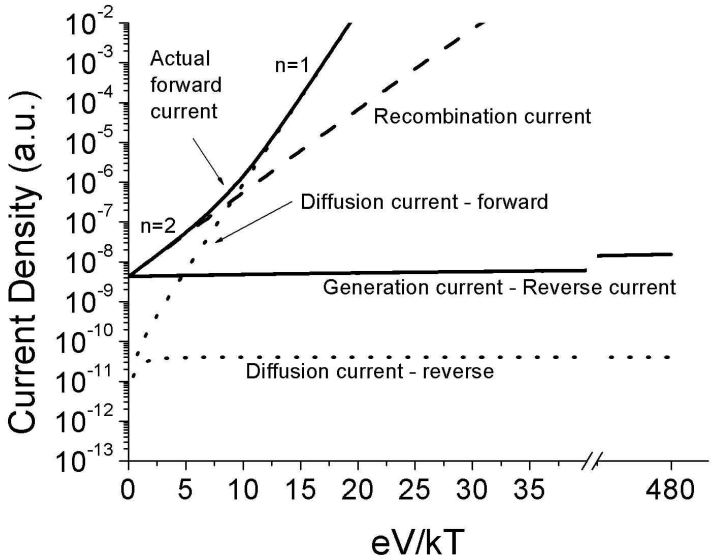


Fig. [3.1]. The effect of diffusion and generation-recombination currents on the actual I-V characteristics of a diode.

3.1.4. Series resistance

The net effect of a resistance in series with a diode is of lowering the actual voltage drop at the junction edges. In this case the actual voltage drop is

$$V_{actual} = V_{applied} - IR_s \qquad (3.26)$$

thus the forward current-voltage characteristics has the form

$$I = I_s \exp[e(V - IR_s)/nkT] \qquad (3.27)$$

3.1.5. I-V of heterojunctions

The current-voltage characteristics of heterojunctions are significantly different from those of homojunctions, because of the more complicated band diagrams (figure [3.2]). In fact, the barrier heights are different for electrons and holes, according to the particular band gap discontinuities proper of the junction. This is a factor which controls the relative magnitude of the hole and electron current, whereas in the p-n homojunction it is only the relative doping level to play a role.

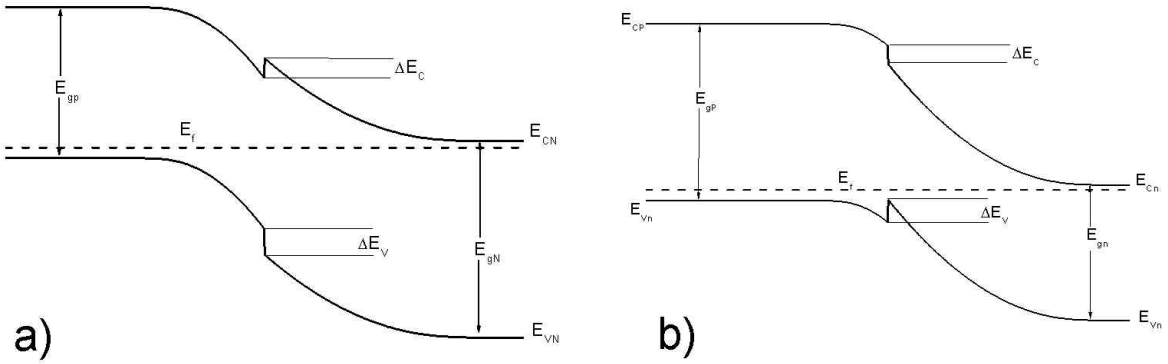


Fig. [3.2]. Band diagrams of: a) a p-N and b) a P-n heterojunction.

This can be exploited, as it is shown in fig. [3.3]. In this picture, the typical band diagram of a III-V LED is sketched. In LEDs the main concern is to have a high concentration of both electrons and holes in the active region (usually intrinsic) of the p-n junction. Electrons and holes must dwell for the longest possible time in the QW region, in order to efficiently recombine emitting a photon. As in III-V semiconductors the diffusion coefficient is much higher for electrons than for holes, electrons tend to occupy the active region for a shorter time. The problem is resolved by inserting an AlGaN layer at the boundary between p-region and active region. As AlGaN has a relatively higher band discontinuity for electrons than for holes, the electron current is significantly limited by the presence of this blocking layer, whereas the hole current is hardly unaffected. This yields two almost equal current component, which means a more efficient recombination process [Schubert].

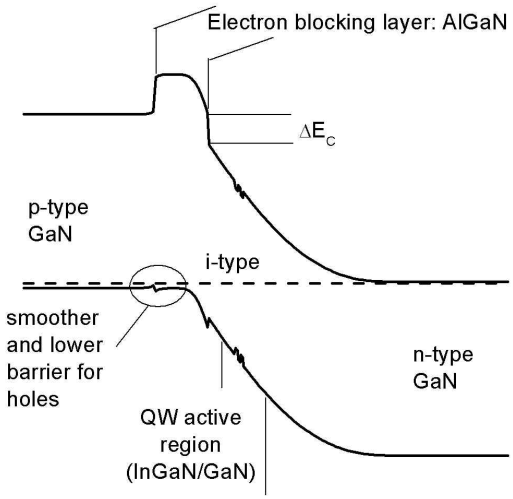


Fig. [3.3]. The band diagram of a nitride-based LED with an electron blocking layer

The presence of band discontinuities in heterojunctions allows one to treat the case similarly to metal-semiconductor, considering the charge transport as a basically thermionic phenomenon, so that the forward current has the form

$$J = A^* T^2 \exp\left[\frac{-E_w}{kT}\right], \qquad (3.28)$$

where E_w is an effective barrier height. The barrier height is increased or reduced by the application of an external bias. In any case, other factors, such as diffusion or tunnelling can play significant roles, according to the junction structure and the operation conditions. Other complications come from the discontinuities in the electron and hole effective masses in the different regions of the junction. The I-V characteristics can still be described with the help of an ideality factor. However, the values that the ideality factor can assume are no longer contained in the interval between 1 and 2. In some cases, for instance in the case of a GaN/SiC n-p junction, ideality factors very close to the unity have been obtained [Torvik98]. In other cases, much larger values of n have been reported.

3.1.6. Anomalously high ideality factors

When the ideality factor is $n>>2.0$, other factors than diffusion or recombination of carriers must be taken into account. One of the factors is tunnelling, as already mentioned. Another model [Shah03] takes into account the formation of potential barriers at every heterointerface, from the metal-semiconductor contacts to the semiconductor-semiconductor junctions. Each of the junctions is characterized by an own ideality factor n_j. The current and the voltage drop $V_j>>kT$ at each junction are given by

$$I = I_{sj} \exp\left[\frac{eV_j}{n_j kT}\right]. \qquad (3.29)$$

The total voltage drop is $V = \Sigma_j V_j$, so that the I-V characteristic of the structure is

$$V = \sum_j V_j = \sum_j \left[n_j (kT/e) \ln I - n_j (kT/e) \ln I_{Sj}\right]. \qquad (3.30)$$

Thus one has, rearranging the terms

$$\ln I = \frac{(e/kT)}{\sum_j n_j} V + \frac{\sum_j n_j \ln I_{Sj}}{\sum_j n_j} \qquad (3.31)$$

As the second summand in the above equation is constant, one obtains an effective ideality factor for the heterostructure given by the sum of the single ideality factors of each junction (p-n junction, unipolar heterojunctions and metal-semiconductor junctions).

3.1.7. Experimental setup

The I-V measurements have been performed with a Keithley 6517 electrometer. The computer program sets the following measurement parameters:

- Bias interval V_{max}, V_{min}
- The scanning step ΔV
- The Δt time interval between bias change and meter reading.

The temperature at which the measurement is performed is controlled by the Lakeshore 330 temperature controller. The typical experimental setup for the I-V characterization is shown in fig. [3.4].

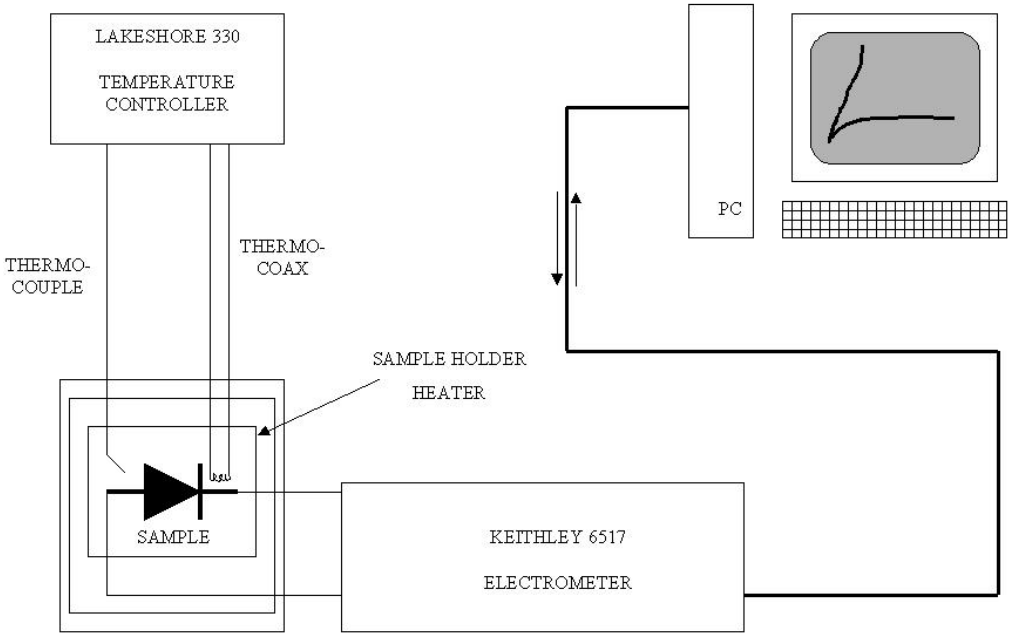


Fig. [3.4]. Block diagram for the I-V characterization

3.2 Capacitance-Voltage (C-V) characterization

A p-n or a Schottky junction in reverse bias has a capacitance, which it will be referred to as depletion capacitance. This capacitance depends on various factors: the doping concentration, the temperature, the doping distribution, the built-in potential. The capacitance-voltage (C-V) characterization is the measurement of the capacitance as a function of the reverse bias, and is a paramount technique in semiconductor material and device characterization, as it gives the possibility of determining the doping density in the semiconductor and also its variations in depth. Moreover, the capacitance may be affected by the presence of electronic states (deep levels) in the gap, thus making it possible to characterize them by means of capacitance-based measurements. This will be illustrated in the next chapter, which deals with deep level transient spectroscopy (DLTS). A convenient situation for C-V characterization is that of a one-sided junction, i.e., a Schottky diode or a p^+-n diode. In the following we will consider only n-type semiconductors as active layers and metal-semiconductor junction. However, the formulae are also applicable to the case of a p^+-n junction.

In this section we will deal with capacitance-voltage measurements in one-sided junction under different conditions of doping and composition. The case of deep levels present will be dealt with after introducing the so-called Shockley-Read-Hall statistics for electron deep states in chapter 4, dedicated to thermal spectroscopy of deep levels.

3.2.1. C-V characterization of a one-sided junction

Figure [3.5] illustrates the spatial variations of energy bands and Fermi levels in a n-type semiconductor in the vicinity of the metal Schottky contact for two different values of applied reverse bias, V_1 and V_2. The distance x is measured from the metal /semiconductor interface, and w_1 and w_2 are the depletion distances for $V= V_1$ and $V= V_2$, respectively. Increasing the absolute value of the reverse bias from V_1 to V_2 causes the transport of negative charge from the semiconductor into the negative terminal of the external circuit. Therefore, the junction can be seen as an electrical capacitor, with a capacitance $C=Q/V$, where Q is the charge stored in the depletion region. In the same way, a small-signal (differential) capacitance $C_s=dQ/dV$ can be defined.

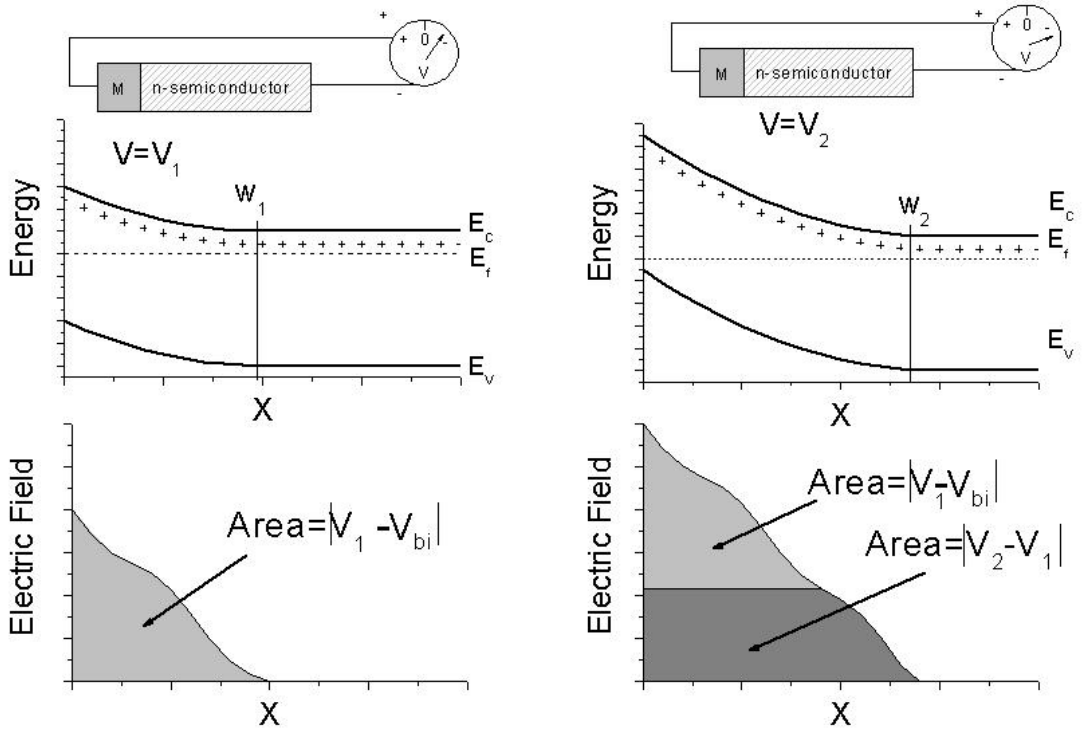


Figure [3.5] Electron energies and electric field values as function of depth *x* in a Schottky-n-type junction for two applied reverse biases $|V_1| < |V_2|$. The respective depletion distances are defined by w_1 and w_2. The general situation with non-uniform donor concentration is depicted here.

When the donor concentration N_d is uniform in a one-sided junction with an applied bias V, the depletion width w is given by

$$w = \sqrt{\frac{2\varepsilon_s}{e}\frac{V_{bi} - V}{N_D}}, \qquad (3.32)$$

then the charge stored in the depletion region of width w is $N_d\, w\, A$, where A is the junction area. Therefore, Q is also given by the following expression:

$$w = A\sqrt{2\varepsilon_s e(V_{bi} - V)}, \qquad (3.33)$$

and the small-signal capacitance C_s is

$$C_s = \frac{dQ}{dV} = A\sqrt{\frac{\varepsilon_s e}{2}\frac{N_D}{V_{bi} - V}} = \frac{\varepsilon_s A}{w}, \qquad (3.34)$$

which is exactly the expression for the parallel plate capacitor of area A and inter-plate distance w.

In the case of a more general doping distribution, i.e. with $N_D(x)$ function of the junction depth, we can find a similar results. The electric field variations due to a space charge density $\rho(x)$ follow Gauss's law

$$\frac{d\xi(x)}{dx} = \frac{\rho(x)}{\varepsilon_s} = \frac{eN_D(x)}{\varepsilon_s} \qquad (3.35)$$

The areas under the respective $\xi(x)$ curves are the voltages $|V_{bi}-V_1|$ and $|V_{bi}-V_2|$. If V_2 is only slightly larger than V_1, one can write:

$$V_1 - V_2 = \Delta V \qquad (3.36)$$

and

$$w_2 - w_1 = \Delta w\,. \qquad (3.37)$$

Then, from fig. [3.5] one has

$$V_1 - V_2 = \xi(w) \times w\,, \qquad (3.38)$$

and

$$\xi(w) = \left.\frac{d\xi(x)}{dx}\right|_{x=w} \Delta w = \frac{eN_D(w)}{\varepsilon_s}\Delta w\,. \qquad (3.39)$$

Now, $eN_D(w)\Delta w = \Delta Q(w)/A$, so that one finds

$$\Delta V = \frac{\Delta Q(w) w}{A\varepsilon_s}\,, \qquad (3.40)$$

which leads to the same expression found in the case of uniform donor distribution

$$C_s = \frac{\Delta Q}{\Delta V} = \frac{\varepsilon_s A}{w}\,. \qquad (3.41)$$

Thus, the capacitance of a one-sided junction is equivalent to that of a parallel-plate capacitor in both cases of uniform and non-uniform doping distribution. This allows the determination of the depletion width w from the measurement of the junction capacitance, and, at the same time, to the determination of the doping profile. In fact, from the expression (3.41), dropping the subscript *s* of the small-signal capacitance for clarity

$$\frac{dC}{dw} = -\frac{\varepsilon_s A}{w^2} = -\frac{C^2}{\varepsilon_s A} \qquad (3.42)$$

one can calculate, using

$$\frac{dV}{dw} = \left(\frac{dC(w)}{dw}\right)\left(\frac{dV}{dC\ (w)}\right) \tag{3.43}$$

and

$$C(w) = \frac{dQ(w)}{dV} = eN_D(w)A\frac{dw}{dV}, \tag{3.44}$$

the quantity

$$C(w) = -eN_D(w)\varepsilon_s A^2 \frac{1}{C^2}\frac{dC}{dV}. \tag{3.45}$$

This eventually leads to the expression for the doping density:

$$N_D(w) = -\frac{C^3}{e\varepsilon_s A^2}\left(\frac{dC}{dV}\right)^{-1} = \frac{2}{e\varepsilon_s A^2}\left[\frac{d}{dV}\left(\frac{1}{C^2}\right)\right]^{-1}. \tag{3.46}$$

This is the procedure by which one can obtain the doping profile by measuring the small-signal depletion capacitance. More rigorously, one should consider that the small-signal capacitance arises from the movement of free electric charge (i.e. electrons in the conduction band in a one-sided n-type junction) caused by the change in the applied reverse bias. Therefore, the quantity $N_D(x)$ should be substituted by the quantity $n(x)$, the electron density in the conduction band at the edge of the depletion region, in all above expressions. If the doping is uniform and there are no deep levels present, the donor density and the free electron density generally coincide at room temperature. However, if the doping distribution varies strongly with x, diffusion of free electrons causes $n(x)$ to be different from $N_D(x)$. In this case, the results of the C-V measurements yield the values of $n(x)$ [Stradling], [Sze], [Rhoderick], [OrtonBlood]. So, from now on, we will refer to the charge density measured by C-V characterization as to the *apparent carrier density (ACD)*, which we will identify using the symbol $n_{CV}(x)$.

3.2.2. C-V characterization of non-uniformly doped junctions

As already mentioned, the effect of a non-uniform doping distribution in the n-type part of the junction of a p^+-n or a Schottky-n junction is that by C-V profiling one does not measure the doping profile, but the free charge density at the edge of the depletion region. We will from now on indicate the charge density measured by C-V (eq. 3.a.15) by $n_{CV}(x)$, while the free-electron density profile in the conduction band, which is defined for a certain applied bias V, will be indicated by $n(x)$.

If $N_D(x)$ has any sharp gradients in the x direction, these gradients are smoothed out in the $n(x)$ distribution because of diffusion effects. In other words, diffusion makes $n(x)$ vary more slowly than $N_D(x)$, which also has the consequence of measuring $n_{CV}(x) \neq N_D(x)$. The scale on which the free charge distribution is smoothed out is typically given by the Debye length L_D:

$$L_D = \sqrt{\frac{\varepsilon_s kT}{e^2 N_D}}, \tag{3.47}$$

in which k is the Boltzmann constant and T the temperature of the sample. In the proximity of an abrupt change in the doping profile occurring at $x=x_a$, $n(x)$ varies approximately as

$$n(x) \propto \exp\left[-\frac{(x-x_a)^2}{2L_D^2}\right] \qquad (3.48)$$

This behaviour is followed also by the measured charge density $n_{CV}(x)$. Figure [3.6] shows the situation of a GaAs Schottky junction with steps in the doping concentration. The band diagram and the conduction band electron profile $n(x)$ have been simulated for the unbiased structure, whereas the $n_{CV}(x)$ has been calculated from the free charge profile simulations at different biases. One can see that both the free and the apparent charge profiles are smoothed out at the points where abrupt doping changes occur. One can also notice how the smoothing length (the Debye length) is smaller in correspondence of the step at x=300 nm, where the doping level is higher, than in correspondence of the step at x=600 nm. Another important remark is about the difference between $n(x)$ and $n_{CV}(x)$: these two quantities differ because the first one is defined as an in-depth profile at a certain bias, while the second one as a differential quantity, obtained by varying the bias applied to the sample.

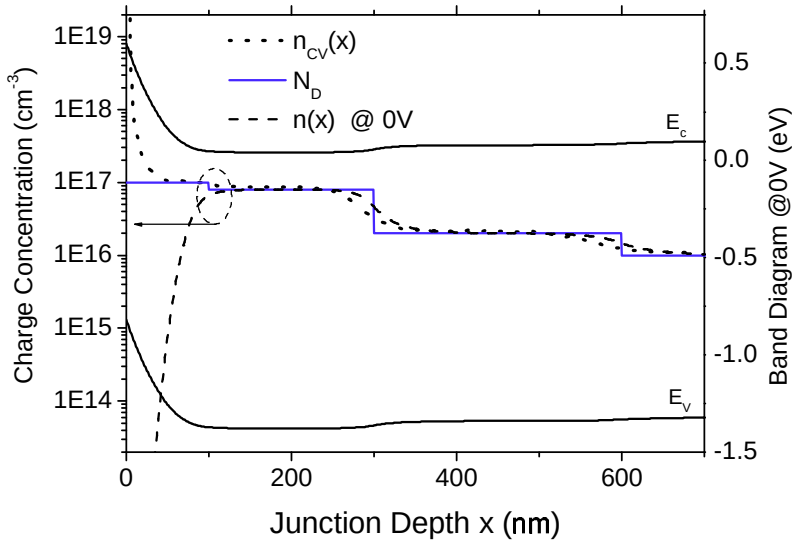


Fig. [3.6] Simulation of band diagram and free charge profile @ 0V and of apparent charge concentration in a Schottky-n GaAs junction with abrupt steps in the doping concentrations.

3.2.3. Effects of quantum confinement of carriers

When the carriers are confined in Quantum Wells, their distribution should be calculated by means of a self-consistent Schrödinger-Poisson (SP) procedure. The results of such a procedure are illustrated in fig. [3.7] for a double quantum well (DQW) nitride-based heterojunction, without consideration of polarization-induced fields. In figure, the band diagram and the distribution of electrons in the conduction band $n(x)$ have been calculated with a self-consistent SP solver for the unbiased structure, and compared with the profile calculated classically, with the only application of Fermi-Dirac statistics for the carrier distribution. Various studies can be found in the literature on the interpretation of C-V characterization on QW heterostructure junctions: it can be demonstrated that this technique is very powerful, as it can give important information about the structure of the junction, evidencing the charge accumulation peaks due to the presence of confined carriers in the QWs [Tschirner96], [Moon98], [Zervos02]. Although a classical approach can be regarded as satisfying in many cases, because it accounts for the main broadening effect of the apparent charge distribution, i.e. the Debye broadening [Tschirner96], the SP procedure has to be considered more

accurate [Moon98], as it also calculates the effects of the broadening related to quantum confinement. As it can be seen from fig. [3.7], the quantum calculation has the main features of smoothing the charge distribution in the wells and raising it in the barrier regions close to the wells. In most cases, both procedures yield very similar band diagrams.

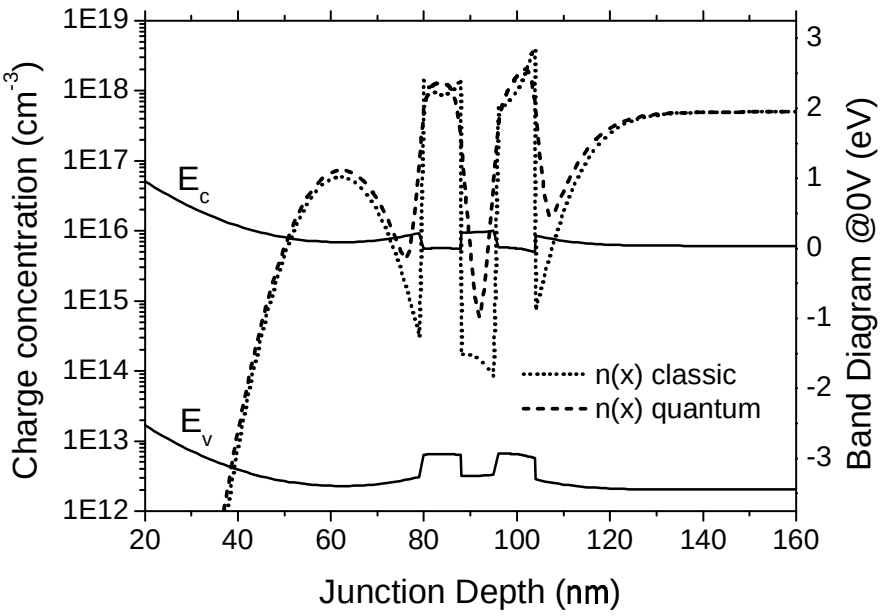


Fig. [3.7]: Simulation of band diagram and conduction band electron concentrations calculated by a classical (dotted line) and a self-consistent Schrödinger-Poisson procedure (dashed line) in a DQW nitride-based heterostructure at zero bias.

The apparent charge density profile $n_{CV}(x)$ is found for QW structures in the same way described in section 3.2.1. A simulation of a typical measurement is illustrated in fig. [3.8]. Here a nitride-based structure is considered. Two $In_{0.1}Ga_{0.9}N$ wells are embedded in a GaN n-type layer, with uniform doping concentration in the bulk and with varying doping concentration in the active QW region. The effect of electric fields arising from piezoelectric charges at the interfaces of layers with different composition is also taken into account. In fig. [3.8].a) the capacitance-voltage and the $1/C^2$-voltage characteristics are shown. It is possible to see in the latter curves the variations of the slope corresponding to the variations of apparent charge concentrations. These variations are then evidenced in fig. [3.8].b), where the $n_{CV}(x)$ profile is shown, together with the $n(x)$ profiles calculated classically and quantum mechanically and with the band diagram for the unbiased structure. The $n_{CV}(x)$ profile is calculated basing on the classical calculation. It is apparent that the information given by the $n_{CV}(x)$ profile must be critically interpreted. First of all, the $n_{CV}(x)$ profile appears as much more broadened than both $n(x)$ profiles, what is mainly due to the fact that in the proximity of the QWs it is not possible to define an edge of the depletion region properly. Secondly, one of the peaks corresponding to the QW on the bulk side is not clearly resolved. Third, the apparent charge peaks are shifted towards the bulk with respect to the conduction band peaks. The accuracy of the information given by the C-V characterization of QW structures varies with all the structure parameters describing the junction: doping profile, In fraction in the QWs, well thickness, barrier thickness, number of QWs, temperature. The effect of some of these parameters on the accuracy of the C-V measurement are studied in a work by Moon et al. [Moon98]. In any case, C-V characterization allows one to visualize the position and the extension of the depletion region. In some cases, if the inter-well distance is large enough, it makes it possible to assess the number of QWs present in the structure.

Experimental C-V characteristics can be fitted by using Schrödinger-Poisson solvers. However, due to the abundance of parameters in the fitting procedure, it is not convenient to extract the structure parameters by fitting the C-V characteristic. Nevertheless, fitting can be used as a guide

for interpreting the experimental data in particular cases, for instance, when a sample exhibits different C-V characteristics before and after certain treatments (current stress, heating, passivation, etc.).

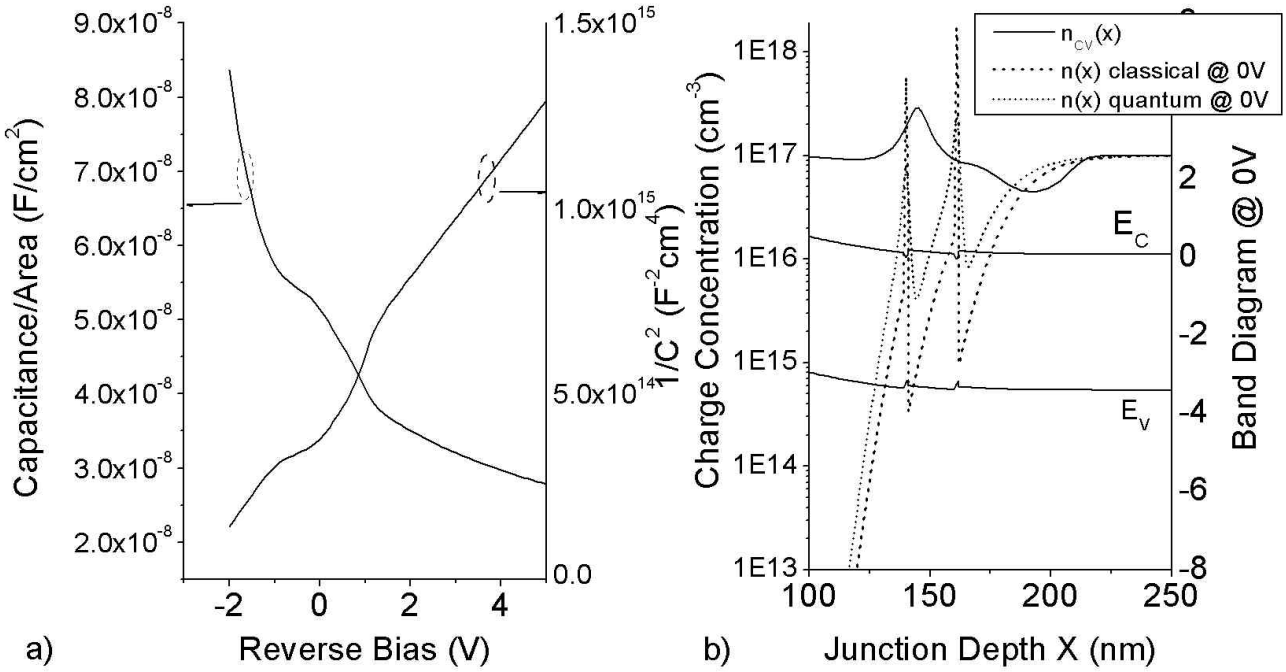


Fig. [3.8]. a) Simulation of C-V and 1/C²-V characteristics of a GaN-$In_{0.1}Ga_{0.9}N$ DQW one-sided junction. b) Simulation of the $n_{CV}(x)$ profile (solid line) and of the conduction band electron concentrations for the unbiased junction, calculated classically (dotted line) and with a self-consistent SP procedure (dashed line). The band diagram of the unbiased junction is also shown.

3.2.4. Series resistance and leakage current

A simplified equivalent circuit for a semiconductor diode (p-n or Schottky) under reverse bias is shown in fig. [3.9]. The standard method of measuring the capacitance consists in determining the component of the junction current that is 90° out of phase with the AC voltage superimposed to the constant DC bias. A resistance in series with the junction capacitance produces an error in the capacitance measurement, as it induces a change in the phase angle. Under most conditions, the analysed samples have a low enough series resistance, so that the capacitance measurement can be considered as accurate. Significant contributions to the series resistance can be due to the contacts (Schottky and ohmic), to the external circuitry or to a high resistivity of the semiconductor layers. The first two factors are usually easy to minimize. However, capacitance measurements of wide bandgap semiconductors can become difficult at low temperatures, due to the so-called freeze-out effect. Shallow dopant levels, in fact, are at a distance of at least 0.1 eV from the respective band, much more than in Si or GaAs: when the temperature drops to values lower than 100 K, the free carriers begin to occupy the dopant states, thus depleting the bands. The net effect is a general increase of the resistivity, and therefore of the series resistance R_s. Thus, at temperatures below 100 K the capacitance measurement of junctions based on wide-bandgap semiconductors becomes inaccurate.

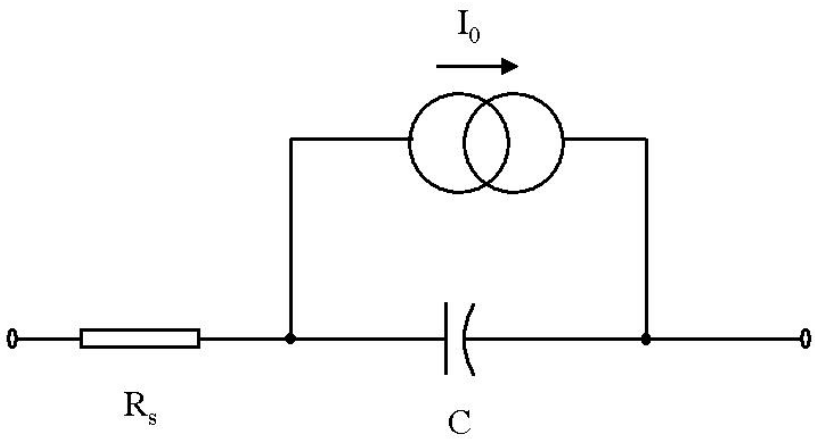


Fig [3.9]. The simplified equivalent circuit for a semiconductor diode during a CV measurement

The formula relating the capacitance C_m measured at a frequency $f_{CAP} = 2\pi\omega$ to the real junction capacitance is the following [Schroder]:

$$C_m = \frac{C}{1+\omega^2 R_s^2 C^2}. \qquad (3.49)$$

From this formula one can see that for high values of the series resistance the measured capacitance is a decreasing function of the real capacitance. This can seriously affect the interpretation of DLTS results, which are based on measurements of capacitance differences, occurring at low temperatures.

3.2.5. Experimental setup

The block diagram of the instrumentation used for the C-V characterization of the samples analysed in this thesis is depicted in fig. [3.10]. The building blocks of the setup are the Lakeshore 330 temperature controller, which makes it possible to perform measurements at different temperatures, the Keithley 230 bias generator and the Keithley 3330 LCZ meter. The LCZ meter has operating frequencies ranging from 120 Hz to 10^5 Hz. Measurements with frequency 1MHz have also been performed by using a Boonton capacitance meter. The measurements is driven by a software, with the following parameters:

- The bias interval V_{max}, V_{min}
- The bias step ΔV
- The time interval Δt between two successive bias values
- The LCZ meter operating frequency f_{CAP}

The same setup can be used for slightly different measurements, such as capacitance-frequency (C-*f*) characterization and admittance spectroscopy.

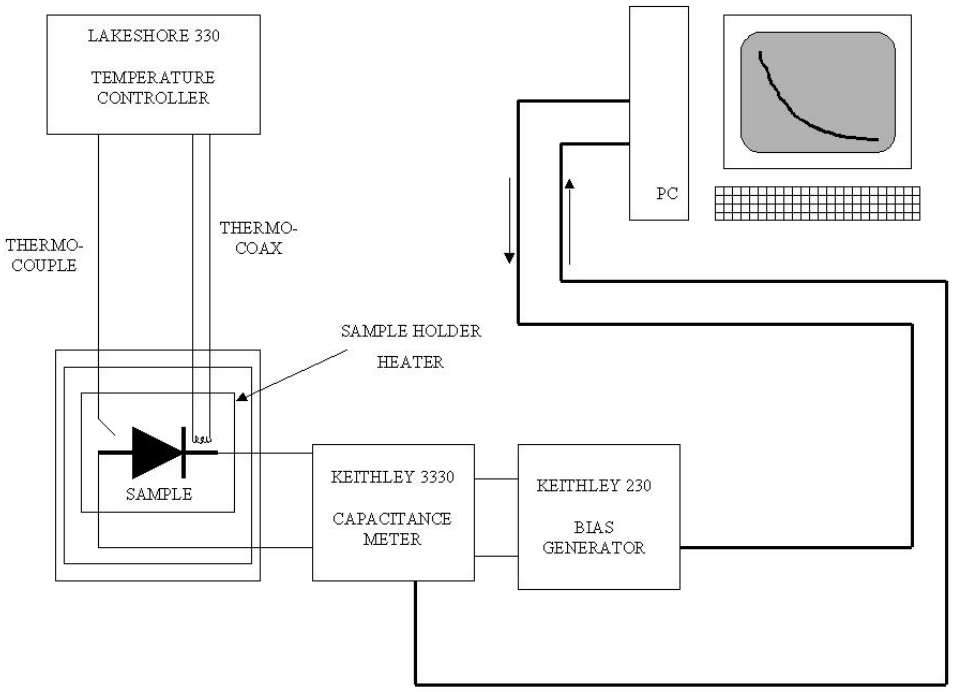


Fig. [3.10]. Block diagram for capacitance-voltage characterization.

REFERENCES

[Fang00] Z.Q. Fang, D.C. Reynolds and D.C. Look, *J. Elec. Mater*. **29**, p.448 (2000).

[McKelvey] J.P. McKelvey, Solid state and semiconductor physics, (Harper, 1967)

[Moon98] C.R. Moon, B.D. Choe, S.D: Kwon, H.K. Shin and H. Lim, *J. Appl. Phys.* **84**, p. 2673 (1998).

[Neamen] D.A. Neamen, *Semiconductor physics and devices*, (McGraw Hill, New York, 2003)

[OrtonBlood] Blood P and Orton J W *The electrical characterization of semiconductors: majority carriers and electron states* (London: Academic Press, 1990)

[Rhoderick] W. H. Roderick, Williams, *Metal-Semiconductor Contacts*

[Schroder] D. K. Schroder, *Semiconductor Material and Device Characterization* (Wiley-Interscience, New York, 1998)

[Schubert] E.F. Schubert, *Light-Emitting Diodes*, (Cambridge University Press, Cambridge, 2003)

[Shah03] J.M. Shah, Y.L. Li, T. Gessmann and E.F. Schubert, *J. Appl. Phys*. **94**, p. 2627 (2003)

[Shockley50] W. Shockley, "*The theory of p-n junctions in semiconductors and p-n junction transistor*" Bell Syst. Tech. J. 435 (1949)

[Stradling] *Growth and characterization of Semiconduc*tors, edited by R.A. Stradling and P.C: Klipstein, Adam Hilger, Bristol, 1990.

[Sze] S. M. Sze, *Physics of semiconductor devices*, Wiley, 1981

[Torvik98] J. T. Torvik, M. Leksono, J. I. Pankove, B. Van Zeghbroeck, H.M. Ng and T.D. Moustakas, *Appl. Phys. Lett.* **72,** p. 1371(1998)

[Tschirner96] B.M. Tschirner, F. Morier-Genoud, D. Martin and F.K. Reinhart, *J. Appl. Phys.* **79**, p. 7005 (1996)

[Zervos02] M. Zervos, A. Kostopoulos, G. Constantinidis, M. Kambayaki and A. Georgalikas, *J. Appl. Phys.* **91** p.4307 (2002)

4 Thermal Spectroscopy

4.1. Deep Levels and theory of capacitance transients

4.1.1. Shallow levels and deep levels. Shockley-Read statistics.

In an ideal semiconductor there are no electron levels within the energy gap. In reality, defects or the proximity of the surface make possible the existence of such levels. A commonly accepted way of categorizing gap levels is to divide them into shallow levels and deep levels. Shallow levels lie in narrow bands of few meV from the conduction or valence bands, while deep levels are energetically more distant from the bands.

A typical example of the first category are the electron states introduced by means of dopant impurities. The Schrödinger equation of the system *crystal + impurity* may, in certain cases, be treated with the so-called effective mass approximation [Davies], leading to a solution for the bound states of the impurity similar to that of the hydrogen atom. The energy of the electron state is thus interpreted as an ionisation energy with respect to the correlated band, i.e. to valence band for holes on shallow acceptor states, and to conduction band for electrons on shallow donor states. This energy is estimated by the first term of a modified Rydberg series:

$$E_{d,a} = \frac{m_{n,p}^{*}\varepsilon_0^2}{m\varepsilon_s^2}13.6 \quad eV\,, \qquad (4.1)$$

where m is the electron mass, $m_{n,p}$* is the effective electron or hole mass, and ε_s is the permittivity of the material.

Deep states have a higher ionisation energy E_t than the corresponding hydrogenic state. Therefore, they are called so because of their deeper position in the energy gap. Deep levels are mostly related to defects in the material: intrinsic defects, impurities, extended defects, etc. Some of these defects are monovalent, i.e., they have only two possible charge states, some other may have more charge states. In the following Shockley-Read statistics is illustrated for just monovalent centres, although it is possible to generalize the theory to multi-valent centres [Look81]. Negative-U centres, for instance, are multi-valent centres which can capture typically two electrons, and for which the second captured electron is more tightly bound than the first one (unlike the common case in which successive electrons bound to a certain centre are more and more loosely bound): as a consequence, they tend to capture and emit two electrons at once, in order to acquire the minimum energy configuration.

Deep levels may be distinguished into donor or acceptor centres according to the charge they acquire in the case of occupation by an electron: donor centres are neutral if occupied by an electron and positively charged ($+e$) if ionised; acceptor centres have charge $-e$ when occupied by an electron and neutral if empty. Multi-valent centres may be amphoteric, assuming donor or acceptor character according to their charge state.

Deep states are strongly localized spatially; as a consequence, there is a strong delocalisation in the momentum space $\boldsymbol{k}$; therefore, the transitions mediated by them couple to a large number of phonons and tend to be non-radiative. In order to define the characteristic parameters of a deep centre, we use the scheme illustrated in fig. [4.1], where the four possible recombination and generation processes for a generic deep level with ionisation energy E_t (with respect to the conduction band) and concentration N_t are shown, along with the respective expressions of transition rate per unit time and volume.

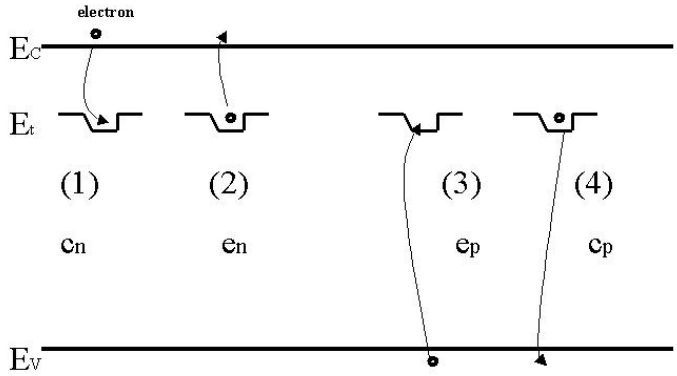


Fig. [4.1] Emission and capture processes at a deep level: (1) electron capture from the conduction band, (2) electron emission to the conduction band, (3) hole emission to the valence band, (4) hole capture from the valence band. The small circle represents an electron.

Indicating as $\sigma_{n,p}$ the capture cross sections of the deep state for electrons and holes, respectively, the capture probabilities are:

$$c_n = \sigma_n < v_n > n \quad \text{for electrons, and} \tag{4.2}$$

$$c_p = \sigma_p < v_p > p \quad \text{for holes} \tag{4.3}$$

where $<v_{n,p}>=(3kT/m^*_{n,p})^{-1/2}$ is the average thermal velocity of the charge carrier, and n, p are the concentrations of free electrons and holes. Thus, capture probabilities are temperature-dependent. The ratio of electron and hole capture probability determines the deep state behaviour. Deep levels are therefore classified as:

- recombination centre, if $c_n \approx c_p$
- electron trap, if $c_n >> c_p$
- hole trap, if $c_n << c_p$

The emission probabilities can be found from the detailed balance of capture and emission at the generic level t at thermal equilibrium. One has for electrons:

$$e_n N_t f(E_t) = c_n N_t [1 - f(E_t)] \tag{4.4}$$

with the level occupation probability

$$f(E_t) = \frac{1}{1 + g^{-1} \exp \dfrac{E_t - E_f}{kT}} \tag{4.5}$$

whereby g is the degeneracy factor and E_f is the Fermi level at temperature T. From eq. (4.2), substituting $n = N_c \exp\left(\dfrac{E_f - E_C}{kT}\right)$ one obtains the emission rate for electrons:

$$e_n = \frac{\sigma_n < v_n > N_C}{g} \exp\left(\frac{E_t - E_C}{kT}\right) \tag{4.6}$$

and an analogous result for holes

$$e_p = \frac{\sigma_p < v_p > N_V}{g} \exp\left(\frac{E_t - E_V}{kT}\right) \qquad (4.7)$$

The properties of the deep centre may thus be categorized basing on the emission properties; one has:

- recombination centre, if $e_n \approx e_p$
- electron trap, if $e_n >> e_p$
- hole trap, if $e_n << e_p$

The dynamics of the thermal generation-recombination process is governed by the continuity equation expressing the time variation of the occupation of the state. Indicating with $n_t(t)$ and with $p_t(t)=N_t-n_t(t)$ the number of captured electrons and holes, respectively, one has:

$$\frac{dn_t(t)}{dt} = (c_n + e_p)\, p_t(t) - (c_p + e_n) n_t(t) = a(N_t - n_t(t)) - b n_t(t) \qquad (4.8)$$

with $a=c_n+e_p$ and $b=c_p+e_n$. If the carrier concentration are steady ($c_{n,p}$ and $e_{n,p}$ constant), it is easy to integrate (4.8) with the initial condition $n_t(0)$ imposed by the experimental conditions. One has then:

$$n_t(t) = n_t(\infty) + [n_t(0) - n_t(\infty)] \exp[-(a+b)t] \qquad (4.9)$$

For $t \to \infty$ the concentration of electrons on level t is:

$$n_t(\infty) = \frac{a}{a+b} N_t = \frac{c_n + e_p}{c_n + e_p + c_p + e_n} N_t \qquad (4.10)$$

Responding to a perturbation which varies their occupation, the electron population evolves exponentially with the following time constant

$$\tau = [a+b]^{-1} = [c_n + e_n + c_p + e_p]^{-1} \qquad (4.11)$$

independently of the concentration of the deep level N_t. Significant simplifications in eqs. (4.9 – 4.11) may be due to the fact that only one of the four capture/emission coefficients is dominant in most experimental conditions [Miller77].

4.1.2. Deep levels and junction capacitance

We consider now, as a first approximation, a Schottky or a p^+-n junction with only one deep state t present, with concentration $N_t << N_D$ uniform in depth, introducing only one level at energy E_t from the conduction band. Figure [4.2] shows, as an example, the n-side of a Schottky or a p^+-n junction at a certain bias in presence of a donor trap (positively charged if empty, neutral if occupied by an electron. In figure one can see the band bending in the space charge region of thickness w associated with the positive charge of the shallow donors. The Fermi level crosses level

E_t at y, yielding different occupation conditions. In the region $0<x<y$ one can make use of the depletion approximation, which yields, together with (4.2, 4.3)

$$c_n = c_p \approx 0 \qquad (4.12)$$

so that emission processes dominate the kinetics of the traps; the continuity equation reads in this case:

$$\frac{dn_t(t)}{dt} = e_p p_t(t) - e_n n_t(t) \approx -e_n n_t(t)\,. \qquad (4.13)$$

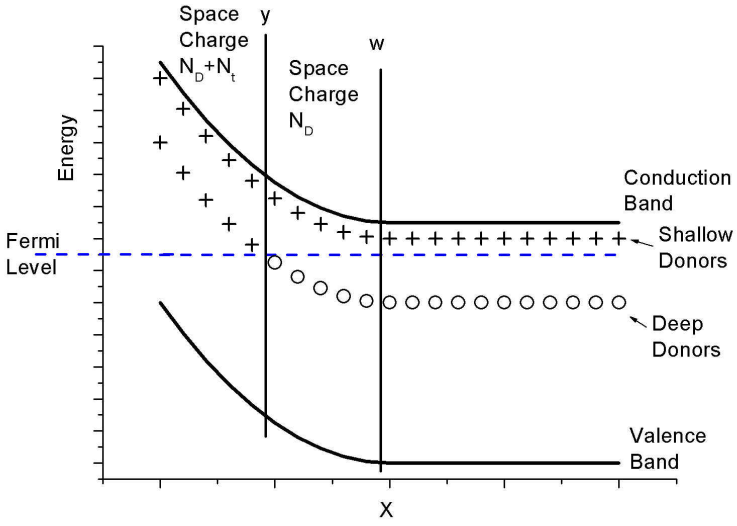


Fig. [4.2] One-sided junction with a donor deep level

Being also $e_n >> e_p$, the steady-state electron population of the trap is given by

$$n_t(\infty) = \frac{e_p}{e_p + e_n} N_t \approx 0\,. \qquad (4.14)$$

Thus, the space charge density associated with traps results equal to $+eN_t$, as non-occupied donor traps are singly positively charged. In the transition region $y<x<z$ capture phenomena should not be neglected, as diffusion of free carriers from the neutral region can yield a non-zero free carrier density, which can be usually described by

$$n(x) = n(w)\exp\left[-\frac{(w-x)^2}{L_D^2}\right], \qquad (4.15)$$

where $L_D = \sqrt{\frac{\varepsilon_s kT}{e^2 N_D}}$ is the intrinsic Debye length. The width of the transition region, where shallow donors are empty but the trap is occupied, can be calculated from the Poisson equation, yielding, for a Schottky barrier

$$\Delta w = w - y = \sqrt{\frac{2\varepsilon_s}{e}\frac{\phi_b - V}{N_D}}\,. \qquad (4.16)$$

The effects of the transition region can be neglected in a first approximation, provided the applied bias is such that

$$V >> (E_f - E_t)/e - \phi_b \ . \qquad (4.17)$$

In the following, this hypothesis will be considered valid. Corrections will be developed, if needed, in the sections dealing with the experimental data. Analogous results can be obtained when considering a deep level of acceptor type or a p-type semiconductor. Further corrections to the theory in cases of heterostructure devices or non-uniform shallow dopant concentration, will be treated in the sections dealing with experimental results.

Table 4.1 summarizes the different possible cases of contribution to the space charge by deep levels in the depletion region.

Table 4.1: Density N_I of charged centres in the depletion region, according to semiconductor type and deep level type.

Deep Level type	metal-*n* or p^+-*n* junction	metal-*p* or n^+-*p* junction
Donor	N_D+p_t	N_A-p_t
Acceptor	N_D-n_t	N_A+n_t

The measurement of the differential capacitance of the junction requires the superposition of a small oscillating voltage ΔV_ω with frequency ω on the reverse bias V. When deep levels are present, the corresponding charge variation is

$$\Delta Q = \Delta Q(y) + \Delta Q(w)\,, \qquad (4.18)$$

where the first term is related to the variation of trap occupation at depth y, and the second term to the variation of the space charge region depth. For the example in fig. [4.2] one can write, considering a small increment

$$dQ = eN_t^- dy + e(N_D^+ - N_t^-)dw\,. \qquad (4.19)$$

The measurement is conditioned by the time constant of the traps, i.e., at high frequency, with

$$\omega >> e_n \approx 1/\tau \qquad (4.20)$$

the response of deep level to the solicitation is too slow, and it remains continuously at the occupation level defined by the polarization V. Therefore, the measurement is sensitive only to the charge variation taking place at the depletion region edge w. If, for instance, only one deep level, of acceptor type, with density N_t and with $E_t<E_C-E_f$, the effective density sensed by the capacitance measurement is N_D-N_t.

A different case is at low frequency: if

$$\omega << e_n \approx 1/\tau \qquad (4.21)$$

is valid, the traps at depth y are able to discharge, so that the contribution to measurement consists of both states at $x=w$ and at $x=y$. In this case, with an acceptor trap, the charge density is N_D-N_t at $x=w$ and N_t at $x=y$. Therefore, the capacitance depends, approximately, on N_D only. In the case of high frequency (4.20) the charge variation is only at the depletion region edge w, which can be estimated, neglecting the transition region, as

$$w = \sqrt{\frac{2\varepsilon_s}{e}\frac{\phi_b + V}{N_I(\infty)}}\,, \qquad (4.22)$$

where $N_I(\infty)$ is the total space charge in the steady state, with contributions from both shallow and deep levels, as in tab. 4.1. Therefore, the differential capacitance is given by [Miller77], [OrtonBlood]

$$C_0 = \frac{\varepsilon_s}{w} = \sqrt{\frac{e\varepsilon_s}{2}\frac{N_I(\infty)}{\phi_b + V}} \qquad (4.23)$$

4.1.3. Capacitance-Voltage measurements in presence of deep levels. Compensation

Deep levels in a semiconducting material may give rise to *compensation* effects, i.e., to a rearrangement of the free carrier concentration in the bands [LookGaAs]. Acceptor-like levels, for instance, subtract a net amount of free charge from the conduction band in n-type materials. The presence of deep levels may significantly affect the capacitance-voltage (C-V) characteristics.

C-V measurements, as described in chapter 3, provide a method to measure the profile of the apparent charge density n_{CV} in a Schottky junction. The apparent charge density n_{CV} is in most cases roughly equal to the free charge density n in conduction band (for n-type material). The differences between these two quantities are due to the presence of deep levels in significant concentration. The present description is limited to the case of deep levels in a one-sided n-type junction. We also assume here, for simplicity, that the concentration of doping and deep levels is uniform in the analysed region. This leads us also to drop the x dependence in the expressions for the apparent charge density n_{CV} and for the conduction band electron density n.

The contribution of deep levels to the measurement of n_{CV} can be classified into two main categories: first, acceptor-like deep levels tend to lower n at the edge of the depletion region and in the neutral region, i.e., they *compensate* the material. Secondly, deep levels (donor- or acceptor-like) in high concentration provide charge oscillations due to trapping and de-trapping of free carriers in that part of the depletion region where their energy level and the Fermi level cross [Miller77], [OrtonBlood].

Figure [4.3] reports an example of a Shockley diagram used for the calculation of n in compensated 4H-SiC. In the illustrated case, we report deep levels introduced by proton irradiation. The Shockley analysis is based on the occupation number of each level according to Fermi-Dirac statistics. An analogous statistics describes the occupation of multi-valent centres [Look81]. The neutrality condition, which is strongly influenced by the presence of deep levels in significant concentration, determines the position of the Fermi level E_f. This, in turn, determines the concentration of free carriers in the conduction band:

$$n(T) = N_C\left(\frac{T}{300}\right)^{3/2}\exp\left[\left(E_f - E_g\right)/kT\right], \qquad (4.24)$$

where T is the temperature in Kelvin, k is the Boltzmann constant, E_g is the energy gap and N_c is the equivalent density of states in the conduction band of the semiconductor.

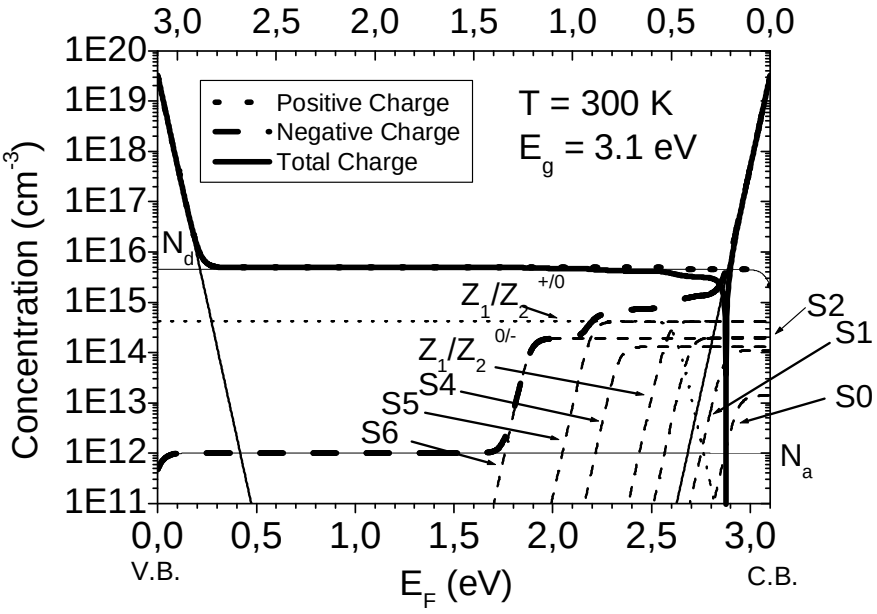


Figure [4.3] Shockley diagram for 4H SiC with a set of deep levels in the upper bandgap.

In order to calculate the expected n_{CV} the response of the deep levels to the oscillating voltage of the capacitance meter must be considered. As is well-known [OrtonBlood], this response is governed by the occupation level of the traps and by their emission rate for electrons:

$$e_{n,t}(T) = \gamma_n \sigma_{n,t} T^2 \exp\left(\frac{E_t - E_C}{kT}\right), \qquad (4.25)$$

where $\sigma_{n,t}$ is the deep level capture cross section for electrons, and E_t the defect enthalpy, and $\gamma_n = \frac{4\pi m_n^* k^2}{gh^3}\sqrt{6\pi}$ is a constant material parameter. Temperature variations allow the monitoring of variations of capacitance, and therefore of apparent charge density due to deep levels. The expected n_{CV} can be modelled as a function of temperature T in the following way:

$$n_{CV}(T) = n(T) + \sum_t N_t \left(1 - \frac{1}{1 + \left(e_t(T)/2\pi f_{CAP}\right)^2}\right), \qquad (4.26)$$

where N_t is the deep level concentration and f_{CAP} is the capacitance meter frequency. The concentration of a negative-U centre must be multiplied by a factor two in order to account for the double electron emission, and similar care must be taken for multi-valent centres. The lorentzian function by which the deep level density is multiplied in the second summand of eq. (4.26) is somewhat arbitrarily chosen, but it has the property of being equal to one if $e_t << f_{CAP}$, and zero if $e_t >> f_{CAP}$. This occurs because when the emission rate is lower than the operating frequency, a net contribution to the measured charge is given by the deep levels in the region where their energy levels cross with the Fermi level [Miller77]. Equation (4.26) is a good enough estimate of n_{CV}, provided the reverse bias is high enough for the so-called λ effect to be neglected [OrtonBlood].

4.1.4. Capacitance transients

Deep levels are characterized according to relations (4.6) and (4.7). By reducing the applied reverse bias from *V* to *V'*, the space charge region width diminishes from *w* to *w'* very fastly, compared to the typical response times of the generation-recombination mechanisms related to deep levels. Therefore, the concentration of free carriers can be regarded as stationary in the region $w'<x<w$. As the free carrier concentration is much higher than under equilibrium conditions (bias *V*), the condition $c_p,c_n>>e_p,e_n$ is valid. The time constant for the charging up of deep levels is then:

$$\tau_c = 1/(c_n + c_p) \quad , \tag{4.27}$$

while the steady state electronic occupation is

$$N_I(\infty) = \frac{c_n}{c_n + c_p} \approx \begin{cases} N_t & \text{for electron traps} \\ 0 & \text{for hole traps} \end{cases} \quad . \tag{4.28}$$

Upon rapidly changing the bias to the previous value V, the occupied traps in the depletion region empty with the following time constant:

$$\tau_d = 1/(e_n + e_p) \quad . \tag{4.29}$$

Neglecting again the transition region, one can write the depletion region width as a function of time

$$w(t) = \sqrt{\frac{2\varepsilon_s}{e}\frac{\phi_b + V}{N_I(t)}} \quad , \tag{4.30}$$

and the high-frequency differential capacitance as

$$C(t) = \frac{\varepsilon_s}{w(t)} = \sqrt{\frac{e\varepsilon_s}{2}\frac{N_I(t)}{\phi_b + V}} \quad . \tag{4.31}$$

Summarizing these results, it is to be noticed that the emission of a majority charge carrier from a trap, either acceptor or donor type, corresponds always to a net increase of the space charge, because in this case a shallow ionized centre is no longer compensated. On the contrary, the emission of a minority charge carrier yields a net decrease of the space charge. Therefore, it may be useful to categorize the deep centres as follows:

- majority carrier traps if $e_{maj} >> e_{min}$
- minority carrier traps if $e_{maj} << e_{min}$

4.1.5. Transient spectroscopy. Majority and minority carrier trap parameters

In the case of a Schottky-*n* or a p^+-*n* junction the width of the space charge region can be decreased by means of a bias pulse in forward direction *V'*; after that, the bias can be set back to the previous value *V*. The electron traps charge up during the pulse and discharge after the pulse, so that the high-frequency capacitance has a trend like that illustrated in fig. [4.4], and is given by

$$C(t)=\frac{\varepsilon_s}{w(t)}=\sqrt{\frac{e\varepsilon_s}{2}\frac{N_I(\infty)-n_t(t)}{\phi_b+V}}\quad , \qquad (4.32)$$

where $n_t(t)$ is the solution of eq. (4.9) with $n_t(0)=N_t$, i.e. $n_t(t)=N_t exp(-e_n t)$. Under the hypothesis of uniform concentrations and with $N_t<<N_D$, the relative variation of capacitance can be approximated as follows:

$$\frac{\Delta C(t)}{C_0}=\frac{C(T)-C_0}{C_0}=\sqrt{1-\frac{n_t(t)}{N_I(\infty)}}-1\cong-\frac{n_t(t)}{2N_I(\infty)}=-\frac{N_t\exp[-e_n t]}{2N_I(\infty)} \qquad .(4.33)$$

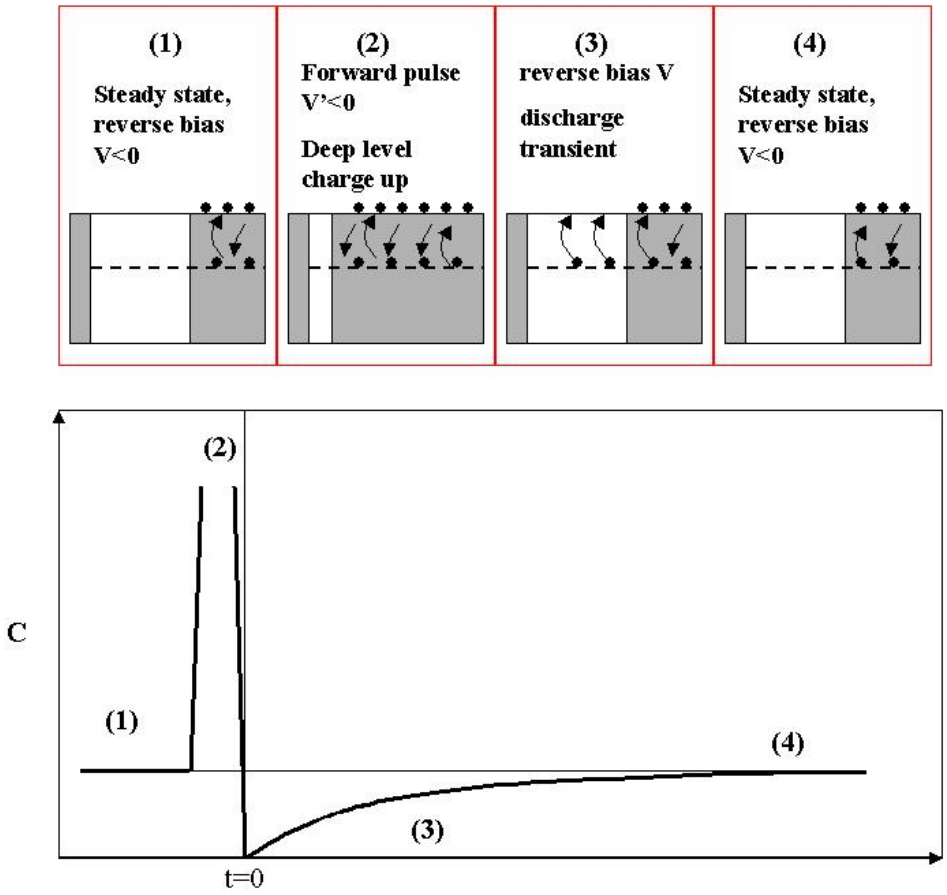


Fig. [4.4] Effect of a forward bias pulse applied on a reversely biased junction in presence of a majority carrier trap.

This quantity, calculated for *t*=0 yield, as a first approximation, the concentration (which is underestimated, if not all the traps are occupied at t=0) of the deep level:

$$N_t \approx -2N_I \frac{\Delta C(0)}{C_0}. \qquad (4.34)$$

In the case of a p^+-n junction it is possible to generate transients due to minority carrier traps; this is not possible in a Schottky junction, which is unipolar. Minority trap-related capacitance transients are generated by injecting minority carriers by means of a forward bias pulse, with $V'>0$, as illustrated in fig. [4.5]. In this case, under the usual conditions of doping concentration and deep level concentration uniformity, and with $N_t << N_D$, the relative variation of capacitance is

$$\frac{\Delta C(t)}{C_0} = \frac{C(t)-C_0}{C_0} = \sqrt{1+\frac{n_t(t)}{N_I(\infty)}} - 1 \cong \frac{n_t(t)}{2N_I(\infty)} = \frac{N_t \exp[-e_p t]}{2N_I(\infty)}. \qquad (4.35)$$

From this relation the concentration of minority carrier traps can be estimated as

$$N_t \approx 2N_I \frac{\Delta C(0)}{C_0}. \qquad (4.36)$$

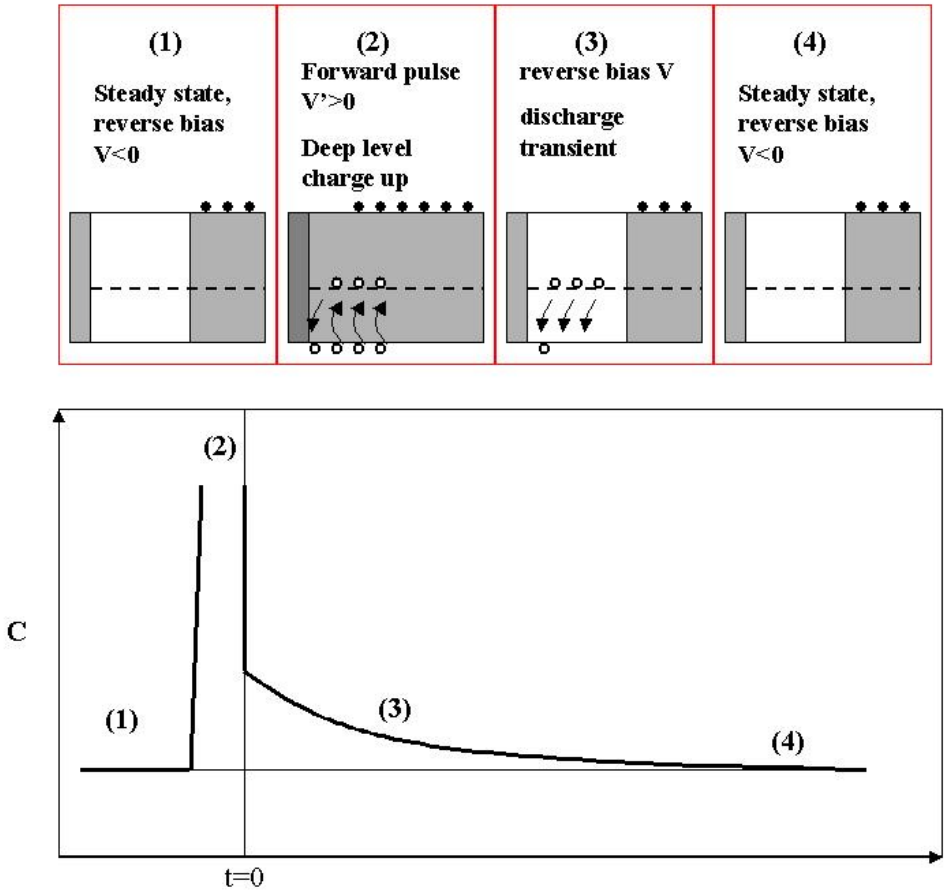


Fig. [4.5] Effect of a forward bias pulse with *V'>0* applied on a reversely biased junction in presence of a minority carrier trap.

From (4.6) and (4.7) one can notice that $e_{n,p}$ are function of temperature. Being $<v_{n,p}> \propto T^{1/2}$ and $n,p \propto T^{3/2}$, one can re-express the emission rates as:

$$e_{n,p} = \gamma_{n,p}\sigma_{n,p}T^2 \exp\left(-\frac{\Delta G}{kT}\right) \tag{4.37}$$

$$\gamma_{n,p} = \frac{4\pi m^*_{n,p}k^2}{gh^3}\sqrt{6\pi} \tag{4.38}$$

where γ is a constant material parameter, $\Delta G = T\Delta S + \Delta H$ is the variation in the Gibbs free energy accompanying the emission of an electron $\Delta H = E_C - E_t$ for electron traps, $\Delta H = E_t - E_V$ for hole traps. In any case it should be noted that E_t is not the exact value of the energy level in the gap, but an approximation, as the energy distance of the deep level from the conduction or valence band corresponds to the variation of the free energy ΔG. Under the hypothesis that $\sigma_{n,p}$ is independent of temperature, one can re-write the previous equation as follows (here in the case of electron traps):

$$\ln\frac{T^2}{e_n} = \ln(\gamma_n\sigma_n) - \frac{\Delta S}{k} + \frac{\Delta H}{kT} \quad . \tag{4.39}$$

The experimental procedure is usually the following: one measures the time constants (or the emission rates) of a single trap at different temperatures, then plots the collected data in an Arrhenius plot [ln (T^2/e_n); 1000/T] (see fig. [4.6]). Thus, a straight line should fit the data. The slope of the line is proportional to the enthalpy associated to the emission of the carrier to the respective band, and the intercept is related to the capture cross section.

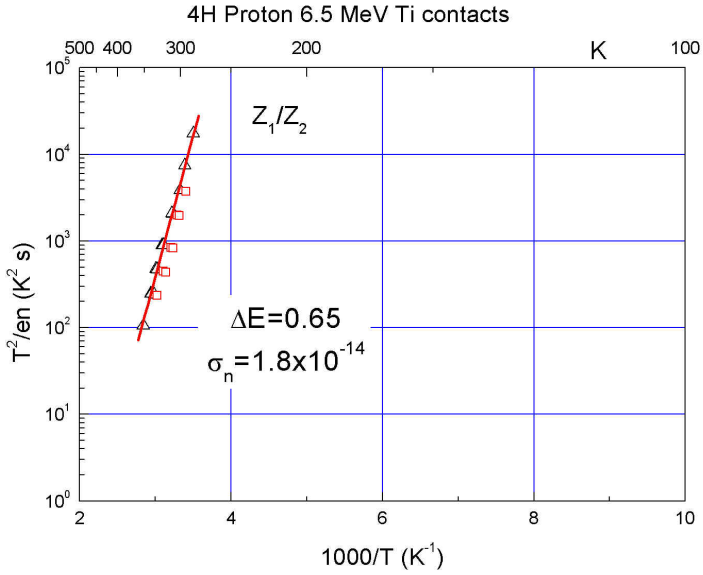


Fig. [4.6] Example of Arrhenius plot for Z_1/Z_2 centre in 4H-SiC. In figure the parameters of the trap are reported (the energy in [eV] units, the capture cross section in [cm^{-2}] units).

4.2.Deep level transient spectroscopy (DLTS)

The origins of DLTS technique go back to the 70's, when this experimental method was developed [Miller77]. This technique is still one of the most used and accurate methods for characterizing electrically active deep levels and determining their parameters. DLTS is based on a temperature scan of the capacitance transient of the analysed sample. The DLTS signal is univocally related to e_n by means of the *rate window* concept, which will be illustrated in the following. The measurement procedure is illustrated in fig. [4.7].

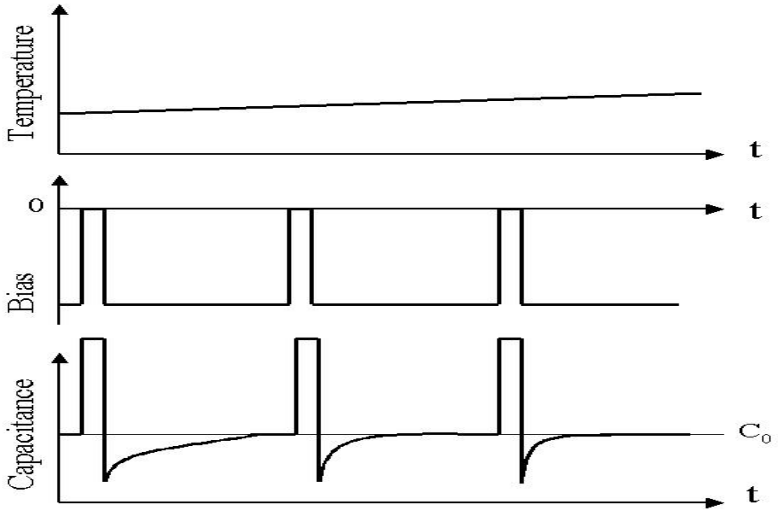


Fig. [4.7] Typical variation over time of significant quantities during a DLTS measurement.

A negative bias is applied to the sample, so that a certain region is depleted of free carriers. At each period P a pulse of duration t_p and amplitude V_p is applied to the junction. The duration of the pulse affects the filling level of the traps, while the amplitude can be related to the region probed by the bias pulse:

- $V_p<V$: in this case an interval inside the space charge region is selected. This can be useful if one is interested in concentration profiles of traps close to the surface or to the junction region, although more sophisticated techniques, such as DoubleDLTS, should be used for that.

- $V_p \approx V + V_{bi}$: the whole space charge region is repleted with free carriers.

- $V_p > V + V_{bi}$: In p-n junctions, minority carriers are injected from the opposite side of the junction, thus making possible the observation of minority carrier traps.

4.2.1. Rate window concept

One way to introduce the rate window concept is by considering the DLTS technique based on the boxcar correlator. In this technique, the capacitance transient subsequent to each bias pulse is sampled at two different times t_1 and t_2, and the difference between the capacitance values at these two instants is considered. This is the DLTS signal, which we consider as a function of temperature T:

$$S(T) = \Delta C(T) = C(t_1, T) - C(t_2, T)$$
, (4.40)

which, under the hypothesis that transients are exponential with inverse time constant $e_n = \sigma_n \gamma_n T^2 \exp(\Delta E/kT)$, is equivalent to the relation

$$S(T) = \frac{1}{2} N_t \{\exp[-e_n(T)t_1] - \exp[-e_n(T)t_2]\}. \quad (4.41)$$

In order to calculate the emission rate as a function of the temperature, one has to consider the point of maximum/minimum signal:

$$\frac{dS(T)}{dT} = \frac{dS(e_n)}{de_n}\frac{de_n}{dT} = 0. \quad (4.42)$$

This is equivalent with finding

$$\frac{dS(e_n)}{de_n} = 0 \quad (4.43)$$

as it is always verified that

$$\frac{de_n}{dT} \neq 0 \qquad \forall T. \quad (4.44)$$

The temperature at which the signal has a maximum (minimum for minority carrier traps) is thus correlated to the emission rate according to

$$e_n = \frac{t_2 - t_1}{\ln \frac{t_2}{t_1}}. \quad (4.45)$$

One sees that e_n is a quantity depending on the constants t_1 and t_2, which are fixed at the beginning of the measurement. The time interval t_2-t_1 is also called rate window, and the procedure is summarized in fig. [4.8].

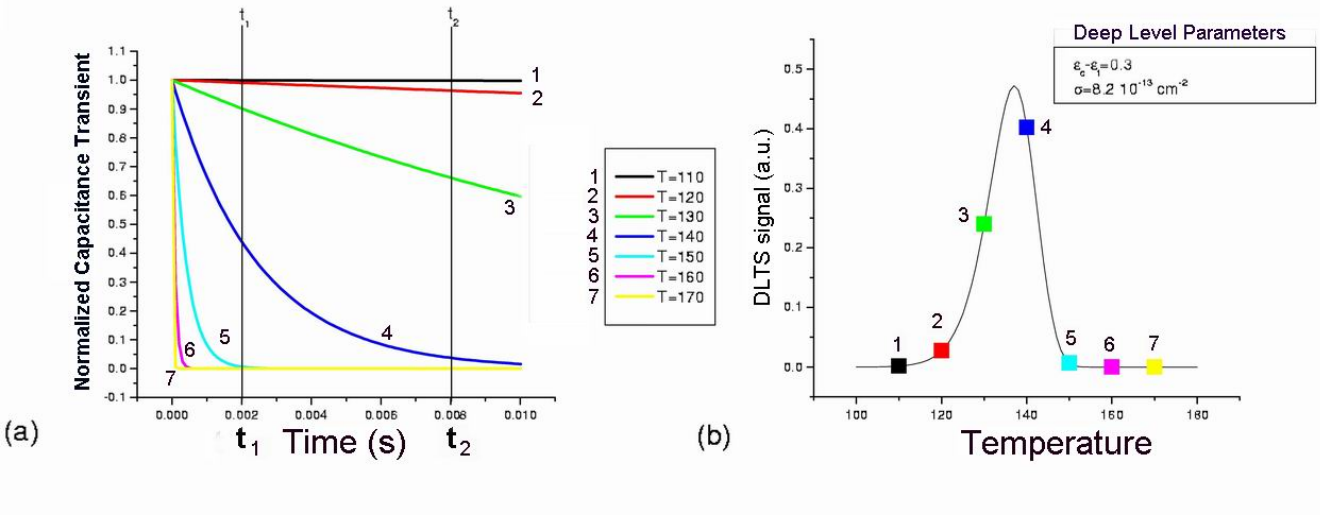


Fig. [4.8]. a) Discharge capacitance transients at different temperatures, with the rate window t_1-t_2 in evidence; b) associated DLTS signal.

4.3.DLTS with non-exponential transients

So far DLTS has been considered under the hypothesis that the emission processes have exponential transients. However, in many cases this hypothesis is insufficient. Many circumstances may lead, in fact, to non-exponential capacitance transients. First of all, eq. (4.36) is an approximation valid when $n_t << N_D$. If the trap concentration is not much lower than the doping concentration, the first part of the transient is no longer exponential. Secondly, there can be two deep levels in the gap with similar parameters; thus, their signal are superimposed to each other, and sometimes one deep level can not be distinguished from the other. More complex cases can occur, i.e. when a particular defect is associated with deep levels having a continuous spectrum of activation energies: in this case the single, discrete level with activation energy E_C-E_t is substituted by a deep level band described by a density of states $n_t(E)$. This last cases can be divided into other particular cases: the defect can distorce the bands if occupied by charge carriers, which also modifies the capture and emission kinetics. Again, the electron states giving rise to the continuous density of states can have a *localized* behaviour, if every level emits independently of the others, or a *bandlike* behaviour, if the captured carriers tend to rearrange very quickly in order to occupy the lowest-energy states of the defect. In the following we will analyse some of these cases, which are very important when considering electron emission from extended defects.

4.3.1. Superposition of several discrete deep levels

If two deep levels have similar parameters, the signal associated to their emission is the superposition of two different transients. The corresponding DLTS signal is the linear combination of the theoretical DLTS signals associated to both levels. If the trap parameters are very similar, it may not be possible to separate the contribution of either level. An example is reported in fig. [4.9], illustrating a simulation of the DLTS given by two levels with similar parameters. For many values of the rate window, the peaks can not be separated. However, one can notice that for the lowest emission rate (most left peak) a shoulder, corresponding to the first trap, appears in the spectrum. In such cases the most convenient procedure is to estimate the trap parameters from Arrhenius analysis (if the peaks are separable), then to double check the consistency of the results by fitting the DLTS signal with a Levenberg-Marquardt algorithm reproducing the DLTS spectrum yielded by two levels.

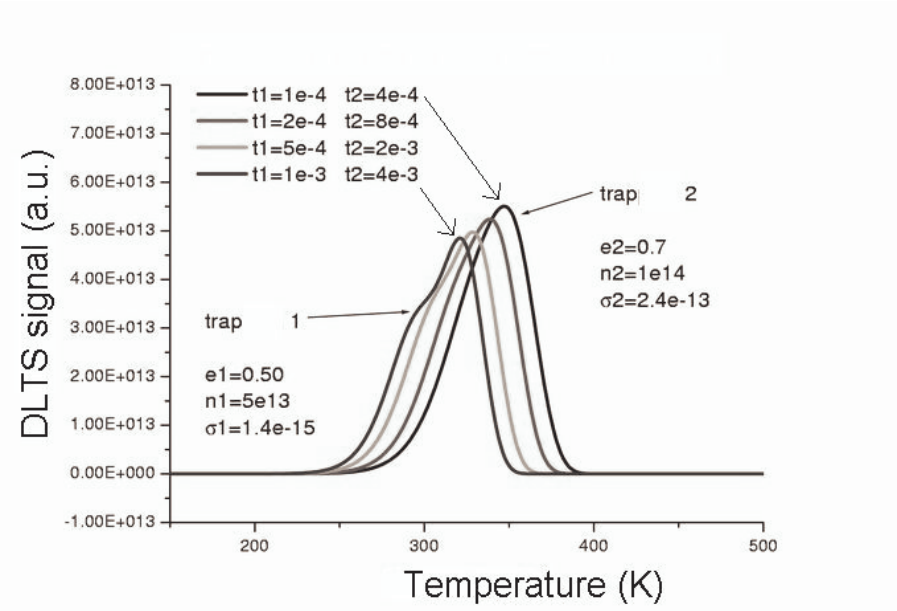


Fig. [4.9] Simulation of DLTS signal from two neighbouring discrete levels. The deep level parameters and the rate windows are given in figure.

4.3.2. Continuous density of states of deep levels

In some cases it may occur that defects of the same type have different energy levels and therefore different energy distance from the conduction band. This occurs, for instance, in the case of point defects in the neighbourhood of a dislocation. A dislocation, in fact, distorces the surrounding lattice, which has the effect that intrinsic defects, such as vacancies or interstitials, or also impurity defects, "sense" a different environment according to their distance from the dislocation. A different case is the Poole-Frenkel effect, where the energy level of defects in presence of different values of an external field (e.g. at different distance from the junction plane in the depletion region of a device) are modified by the external field. Another important case is that of semiconductor alloys: in a ternary semiconductor, such as $Al_xGa_{1-x}As$, Al and Ga are usually distributed quite uniformly in the volume; however, regions with higher Al concentrations can be next to regions with lower Al concentration. As a consequence, the bandgap is no longer uniquely defined for the whole semiconductor: domains with lower bandgap will neighbour domains with higher bandgap. Similarly, the same defect will have lower or higher energy levels, according to the Al concentration in the neighbourhood. The case of AlGaAs and InGaAs ternary alloys has been studied in detail by Omling et al. [Omling85]. In order to account for alloy broadening, a Gaussian density of states was introduced for a single defect type:

$$N_t(E) = \frac{N_{t0}}{\Sigma\sqrt{2\pi}} \exp\left[-\frac{(E-E_{t0})^2}{2\Sigma^2}\right], \qquad (4.46)$$

where N_{t0} is the total concentration of the defect, E_{t0} is the average defect energy level, and Σ is the standard deviation of the energy level. When each defect of this distribution of states emits electrons after the DLTS filling pulse, under the hypothesis that the capture cross-section is independent of the energy, the collected signal is:

$$S(T) = \frac{N_{t0}}{\Sigma\sqrt{2\pi}} \int_0^\infty \{\exp[-e_n(E,T)t_1] - \exp[-e_n(E,T)t_2]\} \exp\left[-\frac{(E-E_{t0})^2}{2\Sigma^2}\right] dE \,. \qquad (4.47)$$

In the case of a Gaussian-like broadening, the best procedure is usually to estimate cross-section σ and average activation energy E_{t0} from the spectra by Arrhenius analysis, then to extract N_{t0} by fitting the spectra with a Levenberg-Marquardt algorithm, keeping the σ and E_{t0} values fixed and having Σ and N_{t0} as fit parameters.

4.3.3. Effect of junction field : Poole-Frenkel effect

Another case where the capacitance transient is non-exponential is when the emission rate e_n depends on the electric field in the space charge region. According to the Poole-Frenkel model [Ganichev00], the potential energy barrier associated to a deep level can decrease in presence of an electric field. In this case the emission rate is enhanced by the presence of an electric field according to the following expression:

$$e_n(\xi) = e_n(0)\exp\left[\frac{1}{kT}\sqrt{\frac{e^3\xi}{\pi\varepsilon_s}}\right] . \qquad (4.48)$$

Thus, the measurement becomes more complicated, as the dependence of the emission rate on the field is also a dependence on the position of the deep level in the space charge region. For high reverse biases, the same deep level emits in different positions with different emission rates, as $n_t(t)$ is given by:

$$n_t(t) = \int_{w_1}^{w_2} \frac{dn_t(t=0,x)}{dx}\exp[-e_n(\xi(x))t]dx , \qquad (4.49)$$

where w_1 and w_2 are the edges of the space charge region before and during the filling pulse, respectively. The Poole-Frenkel effect is important in the characterization of deep levels because it is usually associated with donor-like deep levels in the upper bandgap of n-type semiconductors, whereas acceptor-like levels do not exhibit similar field-induced emission enhancement.

4.3.4. Extended defect with associated potential barrier

The distortion of the bands is a phenomenon occurring very often when an extended defect is present in a semiconducting material. First, an extended defect like a dislocation may produce both compressional and dilatational strain, thus inducing local variations in the energy gap [Pankove]. Secondly, when an extended defect charges up capturing electrons, the charge associated to the electrons is concentrated locally on the defect, so that variations in the potential energy can occur near the extended defect. When electrons are captured on a dislocation line, a depleted region builds up around the defect line, along with an electric field and a potential barrier [Read]. If the trapping of electrons is not too large, i.e. only a fraction $f << 1$ of the defect states is occupied by electrons one can consider the potential difference as proportional to the occupation level:

$$e\Phi = \alpha f . \qquad (4.50)$$

This has an influence on the capture and emission kinetics of free charge carriers. In the portion of crystal close to the defect a variation of the potential causes a variation of the local density of free charge carriers. If we consider $e\Phi$ as the rise of the valence and conduction bands (the Fermi level keeping at a constant value), here the free carrier densities are modified as follows:

$$n = n_0 \exp[-e\Phi / kT] \quad (4.51)$$

$$p = p_0 \exp[e\Phi / kT] \quad (4.52)$$

where n and p are the electron and hole densities, respectively (the 0 subscript refers to equilibrium conditions) [Figielski64]. If, for instance, one considers a discrete level E_t, the equation describing the kinetics of the trap is:

$$\frac{df}{dt} = \sigma\gamma T^2 (1-f) \exp\left[\frac{-\alpha f}{kT}\right] - \sigma\gamma T^2 \exp\left[-\frac{E_t - E_c}{kT}\right] \quad (4.53)$$

The solution of this equation has a logarithmic behaviour f~ln t for $f(t=0) \approx 0$. For larger t, the behaviour becomes exponential. Generally speaking, the states associated to an extended defect should not be described by a discrete level E_t, but by a continuous distribution (density of states) $n_t(E)$. Figure [4.10] illustrates schematically the band distortion associated with an extended defect neutrally charged if empty, and negatively charged if occupied by electrons.

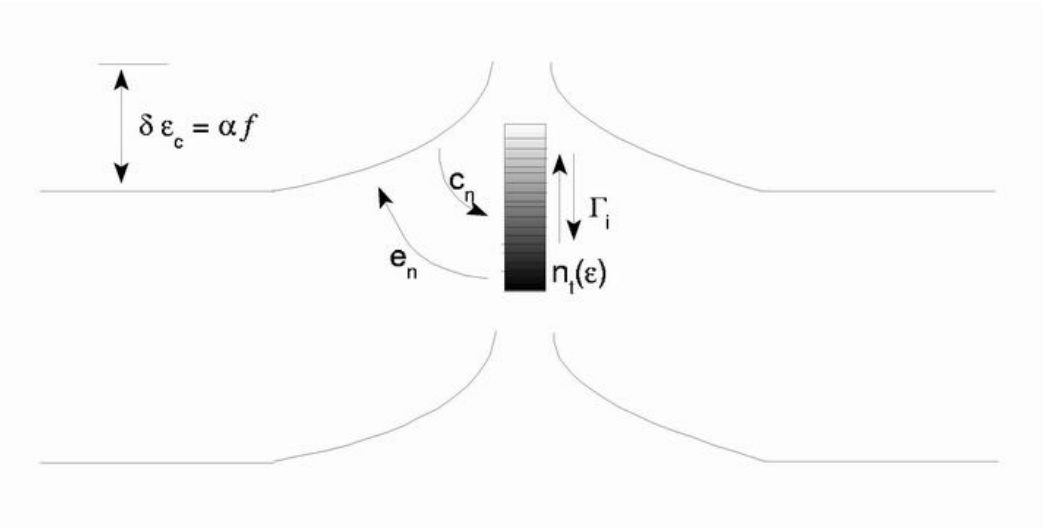


Fig. [4.10]. Band distortion induced by an extended defect

Bandgap electron states can have different behaviours according to the relaxation times of the captured carriers to the minimum energy configuration. If the trap is such that the captured carriers reach the minimum-energy configuration with a speed Γ_i much smaller than the typical emission or capture rates e_n, c_n, the deep states are called *"localized"*. Otherwise, if $\Gamma_i >> e_n$, c_n is large, so that the captures carriers acquire very quickly the minimum-energy configuration, the deep states are called *"bandlike"* [Schroeter95], [Hedemann95].

Localized states

In these states single deep levels give a contribution to the formation of an energy barrier associated to the defect, but do not communicate with each other. An example for such states could be the point defects gathered around a dislocation where a strain field is present. In this case the distance from the dislocation line influences the energy level of the point defect, so that the distribution of the point defects according to the distance has a counterpart in the distribution of their energy levels $n_t(E)$. Indicating with f the occupation of a single level at energy E, the filling kinetics is described by [Hedemann97]:

$$\frac{df(E,t)}{dt} = \sigma\gamma T^2 [1 - f(E,t)] \exp\left[\frac{-\alpha F(t)}{kT}\right] - e_n(E) f(E,t) \quad (4.54)$$

with

$$F(t) = \sum_{E'} f(E', t_p) n_t(E') \exp[-e_n(E')t] \qquad (4.55)$$

where σ is considered as a constant and t_p is the filling pulse width. The system of differential equations can be dealt with numerically in the filling phase of the DLTS period, while the emission kinetics is described by the analytical expression

$$C(t) \propto \sum f(E, t_p) n_t(E) \exp[-e_n(E)t] \,. \qquad (4.56)$$

Among the parameters appearing in the above equation, $f(E,t_p)$ is very sensible to the variation of the duration of the filling pulse over various orders of magnitude. The results of a simulation of DLTS signal from localized states is reported in fig. [4.11]. One can notice that in this interval of pulse widths the peak height grows almost linearly with respect to ln t_p.

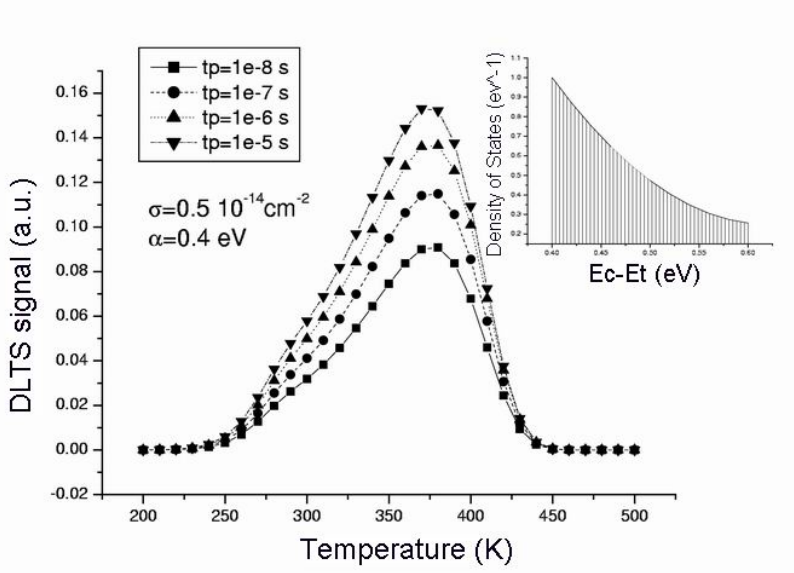


Fig. [4.11]. Simulation of DLTS signal from localized states associated to extended defects. In the inset, the density of states used in the simulation is reported.

Bandlike states

Dislocation cores or metal inclusions may give rise to electron states whose energy levels behave like a band, where the trapped charges tend to quickly occupy the levels with the lowest energy available. For such traps it is convenient to consider the total occupation level

$$F(t) = \sum_{E'} f(E', t_p) n_t(E') \qquad (4.57)$$

Moreover, it is convenient to introduce a quasi-Fermi level E_{qf} describing the trap occupation state in non-equilibrium conditions. The equation for trap kinetics is [Hedemann97]

$$\frac{\partial F(t)}{\partial t} = \sigma \gamma T^2 [1 - F(t)] \exp\left[\frac{-\alpha F(t)}{kT}\right] - R_e(E_{qf}) \,, \qquad (4.58)$$

with

$$R_e(E_{qf}) = \int_0^{\infty} n_t(E) e_n(E) \left[1 + \exp\left[\frac{E + \alpha F(t) - E_{qf}}{kT} \right] \right]^{-1} dE \,. \qquad (4.59)$$

With some approximation it is possible to obtain equations that can be solved numerically. A first approximation is the substitution of the occupation function of the deep levels with a step function. A second approximation is making the hypothesis of a uniform density of states $n_t(E)$ in the interval $\Delta E = E_{max} - E_{min}$. The equivalent emission rate is therefore given by

$$R_e(F) = \frac{1}{\Delta E} \int_{E_{\min}}^{\Delta E F(t) - E_{\min}} e_n(E) dE \,. \qquad (4.60)$$

Finally, the differential equation for trap kinetics is

$$\frac{\partial F(t)}{\partial t} = \sigma \gamma T^2 [1 - F(t)] \exp\left[\frac{-\alpha F(t)}{kT} \right] - \sigma \gamma T^2 \frac{kT}{\Delta E} \exp\left[-\frac{E_C - E_{\min}}{kT} \right] \left[\exp\left[\frac{\Delta E F(t)}{kT} \right] - 1 \right] \qquad (4.61)$$

which is more easily solvable numerically. Also in this case the DLTS signal, as shown in the simulation of fig. [4.12], has a strong dependence on the duration of the filling pulse. The peculiarity of bandlike levels is that the peak temperature shifts towards lower temperature with increasing pulse width t_p. This is due to the fact that when the filling pulse gets longer, levels closer and closer to the conduction band are filled and can emit.

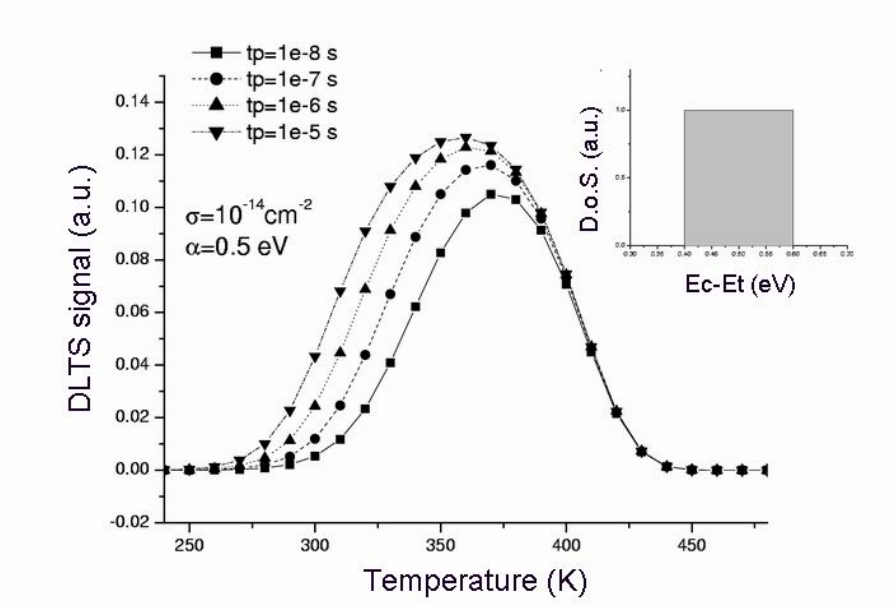


Fig. [4.12]. Simulation of DLTS signal from bandlike states associated to extended defects. In the inset, the density of states used in the simulation is reported.

4.3.5. Characterization of barrier height of the extended defect

When localized electron states associated with an electrically active extended defect are present, one can characterize the barrier height as follows. One needs first to know at which temperature the DLTS signal associated to the analysed deep level is maximum for a particular emission rate. After that, one can stabilize the sample at that temperature and start to collect the DLTS signal no longer as a function of T, but as a function of t_p. This procedure yields characteristics like that illustrated in

fig. [4.13].a). From this plot it is possible to see that the DLTS peak signal is almost proportional to the logarithm of the filling pulse width for a wide time interval (4 decades). After that, the signal tends to saturate. The logarithmic part of the characteristics is well described by the following formula [Omling85], [Wosinski89]:

$$n_T(t_p) = nv_{th}\sigma_n \tau N_T \ln\left[1 + \frac{t_p}{\tau}\right] \qquad (4.62)$$

with

$$\tau = \left(\frac{kT}{q\Phi}\right)\left(\frac{n_{T0}}{N_T}\right)\left(nv_{th}\sigma_n\right)^{-1}. \qquad (4.63)$$

A deeper insight into the microscopic nature of the defect is given by the calculation of the barrier height $e\Phi$, another quantity dependent on the filling pulse width t_p. This quantity is given by

$$e\Phi(t_p) \approx kT\left[\ln(nv_{th}\sigma_n t_p) + \ln\left(\frac{n_T}{t_p(\partial n_T/\partial t_p)}\right) - \ln\frac{n_T}{N_T}\right], \qquad (4.64)$$

where v_{th} is the thermal velocity of carriers σ_n is the capture-cross section for electrons, n_t is the apparent trap density given by the DLTS signal, and N_t is the total trap density at saturation [Cavalcoli 97]. The behaviour of the barrier height as a function of t_p is shown in fig. [4.13].b).

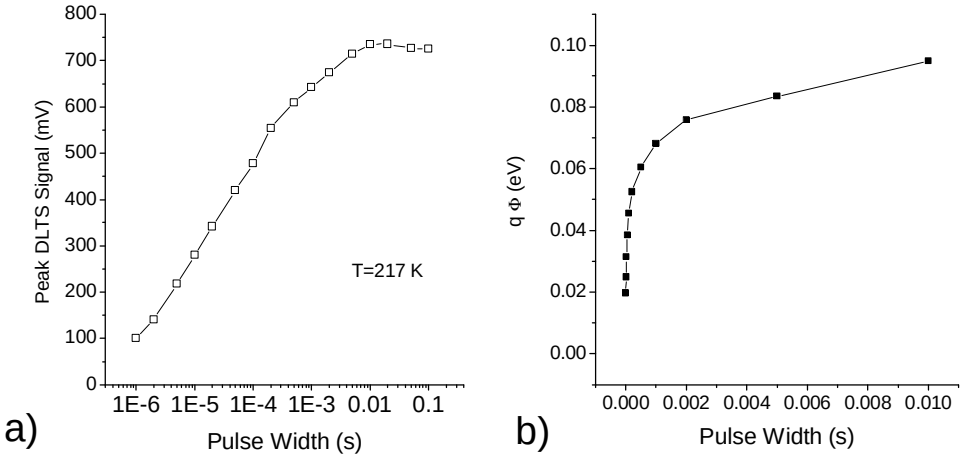


Fig. [4.13]. a) DLTS peak signal at stabilized temperature due to the emission from localized states at extended defect as a function of filling pulse width. b) Barrier height associated with the defect as a function of the filling pulse width.

The characterization of barrier height for bandlike states is somewhat more difficult because the peak temperature shifts with varying filling pulse width.

4.4. DLTS in heterostructure junctions and quantum wells

With extreme care, DLTS can be performed also in heterostructure junctions. In this case, many factors introduce new effects that must be taken into account when interpreting the results of the measurement. An obviously desirable feature of the junction is that it is one-sided. This allows a first determination of the region which is actually probed during the DLTS measurement. The C-V profiling of the junction should always be performed before analyzing a sample by DLTS, in both cases of homo- and heterojunctions. In the latter case the $n_{CV}(x)$ or the $n_{CV}(V)$ profile are even more important, as they give information about the band discontinuities that are present in the junction: thus, the experimentalist can choose appropriate settings of DLTS parameters V_{fill}, V_{rev}, in order to probe the region of interest or to make a DLTS profile along the junction.

The technique of the DLTS profiling, i.e. making successive runs i, i+1,i+2, etc. varying V_{rev}^{i} and setting $V_{fill}^{i+1} = V_{rev}^{i}$, allows one to probe the trap behaviour in the different regions of the heterojunction, and investigate if there are particular effects in correspondence of interfaces between different materials and quantum wells. The choice of the bias parameters is up to the experimentalists, and must keep into account that the probed region must be sufficiently wide to collect a good signal, and sufficiently narrow so that the information collected is localized enough. The example of the quantum well, described in the next section, is interesting from this point of view. Under certain hypotheses, a QW can be regarded as a giant trap localized at a certain junction depth x_w. If we suppose that the QW is sufficiently near the junction, a V_{fill}=0 is sufficient to fill it with electrons, that will give rise to a capacitance transient when the reverse bias V_{rev} is restored. If a filling pulse with lower amplitude is applied, the QW is not filled and thus does not contribute to the DLTS signal. This is an example meaning to show how, under certain circumstances, it is possible to discriminate between emission from a QW and from other deep levels. In the following, the cases of QW emission and of the interaction of the deep level emission with permanent polarization fields will be analysed more in-depth.

4.4.1. *DLTS characterization of a quantum well*

Under favourable conditions, DLTS can be used to study the emission of electrons from a quantum well, which also means the possibility to characterize the quantum well itself. The QW acts similarly to a “giant trap”, an extended defect which can capture and emit electrons. Thus, DLTS can be employed to measure the energy separation between the confined state and the top of the barrier, as illustrated in fig. [4.14].

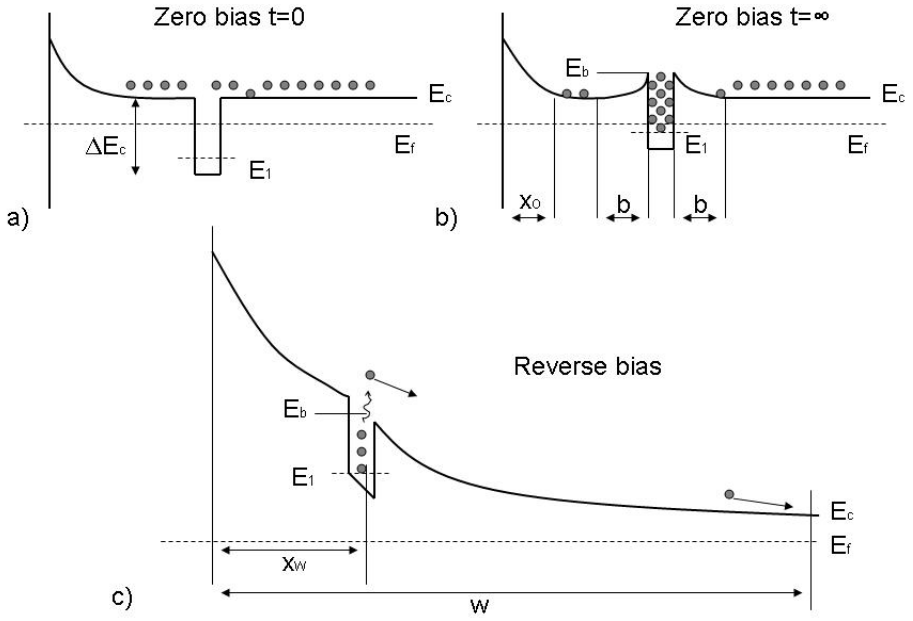


Fig. [4.14]. Conduction band diagrams illustrating filling and thermal emission from a QW: a) immediately after the beginning of the filling pulse; b) at the end of the filling pulse; c) in reverse bias, showing thermal emission.

Provided the quantum well is in an appropriate position, which depends on the doping and on the junction type, DLTS reverse bias and filling pulse can be adjusted so that the QW lies in the depletion region with the reverse bias applied, and can be filled by a forward pulse. In the part a) of the figure, the situation at the beginning of the filling pulse is depicted: electrons from the n-side flow towards the junction. In b) the QW has captured a certain amount of carriers, which also yields a depletion of two regions immediately left and right of the well. This is a strong analogy with the case of the extended defect with a potential barrier associated treated in section (4.3.4). The captured electrons tend to acquire very quickly the minimum energy configuration, filling the lowest subbands first, then the higher ones, which is an analogy with the case of the band-like states at extended defects of section (4.3.4). When the reverse bias is applied, shown in c), the carriers trapped in the QW are emitted thermally and collected by the field, which yields a capacitance transient. Detailed calculations [OrtonBlood] show that, if $L_w << x_w$ the capacitance transient is given by

$$\frac{\Delta C(t)}{C} = \frac{x_w}{w}\frac{L_w n_{w0}}{wN_d}\exp[-e_w t]$$

Where n_{w0} is the initial carrier density in the well, and e_w is the emission rate given by

$$e_w = \sigma_w < v_n > 2\left(\frac{2m_b \pi kT}{h^2}\right)^{3/2}\left(\frac{E_b - E_1}{kT}\right)\exp\left[-\frac{E_b - E_1}{kT}\right]$$

Thus, assuming that σ_w is constant and considering that $<v_n> \propto T^{1/2}$, an Arrhenius Plot of $\ln(e_w/T)$ versus $1/T$ should have slope $(E_b-E_1)/k$.

4.4.2. *Effect of polarization fields in nitride-based MQW heterostructures*

When quantum wells are present in the junction, they can influence the emission from deep levels. This is possible, for instance, in nitride-based LEDs, where a polarization field is present in the QW region due to piezoelectric and spontaneous polarization at the interfaces between different materials. According to the crystal polarity, a net polarization electric field builds up pointing towards the bulk n-side or towards the junction. For crystals grown with Ga-polarity, as most LEDs, the field points towards the junction, as shown in fig. [4.15]. This can introduce new effects during a DLTS measurement. Consider the diagrams in fig. [4.15], which show the sequence of a) the junction at the end of the filling pulse with Vfill=-2V and b) the junction under reverse bias Vrev=-3V. During the filling pulse there is a significant carrier density in the well region, which can suffice to fill the deep traps present. Other traps are filled by the free electrons at the edge of the space charge region on the bulk side, at a distance x~110 nm.

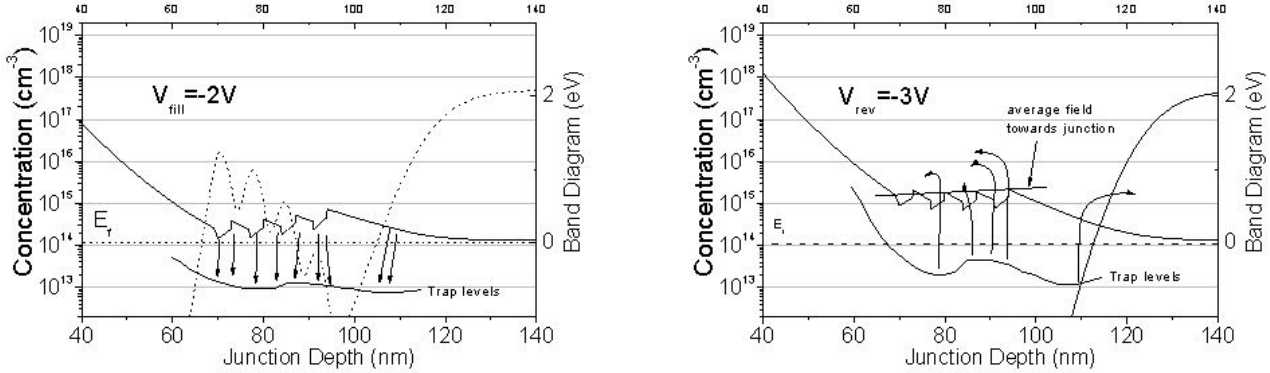


Fig. [4.15]. Effect of permanent polarization fields in a nitride-based MQW heterostructure: a) during the filling pulse, carriers from the conduction band are captured by deep levels; b) when the junction is again reverse-biased, carriers are emitted by the deep levels, but those emitted in the MQW region are collected towards the junction, instead of being collected towards the bulk.

When the reverse bias is re-established, the traps at the boundary of the space charge region emit electrons which are normally collected by the junction field to the n-side bulk region; at the same time a different behaviour characterizes the electrons captured by deep levels in the MQW region. The carriers emitted from these latter traps, once emitted in the conduction band, experience a net average field pointing towards the junction, which has the consequence to leave them confined in the MQW region. As they are not collected to the n-region, they do not contribute to the capacitance transient. Moreover, they possibly give rise to an inversion of the transient from negative to positive, as they are collected towards the junction.

4.5.Experimental setup

The DLTS measurement is performed as follows: the sample temperature varies slowly, while the sample bias varies in a fashion like that described in fig. [4.7], with a fixed period *P*. Moreover, the capacitance of the sample at reverse bias V_{rev}, and the capacitances during the transient at t_1 and t_2 must be measured. In order to achieve this, we make use of the following experimental equipment, which is depicted in fig. [4.16].

4.5.1. Temperature controller and cryogenic apparatus.

This instrument, a Lake Shore DRC91C, allows one to set and control the temperature of the sample. It is interfaced to a computer by means of a NI board and a Visual Basic software. The user

determines the temperature interval and heating rate through the software. For our setup, the heating rate must be lower than 0.083K/s. The temperature interval can vary from 77K, i.e. the liquid nitrogen temperature, to about 700K for SiC devices. Lower maximum temperatures are employed for more delicate materials and devices, such as GaN-based LEDs. The sample is left in a cryogenic apparatus, consisting of a dewar containing liquid nitrogen, and a sample holder. A thermocouple measures the temperature, while the sample holder can be heated by a Thermocoax cable with resistivity 12.5 Ω/cm.

4.5.2. *Impulse generator and capacitance meter.*

The impulse generator and the capacitance meter are, together with the exponential correlators, building blocks of the Sula Technologies Deep Level Spectrometer. The impulse parameters that are set on the impulse generator are:

- the period P
- the reverse bias V_{rev}
- the filling pulse bias V_{fill}
- the pulse duration (pulse width) t_p.

The capacitance meter works at a frequency of 1 MHz, thus being able to reproduce accurately capacitance transients with time constants of the order of some tens of μs. The output of this instrument yield the reverse-bias capacitance C_0, and the capacitance transient, which is further processed by the correlators and boxcar averagers of the instrument.

4.5.3. Exponential correlators and double boxcar averagers.

The deep level spectrometer incorporates four correlators, which play the role of reducing the noise affecting the capacitance transient. The output of each correlator is then processed by a double boxcar averager, which, in turn, yield as output the signal $\Delta C=C(t_1)-C(t_2)$. To each boxcar averager correspond four different values of the couple (t_1,t_2), so that in a single measurement it is possible to collect four spectra speeding up the experimental procedure. The ratio t_1/t_2 is fixed and equal to 2.5. The values of the emission rate en corresponding to the different values of the rate windows range from 1.16×10^4 s^{-1} to 2.33 s^{-1}.

The correlator performs the following operations: the periodic signal consists of the sum of noise $N(t)$ and ideal signal, described by an unknown amplitude A multiplying the function $s(t)=\exp[-e_n t]$. The signal is filtered by means of a weighting function $W(t)=\exp[-t/\tau]$, produced by an exponential function generator with variable time constant. The best measure S of the unknown amplitude A

$$S = \int_0^P [A\exp(-e_n t) + N(t)]W(t)dt \qquad (4.65)$$

occurs just when $s(t)=W(t)$. This provides the best SNR.

The signal is the processed by the double boxcar averager. Here two input channels correspond to two linear gates which sample the signal at instants t_1 and t_2, while a differential block yields at the output the difference of the signals measured by the two gates. The measurement setup is illustrated in fig. [4.16].

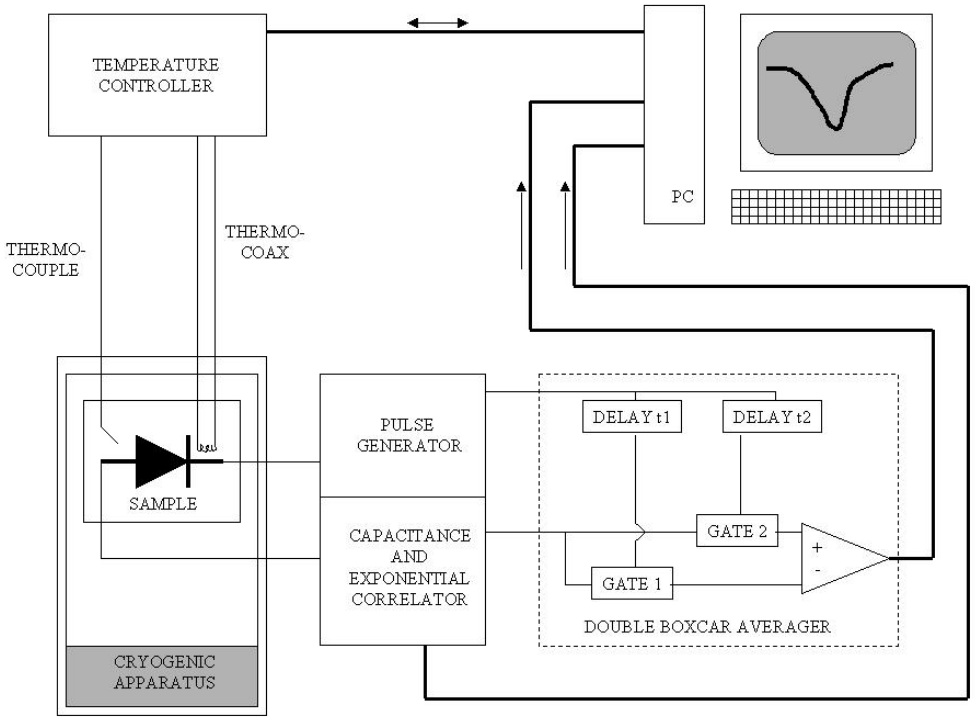


Fig. [4.16] Block diagram of the instrumentation for the DLTS measurements, evidencing the correlator and the double boxcar averager.

REFERENCES

[Cavalcoli97] D. Cavalcoli, A. Cavallini et al., *Phys. Rev. B* **56** (1997) 10208

[Davies] J. H. Davies, *The physics of low-dimensional semiconductors*, (Cambridge University Press, Cambridge, 2000).

[Figielski64] T. Figielski, "Theory *of carrier recombination at dislocations in germanium*", *Phys. Stat. Sol.* **6** p.429 (1964).

[Ganichev00] S.D. Ganichev, E. Zieman, W. Prettl, I.N. Yassievich, A.A. Istratov and E.R. Weber, *Phys. Rev. B* **61**, 10361 (2000)

[Hedemann95] H. Hedemann, *Quantitative Analyse von Kapazitätstransientenspektren ausgedehnter Defekte in Halbleitern*, (Cuvillier Verlag, Göttingen, 1995).

[Hedemann97] H. Hedemann, W. Schroeter, *Journal de Physique III* **7**, p. 1389 (1997)

[Look81] D.C. Look, *Phys. Rev. B* **24** 5852 (1981)

[LookGaAs] D. C. Look, *Electrical characterization of GaAs materials and devices* (New York: Wiley, 1989)

[Miller77] G.L. Miller, D.V. Lang and L.C. Kimerling, *Capacitance Transient Spectroscopy*, Ann. Re. Mat. Sci., p. 377 (1977)

[Omling85] P. Omling, L. Samuelson and H.G. Grimmeiss, *J. Appl. Phys.* **54**, p. 5117 (1983)

[OrtonBlood] P. Blood and J. W. Orton *The electrical characterization of semiconductors: majority carriers and electron states* (London: Academic Press, 1990)

[Pankove] J.I. Pankove, *Optical Processes in Semiconductors*, (Dover, New York, 1971).

[Read] W. T. Read, *Dislocations in Crystals* (McGraw Hill, 1953)

[Schroeter95] W. Schroeter et al., "*Bandlike and localized states at extended defects in silicon*", *Phys. Rev. B* **52**, p. 13726 (1995)

[Wosinski89] T. Wosinski, *J. Appl. Phys.* **54**, p.1566 (1989)

5 Optical Spectroscopy

The term “Optical Spectroscopy” is very generic, meaning a vast category of experimental techniques. These techniques have in common that they deal with the interactions of photons with a sufficiently narrow wavelength interval $\Delta\lambda$ with the electrons in the semiconductor. These interactions may be very diversified: absorption, emission, transformation, modulation of light by a semiconducting material or by a semiconductor device are all examples of interactions between electrons and photons. The diversity of interaction mechanisms is mirrored by the diversity of processes that can take place: formation of excitons, promotion of electrons from valence to conduction band, photon-assisted transitions from defect states to the bands, are some of the processes analyzed in optical spectroscopy. The fundamental issue is having photons of a narrow energy interval interact with the material. This is the key to most of the information about the physical processes, as the energy position of spectral features indicate the energy needed for a certain process to occur. Further, from the shape of the spectral peaks or variations, one can infer information about transition probabilities or energy broadening; from the relative intensities of spectral features one can obtain information about the amount of a physical process in the analyzed part of the material. Moreover, optical spectroscopy results are highly significant when optoelectronic devices are analyzed, as their operation is based on electron-photon interaction.

5.1.Photocurrent spectroscopy in layered structures

Comparison of photocurrent spectra (PCS) of LEDs, which are heterostructure diodes, i.e. devices with a certain degree of complexity in their structure, and of bulk GaN or other semiconductors samples yields striking differences between these two cases. I limit here myself to the description of photocurrent measured while lighting the junction of a diode, i.e. photocurrent as a photovoltaic phenomenon. The case of photocurrent as an increase of photoconductivity of biased homogeneous slices of semiconductor is not taken in consideration here.

Photocurrent spectra of bulk samples tend to be rather simple, showing a single peak in proximity of the wavelength corresponding to the semiconductor bandgap, according to the relation $E_g=h/\lambda$, and other minor peaks possibly related to deep levels and other intragap photo-induced transitions. The complexity of a heterostructure with respect to the bulk material is mirrored by the complexity of its photocurrent spectrum. The presence of semiconductors of different bandgap spatially distributed along the growth direction and the appearance of quantum phenomena, such as the carrier confinement in quantum wells (QWs) introduces in the PCS many new features, due to the presence of regions having different absorption coefficients for light of a certain wavelength. It is possible, for instance, that light penetrates a first layer with a relatively high bandgap, being then absorbed by a second layer with a relatively low bandgap. In other cases, light can be completely absorbed by a first layer, thus not being able to reach the further layers, or, on the opposite, it can pass through every layer of the heterostructure, not being absorbed anywhere. Moreover, a heterostructure is made by layers of different index of refraction. As light can be transmitted or reflected at each interface, *interference* between forward and backward propagating waves can occur, leading to formation of maxima and minima in the field intensity along the structure. This can happen provided the light source is coherent.

In order to generate photocurrent, it is necessary first that light is absorbed by the material. Only if light is absorbed it can generate charge carrier pairs. The generated pairs must then fulfil some conditions to be transformed into photocurrent, i.e., they have to be collected by the electric field present in the junction region before they recombine. Charge carriers can thus be collected either if they are generated in a space charge region, where the electric field is present, or they can diffuse to the space charge region from adjacent regions. As a rule of thumb, it is clear that the further the pair is generated from the junction region, the lower is the probability that it diffuses to the space charge region and is collected by the electric field.

Thus, basically one needs three elements in order to generate a photocurrent: a non-zero light intensity, a non-zero absorption coefficient, and a non-zero electric field in the neighbourhood of the point where the photon is absorbed and an electron-hole pair is created. These are the basic ideas of the model, which will be discussed more in detail in the following.

5.2. Study of the light intensity profile in a semiconductor heterostructure

The problem of the propagation of light in a semiconductor heterostructure, that in first approximation may be regarded as a more or less absorbing isotropic medium, can be quite simply solved in the framework of classical optics with the approximation of the incident light beam as a monochromatic plane wave. This is a good approximation for our experimental set up, in which a quasi-monochromatic beam exiting the slot of a monochromator is directed on the sample by means of an off-axis mirror. A very detailed description of the problem and of its solution is given in the book by Born and Wolf [BornWolf]. The method illustrated is known under the name of "transfer matrix algorithm".

5.2.1. Transfer matrix method for the solution of the problem of light propagation in an absorbing isotropic medium

An electromagnetic monochromatic plane wave of angular frequency ω can be described through the electric field amplitude and direction F_0 and $\hat{\boldsymbol{e}}_0$ respectively, as well as through the wave vector $\boldsymbol{k_0}$. In complex form the electric field is written as

$$\vec{F}_0(\vec{r},t) = \hat{e}_0 F_0 \exp[i(\vec{k}\cdot\vec{r} - \omega t)], \qquad \vec{k}_0 = \frac{\omega}{c}\hat{k}_0 \qquad (5.1)$$

whereby it is valid that $\hat{e}_0 \cdot \boldsymbol{k_0} = 0$ in order to fulfil the homogeneous Maxwell equations.

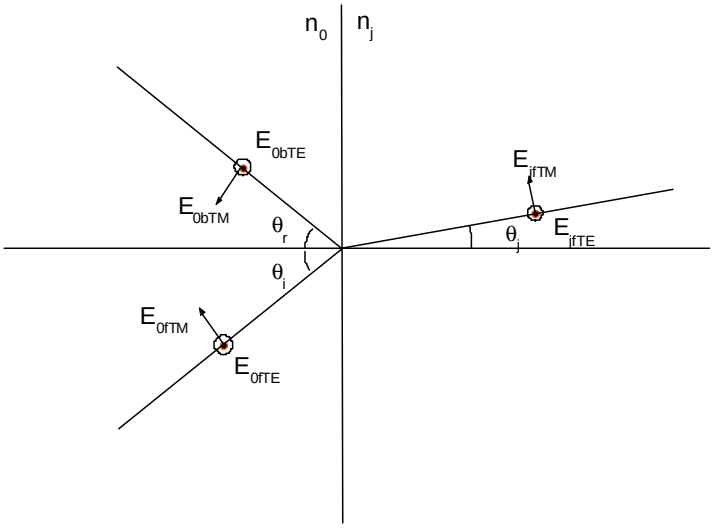


Fig. [5.1] Reflection and refraction at an interface. The plane of incidence is in the plane of the sheet. The electric field components of a TM wave are in the plane of incidence, those of a TE wave are perpendicular to the plane of incidence. Waves propagating backwards from the j-th medium are omitted.

From Maxwell's equations one can also deduce that both the reflected and refracted waves at an interface (Fig. [5.1]) are equally plane waves propagating in the plane of incidence. For isotropic

materials, the angle of reflection is the opposite of the angle of incidence, and Snell's law of refraction is valid ($n_0 \sin(\theta_0) = n_j \sin(\theta_j)$), whereby $n_j = n_j^{ref}(1+i\kappa_j^{att})$ denotes the complex index of refraction of j[th] material. n_j^{ref} is the real index of refraction and κ_j^{att} is the attenuation coefficient. Indicating with $\gamma=f,b$ the field propagating forward or backward, the electric field in the j[th] material is therefore written as

$$\vec{F}_{j\gamma}(\vec{r},t) = \vec{F}_{j\gamma}\exp[i(\vec{k}_{j\gamma}\cdot\vec{r}-\omega t)], \qquad \vec{k}_{j\gamma} = n_j\frac{\omega}{c}\hat{k}_{j\gamma} \tag{5.2}$$

which, once rewritten, yields

$$\begin{aligned}\vec{F}_{j\gamma}(\vec{r},t) &= \exp(-n^{ref}{}_j\kappa_j^{att}\frac{\omega}{c}\hat{k}_{j\gamma}\cdot\vec{r})\vec{F}_{j\gamma}\exp[i(n^{ref}{}_j\frac{\omega}{c}\hat{k}_{j\gamma}\cdot\vec{r}-\omega t)],\\ \vec{F}_{j\gamma}(\vec{r},t) &= \exp(-\frac{\alpha_j^{abs}}{2}\hat{k}_{j\gamma}\cdot\vec{r})\vec{F}_{j\gamma}\exp[i(n^{ref}{}_j\frac{\omega}{c}\hat{k}_{j\gamma}\cdot\vec{r}-\omega t)],\end{aligned} \tag{5.3}$$

thus identifying the absorption coefficient of the j[th] material $\alpha_j^{abs}=2n_j^{ref}\kappa_j^{att}\omega/c$.

Rigorously, one has to be very careful in defining an absorption coefficient α^{abs} for very thin semiconductor layers, such as quantum wells in a heterostructure device. In QWs, in fact, the quantum confinement of the electrons makes it impossible to consider the absorption of carriers as something scaling with well width L_W. In this case, the only physically meaningful way to describe the absorption of photons is through the definition of a coefficient $\gamma_w=\Delta\Phi_0/\Phi_0$, where Φ_0 is the incident flux and $\Delta\Phi_0$ is the flux variation across the well [Blood00]. Keeping this in mind, we will define here the absorption coefficient for a QW as $\alpha_w^{abs}=\gamma_w / L_W$ in order to maintain a notation consistent with the present method.

For the reflection/refraction at an interface placed at $z=z_j$ (z is taken as the direction perpendicular to the interface), one has to solve the problem according to the Maxwell equations, taking into account the polarizations of the electric field perpendicular (TE wave) or parallel (TM wave) to the plane of incidence. This leads to the linear system, which can be written as follows in matrix form:

$$\begin{pmatrix}F_{0\,fTM}(z_j)\\F_{0bTM}(z_j)\\F_{0\,fTE}(z_j)\\F_{0bTE}(z_j)\end{pmatrix} = \frac{1}{2}\begin{pmatrix}1+\zeta_{jTM} & 1-\zeta_{jTM} & 0 & 0\\1-\zeta_{jTM} & 1+\zeta_{jTM} & 0 & 0\\0 & 0 & 1+\zeta_{jTE} & 1-\zeta_{jTE}\\0 & 0 & 1-\zeta_{jTE} & 1+\zeta_{jTE}\end{pmatrix}\begin{pmatrix}F_{jfTM}(z_j)\\F_{jbTM}(z_j)\\F_{jfTE}(z_j)\\F_{jbTE}(z_j)\end{pmatrix}, \tag{5.4}$$

where $\zeta_{jTM}=(n_j\cos(\theta_0) / n_0\cos(\theta_j))$ and $\zeta_{jTE}=(n_j\cos(\theta_j) / n_0\cos(\theta_0))$. From the formula it is apparent that the TE and TM components of the field are independent and can be treated separately.

After entering the film at its interface, the field propagates through the homogeneous film to the next interface. The field at the next interface at $z=z_{j+1}$ is given through the propagation matrix:

$$\begin{pmatrix}F_{jfp}(z_j)\\F_{jbp}(z_j)\end{pmatrix} = \begin{pmatrix}e^{i\delta_j} & 0\\0 & e^{-i\delta_j}\end{pmatrix}\begin{pmatrix}F_{jfp}(z_{j+1})\\F_{jbp}(z_{j+1})\end{pmatrix}, \qquad \delta_j = n_j\frac{\omega}{c}(z_{j+1}-z_j)\cos(\theta_j) \tag{5.5}$$

where p = *TE, TM* respectively.

For the fields emerging from the second interface, after some rearranging of the vector components, one has

$$\begin{pmatrix} F_{0\,fp}(z_j)+F_{0bp}(z_j) \\ F_{0\,fp}(z_j)-F_{0bp}(z_j) \end{pmatrix} = \begin{pmatrix} \cos(\delta_j) & -i\zeta_j^{-1}\sin(\delta_j) \\ -i\zeta_j\sin(\delta_j) & \cos(\delta_j) \end{pmatrix} \begin{pmatrix} F_{0\,fp}(z_{j+1})+F_{0bp}(z_{j+1}) \\ F_{0\,fp}(z_{j+1})-F_{0bp}(z_{j+1}) \end{pmatrix}, \qquad (5.6)$$

where the matrix can be indicated with T_j and called the *transfer matrix of the film j*. If we consider a model of a multilayer, whereby an infinitely thin film of the incident medium lies between each two consecutive films, it is possible to define the transfer matrix of the stack of layers as product of the transfer matrices of each film.

$$T = T_{total} = \prod_j T_j = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}. \qquad (5.7)$$

Once the problem has been set with proper boundary conditions (e.g. no wave propagating backwards in the space after all the layers), one then defines reflectance R_p and transmittance T_p for both polarizations as follows:

$$\begin{aligned} R_p &= \frac{F_{0bp}F^*_{0bp}}{F_{0\,fp}F^*_{0\,fp}} = \left|\frac{t_{11}+t_{12}-t_{21}-t_{22}}{t_{11}+t_{12}+t_{21}+t_{22}}\right|^2, \\ T_p &= \frac{F_{Nfp}F^*_{Nfp}}{F_{0\,fp}F^*_{0\,fp}} = \left|\frac{2}{t_{11}+t_{12}+t_{21}+t_{22}}\right|^2, \end{aligned} \qquad (5.8)$$

where F_{0fp}, F_{0bp} and F_{Nfp} are the incident field, the reflected field and the field transmitted at the last interface, respectively, for the polarization *p*. It is clear that reflectance and transmittance depend on many factors, such as angle of incidence and wavelength, as well as on structural parameters, such as thickness, refraction index and absorption coefficient of the stack of layers considered.

The transfer matrix algorithm can also be used in order to track the amplitude of the electric field inside each layer, which is very important if one has to know the intensity of the field point by point, as it is the case in simulations of carrier generation profiles.

5.2.2. Effect of interface roughness

The effect of surface roughness is a scattering of a certain fraction of light in random directions. The losses from the light beam by scattering at a rough interface between media 1 and media 2 are described by the scattering factors. Thus, in comparison with a smooth interface, the Fresnel coefficient $r_{12}=F_{1b}/F_{1f}$ for the amplitude of the beam reflected from the rough surface back to the medium 1 is reduced by a factor:

$$s_{12}^r = \exp\left[-\frac{1}{2}\left(\frac{4\pi n_1\sigma}{\lambda}\right)^2\right], \qquad (5.9)$$

and the Fresnel coefficient $t_{12}=F_{2f}/F_{1f}$ for the amplitude of the beam transmitted through the rough interface to medium 2 is reduced by a factor

$$s_{12}^t = \exp\left[-\frac{1}{2}\left(\frac{2\pi(n_1-n_2)\sigma}{\lambda}\right)^2\right], \qquad (5.10)$$

where σ is the rms surface roughness of the boundary interface. Both formulas assume (i) $\sigma >> \lambda$, (ii) small correlation length of surface roughness and (iii) normal incidence of light. In the case of oblique incidence the surface roughness σ has to be reduced by the factor $\cos(\theta_i)$ [Poruba00].

5.2.3. Some calculation of reflectance, transmittance and field intensity profile of an absorbing isotropic single layer

As a first set of examples, I start with some simple cases, i.e. a single non-absorbing layer. For a single layer the expressions for transmittance and reflectance take on an analytic form:

$$T = \frac{t^2}{(1-r^2) + 4r\sin^2\frac{\delta}{2}}, \qquad (5.11)$$

$$R = 1 - T$$

where t and r are transmittance and reflectance of the single interface, respectively, and δ is the phase change upon propagation through the layer. The result is valid also for absorbing media, provided the complex index of refraction and complex phase change are taken into account.

As displayed in fig. [5.2].a. and b., transmittance and reflectance are not periodic functions of the film thickness, but become damped oscillating functions. In figures [5.2].a. and [5.2].b. two media with different absorption coefficient are compared. One can see that for the more efficiently absorbing medium of fig. b. the transmittance assumes an exponential decay behaviour tending to zero with increasing layer thickness, whereas the reflectance tends to a constant value. The relation $T+R=1$ is no longer valid, as energy is absorbed by the medium.

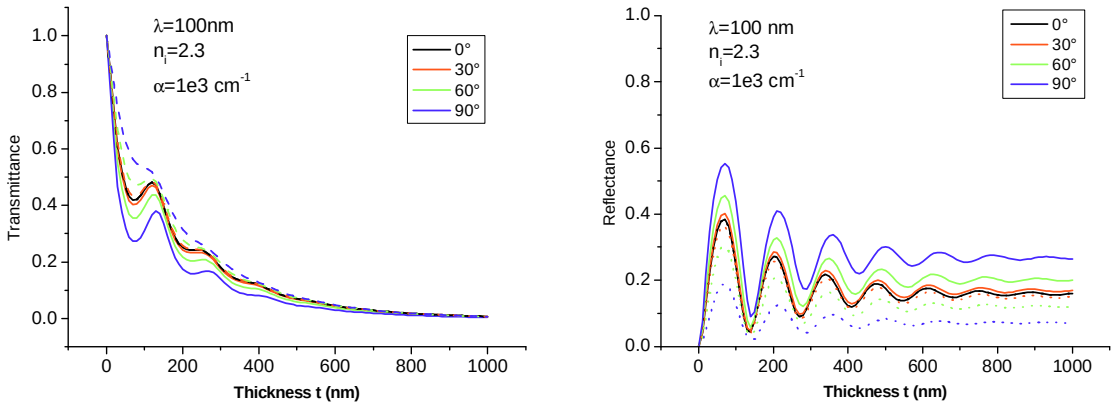


Fig. [5.2]**.a. Transmittance and reflectance of an absorbing single layer. α is given in cm^{-1} units. Continuous lines refer to TE waves, dotted lines to TM waves**

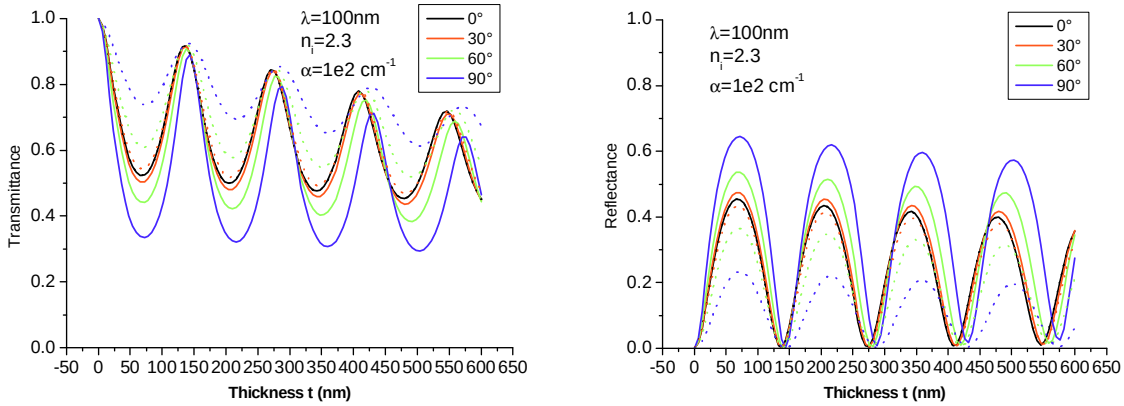


Fig. [5.2]**.b. Transmittance and reflectance of an absorbing single layer. α is given in cm^{-1} units. Continuous lines refer to TE waves, dotted lines to TM waves**

The field profile in an absorbing medium is depicted in the plot in fig. [5.3]. One can notice a region next to the first interface, where only an exponential decay is to be seen, and a region next to the second interface, where many oscillations are present. This is to be explained as follows: the oscillations due to the interference of forward and backward propagating waves can build up only where these two components have similar amplitude. This happens only in the region next to the neighbourhood of the second interface, as this is the only region where both amplitudes (forward and backward propagating) are comparable.

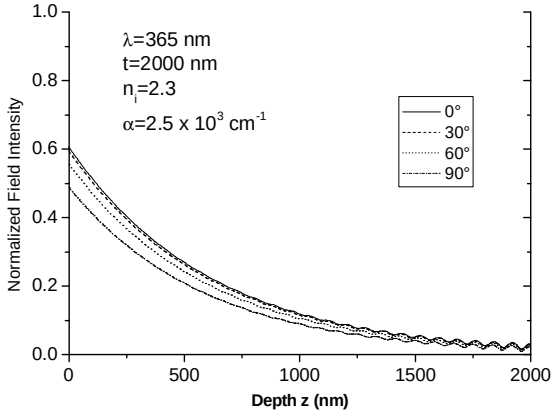


Fig. [5.3] Field profile in an absorbing GaN single layer (TE component)

5.2.4. Some calculation of reflectance, transmittance and field intensity profile of a stack of isotropic layers

In this case the structure we deal with has a higher degree of complexity. Close analytical forms for transmittance and reflectance are no longer possible, although expressions can simplify in some particular cases. As the processes of reflection, transmission, absorption and interference between backward and forward propagating waves were exposed in detail in the previous section, I limit the exposition to some examples of calculations in some particular structures, which can be interesting for clarity purposes. For simplicity's sake, the dependence on the incidence angles will here be dropped too, and we will consider only normal incidence ($\theta = 0$).

After the field profile has been calculated, it is straightforward to calculate the field intensity profile:

$$\bar{I}(z,\lambda) = \frac{1}{2}\frac{\varepsilon(z)}{4\pi}\left(\left|F_{TE}(z,\lambda)\right|^2 + \left|F_{TM}(z,\lambda)\right|^2\right), \qquad (5.12)$$

where $\varepsilon(z)$ is the position-dependent dielectric permittivity of the material.

In Fig. [5.4] we report the field intensity profile for normal incidence on the following structure: GaN – 600 nm; InGaN –400 nm; GaN – 600 nm. GaN Is assumed to have a refraction index n=2.3 (as it is the real case), InGaN is assumed to have a refraction index n=2.6. The incident light has as wavelength in vacuum λ = 600 nm. The periodicity of the optical field intensity inside each layer is $d=2\,\lambda/n$, whereby the factor 2 is due to the fact that the intensity is the square modulus of the field amplitude, which has periodicity $d'=\lambda/n$.

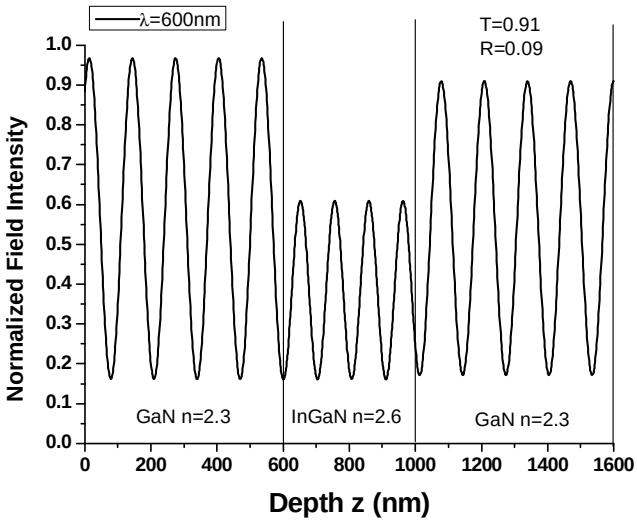


Fig. [5.4] Electric field profile for normal incidence on a GaN/ InGaN/ GaN layer, with no absorption.

5.3. Photocurrent model

5.3.1. Sample structure

The structure of a typical device used in the photocurrent simulations and experiments is depicted in fig. [5.5]. It is a heterostructure-based LED, though the model is also applicable to other device types. The light falls upon the diode from the top, usually normally, although photocurrent runs at different angles of incidence can be performed, in order to check the validity of the model which considers all the effects of internal interference of the optical field. The sample has a total thickness *t*, excluding the top and bottom ohmic contacts. The top surface is set at the coordinate 0 on the *z* axis. The edges of the depletion region, which can be calculated with a Schrödinger-Poisson solver, are set at x_p and x_n. This is the structure to which the procedure illustrated in the previous section is applied in order to calculate the optical field intensity profile and, subsequently, the profile of photo-generated carriers and the photocurrent. Having already illustrated how to calculate the field intensity profile, we describe now the model according to which carriers are generated and can be collected by the junction, thus yielding the photocurrent signal.

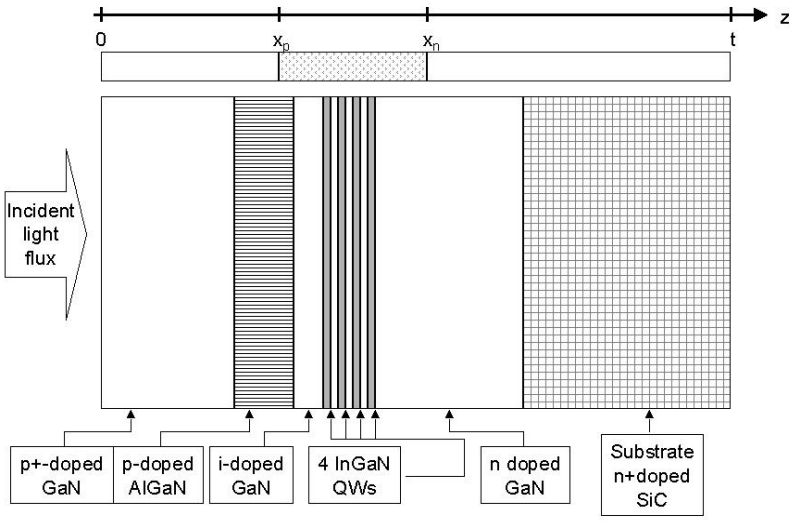


Figure [5.5]. Scheme of the typical LED used for the PC experiments. The relative thicknesses of the different regions are not in scale.

5.3.2. Absorption coefficient dispersion relations

The problem of finding a suitable absorption dispersion relation, i.e., the relation between absorption coefficient and light frequency, is not particularly simple if one is trying to simulate a real device. Absorption dispersion relations are easily found for bulk material as GaN [Ambacher96], [Muth97], although one should keep in mind that even bulk materials can show different absorption properties due to doping, extended defects, and other properties which can vary. An even more difficult issue is to find the proper absorption dispersion for ternary InGaN alloys, where the absorption edge is strongly dependent on the x fraction of In, especially when thin InGaN layers are present in the heterostructure as Quantum Wells (QWs). The absorption edge of a QW is namely dependent not only on the x In fraction, but also on the electric field due to piezoelectric polarization in strained pseudomorphic layers and to the bias applied to the device. The behaviour of the absorption edge of InGaN QWs can be modelled through Franz-Keldysh absorption mechanisms and through electric field-dependent Quantum Confined Stark Effect (QCSE) [Franssen04], for which knowledge of the structural parameters of the QW region, such as the In fraction and the well thickness, is required.

The dispersion relation for the absorption coefficient in bulk GaN was modelled as follows: as for most common direct bandgap semiconductors, the absorption takes place for energies higher than E_g and has the typical square root law [Bube]:

$$\alpha'_{bulk}(E) = A\sqrt{E - E_g}\ . \qquad (5.13)$$

However, it is attested that in GaN there are significant band tailing effects, which make absorption possible also for photons with energy lower then bandgap. This behaviour was modelled by convolving a normal distribution with the absorption coefficient of eq. (5.13); this yielded the absorption coefficient used for the calculation of photocurrent:

$$\alpha_{bulk}(E) = \alpha'_{bulk}(E) \otimes \frac{1}{\sqrt{2\pi}\Delta} \exp\left[-\left(\frac{E}{\Delta}\right)^2\right], \qquad (5.14)$$

where Δ is a measure of the energy broadening of the absorption edge due to band tailing effects.

As far as the absorption from the QWs concerns, one can define an effective absorption coefficient which is proportional to the joint density of states (JDOS) of the QW system, and to a contribution by excitonic absorption (as excitons have higher binding energy in QWs). The joint density of states in a QW has a step-like behaviour [Davies], with a certain amount of broadening, caused either by thermal factors or by composition fluctuations. The step-like density of states can be well reproduced by a set of Fermi functions, whose number has been limited to 2. The exciton contribution is described through a Lorentzian function with a thermal broadening parameter [Bulutay99]:

$$\alpha_{QW}(E) = \frac{2B}{\pi} \frac{\Gamma}{4(E - E_{ex})^2 + \Gamma^2} + \frac{C_1}{1 + \exp\left[-\dfrac{E - E_g}{D_1}\right]} + \frac{C_2}{1 + \exp\left[-\dfrac{E - E_g - E_{2s}}{D_2}\right]}, \qquad (5.15)$$

where B and $C_{1,2}$ are normalization constants, $E_{ex}=E_g-R$ and E_g are the exciton energy and the effective bandgap energy (which, for QWs, corresponds to the energy separation between the hole and electron subbands). R is the exciton binding energy, while E_{2s} is the energy describing the onset of the second step in the JDOS. Γ and $D_{1,2}$ are broadening parameters. The first summand in eq. (5.15) corresponds to the excitonic transitions. The Lorentzian function describing the excitonic transition has a maximum for a photon energy equal to the exciton energy and a broadening parameter Γ which should be roughly equal to kT (homogeneous broadening). Other distribution functions could be used for the exciton transition, such as the Gaussian function. In this case, however, other broadening mechanisms are supposed (i.e. interface roughness, disorder, etc.). The second summand describes approximately the interband transitions of free carriers in bulk material or of electrons and holes in quantum wells.

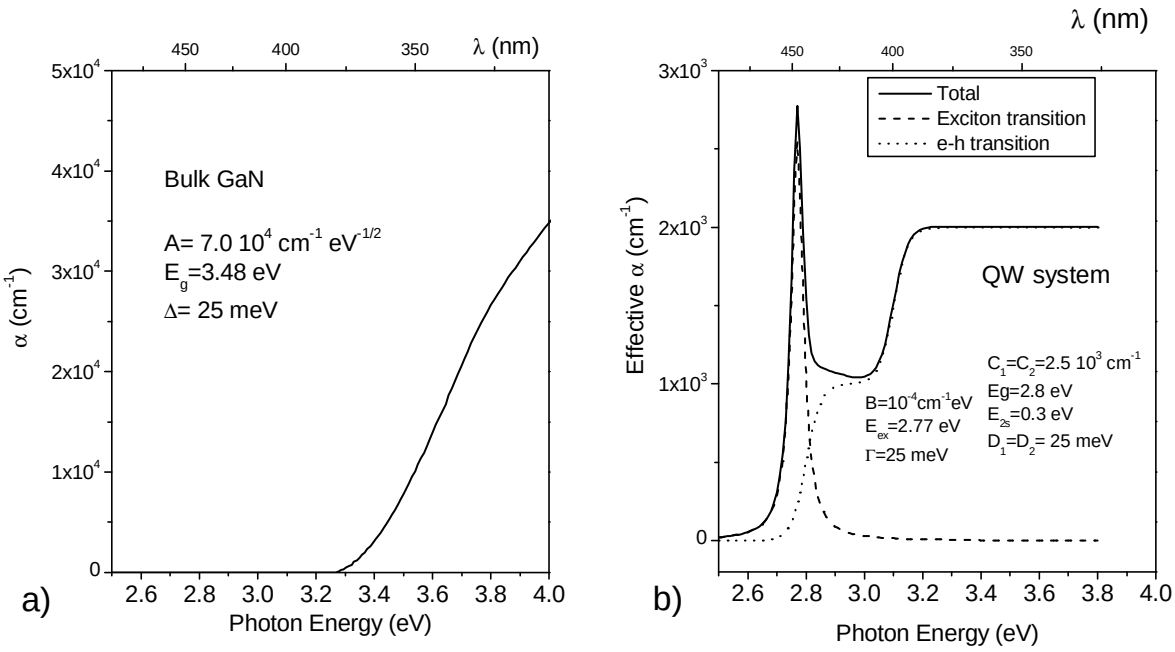


Fig. [5.6]. Example of dispersion relation for the absorption coefficient a) in bulk GaN, b) in a QW system, modelled according to eqs. (5.14-5.15).

The sigmoidal function used is not rigorously physical, but it is mathematically convenient in fitting procedures: for QW transitions, in fact, it reproduces efficiently the Urbach band tail [Martin99], and it mimics the step in the absorption due to the step-like joint density of states. This dispersion relation for the absorption coefficient is then easy to employ in fitting procedures. An example of dispersion relation for both bulk GaN and a QW is given in fig. [5.6].

5.3.3. Refractive index dispersion relations

Refractive index n^{ref} (term by which we intend the real part of the complex refractive index introduced in sec. (5.2), i.e. $n_r = n^{ref}(1+i\kappa^{att})$) and absorption coefficient are connected by a particular relationship. The complex refractive index, in fact, is connected to the complex dielectric function ε_r by the simple relationship

$$n_r^2 = \varepsilon_r = \varepsilon_1 + i\,\varepsilon_2 \qquad (5.16)$$

The real and complex parts of the dielectric function, which is the response function of a system under the sollicitation of an incoming electromagnetic wave, are linked together by the so-called Kramers-Kronig relations [Davies]:

$$\varepsilon_1(\omega) = \frac{1}{\pi}\wp\int_{-\infty}^{+\infty}\frac{\varepsilon_2(\Omega)}{\Omega-\omega}d\Omega \qquad (5.17\ a)$$

$$\varepsilon_2(\omega) = -\frac{1}{\pi}\wp\int_{-\infty}^{+\infty}\frac{\varepsilon_1(\Omega)}{\Omega-\omega}d\Omega\,, \qquad (5.17\ b)$$

where $\omega=2\pi/\lambda$ and c is the speed of light in vacuum.

As a consequence, changes in the dispersion relation of one of these two quantities imply changes in the dispersion relation of the other quantity, in the same spectral interval. In the case of semiconductors, changes in the refractive index have to be expected in the spectral region close to the bandgap of the material.

In the present calculation on optical field intensity along the heterostructure, the important effect of the refractive index is the role it plays in the phase shift of a wave over the thickness of a given layer of material. This phase shift depends is also proportional to the thickness of the material. For thin layers as QWs, there are two facts to consider: one is that, similarly to the case of absorption coefficient, it is not possible to define a refractive index of the layer properly, as the phase shift of a wave across a QW does not scale with the well thickness [Blood00]; secondly, as the QW layer is very thin, the optical path length across the well is very small compared to the optical path length travelled by the wave in the other layers of the heterostructure. Therefore, the following approximation was adopted: the refractive index was kept constant for the QW InGaN layers, and varied according to the dispersion relations found in the literature [Djurisic99], [IOFFEnk] in thicker GaN and SiC layers; The refractive index in the AlGaN layer was assumed equal to that of GaN.

5.3.4. Carrier generation model

Once the optical field profile along the structure is known, one can calculate the generation rate of free electron-hole pairs along the structure. Considering the mean intensity – i.e. energy per unit volume - of the field

$$\bar{I}(z,E)=\frac{1}{2}\frac{\varepsilon(z)}{4\pi}\left(\left|F_{TE}(z,E)\right|^2+\left|F_{TM}(z,E)\right|^2\right), \qquad (5.18)$$

the number of electron-hole pairs generated per unit area is

$$\delta n(z,E)=\bar{I}(z,E)\alpha(z,E)\frac{\eta_i}{E}, \qquad (5.19)$$

where η_i is the intrinsic quantum efficiency. This quantity correspond to the number of electron-hole pairs created by one absorbed photon. If, for bulk transition, one can reasonably assume that the intrinsic quantum efficiency is very close to unity, the same is not true for exciton and quantum well transitions. Excitons created by photons must dissociate in order to form an electron-hole pair able to be collected by the junction field. Similarly, electron-hole pairs created inside a QW must overcome the confining potential of the QW in order to be collected. Presently, we will approximate the internal quantum efficiency approximately equal to one for exciton and QW transitions as well.

5.3.5. Carrier collection

Once the electron-hole pair is created, it must be collected by the junction field. Electrons generated on the p-side of the diode and holes generated on the n-side must diffuse towards the junction, while electron-hole pairs generated inside the space charge region will be immediately collected by the junction field, unless they recombine in the space charge region. The photocurrent can be described, according to a model by Bube [Bube] modified for the present device structure, as a function of the incident photon energy E and of the applied bias V as:

$$J(V,E) = g(V,E)h(V)\,, \qquad (5.20)$$

where

$$g(E,V) = \int_0^{x_p} \delta n(z,E)\exp\left[-\frac{(x_p - z)}{L_n}\right]dz + \int_{x_p}^{x_n} \delta n(z,E)dz + \int_{x_n}^{t} \delta n(z,E)\exp\left[-\frac{(z - x_n)}{L_p}\right]dz \qquad (5.21)$$

is a function describing how many carriers reach the junction for a certain photon energy E and the applied bias V. The dependence on the energy is given by the term δn, while the dependence on the applied bias is given by the fact that x_n and x_p, the depletion region edges, are bias-dependent. The function $h(V)$ is a loss factor which takes into account possible recombination mechanisms at the heterojunctions. If there are N_i interface states per unit area at the junction with capture coefficients equal to β_i, then there is an interface recombination velocity $s_i=\beta N_i$. The form of h is then

$$h(V) = 1/(1 + s_i / \mu\xi(V)) \qquad (5.22)$$

where ξ is the electric field at the junction, $\xi=2(V_{bi}-V)/w(V)$. We see that both g and h increase when the reverse bias $-V$ increases, though the physical reasons are different: the increase of g is due to a higher probability that the generated carrier reach or find themselves in the junction region, whereas the increase of h is due to the higher charge carrier velocity, which diminishes the recombination probability for the carrier collected by the junction.

5.3.6. Example of simulated and experimental photocurrent spectra

Figure [5.7] illustrates the comparison between an experimental photocurrent normalized with respect to the photon flux and a simulation performed with the present model. Together with the total photocurrent simulated spectrum, the contributions to the photocurrent from each single layer are reported. The sample structure is described by the parameters reported in table 5.1. Each material employed has an own dispersion relation for the absorption coefficient, which, in the present case, makes use of the sigmoidal function for the bulk and free-carrier transitions. The bandgap energies are 3.45 eV for GaN and 2.9 eV for the InGaN QWs, both with a broadening parameter equal to 50 meV. The simulated spectrum is not intended as a guide for the understanding of the mechanisms leading to the total photocurrent signal observed, rather than a fit to the experimental spectrum.

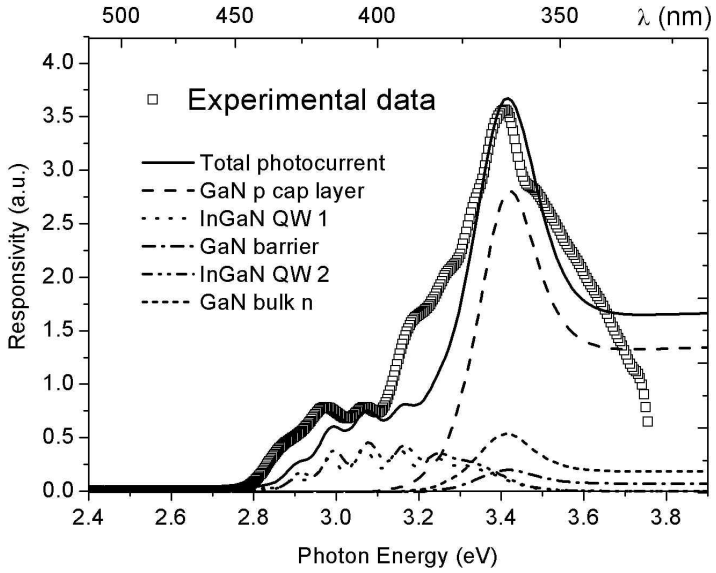


Fig. [5.7]. Experimental and simulated photocurrent of a blue LED.

Table 5.1. Device structure for the simulated photocurrent spectrum reported in fig. [5.7]. The absorption model does not consider here excitonic effects.

Layer	Material	Layer Thickness (nm)	Energy gap E_g (eV)	Broadening D (meV)
1	GaN	200	3.45	50
2	AlGaN	100	4	50
3	GaN	10	3.45	50
4	InGaN	10	2.9	50
5	GaN	10	3.45	50
6	InGaN	10	2.9	50
7	GaN	2900	3.45	50
8	SiC	substrate	3.11	50
Depletion region edges		x_p=200 nm		x_n=385 nm

The spectrum can be described starting from the high-energy side.

a) At high photon energy, photons are absorbed mostly by the GaN p-type cap layer. This yields a significant contribution to the total photocurrent by this first layer, but prevents the light from reaching the QW region in significant amount, which is mirrored by the almost vanishing contribution from the QW layers for energy higher than 3.5 eV.

b) At energy lower than 3.5 eV the absorption of GaN decreases. This allows more light to reach the space charge region. Thus, more photons are absorbed by GaN in the space charge region, yielding a higher photocurrent signal (the maximum is at $E \approx E_g$). At the same time, the QWs, placed in the space charge region, begin to be reached by a higher amount of photons, which yields an increasing contribution by the QWs to the total photocurrent.

c) For photon energies between 3.3 eV and 3.1 eV the absorption by GaN, and thus its contribution to the photocurrent, goes to zero. A higher amount of photons reaches the QWs, whose contribution to the total photocurrent increases.

d) For photon energies between 3.1 eV and 2.8 eV the photocurrent is given exclusively by absorption in the QWs. Due to effects of interference of the optical field in the layered

structure, maxima and minima of field intensity are present in correspondence of the QW region, thus yielding the oscillating pattern of the photocurrent in this energy region.

e) For photon energies lower than 2.8 eV, no photons are absorbed and the photocurrent signal goes to zero.

5.4. Electroluminescence spectroscopy

Luminescence is a term generically identifying a process which has as a final result the emission of photons from a material. The spectral analysis of luminescence consists in the collection of a narrow spectral interval of the light emitted by the analysed sample and in the detection of its intensity by means of a radiation detector. It is a widely used technique in semiconductor physics and technology, as it gives important information on the radiative transition which take place in the materials and devices. It is therefore apparent that this technique is of particular interest in optoelectronic devices like LEDs, where the principles of operation themselves are based on luminescence. In order to excite luminescence in semiconductors, a relatively high concentration of electron-hole pairs must be generated. This can be done most commonly by electrical injection as in Electroluminescence (EL), by optical excitation as in Photoluminescence (PL) or with an electron beam as in Cathodoluminescence (CL). In the present work, electroluminescence was employed in the analysis of the degradation of nitride-based LEDs.

5.2.1. Radiative electron-hole recombination

Excess carriers in semiconductors can be generated in different ways. In the electroluminescence technique, they are injected in the material electrically. The rate at which the carrier concentration in valence and conduction band decreases is denoted by R. The probability for an electron in the conduction band to recombine with a hole in the valence band is proportional to the hole concentration, therefore $R \propto p$. The number of recombination events is also proportional to the electron concentration n. Thus one has

$$R = -\frac{dn}{dt} = -\frac{dp}{dt} = Bnp\,, \qquad (5.23)$$

which is called the bimolecular rate equation, with the proportionality constant B roughly equal to 10^{-11}-10^{-9} cm^3/s for III-V semiconductors.

A forward biased p-n junction is a system in which it is possible to achieve a high concentration of electron-hole pairs. This high concentration is reached in most cases in correspondence of the p-n junction. If a steady state is reached, the excess electrons and holes concentrations are equal, as these particles recombine in pairs:

$$\Delta n(t) = \Delta p(t)\,. \qquad (5.24)$$

The bimolecular rate equation can thus be written as

$$R = B[n_0 + \Delta n(t)][p_0 + \Delta p(t)]\,. \qquad (5.25)$$

In the case of low level excitation, one has

$$R = Bn_i^2 + B(n_0 + p_0)\Delta n(t) = R_0 + R_{excess}\,. \qquad (5.26)$$

If we suppose that the carrier generation rate by injection is switched off at t=0, one obtains the rate equation:

$$\frac{d}{dt}\Delta n(t) = -B(n_0 + p_0)\Delta n(t) \qquad (5.27)$$

for which the solution is

$$\Delta n(t) = \Delta n_0 \exp[-B(n_0 + p_0)t], \qquad (5.28)$$

i.e. an exponential decay with time constant $\tau=[B(n_0+p_0)]^{-1}$. If the excitation level is high, the excess carrier concentration is larger than the equilibrium carrier concentration, $\Delta n >> n_0+p_0$. The bimolecular rate equation is then given by:

$$\frac{d\Delta n}{dt} = -B\Delta n^2 . \qquad (5.29)$$

The solution for this differential equation is

$$\Delta n(t) = \frac{1}{Bt + \Delta n(0)^{-1}} . \qquad (5.30)$$

For this non-exponential decay, one can operatively define a “time constant” as

$$\tau(t) = -\frac{\Delta n(t)}{\frac{d\Delta n(t)}{dt}} = t + \frac{1}{B\Delta n(0)} , \qquad (5.31)$$

which shows that under high excitation conditions the carrier lifetime increases with time. Eventually, low level excitation conditions will be reached and the lifetime will become constant.

5.2.2. EL in quantum wells

Quantum wells provide confinement of carriers to a narrow region of thickness L_w. If the conduction and valence bands in the wells have 2-dimensional concentration of carriers of n^{2D} and p^{2D}, respectively, the effective 3-dimensional concentration will be given approximately by $n^{3D} = n^{2D}/L_w$ and $p^{3D}= p^{2D}/L_w$.

The recombination rate becomes then

$$R = = B\frac{n^{2D}}{L_w}\frac{p^{2D}}{L_w} . \qquad (5.32)$$

This result explains that much higher recombination rates can be found in QWs, as much higher 3D concentrations of carriers can be achieved, so long the envelope wave function scales with the QW thickness.

5.2.3. Deep levels and non-radiative recombination

During non-radiative recombination, the electron energy is converted to vibrational energy of lattice atoms. The reasons why non-radiative recombination occurs are to be found in the presence of certain defects in the material, especially those giving rise to deep levels in the gap. The non-radiative recombination concurs with the radiative recombination, so non-radiative recombination carriers are also called "luminescence killers" [Schubert]. In extrinsic semiconductors and under small deviations from equilibrium, the recombination rate is limited by the rate of capture of minority carriers by deep levels in concentration N_t. In a p-type semiconductor, for instance, the electron lifetime is:

$$\frac{1}{\tau} = \frac{1}{\tau_{n0}} = N_t v_n \sigma_n . \qquad (5.33)$$

and in an n-type semiconductor, the hole lifetime is:

$$\frac{1}{\tau} = \frac{1}{\tau_{p0}} = N_t v_p \sigma_p , \qquad (5.34)$$

where $v_{n,p}$ and $\sigma_{n,p}$ are the thermal velocities of electrons and holes and the capture cross section of the deep level for electrons and holes, respectively. However, it is important to notice that in some cases deep levels do assist radiative recombination. Thus, the minority carrier lifetime calculated by SRH statistics is indeed connected to the radiative recombination rate.

Another important mechanism of non-radiative recombination is Auger recombination. In this process the energy becoming available through electron-hole recombination is dissipated in the excitation of (i) a free electron into higher regions of the conduction band , or of (ii) a hole into deeper regions of the valence band. The recombination rates due to Auger processes are:

$$R_{Auger} = C_p n p^2 \qquad (5.35)$$

and

$$R_{Auger} = C_n n^2 p . \qquad (5.36)$$

In the high excitation limit, the recombination rates becomes proportional to n^3. Auger recombination is effective in reducing the radiative recombination only in the limit of high excitation, as typical Auger coefficients are in the range 10^{-28} – 10^{-29} cm^6/s for most III-V semiconductors [Schubert].

For the purposes of this thesis, it is sufficient to consider the non-radiative recombination due to the deep levels. In the study of the degradation of efficiency of LEDs, in fact, a correlation can be established between a decrease of the optical efficiency of the device, which is observed by comparing EL spectra after successive stress stages, and the increase in concentration of certain deep levels. In other cases, for instance when a radiative recombination centre is introduced, one can notice the increase of both EL and DLTS signals, which must be energetically correlated.

5.5. Experimental setup

5.5.1. Light sources

Light sources are one of the basic building blocks of many optical spectroscopy experiments. There is a huge variety of light sources, according to the type of spectrum, spectral interval of emission, directionality, etc. The spectrum can be discrete, as in Hg lamps, continuous, as in Quartz Tungsten Halogen (QTH) lamps, or a superposition of continuous and discrete spectrum, as in Xenon Arc lamps. For the present photocurrent experiments a continuous, black-body-like spectrum is preferred, and a lamp with emission in the near UV (λ >200 nm) and in the visible part of the spectrum is needed, because the bandgap of GaN is 3.45 eV, corresponding to a wavelength roughly equal to 360 nm. The spectrum of a typical QTH lamp is reported in fig. [5.8].

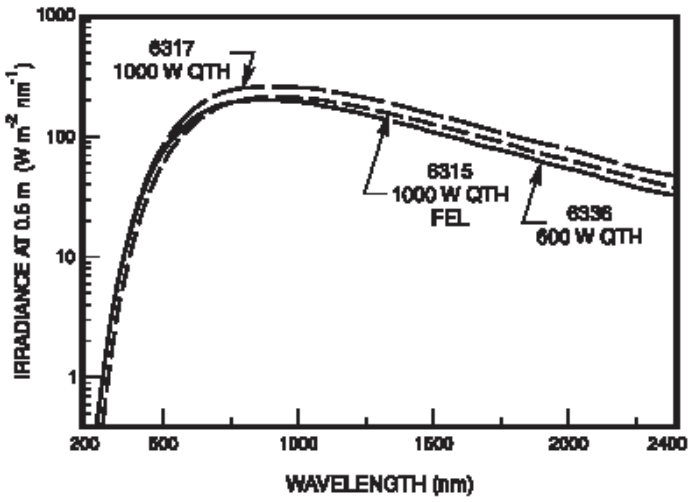


Fig. [5.8]. Irradiance spectrum of QTH lamps with different power

5.5.2. Monochromator

The monochromator is the element which allows the selection of a narrow band $\Delta\lambda$ of wavelengths from a source of radiation. The selected wavelengths can be directed to the sample to analyse, as in absorption or photocurrent spectroscopy, or can be extracted from the emission of the sample, as in luminescence experiments. In the latter case, the monochromator is also called a spectrometer. The monochromator consists typically of a box with an entrance and an exit slit for the light and mirrors and a diffraction grating inside. The principle of operation of the monochromator grating is illustrated in fig. [5.9].

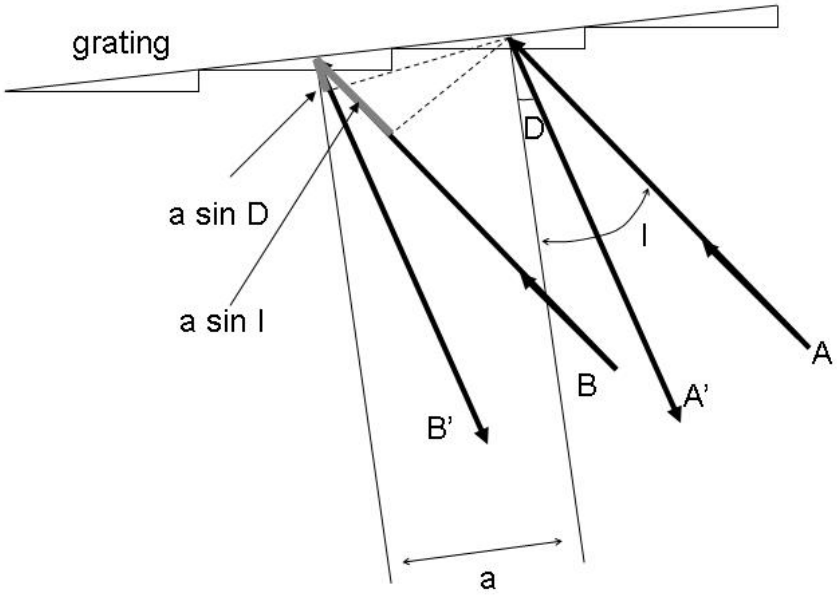


Fig. [5.9]. Scheme illustrating the difference of optical path length between two neighbouring light rays reflected by the grating of a monochromator.

As one can see from the scheme, the condition for constructive interference of neighbouring rays is

$$a\,(\sin D + \sin I) = m\lambda, \qquad (5.37)$$

with m integer. Thus, the angle of incidence I of the light on the grating can be varied in order to obtain the desired wavelength directed on the exit slit. One must consider the fact that, for a certain λ satisfying the grating equation, $\lambda/2$, $\lambda/3$,…, λ/m satisfy it as well. Extreme care is necessary to avoid confusion between harmonics, which in certain circumstances requires the use of filters in order to stop the higher diffraction orders. The efficiency of the grating, i.e. the portion of reflected light, is not equal at all wavelengths. There is a particular wavelength for which the efficiency is maximum, which is called the *blaze* wavelength of the grating. An important parameter of the grating is the dispersive power. This is given by

$$dD/d\lambda = (\sin I + \sin D) / \lambda \cos D \qquad (5.38)$$

For the monochromator employed in the present experiments the average dispersive power is D = 1.96 nm/mm.

The resolution of the monochromator is calculated starting from the following spectra:

- $B(\lambda)$ spectrum of the light source
- $P(\lambda)$ line profile of the monochromator
- $F(\lambda)$ exit spectrum of the system monochromator + source

The total spectrum $F(\lambda)$ of the system is given by

$$F(\lambda) = B(\lambda) \otimes P(\lambda), \qquad (5.39)$$

i.e. the convolution of the source spectrum and the monochromator line profile. The line profile itself is the result of the convolution of different factors:

$$P(\lambda)=P_1(\lambda) \otimes P_2(\lambda) \otimes P_3(\lambda) \ldots \otimes P_n(\lambda) \qquad (5.40)$$

Where the various factors correspond to different monochromator elements: the entrance and exit slits (1 and 2), the grating profile (3), and other minor effects (n). In our case the resolving power is approximately of 1 nm, which corresponds, for a wavelength of 400 nm, to a resolution of the photon energy approximately equal to 10 meV. The total spectra of the system Quartz Tungsten Halogen Lamp + monochromator are shown in fig. [5.10].

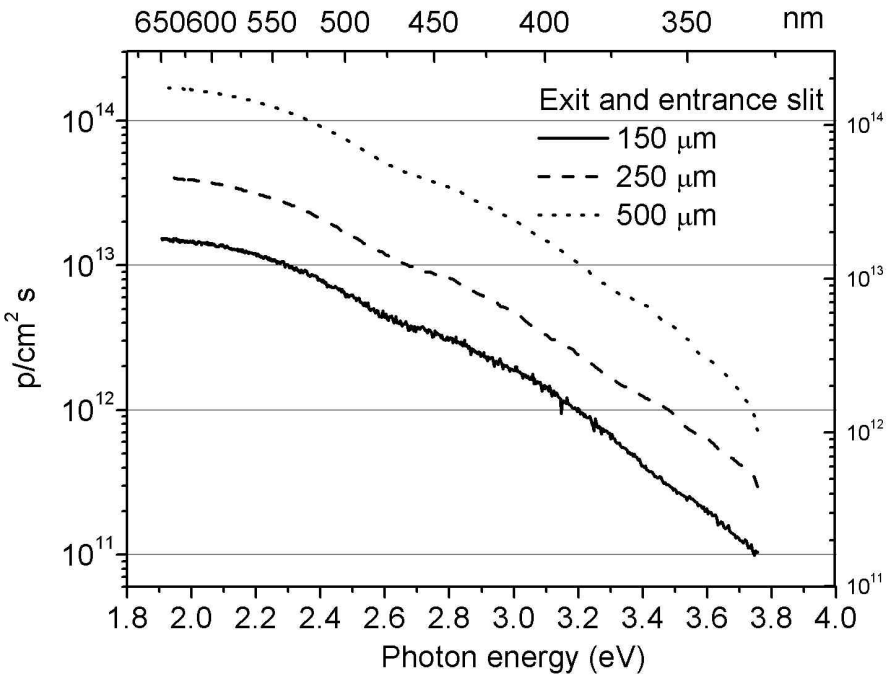


Fig. [5.10]. Spectrum of the system QTH lamp + monochromator with different exit and entrance slit width.

5.5.3. Thermopile detector

The radiation detectors used in the electroluminescence experiments were thermal detectors. This class of detectors works by converting the incident radiation into a temperature rise: this latter change is measured in several ways. In our case, we made use of a thermopile, which is a series connection of a certain number of thermocouple junction. A thermocouple junction consists of two dissimilar metals connected in series. To detect the radiation, one junction is blackened to absorb the radiation. The temperature rise of the junction generates a voltage. An increase in the output voltage is obtained by increasing the number of thermocouple junctions. All "hot" junctions are placed close together and darkened for an efficient collection of the radiation intensity. This constitutes the thermopile (fig. [5.11]). Typical parameters of thermopiles are the time constant of 50-100 ms and DC responsivities of the order of 100-200 μA/W, with a collection area of some squared millimetres. The spectral response of thermal detectors is usually broadband and wavelength-neutral. Should the detector have a more complex spectral response, its spectrum $D(\lambda)$ must be convolved with the spectrometer profile $P(\lambda)$ in order to obtain the total response spectrum.

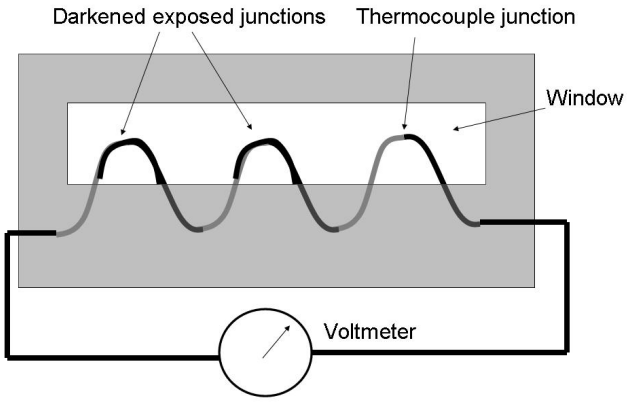


Fig. [5.11]. Schematic drawing of a thermopile detector.

5.5.4. Lock-in amplifier

The lock-in technique is used to measure very small AC signals, even when they are buried in noise. A lock-in is a filter with a very narrow bandwidth, tuned to the frequency of the signal. The filter rejects most of the noise, having Q factors as high as 10^6. In addition to filtering, the lock-in also provides gain, up to 10^9. The basic principle of the lock-in is that the experiment is performed at a fixed frequency, possibly far from the frequency regions where significant noise source can be present. In the present experiments a reference signal is fed to the lock-in by means of a photodiode sensing the on- and off-states of the chopper. The reference signal, as illustrated in fig. [5.12] has the frequency $f_{ref}=\omega_{ref}/2\pi$. This signal enters a Phase-Lock Loop (PLL) circuit which tracks the input signal frequency. The lock-in is then capable to extract the first harmonic of this square wave, in form of a sinusoidal function $V_{ref}=V_L\cos(\omega_{ref}t+\theta_{ref})$ by means of a precision sine converter. The other signal may be either the photocurrent signal from the sample or the signal from the light detector. This latter is the signal to be measured, and will consist of harmonic components of the form $V_s=V_{sig}\cos(\omega_s t+\theta_s)$. The signal $V_{sig}\cos(\omega_s t+\theta_s)$ is amplified by a high gain AC coupled differential amplifier. The output of this amplifier is then multiplied by the reference output in a phase-sensitive detector. This multiplication shifts each frequency component of the input signal by the reference frequency, so the output is given by:

$$V_{psd} = V_{sig}\cos(\omega_{ref}t+\theta_{ref})\cos(\omega_s t+\theta_s) =$$

$$= \tfrac{1}{2} V_{sig}\cos[(\omega_{ref}-\omega_s)t+\theta_s-\theta_{ref}] + \tfrac{1}{2} V_{sig}\cos[(\omega_{ref}+\omega_s)t+\theta_s+\theta_{ref}] \qquad (5.41)$$

The sum frequency component is attenuated by a low-pass filter, and only the difference frequency components within the low-pass bandwidth will pass to the DC amplifier. As the low-pass filter has time constants up to 100 seconds, the lock-in can reject noise which is more than 0.0025 Hz away from the reference input frequency [SR530Man].

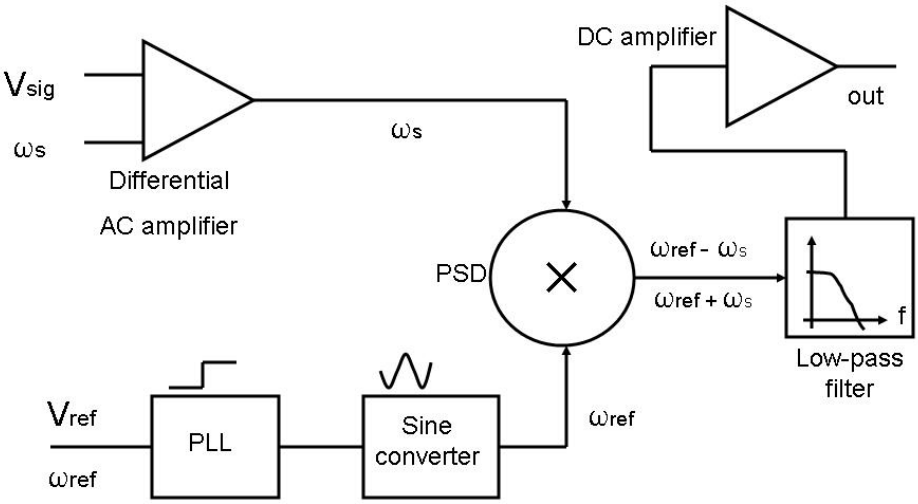


Fig. [5.12]. Block diagram of the lock-in measurement technique

5.5.5. Photocurrent setup

The photocurrent setup is illustrated in the block diagram of fig. [5.13]. The light of a QTH lamp is modulated by a chopper rotating with a particular frequency f_{ref}. A photodiode mounted on the chopper monitors the dark-light period, sending a square-wave voltage signal to the reference input of the lock-in amplifier. The white light enters the monochromator through the S1 slit, is diffracted on the grating and it exits the monochromator through the slit S2 as monochromatic light. The monochromatic light falls then onto the sample giving rise, according to the sample features and the wavelength, to the photocurrent signal, which is modulated with the same frequency f_{ref} of the chopper. The lock-in amplifier, which must be set with a time constant sufficiently higher than the chopper period, analyses the signal yielding a final photocurrent DC signal, which is recorded on the computer together with the wavelength value. The wavelength scan of the monochromator must be sufficiently slow, in order to allow an accurate response of the lock-in (which mostly depends on the low-pass filter time constant). So, if $\Delta\lambda$ is the spectrometer resolution, and τ is the lock in time constant, the scanning rate of the monochromator must be set lower than or approximately equal to $\Delta\lambda/\tau$.

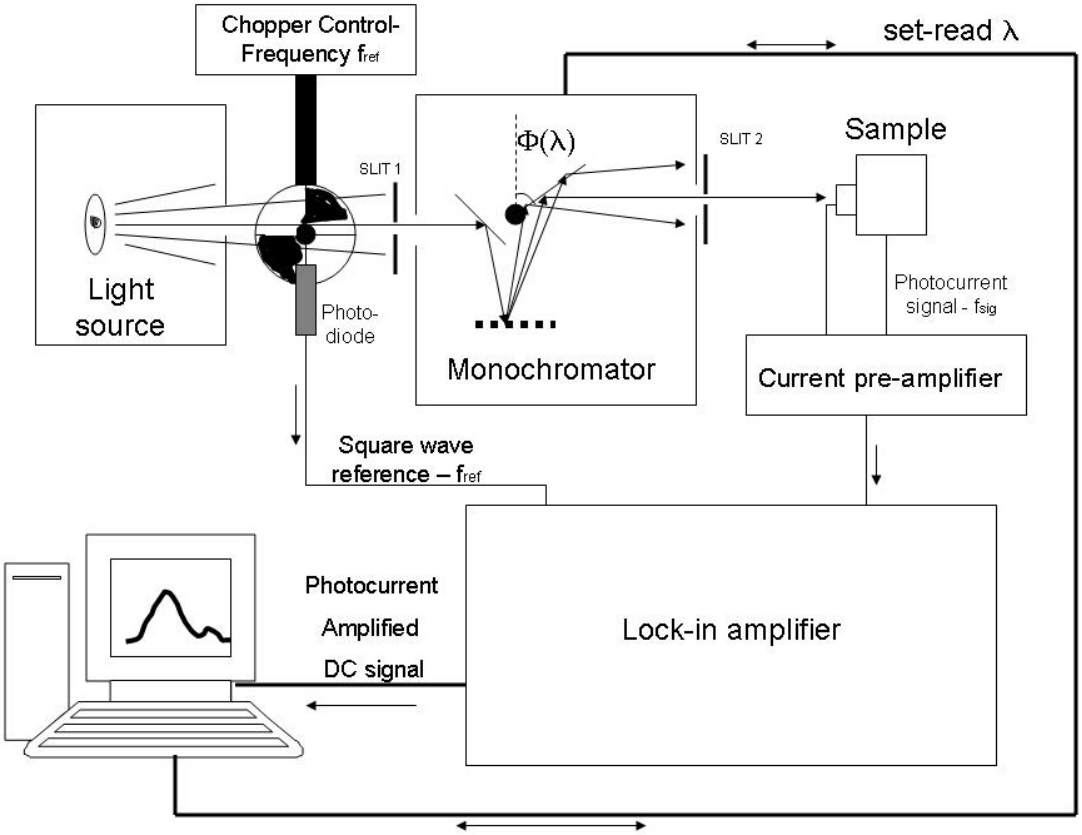


Fig. [5.13]. The block diagram for the photocurrent experiments.

5.5.6. Electroluminescence setup

The electroluminescence setup for the experiments on the LED performed at the University of Bologna is shown in fig. [5.14]. Basically, it is the same setup for the photocurrent, but with the role of the analysed diode changed from detector to light source. The monochromator is here used as a spectrometer. The light emitted by the sample enters the monochromator from slit 2, is diffracted by the grating and the light of determined spectral interval Dl exits the monochromator through slit 1. The light is then chopped with 50% duty cycle and reaches the thermopile detector. The photodiodes provides the reference signal for the lock-in, where the detector modulated signal is elaborated and amplified, in exactly the same way as for PC. This setup is good for samples like LEDs, which have a high enough brightness: much of the emitted light is dispersed in other direction and not collected on the detector: therefore, a minute fraction of the emitted light is detected. Other very efficient EL setups have been used for measurements at the universities of Padova and Parma on the same samples; these setups are designed for an efficient collection of the emitted radiation and for a wide interval of measurement temperatures.

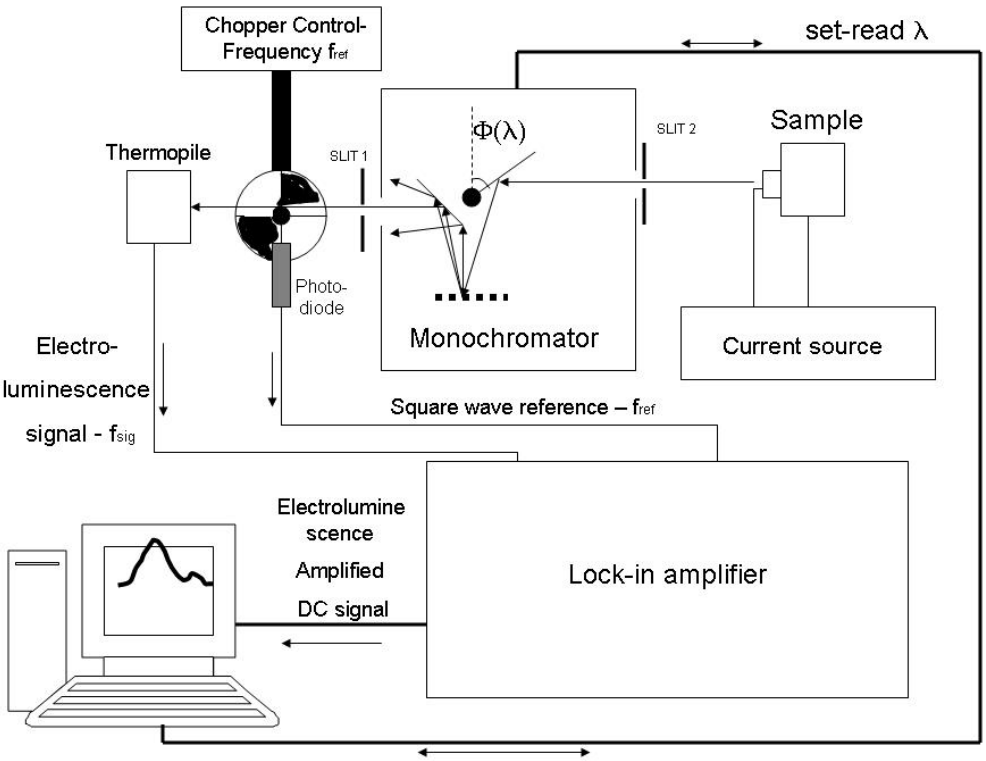


Fig. [5.14]. The block diagram for the electroluminescence experiments.

REFERENCES

[Ambacher96] Ambacher, O., W. Rieger, P. Ansmann, H. Angerer, T.D. Moustakas, M. Stutzman, *Sol. State Commun.* **97(5)**, 365-370 (1996).

[Blood00] P. Blood, *IEEE J. Quantum Electron.* **36**, 354 (2000)

[BornWolf] M. Born, E. Wolf: *Principles of Optics*, Pergamon Press.

[Bube] R.H. Bube, *Photoelectronic properties of Semiconductors* (Cambridge University Press, 1992)

[Bulutay99] C. Bulutay, N. Dagli, A. Imamoğlu, *IEEE J. Quanutm Electron.* **35**, 590 (1999)

[Davies] J. H. Davies, *The physics of low-dimensional semiconductors*, (Cambridge University Press, Cambridge, 2000).

[Djurisic99] Djurisic, A.B., Li E.H., *J. Appl. Phys.* **85**, 2848-2853 (1999).

[Franssen04] G. Franssen et al., *Phys. Rev. B* **69**, 045310 (2004)

[Friel04] I. Friel, C. Thomidis, Y. Fedyunin and T.D. Moustakas, *J. Appl. Phys* **95** p. 3495 (2004).

[Guo94] Guo, Q, Yoshida A., *Jpn. J. Appl. Phys.* ***33*** *(1994), 2453-2456.*

[HaugKoch] H. Haug and S.W. Koch, *Quantum theory of the optical and electronic properties of semiconductors* (World Scientific, Singapore, 1990).

[IOFFEnk] *http://www.ioffe.ru/SVA/NSM/nk/index.html* dispersion relations various materials

[Martin99] R.W. Martin, P.G: Middleton, K.P. O'Donnell and W. Van der Stricht, *Appl. Phys. Lett.* **74**, 263 (1999)

[Muth97] Muth, J.F, J.H. Lee, I.K. Shmagin, R.M. Kolbas, H.C. Casey, Jr., B.P. Keller, U.K. Mishra, S.P. DenBaars, *Appl. Phys. Lett.* **71(18),** 2572-2574 (1997).

[Poruba00] A. Poruba et al., *J. Appl. Phys* **88**, p.148 (2000).

[Schubert] E.F. Schubert, *Light-Emitting Diodes* (Cambridge University Press, 2003)

[SR530Man] Model SR530 Lock-in amplifier, Instruction manual

6 Evolution of defect states in nitride-based Light-Emitting Diodes

The long term behaviour of commercial LEDs is quite good as verified by long term tests of OSRAM products for modifications of designs and processes. The following investigation deals with special test structures with worse degradation behaviour in order to get information about microscopic and macroscopic degradation mechanisms associated with carrier transport. The literature about this topic describes many effects, produced by both charge flux and thermal heating processes, as responsible of LEDs degradation: modifications of the electrical contacts [Barton99], changes in the local In concentration in the QW [Youna03], [Chuo01], formation of radiative and non-radiative centres [Egawa97], [Cao03], changes in the charge-injection mechanisms across the cladding and barrier layers (pure tunnelling, trap/phonon assisted tunnelling, thermal emission etc.) [Polyakov02], aggregation or breaking of complexes mainly involving hydrogen impurities and magnesium p-dopant [Pavesi04]. In the last years, the failure modes responsible for the output optical power reduction have been intensively investigated, mainly by injecting electrical currents in order to simulate the working applicative conditions. However, in most of the cases high values of the drive current have been used in order to perform the so called “accelerated tests”. Many experimental data are therefore masked by contributions of thermal effects, arising from Joule dissipation of the current itself. In this situation the understanding of the degradation mechanisms results to be difficult due to the overlapping of two different aging inputs. The purpose of the present work is the aging of test structures under current values (20 mA, corresponding to a current density of 32 A/cm^2) comparable or lower than those employed in the operation modes, in order to exclude any thermally-induced degradation (thermal maps have shown that the junction temperature never exceeds 60°C during stress test).

6.1. Samples, treatment and experimental method

Test structures, similar to those illustrated in section (2.1.1), grown by metal-organic vapour phase epitaxy on silicon carbide substrates have been investigated. The device vertical structure consists of a buffer layer, a 2mm thick Si-doped GaN layer, an active region with an InGaN/GaN multiple quantum well system, a 100 nm thick heavily Mg-doped AlGaN cladding layer and a 200 nm Mg-doped GaN contact layer. The chips were covered with a SiN passivation layer. Squared chips with 250 µm sides, bonded on metallic packages TO18 with a bi-component epoxy resin acting as ohmic cathode contact, have been studied at an initial unstressed stage and at different steps of a 20 mA (32 A/cm^2) DC-aging treatment. This was performed by biasing the samples with Keithley 220 current source, at room temperature in air, up to a total time of 100 hours. The number of steps and the total stress time was not the same for all analyzed LEDs. The sample name, the stress steps and the experimental techniques employed for the analysis of each individual sample at every stage are summarized in table 6.1.

Tab. 6.1 stress steps and the experimental technique employed for the analysis of each individual sample at every stress stage.

Sample name	Stress stages	Experimental techniques
HP	15 min, 30 min, 60 min	I-V, C-V, DLTS
HG	2 h, 5 h, 10 h, 20 h, 50 h, 100 h	I-V, C-V, DLTS
HJ	15 min, 30 min, 60 min, 2 h, 5 h, 10 h, 25 h, 50 h, 100 h	PC, EL

The analysis was performed in the following way: InGaN-based test structures for blue LEDs have been electrically and optically characterized before any aging treatments (untreated samples) and then submitted to low-value DC current aging and examined at different aging steps. Results concerning: electrical current-voltage (I-V) characteristics and capacitance-voltage (C-V) measurements, Deep Level Transient Spectroscopy (DLTS) analyses, monitoring the traps present in the devices, the changes of their concentration and the eventual growing up of new centres during the aging treatments, optical device emission and absorption, studied by complementary techniques of ElectroLuminescence (EL) and PhotoCurrent (PC) spectroscopy are discussed.

6.2. I-V characterization

The current vs. voltage (I-V) measurements were carried out at room temperature. Measurements at higher temperature would have yielded further elements about the transport mechanisms, but were accurately avoided in order to introduce no thermally activated degradation factors. Before stress, the individual devices (coming from the same wafer) showed identical forward I-V curves, while the reverse current was slightly different from sample to sample, possibly due to different dislocations densities [Cao04][Fang00][Li04]. Generally speaking, the stress did not modify the forward region of the I-V curves, inducing however an increase in reverse current. The magnitude of this reverse current increase was different for each LED, as it was more pronounced in sample HJ, stressed with steps 2h, 5h, 10 h, 20 h, 50 h, 100h, than in sample HP, stressed with steps 15 min, 30 min, 60 min. This suggests that the current increase is due to mechanisms depending on sample quality, such as generation/propagation of threading dislocations in the active region [Fang00] and/or modifications of the surface states at the mesa sidewalls.

6.2.1. LED HP

Figure [6.1] reports the I-V characteristics of sample HP, stressed with steps 15 min, 30 min, 60 min. The forward characteristics, which remained unaltered over the stress stages, show a complex behaviour: there is an initial plateau from 0 V to 1 V, where the current grows extremely slowly, remaining to levels of few picoamperes. Then, one can identify three regions of exponential growth, well described by the expression

$$I = I_s \exp [qV/nkT] \qquad (6.1)$$

with different ideality factors for each region (table 6.2). The reverse current is best fitted assuming the soft-breakdown characteristics:

$$I = B\, V^m \qquad (6.2)$$

where the exponent m grows slowly from 6.7 to 7.2 during the stress. However, the reverse characteristics do not appear to change significantly.

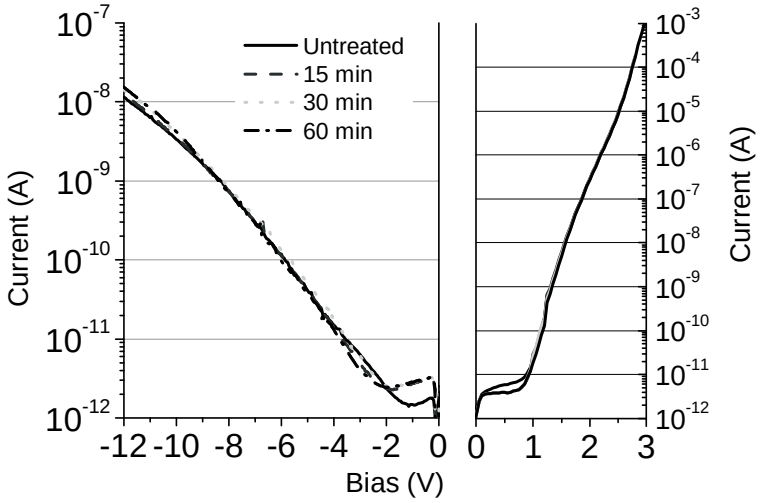


Fig. [6.1]. The I-V characteristics of sample HG at progressive stress stages. The reverse and the forward characteristics are plotted on different scales for both x and y coordinates.

6.2.2. LED HG

Figure [6.2] reports the I-V characteristics of sample HG, stressed with steps 2h, 5h, 10h, 20h, 50h. As in the case of LED HP, the forward characteristics remained unaltered over the stress stages, and can be described according to the exponential expression used in the previous case with different ideality factors for each region (table 6.2). The reverse characteristics follow the soft-breakdown model, but the evolution is different than in HP. The exponent *m* remains confined between 5 and 6, with no trend over the stress steps, showing that the transport mechanism remains unchanged over time. The reverse current intensity, which is about one order of magnitude lower than in HP, grows over the stress steps. We will return on this hypothesis later, when dealing with further observations. Fig. [6.3] shows the increase of the reverse current level as a function of stress time, calculated as the average value of the reverse current in the interval [-11, -6] V and normalized to the value in the untreated sample. The trend reminds of a logarithmic evolution in time.

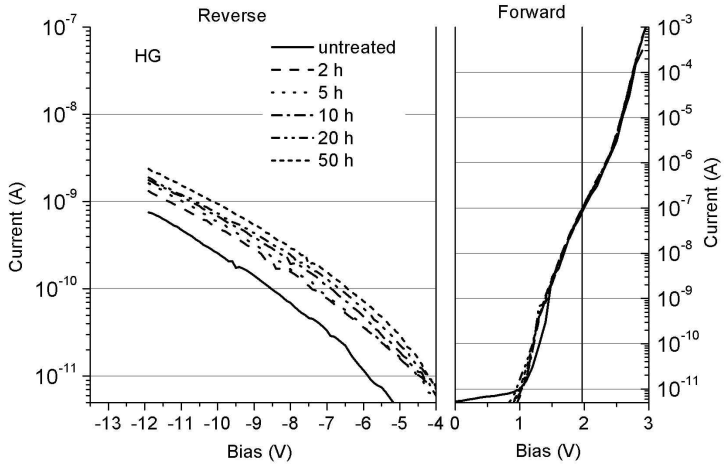


Fig. [6.2] The I-V characteristics of sample HG at progressive stress stages. The reverse and the forward characteristics are plotted on different scales for both x and y coordinates.

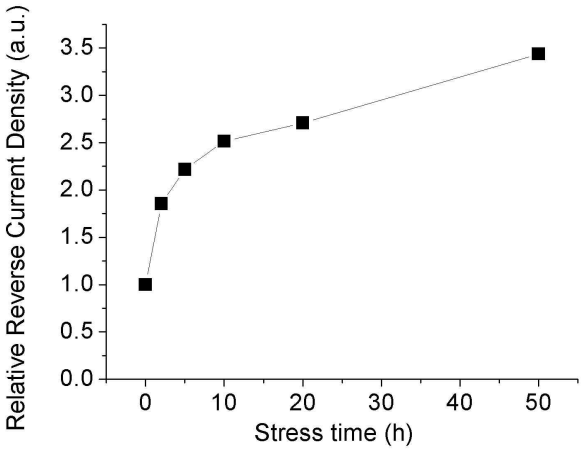


Fig. [6.3] The relative variation of the reverse current (calculated as an average on the interval [-11,-6] V and then normalized) in sample HG as a function of stress time.

6.2.3. Transport mechanisms

Table 6.2 reports the ideality factors and extrapolated zero-bias currents calculated for both HP and HG LEDs. Both samples, in fact, have identical forward characteristics, which do not change with stress. The very high values of the ideality factor in all three considered bias regions can be explained either with the tunnelling model [Sze] or with the heterojunction multi-barrier model [Shah03], section (3.1.6).

Table 6.2 Ideality factors and extrapolated zero-bias currents for the analyzed LEDs

Forward bias region (V)	Ideality factor n	Extrapolated zero-bias current I_s (A)
0.8 – 1.7	4.1±0.1	$(2.6\pm0.2) \times 10^{-15}$
1.8 – 2.4	5.4±0.2	$(1.7\pm0.4) \times 10^{-13}$
2.6 – 3.0	3.9±0.2	$(1.7\pm0.4) \times 10^{-16}$

6.3. C-V characterization

Capacitive measurements were carried out by means of a Keithley 3300 LCR Meter; The apparent charge profile of the devices can be extracted from capacitance vs. voltage (*C-V*) measurements, as described in chapter 3. The model used for these LEDs assumes the junction to be unilateral, with the p-side more heavily doped ($p>>n$) with respect to the n-side: for this reason the profile extracted from the C-V measurements is referred to the n-side. It is worth noticing that this assumption introduces a slight approximation in the determination of the apparent depths and levels, since the charge concentration at the p-side is not infinite, and the SCR extends both at the p and at the n side. However, it is a good approximation for this kind of devices, where usually the p region is much more doped than the n region. The C-V measurements were carried out at 100 kHz or 1 MHz. Here we show the data obtained with a 1 MHz frequency, which allowed us to obtain a good signal to noise ratio and to reduce the contribution of the diffusion capacitance on the measured capacitance for forward bias values, providing reliable junction capacitance measurements also for low positive voltages [Lucia93].

6.3.1. LED HP

In Fig. [6.4] are shown the results of the C-V measurements carried out before and during the ageing test at 1 MHz on sample HP. The solid curve corresponds to the untreated sample. As can be noticed, the junction capacitance shows a step-like behaviour. With the decrease of the reverse bias the boundary of the space charge region is swept through the barriers and the wells: as it moves inside the wells, the capacitance changes very slowly with voltage, due to the large amount of carriers [Ershov01] [Moon98]. This behaviour is characteristic of quantum-well structures.

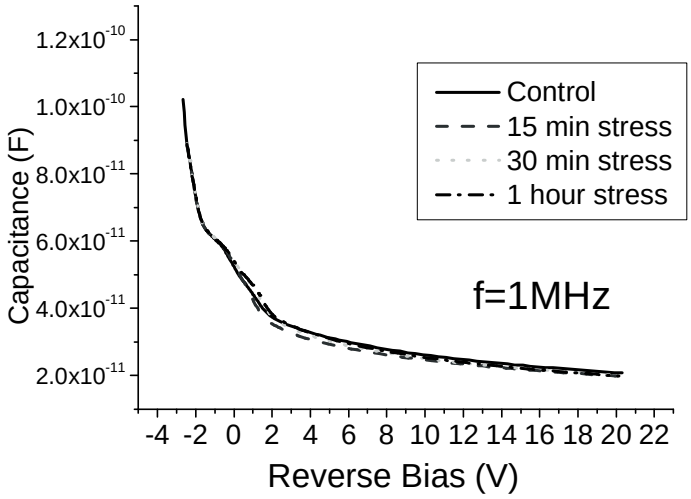


Fig. [6.4] Capacitance-Voltage characteristics of sample HP.

The results of the C-V measurements were processed in order to obtain the $n_{CV}(x)$ and $n_{CV}(V)$ profiles, which are shown in Fig. [6.5]. The solid curve and the squares are referred to the untreated samples. The junction is considered to be in x=0. The spatial dependence of this charge profile suggests that during voltage sweep the SCR boundary is moved through the active region (the peaks at 80 and 100 nm, corresponding to +1V and -0.5V, respectively) and the n-side (for greater depths) [Polyakov02], [Schroder], [Fang00], [Ershov01], [Moon98]. Changes in the apparent charge density occur.

- After the first stress step (15 min) the apparent charge density decreases both in the active region and in the bulk. The peak on the right side of the QW region is almost no longer resolvable.

- After the stress stages of 30 min and 60 min, the apparent charge density tends to increase, both on the right side of the QW region and in the bulk. The increase in the QW region is more visible, as the peak at 100 nm becomes visible again and its height grows.

These changes are also visible in the bias-profile of the apparent charge. The bias profiles have been reported because of their connection with the DLTS measurements. DLTS measurements, in fact, have been performed in the following way, with a determined set of reverse bias – filling pulse values for all stress stages. Due to the modifications in the charge profile, the depletion region edge corresponding to one reverse bias at a certain stress stage is not the same depletion region edge at the same reverse bias at the successive stress stage. Thus, the region probed by DLTS after successive stress stages is not exactly the same.

The observed variations of the charge profile may be related to the generation or modification of levels or interfacial states near the border of the active region, which could locally change the charge profile, modify the capture efficiency of the quantum-wells, and/or contribute to the

generation of non-radiative paths in the bandgap. Also the slight and uniform n_{CV} increase detected for depths greater than 150 nm after the second stress stage possibly reflects a generation and propagation of defects.

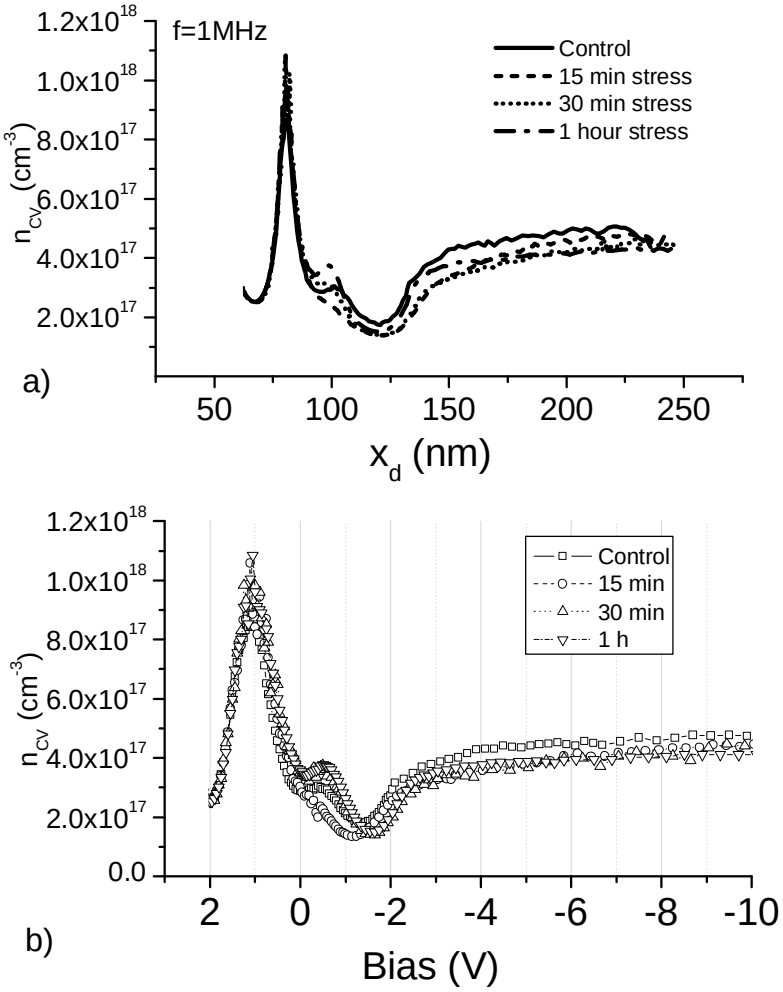


Fig. [6.5]. The profile of apparent charge distribution n_{CV} as a function of a) junction depth and b) reverse bias at different stress stages in sample HP, measured at 1MHz frequency.

6.3.2. LED HG

The features of the C-V characteristics of sample HG, shown in fig. [6.6], are very similar to those of sample HP. The same step-like behaviour, typical of quantum well structures, can be noticed in these curves, measured at 1MHz.

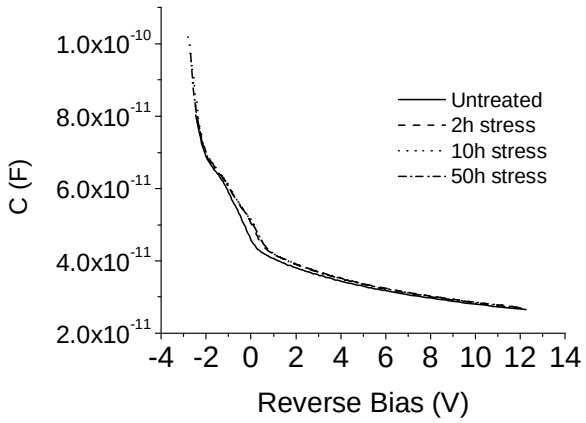


Fig. [6.6] Capacitance-Voltage characteristics of sample HG.

The depth- and bias-profiles of the apparent carrier density n_{CV} are reported in fig. [6.7]. The peaks of the QW active region are visible at depth 75 nm and 90 nm, corresponding to +1.5V and 0.5V, respectively. A region slightly depleted of carriers is then found between 100 and 130 nm. At higher depth is the bulk n-region, showing a concentration gradient most likely due to a doping gradient.

The evolution of the charge profile is the following: stress produces an overall increase in the apparent charge concentration, mostly localized at the interface between the QW region and the depleted region between QWs and bulk. The peak on the right side of the QW region becomes higher. The variation of the charge concentration occurs mainly in the first stress stage (2h), then the increase becomes slower. The charge concentration in the bulk remains constant over all stress stages. The bias profiles indicate that, after the first stage of stress, the same reverse bias corresponds to the same depletion region edge. This will have the consequence that DLTS measurements made with the same parameters (reverse bias – filling pulse) after the first stress stage will probe the same region of the device.

There are some differences and similarities between HP and HG, both as the structural parameters and the degradation behaviour concerns. We can already list some similarities and differences with sample HP, starting from the structure. The similarities are:

- There is a quantum well region, which is about 40 nm thick.
- There is a depleted region between the QW region and the bulk, due to the presence of the QWs that bend the conduction band upward when filled with electrons.
- There is a bulk n-region

The most important of these similarities is surely that regarding the QW region: this is the device active region, which determines the light emission properties of the LED. However, the differences are also relevant:

- Sample HG is overall more doped: the doping density on the n bulk side is about $4 \times 10^{17} cm^{-3}$ in HP, and $5\text{-}6 \times 10^{17} cm^{-3}$ in HG.
- There is a higher doping gradient on the bulk side of HG than in HP.
- In HG the active region is closer to the junction region

This set of observation is consistent with the fact that both samples come from the same wafer, as the arrangement of the layers and of the functional regions is the same. However, as it often occurs, different parts of the same wafer can be affected by doping non-uniformity. The doping variations are minimal, and do not yield significant modifications in the device operation. However,

they can yield some differences in the study of complex mechanisms as those involved in the degradation.

The evolution of the charge concentration profiles is similar in both cases: with exception of the first stress stage of HP, the charge tends to increase. This increase is much more localized in sample HG, where it is the only QW-related peak at *x*=90 nm to be involved, while in HP the charge increases both in the QW region and in the bulk.

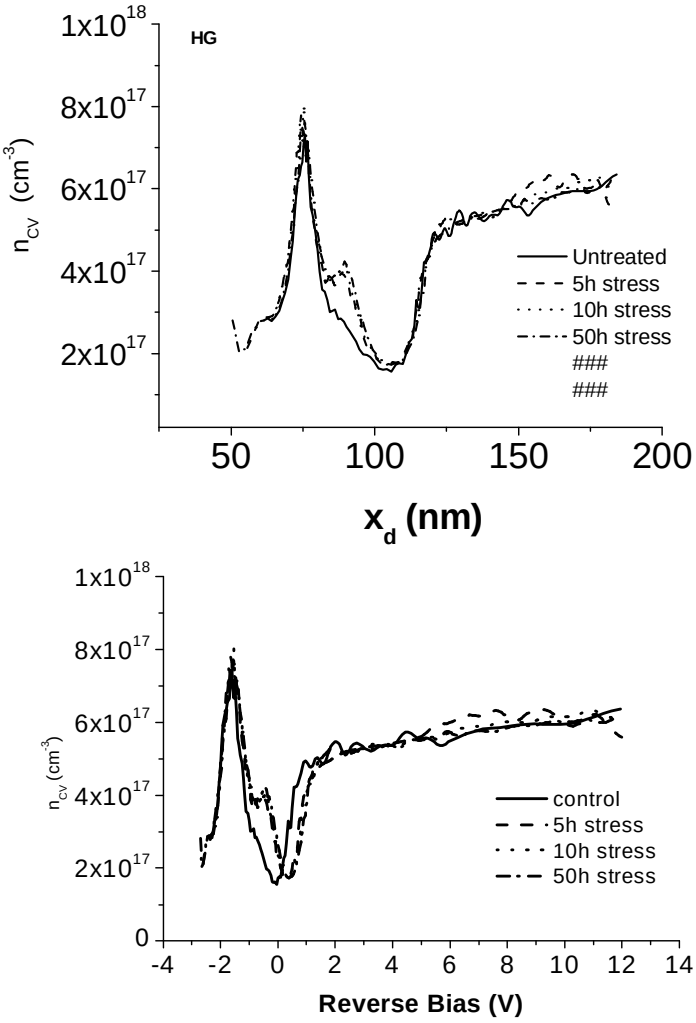


Fig. [6.7]. The profile of apparent charge distribution n_{CV} as a function of a) junction depth and b) reverse bias at different stress stages in sample HG.

6.3.3. Fit of the apparent charge profiles

In the following table 6.3 and figure [6.8] the results of the fit of the experimental n_{CV} profiles are reported. The fitting procedure was performed by means of freely downloadable Schrödinger-Poisson solvers like 1DPoisson [1DPoisson] or Band Engineering [BandEng]. The present results were obtained with Band Engineering. This latter software is designed for simulation of 1-dimensional nitride-based devices, and provides the possibility of calculating band diagrams, charge densities, and C-V characteristics. In case of heterostructures, there is the possibility of keeping into account the electric fields arising at the interfaces between different materials due to piezoelectric and spontaneous polarizations. The calculation of the charge density in the quantum wells can be performed either classically or solving the Schrödinger equation for the electron envelope function. The C-V characteristics can be calculated by approximating the p-n junction with a Schottky

junction with appropriately high potential barrier (roughly equal to the diode built-in potential). This procedure is questionable when calculating current densities and recombination dynamics, as the bipolar junction is approximated with a unipolar one, but it is acceptable in the case of a C-V characteristics of a reverse-biased junction, where the only important role is played by the electrons in the conduction band at the edge of the depletion region.

As is possible to deduce from the large amount of different parameters listed in the table, it would be nonsense to extract the exact structural parameters of the device from such a fitting procedure: it can be shown, for instance, that the same charge profile can be generated with a similar structure having only two quantum wells. However, the procedure can be useful in the interpretation of the trends which emerge during the stress treatment.

The parameters related to layer thickness and In fraction in the wells has been kept constant, and the only parameters to be varied were the doping levels in the layers. It turned out that the modifications occurring to the C-V characteristics of HG were well explained by the increase in the doping density in the right side of the QW region (Barrier 3, Well 3, Barrier 4, Well 4). This doping increase, which could be related to accumulation of shallow Si donors in the QW region as well as to the generation of donor-like deep levels, is consistent with other observations which will be described in the following. An equally good description could not be obtained by varying the In fraction or the QW thickness, which would correspond to a mechanism of In diffusion from the QWs to In-poorer regions.

Another interesting result of the C-V simulation is the need of postulating an active region less doped than the bulk region: this is common practice in LED technology, because it prevents the active region from being too close to the p-n junctions, which would yield unwanted carrier spill-over to the confinement regions [Schubert].

Table 6.3. Parameters used for the fit of the apparent charge profile of LED HG before and after the stress treatment.

Heterostructure region (n-side)	Thickness d (nm)	In fraction x	Doping before stress (cm^{-3})	Doping after stress (cm^{-3})
Barrier 1	70	0	5e17	5e17
Well 1	3	0.09	2e17	2e17
Barrier 2	5	0	2e17	2e17
Well 2	3	0.09	2e17	2e17
Barrier 3	5	0	**0**	**2e17**
Well 3	3	0.09	**0**	**2e17**
Barrier 4	5	0	**0**	**2e17**
Well 4	3	0.09	**0**	**2e17**
Substrate 1	10	0	0	0
Substrate 2	10	0	2e17	2e17
Substrate 3	140	0	5e17	5e17

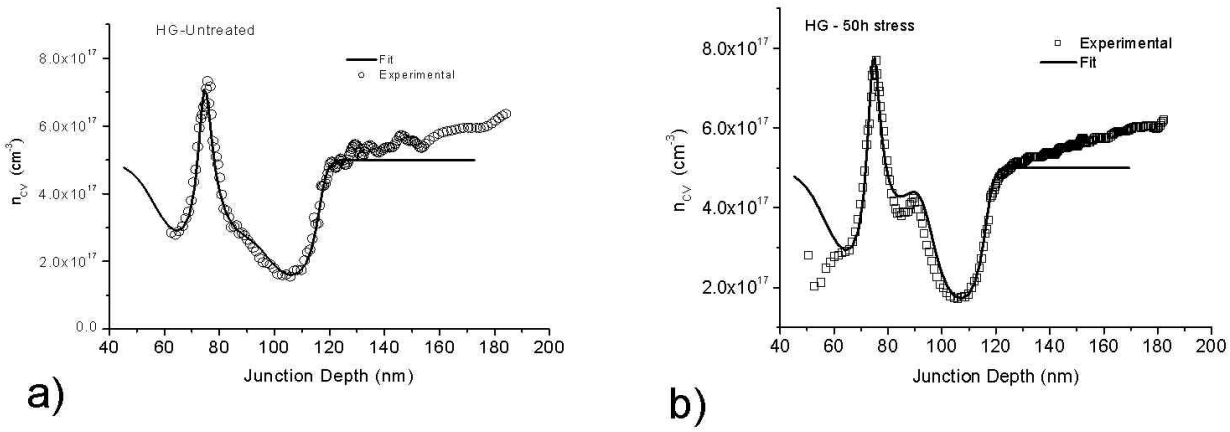


Fig. [6.8]. Fit of the depth profiles of apparent charge density in sample HG with the software Band Engineering; a) untreated diode; b) diode after 50 h stress at 20 mA.

6.4. DLTS characterization

Deep Level Transient Spectroscopy was employed in the study of degradation mechanisms of LEDs under DC current operation as a tool for determining whether certain deep levels play a role in this process. The capabilities of this technique in studying this problem were limited by some factors. First, DLTS has the intrinsic limitation that it can probe only the upper part of the bandgap in one-sided p^+-n junctions. Secondly, the investigated bandgap region is even narrower, as it was not possible to reach temperatures higher than 300 K, in order to avoid thermally activated degradation mechanisms. This prevented us from detecting levels with activation energy higher than 0.4 eV. Third, the device studied is a QW device with very large charge variations along the conduction band, which makes an accurate determination of concentration of deep levels impossible in the active device region. Thus, the concentration of deep levels can be calculated just when the polarization values used probe only a portion of the bulk region with nearly constant doping concentration. Another problem is connected, as already mentioned, with the variations of the apparent charge concentration during the stress stages. DLTS was always performed with a determined set of parameters (reverse bias-filling pulse, [V_{rev}, V_{fill}]) which were the same for all stress stages. In sample HP, where strong variations in the charge profile occurred, this means that the regions probed with the same [V_{rev}, V_{fill}] after successive stress stages were slightly different, possibly giving different information not because of transformation of the defect state, but because the response of the defect was different due to the different position along the junction.

It has been pointed out in chapter 4 that DLTS could be applied to the investigation of carrier emission from a QW. This would make it possible to measure the activation energy of the QW, which is closely related to the energy level of the confined carriers – an important parameter for the device. However, this measurement is not straightforward, being possible only under very particluar conditions. Schmalz et al., for instance [Schmalz96], argue that the studid sample could evidence QW emision provided it has an adequate structure (such as a mesa structure or island QW layers) preventing lateral carrier lekage during bias pulses. This is not the case in the present type of device. Nevertheless, DLTS proves to be a powerful technique even in this case. Heterostructure junctions are often used in the research, as they provide a source of misfit dislocations with controlled density [Chretien96], [Yastrubchak04]. The emission of carriers from dislocations can well be characterized by DLTS. In our case the heterostructure junction is due to the design of the device, which exploits the QW confinement for a more efficient light generation, and the dislocations found eare mostly threading dislocations, a common defect in nitride materials. The situation is favourable for the investigation of possible electron states related to this line defect. The

determination of the defect activation energy is still accurate, as well as the determination of the order of magnitude of the apparent capture cross section. The concentration of the main deep levels in the bulk region has been calculated as well. Moreover, the dependence of the features of DLTS spectra on the filling pulse amplitude does not change in a QW structure, so that one still has the chance of discriminating between point defects and extended defects.

6.4.1. LED HP

For the analysis of LED HP, which was stressed over only 4 stages (0 min, 15 min, 30 min, 60 min) a large amount of DLTS runs was performed at each stage. With constant filling pulse duration, the following polarizations [V_{rev}, V_{fill}] were used, in order to probe different regions of the device:

- [-1V,+1V] probing the whole QW region
- [-0.5V,+0.5V] [-1V, 0V], [-1.25V, -0.25V] probing the right side of the QW region
- [-2V, -1V], [-3V, -2V] probing the depleted region between bulk and QWs
- [-4V, -3V] probing the bulk region.

These polarizations were initially chosen from the analysis of the bias profile of n_{CV} in the untreated diode. Some of the results are shown in fig. [6.9].

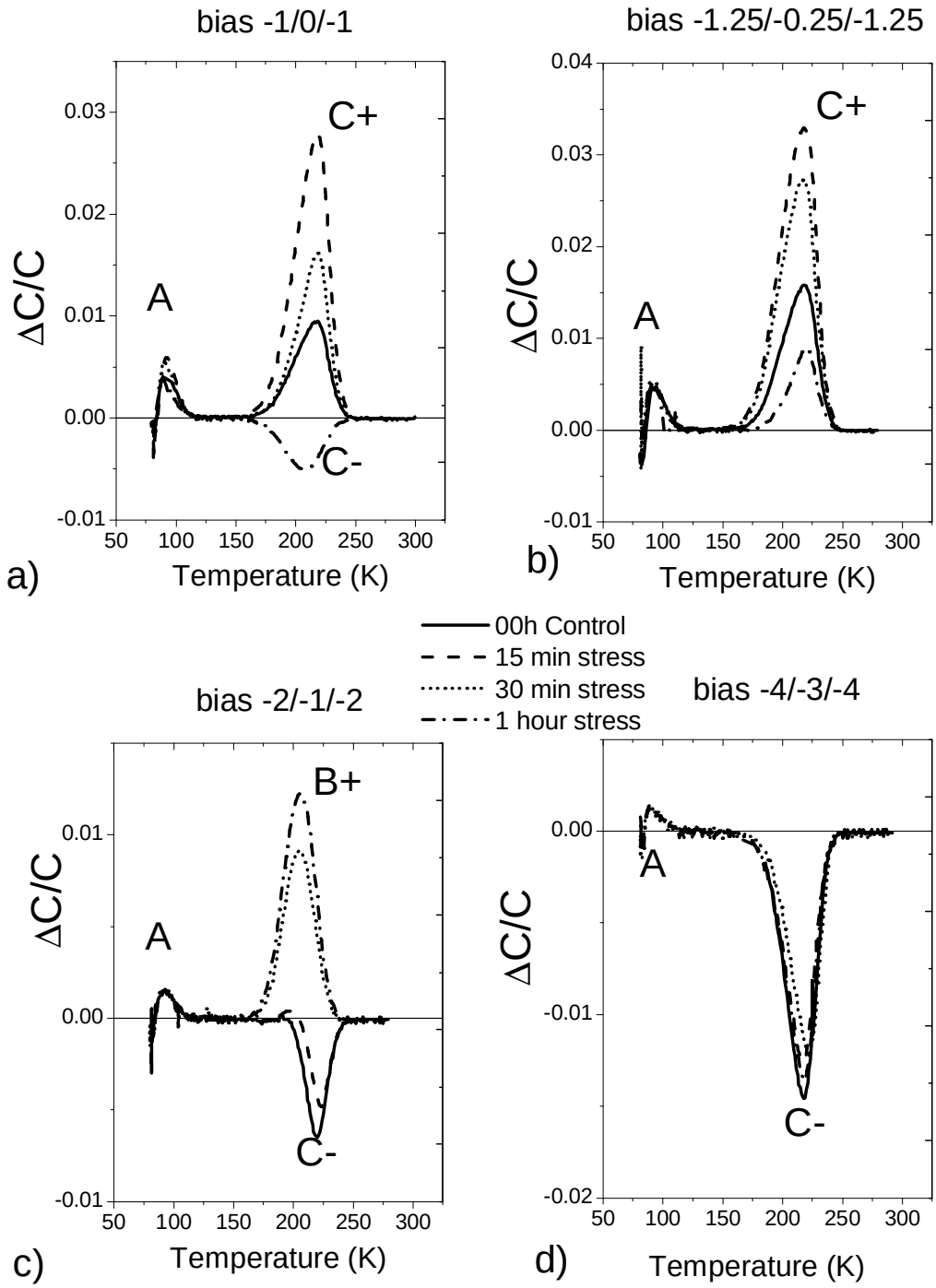


Fig. [6.9]. DLTS spectra of LED HP after progressive stress stages: a) V_{rev}=-1V, V_{fill}=0V; b) V_{rev}=-1.25V, V_{fill}=-0.25V; c) V_{rev}=-2V, V_{fill}=-1V; d) V_{rev}=-4V, V_{fill}=-3V. The filling pulse duration is t_p=1ms and the emission rate is e_n=46.5 s^{-1}.

The [-1V, +1V] polarization yielded very broad spectra, are not shown here, integrating all the information coming from the QW region. Figures [6.9] a) and b), reporting spectra from the right side of the QW region, are somewhat similar, as it is expected due to the relatively small difference in the polarization parameters. Three different peaks were detected in these conditions: a positive peak A at temperature lower than 100 K, a positive peak C+ at T=220 K and a negative peak C– at the same temperature. After Arrhenius analysis, peaks C+ and C- turned out to have the same deep

level parameters, as it is visible in fig. [6.10] and listed in table 6.4. In these runs, peak A is stationary over all stress stages, while peak C evolves: it is initially positive, and its amplitude decreases after the first stress stage, becoming even negative in a). After the second stress stage, C grows, and its amplitude gets larger than in the untreated sample. Another increase occurs after the third stress stage.

Figure [6.9] c) reports spectra from the depleted intermediate region. Peak A remains stationary, while the other feature at T =200-220K evolves. Initially there is the negative peak C–, which decreases in amplitude after the first stress step, then becomes the positive peak B+. The parameters of peak B are listed in table 6.4. Its activation energy E_c-E_t=0.35eV is very close to that of peak C–. Peak B+ also slightly grows after the third stress stage (60 min).

Finally, fig. [6.9] d) reports the spectra from the bulk region. In the bulk region the concentration of peak C– is rather stable during stress, and is calculated as $N_t \approx 9 \times 10^{16}$ cm^{-3}, which is significantly high if compared with the donor doping concentration, $N_d \approx 3.5 \times 10^{17}$ cm^{-3}.

The Arrhenius Plot and the table report another deep level, labelled AB-: this level was detected at a polarization not shown in figure, and is reported here for completeness.

The deep levels detected in this sample at different stress stages do not have an evolution in only one sense: they tend to evolve in one direction after 15 min, and in the other direction after the successive stress stages. This is not straightforward to explain, but it is consistent with the observations made in the C-V characterization of the diode, which showed an overall charge decrease after 15 min, and an increase later. This could be explained by two different mechanisms for the re-distribution of the charge in the device: a first one at very short stress times, and a second one at longer times. The relationship between C-V and DLTS and the interpretation of the defect nature of the detected deep levels will be addressed to in a later section.

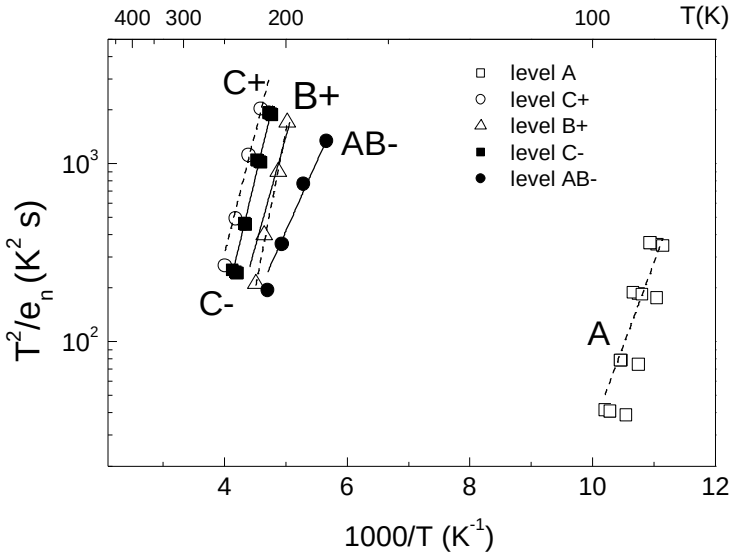


Fig. [6.10]. Arrhenius Plot reporting the signatures of the deep levels detected in LED HP. Full symbols refer to negative peaks (majority carriers), empty symbols to positive peaks.

Table 6.4 Parameters of the deep levels detected in LED HP

Deep Level	Enthalpy (eV)	σ_{inter} (cm^2)	N_t (bulk) (cm^{-3})
A	0.18 ± 0.03	3×10^{-14}	-
B+	0.35 ± 0.02	3×10^{-16}	-
AB–	0.15 ± 0.02	9×10^{-21}	-
C+	0.27 ± 0.08	6×10^{-19}	-
C–	0.29 ± 0.02	2×10^{-18}	9×10^{16}

6.4.2. LED HG

LED HG was stressed over 6 stages (0h, 2h, 5h, 10h, 20h, 50h). With constant filling pulse duration, the following polarizations [V_{rev}, V_{fill}] were used, in order to probe different regions of the device:

- [-0.5V, +0.5V], probing the right side of the QW region
- [-1V, 0V], probing the depleted region between bulk and QWs
- [-2V, -1V], probing the edge of the bulk region

The evolution of the spectra over time is illustrated in fig. [6.11].

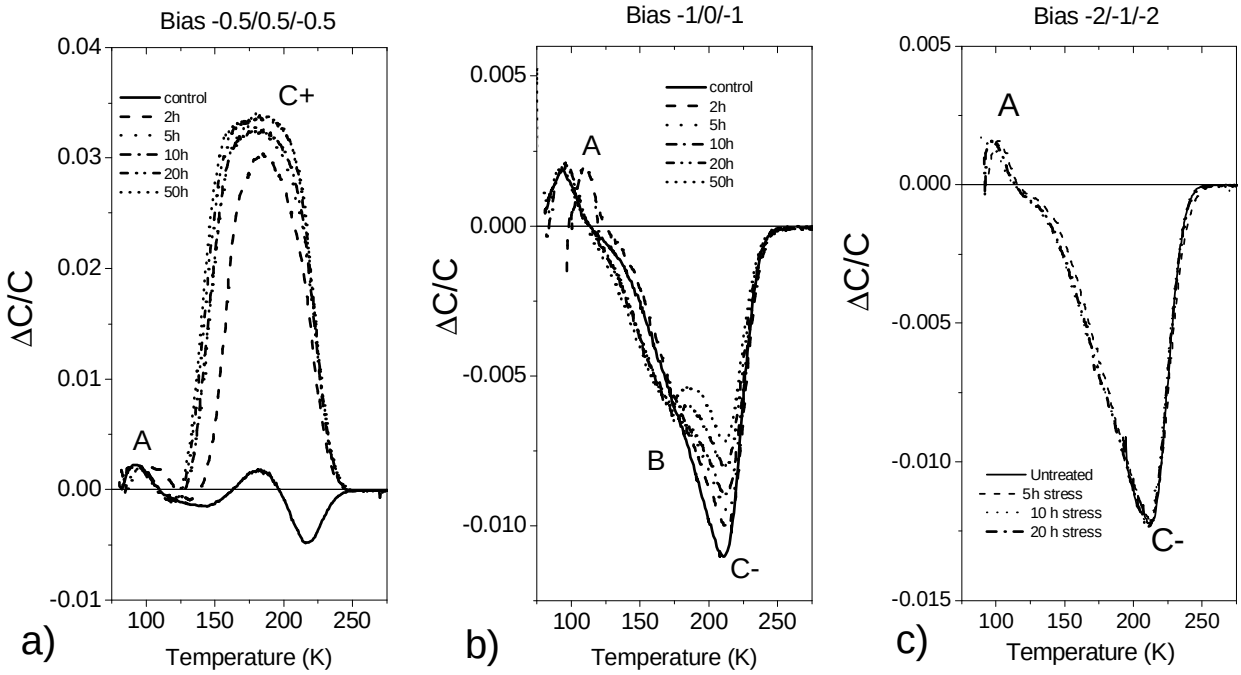


Fig. [6.11]. DLTS spectra of LED HG after progressive stress stages: a) V_{rev}=-0.5V, V_{fill}=+0.5V; b) V_{rev}=-1V, V_{fill}=0V; c) V_{rev}=-2V, V_{fill}=-1V. The filling pulse duration is t_p=1ms and the emission rate is e_n=46.5 s^{-1}.

Figure [6.11] a) reports the spectra collected with polarizations [-0.5V, +0.5V], which corresponds to probing the right side of the QW region. Two peaks are detected. Peak A is positive and does not show any evolution with stress. The feature at T=200K changes significantly with stress time. In the untreated sample, the DLTS signal appear to be the superposition of a positive and a negative signal, yielding a not well resolved spectrum. The spectrum becomes more definite after the first stress step, with a broadened positive peak, labelled C+. This peak increases in

amplitude and broadens over stress time. There is much difference between the spectrum of the untreated sample and those of the stressed sample: this difference is present also in the C-V characterization, and can be due more to the shift of the region explored with the same bias than to the evolution of the deep level itself.

In fig. [6.11] b) the spectra collected with polarization [-1V, 0V] are shown. In the unstressed sample, a positive A and a broad negative peak C– are detected, the parameters of which are the same as in peaks A and C+ shown in fig. [6.11] a) and are listed in table 6.5. After the first stress steps, the amplitude of peak C– decreases and a shoulder on its low-temperature side becomes visible. After 50 hours stress, the initial peak C– clearly splits in two, the aforementioned shoulder becoming the peak here labelled B (Table 6.5). Meanwhile, peak A does not undergo any modification. Peak C– undergoes a significant amplitude reduction, about 40% after 50 h stress, while the amplitude of peak B remains almost constant after the first stress steps, and then slightly decreases after 10 h stress. The amplitude reduction of peak B does not exceed 12% with respect to its initial value even if it is difficult to state how this change is connected to the modifications occurring to peak C–, whose low-temperature tail partially overlaps peak B.

Finally, fig. [6.11] c) reports the spectra collected at the edge of the bulk region. The positive peak A and the broad peak C– are present, but there is no evolution over stress time. An interesting correlation can here be established with the C-V characterization, which indicates that there is no change of the charge concentration in this region.

From the Arrhenius plot (Fig. [6.12]) it results that the activation energies of all these levels are in the range 170÷180 meV, whilst the capture cross sections significantly differ from each other. It is worth noting that the negative peaks B and C– are expectedly associated with majority carrier traps, while the positive sign of peak A does not mean that this peak is related to minority carrier traps. In fact, it was detected in a temperature interval where the high value of the series resistance can significantly affect the capacitance. The parameters of peaks C+ and C– are the same, and extreme care must be taken in the interpretation of this result. The positive peak C+ is detected with a positive bias pulse, which makes possible the emission from minority carrier traps. However, the example of sample HP shows that positive peaks are found also when the filling pulse is negative, and the emission from minority carrier traps is suppressed. Therefore, it is possible that in these diodes there is some mechanism which causes the inversion of the capacitance transient, and that peaks C+ and C– are related to the same deep level. This hypothesis will be analyzed more in detail in the next section, where the sign of the emission of peak C will be correlated with the C-V characteristics. The concentration of peak C– is calculated in the bulk region as $N_t = 10^{17}$ cm^{-3}, which is about one fifth of the donor doping concentration.

These DLTS results could give rise to three different interpretations: i) the decrease of peak C–, as well as the increase of peak C+ is ascribed to a modification of the concentration of the related deep level, possibly due to the dissolution of the relevant defect; ii) the decrease of peak C– and the increase of C+ is due to the stress-induced introduction/generation of other deep levels yielding a positive DLTS signal as from minority carrier traps; iii) the modifications of the shape of peak C are induced by the introduction/generationof other deep levels undetected by DLTS lying eihter higher than 0.15 eV or deeper then 0.4 eV in the gap, which modify the n_{CV} concentration profile.

Comparing the results of DLTS with the n_{CV} profile, we can establish a connection between the net increase in the apparent charge density in the bulk side of the active region and the amplitude changes of peaks C+/– . In fact, DLTS spectra evolve over time only in those device regions where the apparent charge profile evolves. Unfortunately, due to the device configuration (a heterostructure with Quantum Wells, having a strong non-uniformity of charge distribution) and to the polarizations used (aimed to exactly probe the bulk side of the active QW region), it is not possible to compare quantitatively the variations detected by C-V and DLTS. However, the dynamics appears to be quite similar for both ACD profile and trap C amplitude.

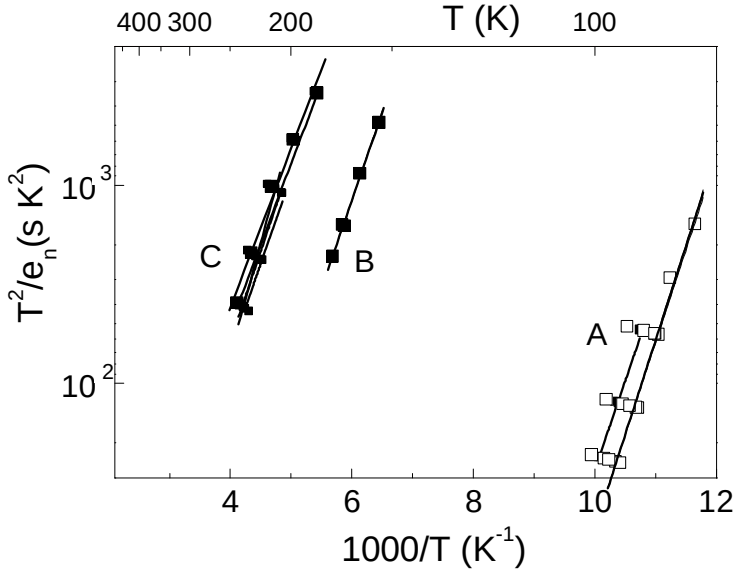


Fig. [6.12]. Arrhenius Plot reporting the signatures of the deep levels detected in LED HG. Full symbols refer to negative peaks (majority carriers), empty symbols to positive peaks.

Table 6.5 Parameters of the deep levels found by DLTS in the active region of LED HG

Deep Level	Enthalpy (meV)	σ_{inter} (cm^2)	N_t (bulk) (cm^{-3})
A	180	9×10^{-14}	-
B	180	5×10^{-19}	-
C	170	10^{-20}	1.0×10^{17}

6.4.3. Correlation between DLTS and CV

As already mentioned in the previous sections, strong correlations emerge between the observations made by DLTS and by C-V. First, we noticed that the DLTS spectra change with stress time only in the regions where also the apparent charge profile changes. Secondly, a connection between the amplitude of peak C, assuming that peaks C+ and C- are related to the same deep level, and the apparent charge profile is evident. This connection has been established as follows in sample HP, and similar correlations could be observed in other samples. During the DLTS runs, the sample was stabilized at the temperature T=217 K where peak C has its maximum for the emission rate e_n=46.5s^{-1}; the reverse bias V_{rev} was varied, keeping the filling pulse $V_{fill}=V_{rev}$+1V. For each V_{rev} value, the DLTS signal was recorded. In this way, for each stress stage, a bias profile of the amplitude of peak C was obtained. This was compared with the bias profile of the apparent charge density n_{CV}, and the results are shown in fig. [6.13]. It was observed that:

- In the bulk region, the DLTS signal from C is negative and roughly constant
- In the depleted region between bulk and QWs, the DLTS signal from C becomes rapidly positive, reaching a maximum in correspondence of the minimum of the apparent charge n_{CV}.
- Approaching the QW region, the DLTS signal from C becomes negative again, reaching a minimum in correspondence of the rightmost QW peak.
- The amplitude profile of peak C follows the n_{CV} profile in its evolution at successive stress stages.

These observations support the hypothesis explaining the occurrence of a positive DLTS peak with negative filling pulse as an inversion of the capacitance transient due to factors which are still unknown, but which might be related to the complexity of the heterostructure diode, and/or to the presence of an electric field directed towards the junction in the QW region, arising because of strain-induced polarization. Further investigation is needed to clarify this phenomenon.

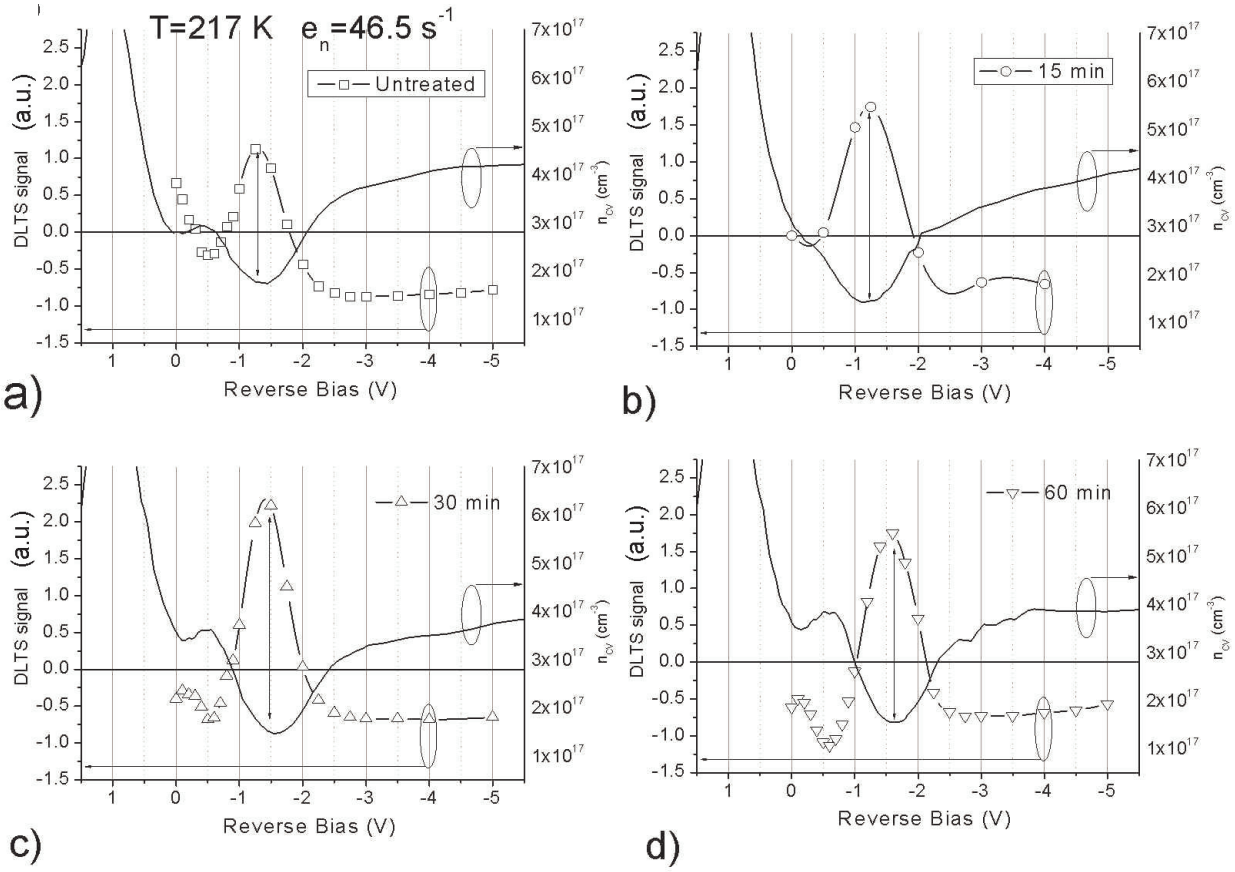


Fig. [6.13]. Study of the correlation between the DLTS amplitude of peak C and the apparent charge profile in LED HP at different stress steps: a) untreated, b) 15 min, c) 30 min, d) 60 min. The emission rate is e_n=46.5 s^{-1} and the sample was stabilized at T=217 K.

6.4.4. Deep Level A

Another positive DLTS peak is that related to deep level A. This peak is found in all samples, with an activation energy equal to 0.18 eV, and it is positive for all polarizations used, even when probing the bulk region. The positivity of this peak is an artefact, and it is in no way related to emission from minority carrier traps. The signal is positive, which is well explained if one considers the temperature value at which the peak appears. In GaN, charge carrier freeze-out begins at relatively high temperatures, around 100 K. When this phenomenon occurs, the series resistance of the diode increases dramatically, thus strongly affecting the measurement of the capacitance: the measured capacitance becomes a decreasing function of the real capacitance. When measuring a capacitance transient, this is equivalent with transient inversion. So, the real capacitance transient associated with A is negative. This is still not enough to conclude that peak A is related with a shallow trap, as the transient, occurring during the freeze-out, could also be related to emission from shallow donor levels. Finally and most importantly, it must be excluded that this peak is related to QW emission, because it is also found when the polaization is such that only a portion of the bulk region is probed.

6.4.5. Deep Level C: potential barrier and density of states

Level C is the broad peak appearing in sample HG with an activation energy E_c-E_t=0.18 eV and in sample HP with an activation energy E_c-E_t=0.30 eV with an emission rate e_n=46.5 s^{-1} at T≈220 K. Despite the significant difference in activation energy, the peak exhibits similar behaviour in both samples. The inversion of the peak when approaching the QW region from the bulk is found in both samples. Moreover, level C shows in both samples a peculiar capture kinetics, which is typical for a deep level associated with extended defects. This observations are consistent with previously reported analyses of emission by dislocation-related deep levels in this temperature interval [Soh04], [Yastrubchak04]. An in-depth analysis on the properties of level C was then carried out. The results presented here refer to sample HP. Figure [6.14] reports the absolute value of the amplitude of DLTS peak C as a function of the filling pulse for different biases and stress stages. As it is visible from the figure, in all sets of data the amplitude behaves like a linear function of the logarithm of the filling pulse for at least 3-4 decades, though with different slopes. For t_p>0.01s, the amplitude starts to saturate.

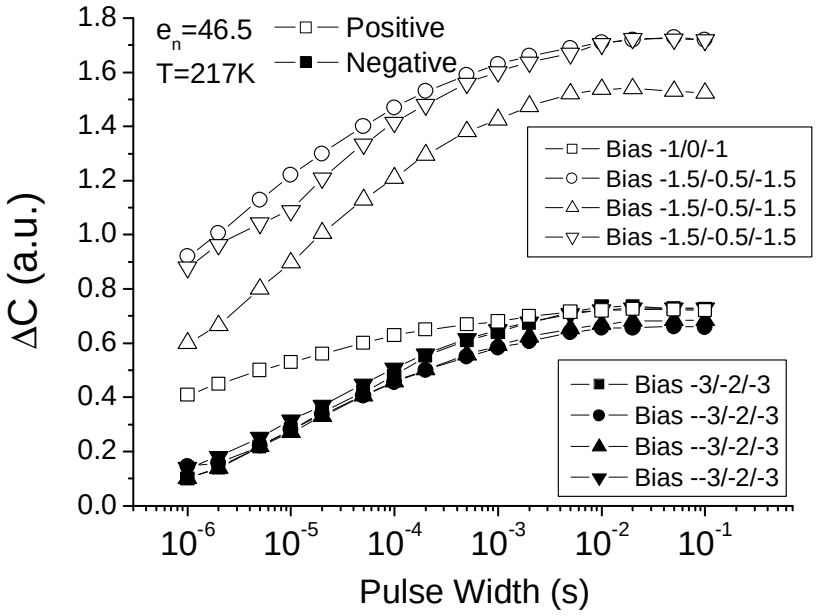


Fig. [6.14]. Dependence of the amplitude of DLTS peak C on the filling pulse width, at the polarizations [V_{rev}, V_{fill}] reported in the legend. Full symbols refer to polarizations at which the C peak was negative, empty symbols to polarizations at which the C peak was positive. Different symbol shapes refer to different stress stages.

From the above curves it was possible to calculate the barrier height, as described in section (4.3.5) by means of the following formula:

$$e\Phi(t_p) \approx kT\left[\ln(n v_{th}\sigma_n t_p) + \ln\left(\frac{n_T}{t_p(\partial n_T/\partial t_p)}\right) - \ln\frac{n_T}{N_T}\right]. \qquad (4.64)$$

The filled trap concentration n_T was calculated from the $\Delta C/C$ signal, and the total concentration of the deep level N_T was calculated from the DLTS signal obtained for t_p=1s, although this can be an approximated quantity because in the case of defects associated with potential barrier saturation can occur before all defect states are filled. The capture cross section of the defect σ_n was calculated from the Arrhenius Plot: this is an approximation too, because this capture cross section incorporates an entropic pre-factor (section (4.1.5)), but it represents only a

minor uncertainty contribution to the total calculation. The electron concentration in the conduction band n was calculated as the average n_{CV} concentration in the bias range [V_{rev}, V_{fill}] used. The results for the potential barrier associated to deep level C are reported in fig. [6.15], for several biases and stress stages.

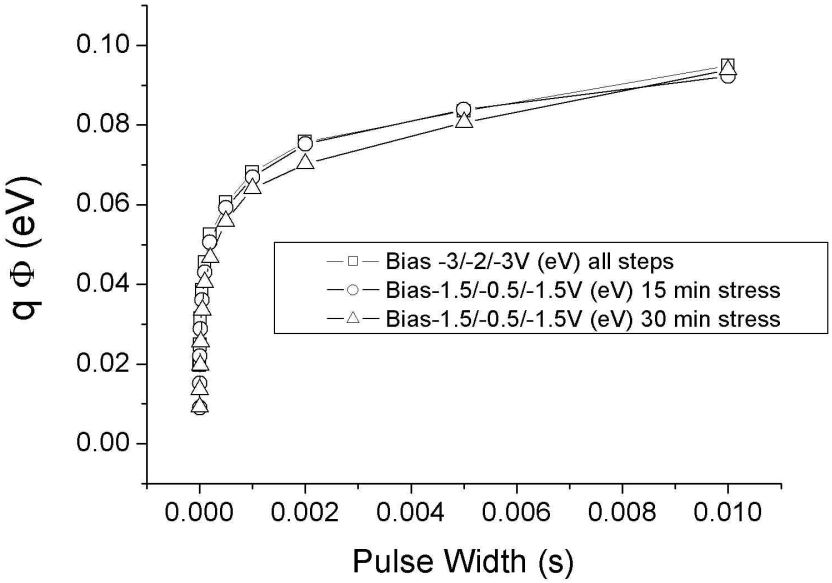


Fig. [6.15]. Barrier height of deep level C as a function of the filling pulse width, for different biases and stress stages.

The result show that the barrier height is approximately the same for different polarizations and stress stages, supporting the hypothesis that level C is the same in the different regions probed by DLTS, that its capture kinetics does not change with the sign of the peak, and that the defect does not change with stress time. The capture kinetics of the defect is different in different device regions because the free carrier concentration, and thus the doping concentration, is not uniform. The maximum barrier height is lower than 0.1 eV.

Once assessed that the deep level has a potential barrier associated, the next step is the analysis of the shape of the DLTS peaks at increasing filling pulse widths, in order to distinguish between localized and bandlike states. For bandlike states, a shift of the peak maximum towards lower temperature is expected with increasing pulse width, while for localized states, the peak should remain at the same temperature. The results shown in fig. [6.16] indicate that the deep level C has a localized nature.

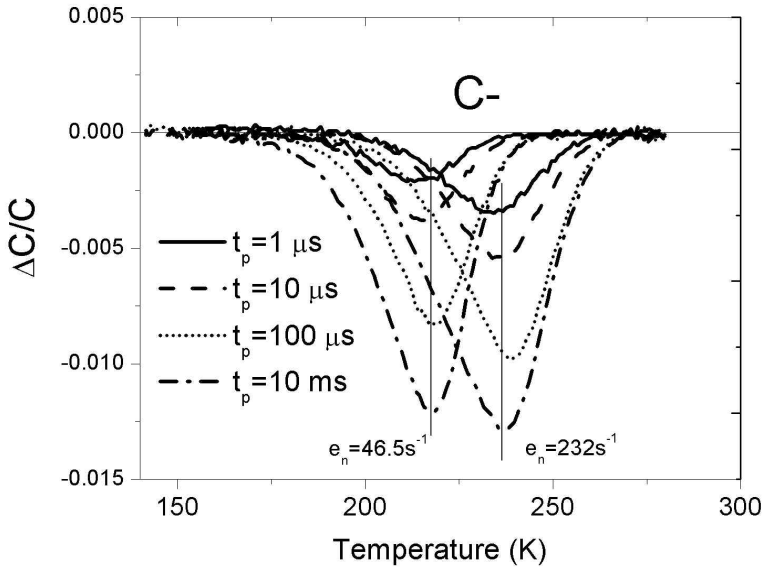


Fig. [6.16]. DLTS spectra of peak C recorded at different pulse widths in sample HP. The parameters used are V_{rev}=-3V and V_{fill}=-2V, corresponding to the bulk region of the device.

After assessing the localized nature of deep level C the DLTS peak was fitted with several models. The procedure was as follows. From Arrhenius plot, we calculated the enthalpy and the capture cross section. Then, we postulated a density of states $N_t(E)$ for the defect equal to:

$$N_t(E)=A\,(E_t^{fit}-E)^{-\alpha}, \qquad (6.3)$$

where the exponent α was kept constant in order to keep the number of fit parameters as low as possible in the Levenberg-Marquardt fitting procedure. The only fitting parameters that were allowed to vary were E_t^{fit} and the proportionality constant A. This particular form of the density of states is used after the example of Hedemann [Hedemann95], who used a similar density of states with α=1/2 for fitting DLTS emission from dislocation states in Si. In his case the choice of the exponent was justified by the argument that in 1-D systems such as states on dislocation lines, the density of states should be similar to that of 1-D free electrons. Here we have tried a similar expression, because it is likely that our defect is related to dislocations, but we used different values for α (namely, 0, 0.5, 1, 1.5, 2) for the achievement of the best fit. The DLTS signal was then calculated as

$$\frac{\Delta C}{C}=f(\lambda)\int_0^{+\infty}N_t(E)\left(\exp\left[-e_n(E,T)t_1\right]-\exp\left[-e_n(E,T)t_2\right]\right)dE \qquad (6.4)$$

Where the constants t_1 and t_2 are the sampling times defining the rate window of the measurement system, $f(\lambda)$ is a scaling factor depending on the bias values used, and $e_n(E,T)$ has the following form :

$$e_n(T,E)=\gamma\sigma T^2\exp[-(E-E_c)/kT], \qquad (6.5)$$

for which the capture cross section was assumed as constant and set equal to that measured from the Arrhenius plot. In fig. [6.17] the normalized density of states yielding the best fit of the DLTS signal is plotted. The best fit was achieved by setting the exponent α equal to 2. For this value of α, the E_t^{fit} value was $E_t^{fit}=E_c$-0.315 eV.

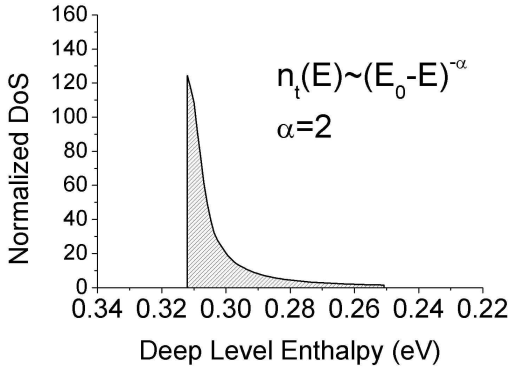


Fig. [6.17]. The normalized density of states yielding the best fit for the DLTS signal of deep level C.

Figure [6.18] reports four examples of fitting of C level, with four different emission rates (i.e., with four different $[t_1,t_2]$ pairs). The emission rates vary from 232.5 s^{-1} to 23.25 s^{-1}. The experimental data are compared with the fit using the density of states with exponent α=2 and with the fit using the model of the single level. It is interesting to notice that in all four cases the best fit parameter E_t^{fit} assumed the same value, which is a strong hint about the goodness of the fitting procedure.

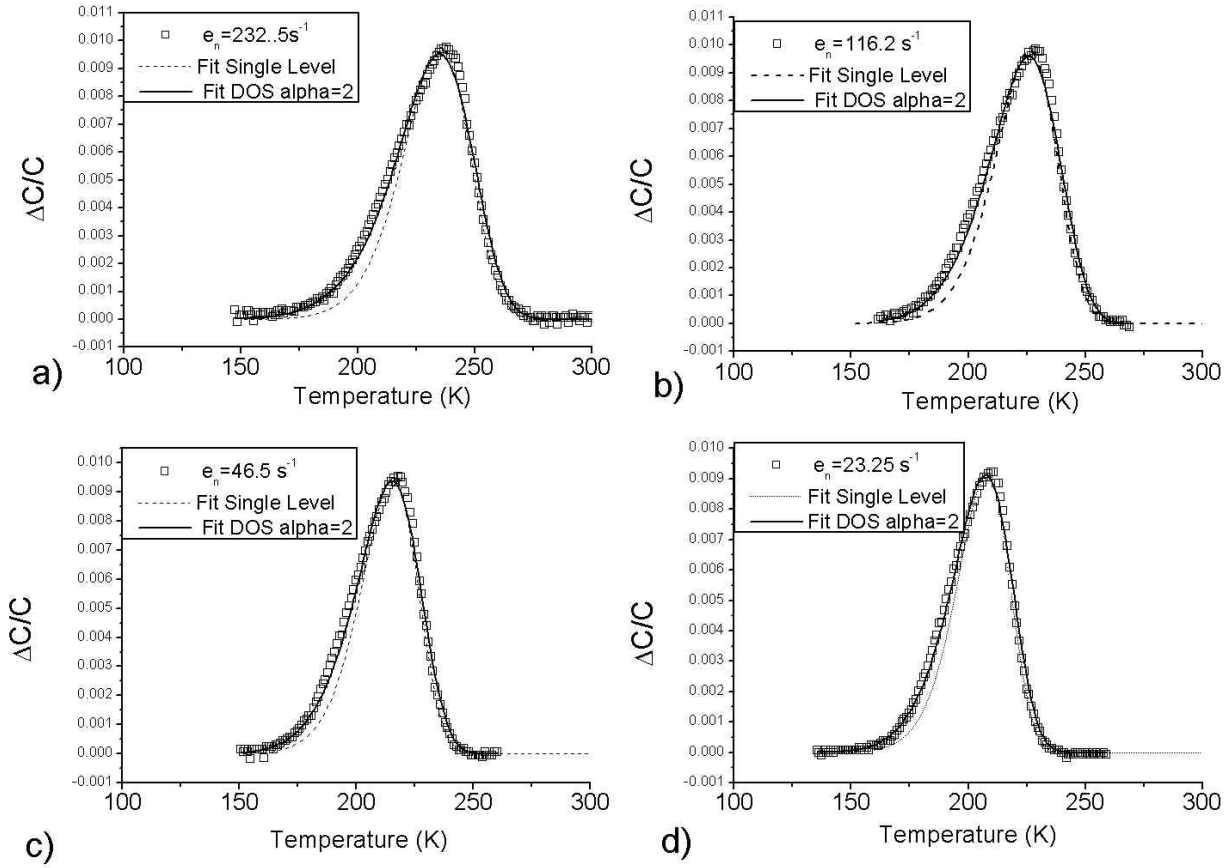


Fig. [6.18]. Best fits of the experimental DLTS spectra at four different emission rates: a) e_n=232.5 s^{-1}, b) e_n=116.2 s^{-1}, c) e_n=46.5 s^{-1}, d) e_n=23.25 s^{-1}. Squares correspond to experimental data, dotted lines to the fit from the single level model, solid lines to the fit with the density of states (eq. 6.3) with the exponent α=2.

The set of observations made on this peak can be summarized as follows:

- The amplitude of peak C changes with stress time only in the regions where also the C-V characteristics change.
- With relatively low filling pulse amplitude (1V) the sign of peak C changes depending on the reverse bias, and its amplitude is strongly correlated with the n_{CV} profile.
- The filling kinetics of deep level C has a logarithmic dependence on the filling pulse width, and it has a potential barrier of about 0.1 eV.
- The shape of DLTS peak C is best fitted with a continuous density of states. The temperature position of the peak does not vary with the filling pulse (only the amplitude does).

The first two observations are linked with the evolution of peak C and with the DLTS analysis performed on a portion of semiconducting material which has a strongly non-uniform conduction band charge density. In sample HP it was impossible to determine whether the changes of C were intrinsic changes or apparent changes, due to the rearrangement of the charge concentration in the conduction band or of the dopant atoms. In sample HG, however, the variations affecting peak C were significant even when the charge concentration from the C-V profile was almost constant, thus suggesting that this deep level has a participation in the degradation process. The third and fourth items of the conclusions give us information about the microscopic nature of the defect related to deep level C. The information is obviously incomplete, as it was not possible to determine, e.g., whether the level was related to an impurity rather than to an intrinsic defect, but they represent a strong hint about a role played by extended defects in the electronic activity of the level. The barrier height is relatively low, about 0.1 eV, and the width of the distribution of electron states used for the fit of the DLTS shape is narrow: this suggests that we are dealing with a distribution of point defects in the neighbourhood of dislocation lines. These defects may be preferably positioned near a dislocation, acquiring slightly different activation energies. This explains the formation of a potential barrier upon charging up, and also their nature of localized levels.

6.4.6. Summary of DLTS characterization

DLTS characterization of the anlyzed LEDs led to some interesting results, despite the limitations that affected this technique applied to heterostructure junctions. The main feature of DLTS spectra collected with temperature varying from 80 K to 400 K is the peak labelled C, which is found in both samples, though with slightly different enthalpy. Peak C was found to vary in amplitude in the device regions where C-V characterizatin evidenced a variation of apparent charge concentration; moreover, the sign of the peak is strongly correlated with the n_{CV} profile, because C, otherwise a negative peak (i.e. related to a level emitting majority carriers), becomes positive when the DLTS reverse bias probes the region depleted of free carriers between the QW region and the bulk. The mechanism governing this anomalous behaviour has not been clarified yet, but it is possibly connected with the complex structure of the device and/or with the piezoelectric fields which are expected to build up in the active region of the device.

The filling kinetics and the shape of this peak have been interpreted according to the theory of emission from extended defects, showing that C is related with localized states, possibly point defects gathering in the neighbourhood of threading dislocation lines. The presence of a significant amount of electron states related to dislocations was already evidenced in the I-V characterization, with a soft-breakdown reverse characteristics. However, the evolution of level C seems to be an effect, rather than the reason, of the degradation process.

6.5. Optical spectroscopic characterization

6.5.1. Electroluminescence characterization: device aging

Spectroscopic measurements of emission spectrum and emission power of the analyzed LEDs were carried out both at the Physics Department of the University of Bologna and at the Department of Electrical and Information Engineering of the University of Padova. In this latter laboratory, dedicated tools for the characterization of emission power of the LEDs were employed, and the results illustrated in this section refer to those measurements, carried out by the group of Padova using an Emission Microscope (PHEMOS 1000). Changes in emission intensity as a function of time during the 20 mA aging test, measured at two different current levels (1 mA and 20 mA), are shown in Fig. [6.19]. The intensity loss with respect to the unstressed value strongly depends on the drive current. A similar behavior is observed at all the investigated temperatures. The low-current regime is clearly more sensitive to the aging effects: e.g. at RT the output optical power reduction is limited to 7÷10% for drive currents of 10÷20 mA whereas it rises to about 20% for I=1 mA. This is consistent with the attribution of the intensity loss to an aging-induced generation of non-radiative defects, creating a recombination channel which is saturated and becomes therefore less and less effective when the drive current increases [Pursiainen01]. The decrease took place in the initial 50 hours. In the successive 100 hours a slight recovery of the emission curves was observed.
The emission intensity vs. forward current curves (L-I curves) can be divided in three zones [Cao03]: for current I< 2mA the L-I characteristic is non linear, as a result of the non-radiative component. For intermediate current values (2 < I < 10 mA) the curve is almost linear, as a result of the radiative recombination dominating in the active layer. For high current levels (I> 10 mA) a sub-linear zone can be identified, which could be due to heating effects that worsen the quantum well efficiency.

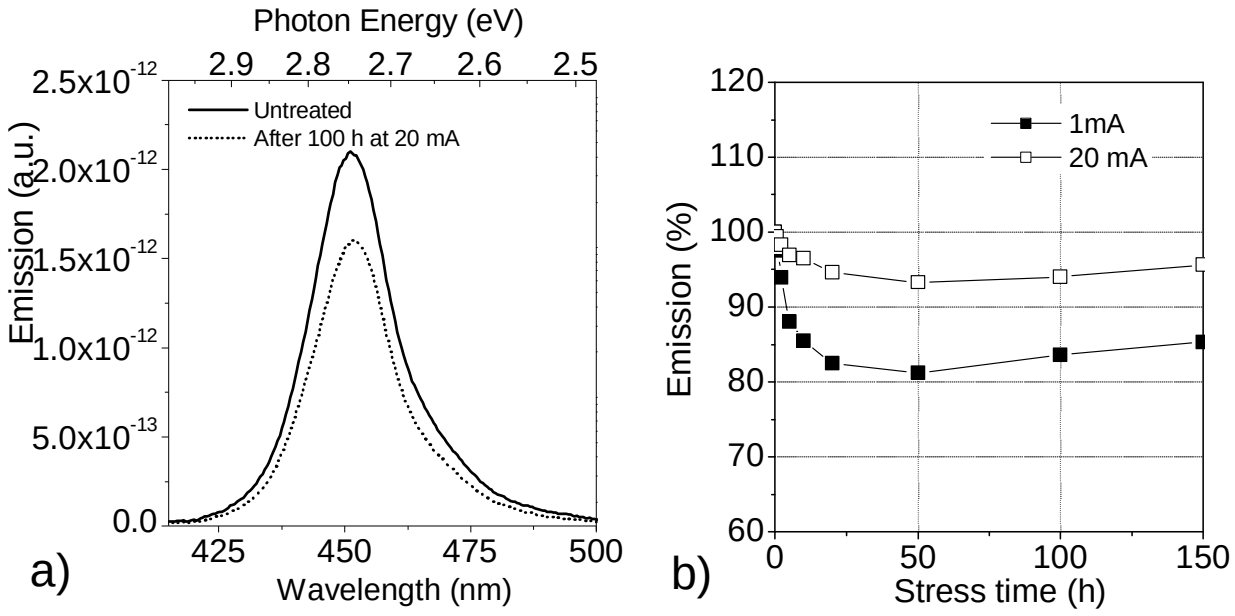


Fig. [6.19]. Optical emission characterization of DC degradation: a) emission spectra of the untreated (solid line) and of the 100 h stressed sample (dashed line); b) relative intensity of emission power as a function of stress time for 1 mA (full squares) and 20 mA (empty squares) drive current.

6.5.2. Photocurrent characterization: device aging

As photocurrent (PC) spectroscopy measurements are, in our setup, very sensitive to even minimal shifts of the sample from its original position, we analyzed the sample (labelled as HJ) in

this way: the sample was put in the sample holder, the PC measurements performed at different applied biases, then the sample was stressed at 20 mA DC without taking it off. After that, the new set of measurements was performed, and so on. The stress steps were the following: 15 min, 30 min, 60 min, 2 h, 5 h, 10 h, 25 h, 50 h, 100 h. This ensured a constant alignment of the sample with the spectroscopic system, and that the observed changes in the PC intensity were actual changes due to internal modifications in the sample, and not to other factors. Other samples have been analyzed only before and after the whole stress treatment: HG and HP, for instance, showed the same spectral features of HJ, as is expected from samples from the same wafer. PC spectroscopy has been performed in the spectral interval from 330 nm to 600 nm. In fig. [6.20] a) the spectra of an LED at different stress stages are reported; the sample was here left unbiased. Only some of the spectra are shown, for clarity. The spectra show a broad band for wavelengths lower than 450 nm. This band corresponds to the absorption by the multi-quantum-well structure, in the interval from 450 to 370 nm, and by the GaN layers in the space charge region, corresponding to the peak at about 360 nm. Moreover, the broad band has a complex structure with more quasi-periodically spaced peaks. This quasi-periodic series can be shown to derive from interference effects due to the multi-layer structure of the device: at certain wavelengths, the optical field interferes constructively in the space charge region, thus yielding a higher PC signal. On the contrary, when the optical field interferes destructively, local PC minima are observed. More information about the absorption mechanism will be extracted from the normalized spectra, which have been fitted with the procedure illustrated in chapter 5.

The shape of the spectra keeps similar at all stress stages, but there is a slight decrease of the signal with increasing stress time. This suggests that the mechanism of absorption of light is not affected by stress, but that the photo-generated carrier are less efficiently collected. This is possibly due to a decrease of the mobility related to the generation of deep levels or to the migration of doping impurities, which can be responsible for the modifications of the n_{CV} profile observed by C-V characterization and of the DLTS signal from peak C. The relative variation of photocurrent at wavelength λ=417 nm is shown in fig. [6.20] b), and it shows that the major variation takes place during the first stress stages. The behaviour of PC is thus similar to that of the EL, but in the device analyzed by EL a slight recovery was observed after 50 h, which was not observed here. This possibly depends on specific features of the individual sample.

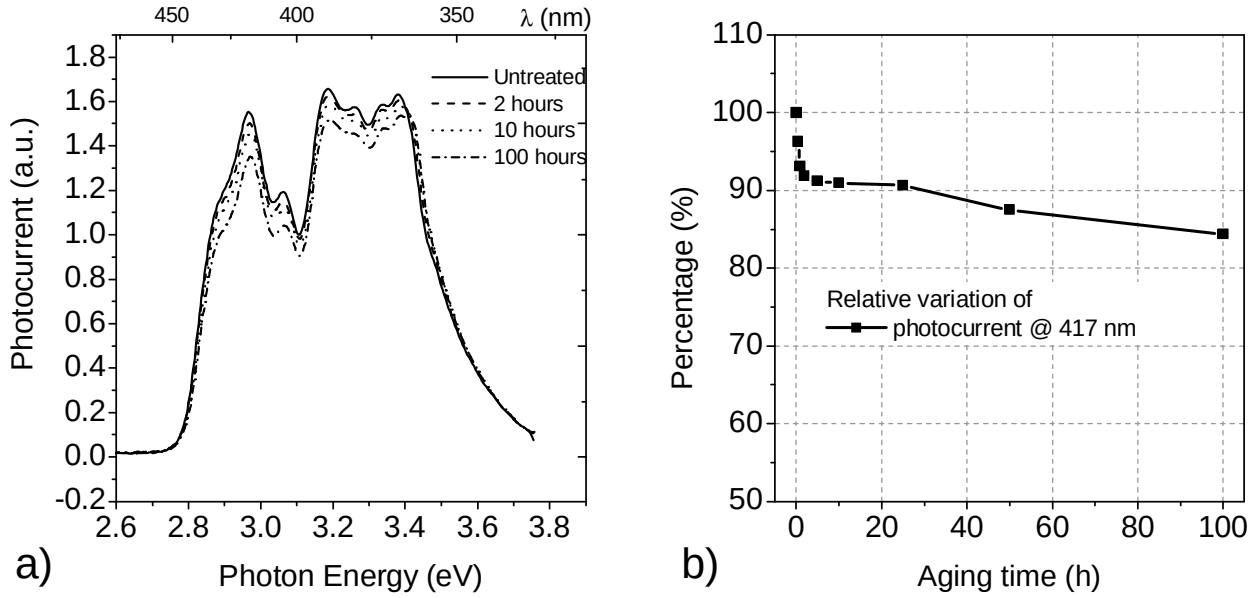


Fig. [6.20] Evolution of photocurrent with stress time at 20 mA: a) photocurrent spectra of the analyzed LED at successive stress stages; b) relative variation of the photocurrent peak at 417 nm. The sample is here left unbiased.

6.5.3. Fitting of responsivity spectra: device structure and absorption mechanisms

Employing the model illustrated in section 5.3 we will be able to explain many details of the photocurrent spectra of the analyzed LEDs. A cautionary note is here needed. The model is based on two main categories of parameters: there are physical parameters, such as those describing the dispersion relations for the various materials, and there are structural parameters, which pertain the particular structure under investigation. Moreover, the boundary between these two categories is not sharp, as structural parameters, such as the width and composition of a quantum well, influence the dispersion relations in the quantum well.

Preliminarily to the comparison of the experimental and simulated spectra, it is convenient to normalize the photocurrent spectrum by the incident light flux, thus obtaining a responsivity spectrum. We will refer from now on to responsivity spectra.

The main issue in the analysis of the responsivity spectra by means of this model, as it was for the fitting of the C-V characteristics, is the insufficient degree of knowledge of the structural parameters of the device. Some of these parameters could be assessed from the results of C-V characterization, but some others (e.g. the width and composition of the QWs, as well as the thickness of the GaN cap and bulk layers) remained unknown. This is not an ideal situation for attempting a quantitative simulation of the photocurrent spectrum as a fit of the experimental data, because the number of unknown fitting parameters is high, affecting thus the validity of the fitting procedure.

However, the model can be applied to the experimental data, in order to extract some of the structural parameters with a certain degree of likeliness. This is the way one can proceed:

- The dispersion relation for the absorption coefficient of GaN was modelled as proportional to the bulk joint density of states of the material, which is close to some dispersion relations found in the literature [Dingle71][Ambacher96]. This dispersion relation was convolved with a gaussian distribution, in order to account for band tailing effects. According to the previous references and to measurements of photocurrent spectra in GaN epitaxial layers [Polenta04], the band tailing Δ for bulk GaN was estimated to be around 0.1 eV. No exciton effects were considered, as these are not observable at room temperature in bulk GaN [Binet96].
- The dispersion relation for the QWs was proportional to the joint density of states for a QW, thus having a typical step-like behaviour. The steps were described through a Fermi distribution, which accounts for both thermal broadening and band-tailing effects. The effect of excitons was here considered, as these are more effective in QWs than in bulk, and they are often visible even at RT in absorption spectra. However, the best fits were obtained by neglecting the exciton absorption peak. The effective energy gap of the QW system was then determined as a fitting parameter.
- Once established the dispersion relations, the structural parameters were chosen partly from the C-V characterization, partly from the most common design parameters for high internal efficiency LEDs [Schubert]. These have typically a MQW system as an active region (here we chose the number of 4 QWs), consistently with our C-V characterization. On the p-side, there is an AlGaN electron blocking layer preceded by a GaN cap layer. Typical thicknesses of these layers are of the order of 100-50 nm and 200-100 nm, respectively. The thickness of the GaN bulk layer on the n-side remains unknown, but its approximated value can be determined from the period of the interference fringes on the photocurrent spectrum generated by absorption in the QWs.

In fig. [6.21] the results of a simulation of photocurrent spectrum are showed and compared with the experimental data normalized by the spectral light intensity of the system *source + monochromator*, the absolute value of the simulated photocurrent being normalized to the photocurrent peak related to bulk GaN absorption. The contributions from the different layers to the photocurrent are also shown. Most of the photocurrent signal comes from the bulk GaN and from

the QWs. The cap layer does not contribute for two reasons: first, the depletion region does not extend to the GaN cap layer because of higher doping on the p-side; secondly, diffusion of electrons (minority carriers) from this layer to the active region is prevented by the interposed AlGaN blocking layer (which during LED operation blocks the transport of electrons in the opposite sense). Minor contributions come from the GaN barrier layers in the active region.

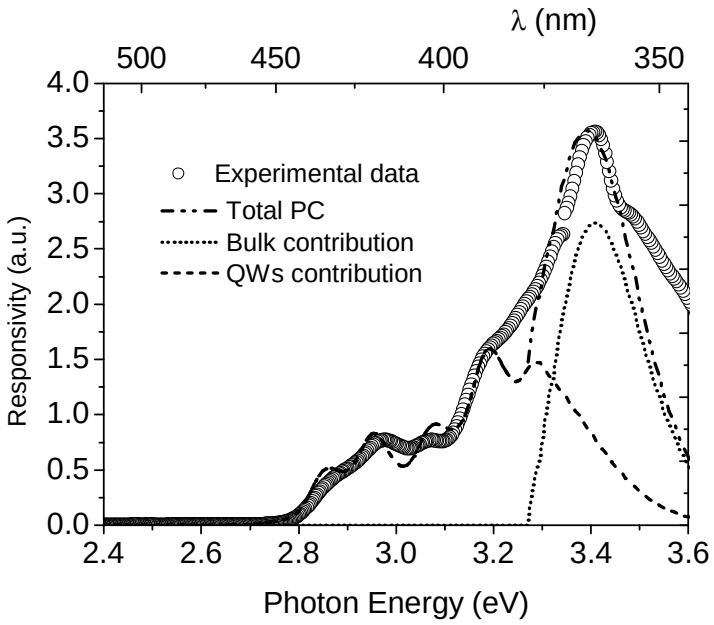


Fig. [6.21]. Comparison between experimental responsivity spectrum and simulated spectrum. The contributions of different layers are distinguished by different line styles.

The structural parameters used for this simulation are listed in table 6.6. They are fairly consistent with those used for the simulation of the C-V characteristics. The x fraction for InGaN QW layers is not given here, because it is calculated from the fitting value of E_g of the dispersion relation for the absorption in the QW. AlGaN is considered as a non-absorbing material in the analyzed energy range. There is no warranty that the given parameters are the actual ones, but they give us confidence that the procedure is based on sound hypotheses.

Table 6.6 Structural parameters for the simulation of the responsivity spectrum shown in fig. [6.21].

Layer	Material	Description	Thickness (nm)
1	GaN	p^+-type cap	200
2	AlGaN	*e*-blocking layer	70
3	InGaN	QW1	5
4	GaN	barrier 1	5
5	InGaN	QW2	5
6	GaN	barrier 2	5
7	InGaN	QW3	5
8	GaN	barrier 3	5
9	InGaN	QW4	5
10	GaN	n-type bulk	2200
11	SiC	n^+-type substrate	2000
Depletion region edges:		x_p=200 nm	x_n=345 nm
Hole diffusion length:		L_p=70 nm	

The dispersion relations are plotted in fig. [6.22], and the parameters used (as described in eq. 5.13) are listed in table 6.7. It is interesting to notice that the best fit of the responsivity spectrum is obtained by neglecting any exciton effect. Due to rather high binding energy, exciton absorption peaks are in many cases visible in nitride-based QW systems even at room temperature [Bulutay99][Friel04]. The reason why this does not happen in our LEDs could be due to different factors, such as a high piezoelectric field in a sufficiently wide QW, which would prevent the formation of excitons and thus their participation to the absorption process. It would be interesting to perform photocurrent measurements at lower temperatures, as this would clarify whether and under which conditions excitons are observable.

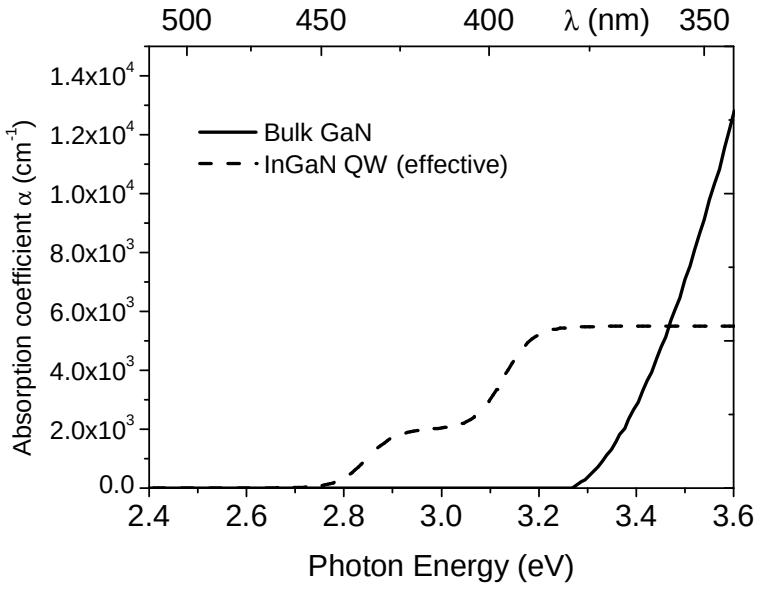


Fig. [6.22]. The dispersion relations for the absorption coefficients of bulk GaN and of InGaN QW used in the simulation of the experimental responsivity spectrum.

Table 6.7. Parameters describing the dispersion relation for the absorption coefficient in both GaN bulk and InGaN QW layers

Simulation parameters as from eqs. (5.14, 5.15)					
A	=	7.0×10^4 cm^{-1} eV$^{-1/2}$	C_2	=	3.25×10^3 cm^{-1}
Δ	=	27 meV	D_2	=	27 meV
$E_{g,bulk}$	=	3.52 eV	E_{2s}	=	0.28 eV
C_1	=	2.0×10^3 cm^{-1}	B	=	0
D_1	=	25 meV	Γ	=	-
$E_{g,QW}$	=	2.85 eV	E_{ex}	=	-

The simulation reproduces well the experimental photocurrent as far as the following elements concerns:

- the individual contribution of the different materials and layers
- the step-like absorption spectrum of the QW system
- the interference pattern of the photocurrent from QW absorption

Some features, however, are not as well reproduced:

- the interference pattern is more definite in the simulation than in the experiments: several factors, as spectral resolution (which is, however, quite high, about 1 nm, corresponding to about 0.01 eV) and a finite (non-zero) surface and interface roughness can explain this discrepancy
- the high-energy part of QW absorption; this can be due to more pronounced band-tailing effects in GaN which are not taken into account in the present dispersion relations;
- the high-energy part of GaN absorption; this part of the spectrum was normalized by the emission spectrum of the system *lamp + slits + monochromator*; in this part of the spectrum the light intensity is very low, and can be more affected by experimental error, so that this part of the normalized spectrum is affected with a significant uncertainty; this does not clear all doubts, but could be a reason for the discrepancy. Alternatively, the spectrum could be better fitted by inserting in the absorption coefficient an excitonic contribution, and by diminishing the constant A describing the bulk absorption strength. In order to distinguish between these two hypotheses, photocurrent measurements at low temperature would be useful, as in this case the excitonic contribution would increase.

6.5.4. Photocurrent dependence on applied bias

The variations in the PC spectra collected at different bias values can be well explained in the framework of model introduced for the simulation. Fig. [6.23] reports three spectra collected at V_{rev}=0V, V_{rev}=-1V and V_{rev}=-5V, respectively. The most visible change in the spectra is the increase in the GaN-related peak at E~3.45 eV. This is due to two main factors: the first factor is the increase in the depletion region width, which increases the amount of charge carriers generated within or in the vicinity of the depletion region; the second factor is the higher carrier velocity during collection, which lowers the probability of recombination at the junction. Provided the diffusion length is small in comparison with the depletion region width, the increase of the peak could be explained only through the widening of the depletion region: comparing these PC results, with the C-V characterization of HJ, one finds that the GaN-related peak height scales well with the measured depletion region width x_d.

Changes affect also the part of the spectra due to QW absorption. Starting from the lowest energies, we analyze first the QW-related absorption edge. As it is highlighted in the inset of fig. [6.23], there is a slight blue shift of the absorption edge with increasing reverse bias. This is due to the Quantum-Confined Stark Effect (QCSE) [Miller84], combined with the presence of strain-induced electric fields opposed to the built-in field [Chichibu], [Perlin], [Franssen04], [Renner02], [Zhang04]: at 0V, the separation in energy between the electron subband in the conduction band and the hole subband in the valence band of the QW have a certain energy separation, which is lower of that one would see in the case of a vanishing field inside the QW. When the reverse bias increases, it compensates the strain-induced field, yielding a higher energy separation of the two subbands. The same is true for transitions between higher subbands, although the shift is lower than for the ground-state subbands.

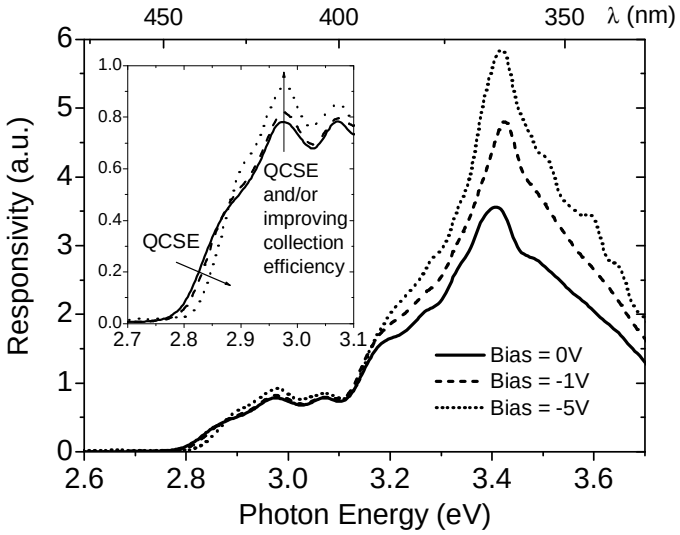


Fig. [6.23]. Photocurrent spectra at different applies bias: solid line, V_{rev}=0V; dotted line, V_{rev} =-1V, dotted line: V_{rev} =-5V. In the inset, the quantum well absorption edge and its variations with applied bias are highlighted.

Another consequence of the QCSE could be the increase in the PC signal due to the QW absorption: when the application of the reverse bias compensates the strain-induced electric field, the overlap of the subband envelope functions increases, yielding a higher transition probability. However, the increase of photocurrent could also be due to the improved collection efficiency due to the higher carrier velocity in the depletion region, which in turn lowers the recombination probability. The higher part of the absorption spectrum of the QWs, starting at E~3.1 eV shows a step in the spectrum, due to the step in the joint density of states of the QW system. Here one can notice another increase in the photocurrent with increasing reverse bias. This increase could be partly due to the superposition of the GaN-related peak, but this can not explain also the increase present for energy E~3.25 eV, which is clearly due to the QW absorption only. The most likely way to explain the PC increase at this energy value is rather by the improvement in collection efficiency due to a lower recombination probability than to the QCSE. Further and more in-depth calculations should be performed, in order to achieve a quantitative assessment of the role played by these different factors.

6.5.5. Comparison between EL and PC: Stokes shift, emission mechanism and In fraction in the quantum wells

The comparison between the responsivity and the electroluminescence spectra of sample HJ is reported in fig. [6.24]. What is apparent in the plot, is that the maximum of the emission takes place at an energy which is significantly lower than the bandgap energy extracted from the responsivity curve, which is related to absorption. In our case the absorption related bandgap energy has the value $E_{g,QW}$=2.85 eV, while the maximum of emission occurs for an energy E_{emi}=2.75 eV. This difference, in our case S = 0.1eV, is well known in the literature under the name of Stokes shift, and has been studied in detail for nitride-based alloys both in the case of epitaxial layers [O'Donnell99a],[O'Donnell99b],[O'Donnell01],[Wu02a] and quantum well systems [Berkowicz99], [Chichibu99]. The Stokes shift is zero for GaN [O'Donnell99b] and InN [Wu02b], and tends to reach a maximum value for equal fractions of In and Ga in the InGaN alloys. In the QWs, there is a further dependence on the well width [Berkowicz99]: the larger the well, the larger the shift. These observations have been explained either in terms of composition fluctuations in the alloy, or of a consequence of the piezoelectric field leading to the QCSE in the case of QW systems.

Composition fluctuations can induce band-tailing effects, as we too observed, and to the localization of carriers in potential minima. Localized carriers can give rise to the formation of excitons, which recombine emitting a photon of energy lower then bandgap. The states responsible for the emission are not equally visible by absorption spectroscopy, because they are steadily occupied and thus not available for the absorption process, as in the Burstein-Moss shift of the optical bandgap energy in degenerate semiconductors [Wu02a].

In the analysis of emission and absorption from epitaxial layers, O'Donnell found a value for *S* which is about 0.2 eV for a PL emission peak at 2.75 eV [O'Donnell99b]. The shift we observe is about the half. Care must be taken when comparing the result of a QW, like in our case, with an epitaxial layer, as differences between the two cases can arise because of the quantum size effect. The difference could also be explained by improved material quality, leading to a reduced role of composition fluctuations.

From the bandgap value extracted from the absorption spectrum we can estimate the x fraction of In in the QWs. In the works by Pereira et al. [Pereira01] and Shan et al. [Shan98] an absorption edge of 2.85 eV corresponds to an In fraction of about 0.15. These observations were performed on epitaxial layers, but the value x=0.15 can represent a good estimate also for our QW system, as the quantum size effect tends to raise the effective bandgap value, and the piezoelectric fields tend to lower it, thus yielding two opposite contributions.

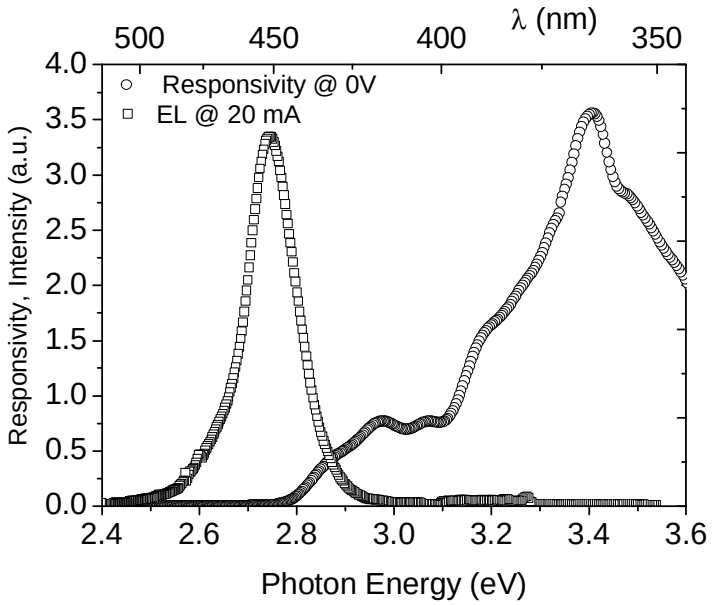


Fig. [6.24]. Electroluminescence (squares) and responsivity (circles) spectra of one of the analyzed LEDs, showing the Stokes shift between emission peak and absorption edge.

6.5.6. Summary of optical characterization

The analysis of the aging effects in these nitride-based LEDs manufactured as test structures led us to the following results. The first result is that there is a degradation of the external quantum efficiency of the device during the first few hours of operation at 20 mA DC. This degradation is stronger when the device is driven at lower current levels: this suggests that the quantum efficiency is lowered by the introduction of non-radiative recombination paths, most likely related to deep levels, in the active region. This observation is consistent with the previous results from C-V and DLTS characterization; it is also indicated that it is the *internal* quantum efficiency to be affected by DC current stress. A slight recovery was observed, however, after 50 h operations, which still must be clarified.

Secondly, photocurrent spectroscopy indicates that DC stress also affects the internal quantum efficiency of the device as a light detector. This is due to the same phenomenon leading to the degradation of the emission, i.e., the introduction of deep levels. Deep levels, in fact, may cause recombination of charge carriers in the active region, thus preventing them from being collected: this leads to a net diminution of the mobility and to a lower photocurrent signal. This diminution was proportional to the photocurrent signal in all spectral regions, thus indicating that it is rather connected with carrier transport than with carrier generation.

Photocurrent was also employed to achieve a deeper insight in the device structure and operation. The spectra were simulated keeping into account all the internal reflections of the optical field in the layered structure, thus explaining the modulated structure of the photocurrent spectrum due to absorption by the QW system. The device structure used for fitting the photocurrent spectra was consistent with that used for the fitting of C-V profiles. Due to the high number of fitting parameters used in both fitting procedures, we do not consider them as the actual structural parameters; however, we believe that they simulation parameters should not be very different from the actual ones.

The dispersion relations for the absorption coefficient used in the simulation showed that the In fraction used in the QW is around 0.15. The role of excitons was found to be negligible at room temperature, possibly due to the strong piezoelectric field building up across the QWs. The effects of this piezoelectric field were also visible in the blue shift of the QW absorption edge with increasing reverse bias.

6.6. Discussion and conclusions

Fig. [6.25] summarizes all the relative variations of physical parameters observed during low current aging with all experimentla techniques employed, with the exception of the reverse current. It shows that there is a strong correlation between all the observed phenomena. This leads us to hypotize that aging is driven mostly by the introduction in the active region of electrically active defects. The increase of defect concentration in the active region explains the following experimental observations:

- the decrease of the internal quantum efficiency in emission, because they provide non-radiative recombination paths;
- the decrease of the internal quantum collection efficiency, as they lower the lifetime fo charge carriers and therefore their mobility
- the increase of the reverse leakage current
- the increase of apparent charge concentration in the active region, provided the introduced/generated defects are associated with donor-like levels;
- the modifications of the DLTS spectra; these modifications could derive from the above described increase inthe apparent charge profile, or could themselves represent the introduction of the defect.

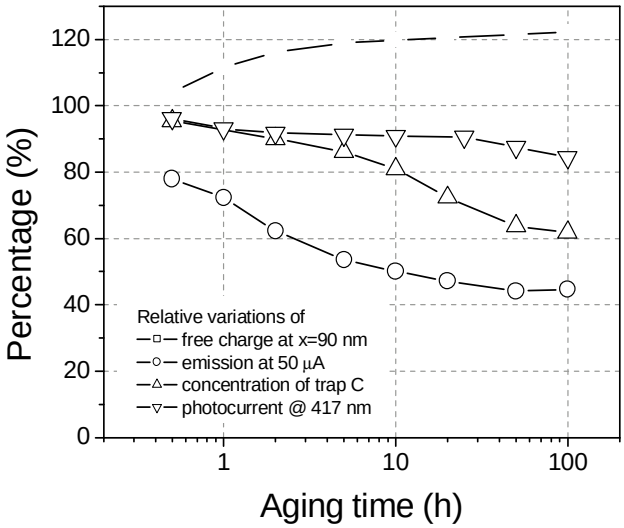


Fig. [6.25]. The relative veriation of different physical parameters observed with the employed experimental techniques during the aging of the devices.

The increase in the reverse current during the DC stress has not been included in fig. [6.25], as the relative variation of the reverse current amounts to about 300% after 50 h stress. Nevertheless, the evolution of this parameter is similar to the evolution of the other parameters, with a rapid variation during the first 5 h, and then a slower increase. The increase of the reverse current can be put in relationship with an increase of the electrical activity of defects associated to dislocations. This could represent a connection between the characterization of electrical transport and the capacitance spectroscopic observations, which evidence that the level changing in concentration is associated to extended defects.

6.6.1. The problem of the identification of the defect(s) responsible for the degradation under DC current stress

The main feature detected by DLTS, which was the only technique employed capable to detect the trace of single deep levels is peak C. This peak has an enthalpy E_c-E_t=0.28 eV, and shows a behaviour typical of deep levels associated with extended defects. Furthermore, this peak changes its shape in the device region where the apparent charge concetntration is strongly affected by stress. However, this is not enough to conclude that it is the generation of defects associated to C to worsen the emission characteristics of the device. The modifications occurring to C may well be driven by the modifications of the apparent charge concentration, which, in turn, may be due to other defect levels lying deeper in the gap. So far, we were not able to detect other deep levels in significant concentration for temperatures between 240 K and 400 K. Thus, if other defects are introduced in the gap and provide non-radiative recombination paths, their activation energy can be estimated as 1 eV from the conduction band or more. In this case, such deep levels would be somewhat difficult to detect by DLTS; in any case, it would be impossible to monitor their evolution during DC stress, as the device heating could activate further generation/migration mechanisms, which would lead to spurious effects.

6.6.2. Possible generation/migration mechanism

The mechanism governing the degradation of DC stressed LEDs is still under study and will be investigated more in detail in the future work. However, we can distinguish between three main categories of generation/migration mechanisms involving deep levels:

Thermally activated mechanisms.

It is known that raising the temperature the diffusivity of all defects and impurities increases. This category is excluded for the present analysis, because the low current levels used yielded a very low device heating, such that the temperature in the active region of the device never exceeded 60 °C. However, the same analysis performed here on DC current stresse devices will be performed on thermally stressed devices, so that the comparison between the two cases could yield significant insight for the interpretation of the present and the future results.

Current-activated mechanisms.

In the past, the degradation of current-stressed GaN-based LEDs has been described to a model of defect formation by hot electrons injected in the quantum wells [Craford94], [Manyakhin98]. According to this model, carriers being captured in the QW can lose a significant amount of kinetic energy, exceeding the defect formation energy. Thus, they would be able to generate a high amount of defects in the active region of the device. This sort of generation is localized, taking place in proximity of the QWs.

Fermi level-activated mechanisms.

It is known that the concentration of intrinsic charged defects in a semiconductiong material is strongly dependent on the Fermi level position, which in turn is dependent on the doping and on the free carrier concentration [Longini56], [Walukiewicz89], [Walukiewicz94]. Generally speaking, cherged intrinsic defects tend to increase with increasing free carrier concentration, tending to self-compensate the material. Moreover, they can enhance the diffusivity of other defect species, such as impurities, through a Frank-Turnbull mechanism involving vacancies [Frank56] or a *kick-out* mechanism involving interstitials [Gösele81]. It could be possible that in a forward current-stressed LEDs the high concentration of free carriers in the junction region (both electrons and holes) have the net effect of raising the concentration of intrinsic charged defects, which then play a role diffusing themselves or assisting other species in the diffusion process. This mechanism is of rather localized nature, as high free carrier densities are reached only in the active region close to the junction. In order to attempt a quantitative estimate of the relevance of this mechanism one should have a good knowledge of the parameters governing the concentration of intrinsic charged defects in GaN and AlGaN and of the device structure, as further, more complex phenomena occur in the diffusion of defects in heterostructures [Chen99].

6.6.3. Summary

The low current stress of the analyzed LEDs induced in them a decrease of the optical efficiency, which was more prominent during the first 20 hours of stress, and enhanced for low current values. The degradation affecting both EL and PC spectra proved to be connected with a decrease of the internal quantum efficiency of the device[1]. The capacitive measurements showed that stress induced an apparent charge increase in the whole investigated region, particularly pronounced at the interface between the active region and the n-side, where during ageing a peak in the n_{CV} profile grew (at 100 nm). The DLTS analysis, performed in the same region of this n_{CV} peak (~100 nm), showed, as a consequence of stress, significant modifications in the properties of traps at 0.18-0.28 eV (traps B and C). These localized modifications are supposed to be related to the generation of non-radiative paths, which lower the efficiency of the devices, especially at low current levels: at high current levels, the non-radiative paths could saturate, thus leading to the

[1] Cathodoluminescence (CL) measurements, performed at the University of Parma in the framework of the present project, confirmed this trend: moreover, the CL characterization carried out at different temperature levels indicated that the QW loss measured at high temperatures was stronger than at low temperatures. The greater OP loss shown at high temperatures by the CL measurements can therefore be attributed to the increase, as the sample is heated, of the non-radiative recombination rate.

minor OP loss shown by the EL measurement. The most important modifications of the apparent charge profile occur at the boundary between active and bulk regions, but we can not conlude that the generation/modification of trap levels takes place only at the active region boundary: an extended defect generation, in fact, has been observed in some of the analyzed samples, where the n_{CV} changed in all the region analysed by the C-V measurements[2]. As previously described [Craford94], [Manyakhin98], the generation/propagation of defects even at these low current densities could be related to a process of defect generation by hot electrons, or, alternatively, to the enhanced diffision of impurity atoms assisted by a Fermi-level driven increase of charged intrinsic defects [Walukiewicz94]. The DLTS analysis doesn't explain all the n_{CV}, EL and PC modifications: the generation of other defects with very deep nature can be supposed, so that they are not detectable within the explored DLTS range, set by the need of avoiding exceeding device heating. Finally, the modifications of the traps at the interface of the active layer (possibly contributing to tunnel-injection of the carriers inside the QWs) could also be related to changes in the capture efficiency of the active region, and thus contribute in lowering the overall LEDs efficiency. This hypothesis has to be confirmed by further measurements, with the aim of characterizing ageing-induced variations of the transport mechanisms.

[2] In this case, the CL decrease didn't affect only the QW line, but also the GaN exciton and yellow bands.

REFERENCES

[1DPoisson] G. Snider, *1DPoisson*, downloadable at www.nd.edu/~gsnider

[Ambacher96] O. Ambacher, W. Rieger, P. Ansmann, H. Angerer, T.D. Moustakas and M. Stutzmann, *Sol. State. Commun.* **97**, 365 (1996)

[BandEng] M. Grundmann, *BandEng*, downloadable at www.ece.ucsb.edu/mgrundmann/bandeng.htm

[Barton99] D.L. Barton, M. Osiński, P. Perlin, P.G. Eliseev and J. Lee, *Microel. Reliab.* **39**, 1219 (1999)

[Berkowicz99] E.Berkowicz, D. Gershoni, G. BAhir, A.C. Abare, S.P. DenBaars and L.A. Coldren, *Phys. Stat. Sol.* (b) **216**, 291 (1999)

[Binet96] F Binet, J.Y. Duboz, E. Rosencher, F. Scholz and V. Härle, *Phys. Rev. B*, **54** 8116 (1996)

[Cao03] X.A. Cao, P.M. Sandvik, S.F. LeBoeuf, and S.D. Arthur, *Microelectron. Reliab.* **43**, 1987 (2003)

[Cao04] X.A. Cao, J.A. Teetsov, F. Shahedipour-Sandvik, and S.D. Arthur, *J. Cryst. Growth* **264**, 172 (2004)

[Craford94] M.G. Craford and F.M. Steranka, Encycl. *Appl Phys***. 8**, 485 (1994)

[Chen99] C.H. Chen, U.M. Gösele and T.Y. Tan, *Appl. Phys. A* **68**, 9 (1999), *Appl. Phys. A* **68**, 19 (1999) and *Appl. Phys. A* **69**, 313 (1999)

[Chichibu99] S.F. Chichibu, A.C. Abare, M.P. Mack, M.S. Minsky, T. Deguchi, D. Cohen, P. Kozodoy, S.B. Fleischer, S. Keller, J.S. Speck, J.E. Bowers, E. Hu, U.K. Mishra, L.A. Coldren, S.P. DenBaars, K. Wada, T. Soda and S. Nakamura, *Mater. Sci. Eng.* B**59**, 298 (1999)

[Chuo01] C.C. Chuo, C.M. Lee, and J.I. Chyi, *Appl. Phys. Lett.* **78**, 314 (2001)

[Dingle71] R. Dingle, D.D. Sell, S.E. Stokowski and M.I. Ilegems, *Phys. Rev. B* **4,** 1211 (1971)

[Ershov01] M. Ershov, B. Yaldiz, A. G. U. Perera, S. G. Matsik, H. C. Liu, M. Buchanan, Z. R. Wasilewski, and M. D. Williams, *Infrared Phys. Technol.* **42**, 259 (2001)

[Fang00] Z. Q. Fang, D. C. Reynolds, and D. C. Look, *J. Electron. Mat.* 29, 448 (2000)

[Frank56] F.C. Frank and D. Turnbull, *Phys. Rev.* **104**, 617 (1956)

[Franssen04] G. Franssen, ...P. Perlin, *Phys. Rev. B* **69**, 45310 (2004)

[Gösele81] U. Gösele, F. Morehead, W. Frank and A. Seeger, *Appl. Phys. Lett*. **38**, 157 (1981)

[Li04] D. S. Li, H. Chen, H. B. Yu, H. Q. Jia, Q. Huang, and J. M. Zhou, *J. Appl. Phys.* **96**, 1111 (2004)

[Longini56] R.L. Longini and R.F. Greene, *Phys. Rev.* **102**, 992 (1956)

[Lucia93] M. L. Lucia, J. L. Hernandez-Rojas, C. Leon, and I. Mártil, *Eur. J. Phys.* **14**, 86 (1993)

[Manyakhin98] F. Manyakhin, A. Kovalev and A.E. yunovich, *MRS Internet J. Nitride Semicond. Res.* **3,** 53 (1998)

[Miller 84] D.A.B. Miller, D.S. Chemla, T.C. Damen, A.C. Gossard, W. Wiegemann, T.H. Wood and C.A. Burrus, *Phys. Rev. Lett.* **53**, 2173 (1984)

[Moon98] C. R. Moon, B. D. Choe, S. D. Kwon, H. K. Shin, and H. J. Lim, *J. Appl. Phys.* **84**, 2673 (1998)

[O'Donnell99a] K.P. O'Donnell, R.W. Martin, S. Pereira, A. Bangura, M.E. White, W. van der Stricht and K. Jacobs, *Phys. Stat. Sol. (b)*, **216**, 141 (1999)

[O'Donnell99b] K.P. O'Donnell, R.W. Martin and P.G. Middleton, *Phys. Rev. Lett.* ***82***, 237 (1999)

[O'Donnell01] K.P. O'Donnell, R.W. Martin, C. Trager-Cowan, M.E. White, K. Esona, C. Deatcher, P.G. Middleton, K. Jacobs, W. van der Stricht, C. Merlet, B. Gil, A Vantomme and J.F.W. Mosselmans, *Mat. Sci. Eng.* **B82**, 194 (2001)

[Pereira01] S. Pereira, M.R. Correia, T. Monteiro, E. Pereira, E. Alves, A.D> Sequeira, N. Franco, *Appl. Phys. Lett* **78**, 2137 (2001)

[Polenta04] A. Castaldini, A. Cavallini and L. Polenta, *Appl. Phys. Lett*. **84**, 4851 (2004)

[Polyakov02] A.Y. Polyakov, N.B. Smirnov, A.V. Govorkov, J. Kim, B. Luo, R. Mehandru, F. Ren, K.P. Lee, S.J. Pearton, A.V. Osinsky, and P.E. Norris, *J. Appl. Phys.* **91**, 5203 (2002)

[Pursiainen01] O. Pursiainen, N. Linder, A. Jaeger, R. Oberschmid, and K. Streubel, *Appl. Phys. Lett.* 79, 2895 (2001)

[Renner02] F. Renner, P. Kiesel, G.H. Döhler, M. Kneissl, G. Van de Walle and N.M. Johnson, *Appl. Phys. Lett.* **81**, 490 (2002)

[Schroder] D. K. Schroder, *Semiconductor Material and Device Characterization* (Wiley-Interscience, New York, 1998)

[Shan98] W. Shan, W. Walukiewicz, E.E. Haller, B.D. Little, J.J. Song, M.D. McCluskey, N.M. Johnson, Z.C. Feng, M. Schurman, R.A. Stall, *J. Appl. Phys*. **84**, 4452 (1998)

[Soh04] C.B. Soh, S. J. Chua, H.F. Lim, D.Z. Chi, W. Liu and S. Tripathy, *J. Phys.: Condens. Matter* **16**, p. 6305 (2004)

[Walukiewicz89] W. Walukiewicz, *Appl. Phys. Lett.* **54**, 2094 (1989)

[Walukiewicz94] W. Walukiewicz, *Phys. Rev. B*, **50**, 5221 (1994)

[Wu02a] J. Wu, W. Walukiewicz, K.M. Yu, J.W. Ager III, E.E. Haller, H. Lu, W.J. Schaff, Y. Saito and Y. Nanishi, *Appl. Phys. Lett.* **80**, 3967 (2002)

[Wu02b] J. Wu, W. Walukiewicz, K.M. Yu, J.W. Ager III, E.E. Haller, H. Lu, W.J. Schaff, *Appl. Phys. Lett.* **80**, 4741 (2002)

[Yastrubchak04] O. Yastrubchak, T. Wosiński, A. Makosa, T. Figielski, S. Porowski, I. Grzegory, R. Czernecki and P.Perlin, *Eur. Phys. J. Appl. Phys*. **27**, 201 (2004)

[Youna03] C.J. Youna, T.S. Jeonga, M.S. Hana, J.W. Yanga, K.Y. Lima, and H.W. Yu, *J. Cryst. Growth* **250**, 331 (2003)

[Zhang04] X.H. Zhang, W. Liu and S.J. Chua, *J. Cryst. Growth*, **268**, 521 (2004)

7 Evolution of irradiation-induced defect states in low-temperature annealed 4H-SiC

The study of defects introduced in a controlled way is a powerful tool to relate the properties of the defects themselves to the device performance. As well known, particle irradiation is a straightforward tool to generate simple defects in a controlled way. For this reason it is widely used in silicon and recently also in silicon carbide [Lebedev99] in order to investigate the origin and, whenever possible, the structure of the defects.

The present analysis deals with the deep energy levels associated with defects induced by irradiation with protons and electrons in 4H SiC, the doping density of which is in the range of $n=2\div6 \times 10^{15}cm^{-3}$. Deep level enthalpy, concentration and apparent capture cross-section have been monitored for both irradiation types by changing the particle fluence, and put in relationship with the observed compensation effects. Some conclusions about the nature of the defect introduced have been drawn. The charge transport properties of the material, a key point for device performance, have also been analyzed.

Another important tool in the study of defects is the thermally induced modification of their properties, i.e., the annealing. In the last few years, particular attention has arisen around defect annealing phenomena taking place at low temperature ($\leq$ 500 K) in SiC polytypes irradiated with different kinds of particles. Early experimental evidence of annealing stages at T~325K is reported for electron irradiated 3C-SiC [Itoh89], followed then by studies on electron, neutron and heavy ion irradiated 6H-SiC [Hemmingsson02], [Chen02] and 4H-SiC [David02], [Alfieri03], [Nielsen03], [Zhang02] with different experimental techniques.

An intuitive and straightforward interpretation of annealing phenomena might be the occurrence of the thermally induced annihilation of defects, in a way similar to the recombination of Frenkel pairs. According to theoretical works [Posselt01], [Gao02], [Gao03], some defects created by ion irradiation have migration energy pathways with energy barriers lower than 1 eV, accounting for the activation of the phenomenon even at relatively low temperatures.

Another model, earlier used to explain thermally- and bias-induced modifications of deep level transient spectra in silicon [Chantre87], takes into account a *configurational metastability* of the defect associated with the deep level detected. This model has been developed for SiC by Hemmingsson et al. [Hemmingsson02] studying deep level spectra of electron irradiated 6H-SiC. The present analysis deals with the observation of a low temperature annealing in 8.6 MeV electron irradiated 4H-SiC Schottky diodes, and interprets the results in the light of the above said models.

7.1. Samples and treatment

7.1.1. Schottky diodes and irradiation conditions

Proton irradiation

The samples examined were 4H-SiC epilayers grown by Chemical Vapor Deposition (CVD) and provided by CREE Inc. The doping concentration of the epitaxial layer in the proton irradiated diodes was $N_d = 5 \times 10^{15}cm^{-3}$. This layer, 7 μm thick, was grown on a 400 μm thick 4H-SiC substrate with doping density $N_d = 10^{19}$ cm^{-3}. The declared micropipe density was 16-30 cm^{-2}. The Ti Schottky contact was 1000 Å thick, annealed at 400 °C in N_2 atmosphere with an area of $1mm^2$. Metal oxide field plates were synthesized in order to prevent the electric field crowding at the contact edges. The ohmic contacts were obtained by growing Ti/Ni/Ag multilayer (with thickness of 1000 Å /2000 Å/2 μm respectively). The diodes were irradiated at room temperature with 6.5 MeV

protons at fluences ranging from $\Phi=10^{11}cm^{-2}$ to $\Phi=3.2x10^{13}cm^{-2}$ (Table 7.1). From now on, we will refer straightforwardly to the irradiation fluences, for the sake of clarity.

Electron irradiation

The electron irradiated samples consisted of CVD-grown 30 μm thick epitaxial layers with doping density N_d= 2 x 10^{15} cm^{-3}, a buffer layer with doping density $N_d=10^{18}$ cm^{-3} and a substrate with doping density $N_d=10^{19}cm^{-3}$. The substrate micropipe density was 10 cm^{-2}. The circular Au Schottky contact was 1500 Å thick with a diameter of 2mm and one guard ring. The ohmic contact was obtained by deposition of a 1000 Å thick multilayer of Ti/Pt/Au, followed by annealing at 500 ºC for 30 s in Argon atmosphere. These samples were irradiated at 26 °C with electrons from a LINAC with energy equal to 8.2 MeV at fluences ranging from 4.75 x 10^{13} cm^{-2} to 9.5 x 10^{14} cm^{-2} (Table 7.1).

The transport properties and the electronic levels associated with the defects were analyzed by current-voltage characteristics (I-V) and deep level transient spectroscopy measurements (DLTS) up to 550 K, respectively. The deep level enthalpy with respect to the conduction band, E_T, the capture cross-section σ_t and the deep level concentration N_T have been measured for each level. In the same temperature range the free carrier concentration n was measured by capacitance-voltage characteristics (C-V) in the as-prepared and irradiated samples in order to monitor the compensation effects by intrinsic irradiation-induced defects. DLTS measurements were carried out by means of a SULA double boxcar spectrometer with exponential correlator and I-V characteristics were collected by a Keithley 6517 electrometer. The capacitance meter was a Keithley 3330, operating at a frequency f_{cap} =100 kHz.

Table 7.1 Irradiation conditions

Proton 6.5 MeV		**Electron 8.2 MeV**	
Sample	Fluence Φ [cm^{-2}]	Sample	Fluence [cm^{-2}]
A4	-	I4	-
A2	10^{11}	G7	4.75 x 10^{13}
A1	10^{12}	A4	2.38 x 10^{14}
A3	3.2 x 10^{13}	G4	9.50 x 10^{14}

7.1.2. Thermal treatments (annealing)

The annealing behaviour of the deep levels, was in a first time observed during the Deep Level Transient Spectroscopy (DLTS) characterization and monitored by running the DLTS up to gradually increasing temperature, varying the temperature at a rate of 0.07 K/s. Thus, we observed two distinct annealing stages, one in the interval 360 – 400 K (87 – 127 °C), the other in the interval 400 – 470 K (127 – 197 °C). Subsequently, we performed thermal treatments on another set of samples (isochronal annealing) at the T=100 °C, T=200 °C , T=300 °C and T=450 °C, each 4 h long, which brought further changes to the DLTS spectra. Capacitance-voltage (C-V) characterization was carried out before irradiation and, after irradiation, before and after annealing, in order to monitor the free charge carrier density. Using both techniques, DLTS and C-V, in a complementary way, allowed us to interpret the annealing stages as a combination of defect annihilation and defect configuration change.

7.2. Effect of irradiation on the analyzed samples

7.2.1. DLTS of proton irradiated diodes

As far as samples irradiated with protons are concerned, a deep level S0, with enthalpy E_C-E_T =0.18 eV and concentration equal to ~10^{13} cm^{-3}, was detected in the as-prepared sample. Upon proton irradiation further levels appeared, i.e. S1 with E_C-E_T =0.20 eV, S2 with E_C-E_T =0.40 eV, S3 with E_C-E_T =0.72 eV, S4 with E_C-E_T =0.76 eV and S5 with E_C-E_T =1.09 eV, the concentration of which was found to increase linearly with the irradiation fluence Φ. The density of level S0 remained constant at N_T≈10^{13} cm^{-3} whatever the fluence was. The parameters of the deep levels and the comparison with data from literature are given in Table 7.2, while the evolution of the spectra with respect to the irradiation dose is shown in Fig. [7.1]. In the diodes irradiated with the highest fluence, 3.2 x 10^{13} cm^{-2}, the total deep level density results comparable to the doping density, though somewhat lower, so that compensation effects should be expected. These effects were confirmed by subsequent C-V measurements, as reported below.

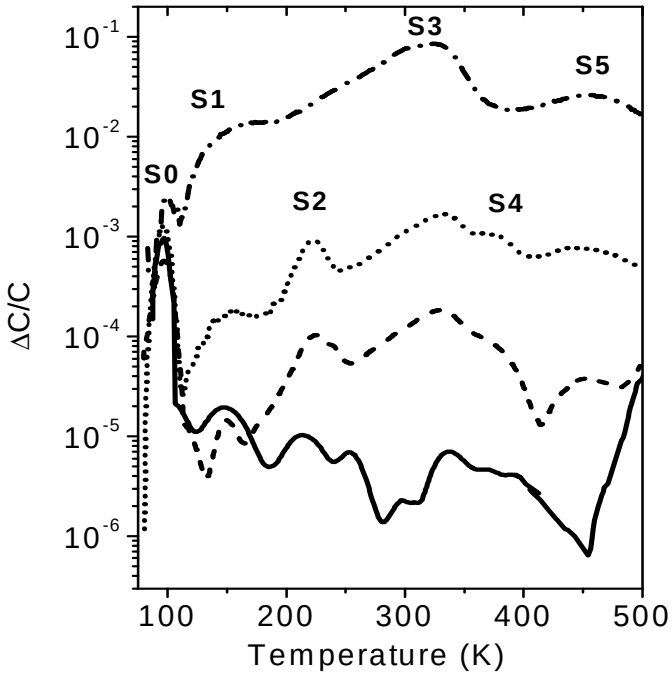


Fig. [7.1]. Evolution of the DLTS spectra of Proton irradiated diodes. Continuous line: as-prepared diode; dashed line: diode irradiated with Φ=10^{11} cm^{-2}; dotted line: diode irradiated with Φ=10^{12} cm^{-2}; dashed-dotted line: diode irradiated with Φ=3.2 x 10^{13} cm^{-2}. Emission rate is e_n=232 s^{-1}.

Table 7.2. Enthalpy E_C-E_T, concentration N_T, apparent capture cross-section σ_t and introduction rate $\eta_T = N_T/\Phi$ of the electronic levels observed in 4H-SiC irradiated with a proton fluence of 3.2 x 10^{13} cm^{-2}.

Trap label	E_C-E_T [eV]	N_T [cm^{-3}]	σ_t [cm^2]	η_T [cm^{-1}]	Comparison with published data Deep Level	Attribution
S0	0.18	1.0 x 10^{13}	2 x 10^{-14}	~ 0	Ti [a], P_1/P_2 [b]	Ti impurity [a,c] primary defects[b]
S1	0.20	6.3 x 10^{13}	7 x 10^{-18}	3.2	-	-
S2	0.40	1.3 x 10^{14}	1 x 10^{-15}	4.1	EH1[d]	-
S3	0.72	3.5 x 10^{14}	2 x 10^{-14}	11.0	Z_1/Z_2 [a,b],EH2 [d]	N+I_C [e]
S4	0.76	1.1 x 10^{14}	1 x 10^{-14}	3.5	EH3 [d]	-
S5	1.09	7.7 x 10^{13}	5 x 10^{-13}	2.4	RD1/2[a], EH5 [d]	V_C+V_{Si}[b]

[a] ref. [Dalibor97], [b] ref. [Lebedev00], [c] present work, [d] ref. [Hemmingsson97], [e] ref. [Eberlein03].

7.2.2. DLTS of electron irradiated diodes

In diodes irradiated by electrons, the deep level generation occurs with characteristics quite similar to proton irradiation (Fig. [7.2]). Also in these epilayers, a shallower level S0 (enthalpy E_C-E_T =0.15 eV) is found in the as-prepared material. Slightly differently from the previous case, we were able to detect also a deeper level, S5*, with E_C-E_T =0.89 eV. Upon irradiation DLTS detected the levels S1* with E_C-E_T =0.23 eV, S2 with E_C-E_T =0.39 eV, S3 with E_C-E_T =0.5/0.65 eV, S4 with E_C-E_T =0.76 eV. The parameters of the relevant deep levels are reported in Table 7.3. It is worth noting that some of the deep levels increasing dramatically in concentration in the irradiated diodes are present also before irradiation (continuous lines in Fig. [7.2]) even though with density close to the resolution limit of the experimental setup. The concentration of levels S1* to S5* is linearly dependent on the irradiation dose, whereas the concentration of S0 is independent of it.

Regarding the electron irradiated diodes, a low temperature annealing was observed, the interpretation of which is dealt with in the following sections. The annealing process consists of two stages. In the first stage, from 360 to 400 K, the amplitude of a DLTS peak labelled S2A (Table 7.3) decreased dramatically while that of peak S3 slightly increased. In the temperature interval where S2A was detected, a minor peak, labelled S2, with equal enthalpy appeared. In the second stage, from 400 to 470 K, the shape of peak S3 sharpened, its amplitude slightly decreased and its enthalpy changed from 0.5 eV to 0.65 eV, taking on the parameters of the peak Z_1/Z_2 known from the literature [Dalibor97], [Eberlein03]. In Fig. [7.2] only the spectra recorded after annealing are shown, since in the present work we consider important the deep level parameters found after both annealing stages, just as the analyzed proton-irradiated diodes had been previously annealed before the DLTS analysis.

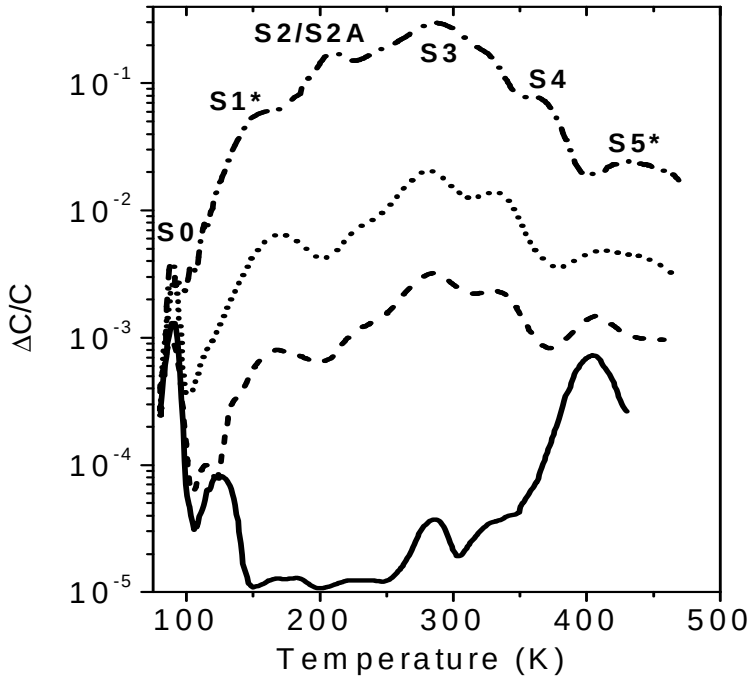


Fig. [7.2]. Evolution of the DLTS spectra of electron irradiated diodes. Continuous line: as-prepared diode; dashed line: diode irradiated with Φ=4.75 x 10^{13} cm^{-2} ; dotted line: diode irradiated with Φ=2.37 x 10^{14} cm^{-2} ; dashed- dotted line: diode irradiated with Φ=9.50 x $10^{14}$$cm^{-2}$. The emission rate is e_n=46.5 s^{-1}.

Table 7.3. Enthalpy E_C-E_T , concentration N_T, apparent capture cross-section σ_t and introduction rate $\eta_T = N_T/\Phi$ of the electronic levels observed in 4H-SiC irradiated with an electron fluence of 9.5 x 10^{14} cm^{-2}.

Trap label	E_C-E_T [eV]	N_T [cm^{-3}]	σ_t [cm^2]	η_T [cm^{-1}]	Comparison with published data Deep Level	Attribution
S0	0.15	1.4 x 10^{13}	6 x 10^{-16}	~ 0	Ti [a], P_1/P_2 [b]	Ti impurity [a,c] primary defects [b]
S1*	0.23	5.5 x 10^{13}	9 x 10^{-16}	0.06	-	-
S2A	0.33	6.9 x 10^{14}	7 x 10^{-16}	0.73	-	I_C [d]
S2	0.39	4.0 x 10^{14}	2 x 10^{-15}	0.42	EH1 [d]	-
S3	0.65/ 0.50	4.2 / 5.4 x 10^{14}	10^{-16} / 10^{-14}	0.44	Z_1/Z_2 [a,b], EH2 [d]	N+I_C [e]
S4	0.75	2.6 x 10^{14}	6 x 10^{-15}	0.27	EH3 [d]	-
S5*	0.89	4.6 x 10^{13}	7 x 10^{-15}	0.09	RD1/2 [a], EH5 [d]	V_C+V_{Si} [b]

[a] ref. [Dalibor97], [b] ref. [Lebedev00], [c] present work, [d] ref. [Hemmingsson97], [e] ref. [Eberlein03].

7.2.3. Analysis of introduction rates and deep levels

From the aforementioned results, we can assess that irradiation with either particle type generates almost the same set of electron traps in n-type epilayers, as is also stressed by the Arrhenius plot (Fig. [7.3). The introduction rate $\eta_T = N_T/\Phi$ is shown in Fig. [7.4] for levels with similar signatures. It is to be noted that η_T is about one order of magnitude higher in the case of

proton irradiation, i.e. $\eta_{T,proton}/\eta_{T,electron}$~ 10 for levels S2, S3 and S4, whilst in the case of levels S1/S1* and S5/S5* $\eta_{T,proton}/\eta_{T,electron} \geq 30$. This strongly suggests that i) defect generation occurs mostly according to the same mechanism in both cases since the same defects are generated and ii) most of the deep levels detected by DLTS should be either of intrinsic nature or very simple defect complexes involving shallow donor impurities. The former result is not obvious, given the significant difference in the impinging particle masses. The latter result comes from the observation that the more complex the defect configuration is, the more different should be the introduction kinetics in either impinging particle type.

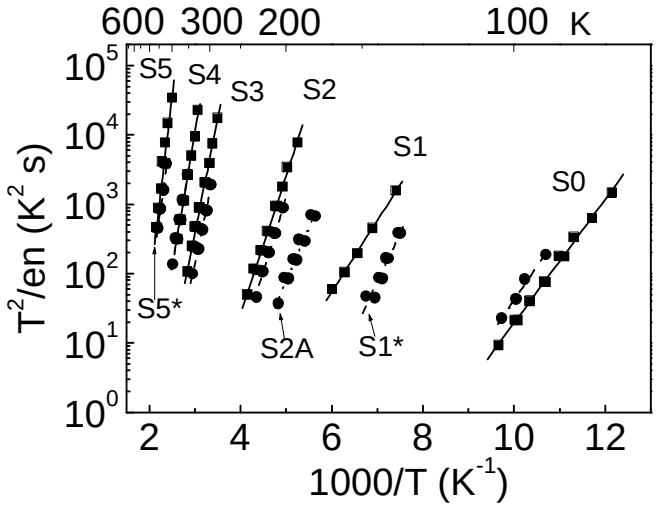


Fig. [7.3. Arrhenius plot of proton and electron irradiated samples, showing that similar sets of traps are present for irradiation with either particle type. Squares refer to proton irradiation, circles to electron irradiation. Solid lines represent the fit for proton irradiation, dashed lines the fit for electron irradiation.

From Fig. [7.4] we also observe that the introduction rate is the smallest for both levels S1*/S1 and S5*/S5. This is also the case in which the ratio $\eta_{T,proton}/\eta_{T,electron}$ is the highest. Although the identification of the defects giving rise to deep levels is in many cases still dubious, some can be advanced. Some levels have been extensively studied in the past few years due to their particular annealing behaviour or to other characteristics, such as the negative-U behaviour [Hemmingsson98], that made them particularly suitable to a more in-depth analysis. We list in the following our findings on the detected levels (Figs. [7.1],[7.2], [7.3, [7.4]).

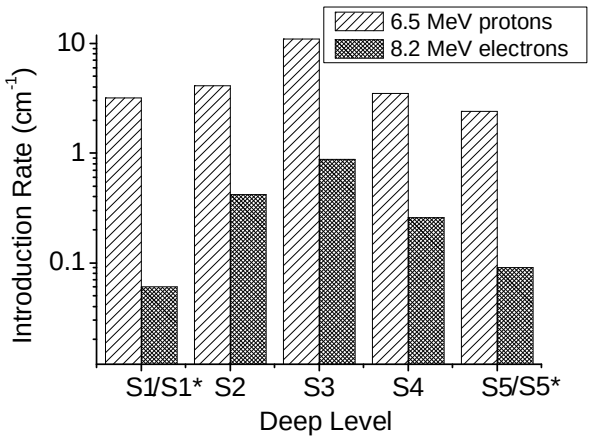


Fig. [7.4]. Introduction rate η_T of deep levels introduced by proton (light columns) and electron (dark columns) irradiation.

Level S0. The behavior of S0 (E_t = 0.15/0.18 eV), which is the only level maintaining a constant concentration throughout the whole range of fluences, suggests a relationship of this level with impurity atoms, which are not affected at all by irradiation. Literature data indicate that S0 is to be put in relationship with an acceptor level originated by the Ti impurity, for which two levels with enthalpy E_C-E_T = 0.12 eV and E_C-E_T = 0.16 eV are attested [Lebedev99], [Dalibor97], [Achtziger97]. The capture cross-section measured in the proton irradiated diode (σ_t=2x10^{-14}cm^2) is significantly larger than that measured in the electron irradiated diode (σ_t =6x10^{-16}cm^2). This is probably due to the large uncertainty associated with the determination of the capture cross-section by extrapolation to T=∞ of the Arrhenius plot fit. This may also explain the difference between our values and that found by Dalibor et al. (σ_t =1x10^{-16}cm^2) [Dalibor97]. Relying more on the enthalpy values, we are therefore confident that S0 has to be associated with the Ti impurity. Indeed, Ti may unintentionally contaminate SiC during the growth due to the crucible graphite parts involved [Kawasuso02].

Levels S1, S1*. In both cases of proton and electron irradiation a deep level is found in the temperature region T ~ 150K. However, some facts suggest that this level is not to be identified with the same defect in one or the other case. First, the deep level parameters are different enough (Tables 7.2, 7.3) to yield two well distinguishable signatures on the Arrhenius plot. Second, the ratio of the introduction rates for proton and electron irradiation here assumes the value of about 60, thus deviating significantly from the value of about 10 found for all levels with the same deep level parameters except for S5. Thus, we conclude that S1 and S1* should not be identified with the same level, i.e., electrons and protons introduce two distinguished complex defect configurations with close enthalpy.

Level S2, S2A. The peculiarity of level S2A (E_C-E_T = 0.33 eV) is its low-temperature annealing out, which will be dealt with in the following sections. The amplitude of the peak related to this level in fact decreases dramatically in the DLTS spectrum after the sample enters the temperature interval 360 – 400 K. Annealing out of a level lying close in enthalpy, labelled EH1 in literature, was previously reported in 2.5 MeV electron irradiated samples, although at a higher temperature (750 °C) [Hemmingsson97]. Level S2A may be related to the carbon interstitial I_C [Gao03], [Zhang02]. After the annealing, level S2 is still present in the DLTS spectrum of electron-irradiated diodes, presenting an enthalpy close to S2A and a higher capture cross-section (Fig. [7.3). This level matches up very well with level S2 found in the proton-irradiated samples and with the EH1 level.

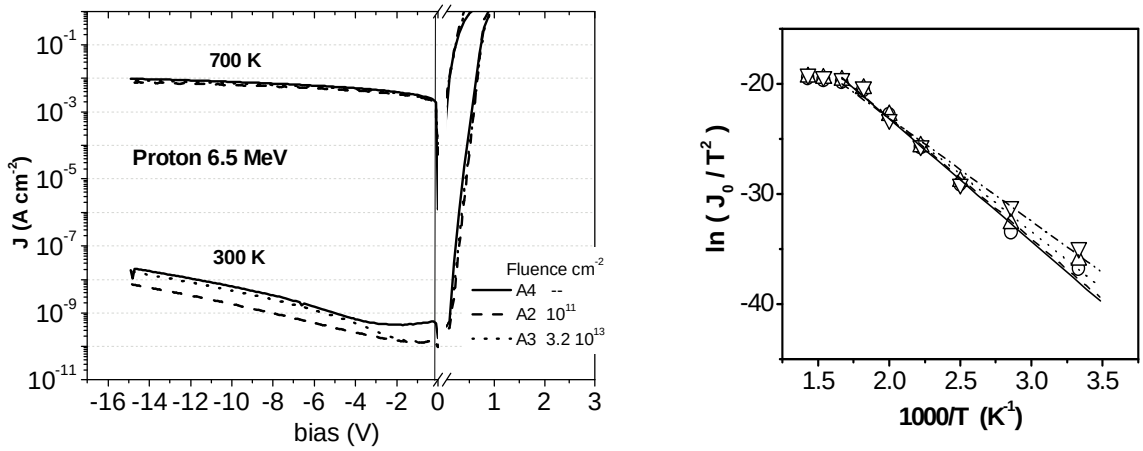
Level S3. It is known that in the region of the DLTS spectrum around 300K a dominating peak with activation energy E_C-E_T = 0.6/0.7 eV appears on irradiation [Lebedev99], [Dalibor97], although it is often found also in as-prepared samples [Fang01]. It is generally believed to be a set of discrete levels with similar energy, some of which are unstable and change configuration after: i) months at room temperature [Doyle98], ii) thermal treatments [Hemmingsson98], iii) combined action of thermal annealing and reverse bias [Martin04]. The level Z_1/Z_2, which is the most stable of these levels, (called S3 in the present work) has been in the last few years the most investigated electronic level both from the experimental and the theoretical point of view. According to recent studies Z_1/Z_2 is a negative-U center [Hemmingsson98] and is related to a di-carbon interstitial next to a nitrogen atom [Eberlein03], [Pintilie02] or, alternatively, to an intrinsic defect complex possibly involving hydrogen [Storasta04]. Our most recent results [SST06] confirm through analysis of compensation in proton-irradiated 4H-SiC Schottky diodes that nitrogen is not involved in the defect structure: the deep level is therefore related to intrinsic defects.

Level S4. From the analysis of the Arrhenius plot and of the introduction rates, it is apparent that also level S4 is associated with the same defect in both cases of proton and electron irradiation. This level peak, significantly overlapping part of the dominating peak associated with level S3, should also be associated to an intrinsic defect or to a small complex.

Level S5, S5*. These levels are part of a group of levels denominated in literature as "radiation damages" (RD) [Dalibor97], [Lebedev00], so as to stress their irradiation-induced origin. This was confirmed by our observations, as we found the concentration of both levels S5 (E_C-E_T =1.09eV) and S5* (E_C-E_T =0.89 eV) to grow linearly with the irradiation dose. However, we observed that level S5* is sometimes present in significant amount also in as-prepared samples. The microscopic structure of these defects has been studied in previous works [Lebedev00]. According to those authors, RD1/2 (E_C-E_T =0.96 eV), quite close to the levels here denominated S5 and S5*, is related to a defect complex involving a carbon and a silicon vacancy V_C+V_{Si}. The attribution to a small defect complex is in agreement with the fact that this level can be found also in non-irradiated samples. The ratio of the introduction rates between proton and electron irradiated samples is about 30, much higher than the value 10 found for the pairs of levels matching up together in the Arrhenius analysis. From this it turns out that S5 and S5* could be associated to different defects or to the same complex defect presenting very different generation kinetics depending on the impinging particle type.

7.2.4. Evolution of I-V characteristics

The defects expectedly introduced in the material by high-fluence irradiation affect the charge transport properties. The high radiation hardness of 4H- SiC manifests itself as an inertness of the charge transport characteristics of the material with respect to the irradiation dose. In Fig. [7.5] a) the I-V characteristics of the proton irradiated diodes do not show any significant difference between the as-grown diode and the diode irradiated with the highest dose. The Richardson's plot of the temperature-corrected diode saturation current J_0/T^2 vs. inverse temperature $1/T$, shown in part b) of the same figure, demonstrates that the activation energy for the charge transport keeps nearly constant within the range 0.80 – 0.96 eV. The same inertness has been found in the electron irradiated samples, whereby for fluences up to 9.5 x 10^{15} cm^{-2} no increase of the leakage current was found for reverse bias in the range 0 – 200 V.



a) b)

Fig. [7.5]. a) I-V characteristics of samples at different irradiation fluences and temperatures. The lower set of curves refers to measurement performed at 300 K, the upper set at 700 K. b) Richardson's plot and its linear fit for the proton irradiated diodes. Squares and continuous line: diode A1; circles and dashed line: diode A2; upwards triangles and dotted line: diode A3; downwards triangles and dotted-dashed line: diode A4.

In Fig. [7.6] the CCE characteristics of electron-irradiated diodes, tested by irradiating the Schottky contact with 4.14 MeV α-particles, are reported. Due to the introduction of defects, with bias $V_R \leq$ 160 V the CCE decreases with increasing radiation dose, with data showing an initial fast decrease

after irradiation at a fluence of 4.75x10^{13} cm^{-2}, followed by a slower fall rate at higher fluences. In the diode depletion region corresponding to this bias field the charge generated by the α-particles is collected by diffusion. Therefore, this result indicates that the radiation-induced defects have a significant influence on the charge diffusion length. On the contrary, the CCE saturates at 100% independently of the radiation dose at reverse bias ≥ 160 V, i.e. above the reverse bias threshold corresponding to a depleted region equal to the α-particle range, R ~ 12μm. Trapping centres, therefore, turn out to be much less effective when the charge collection occurs almost exclusively in drift regime.

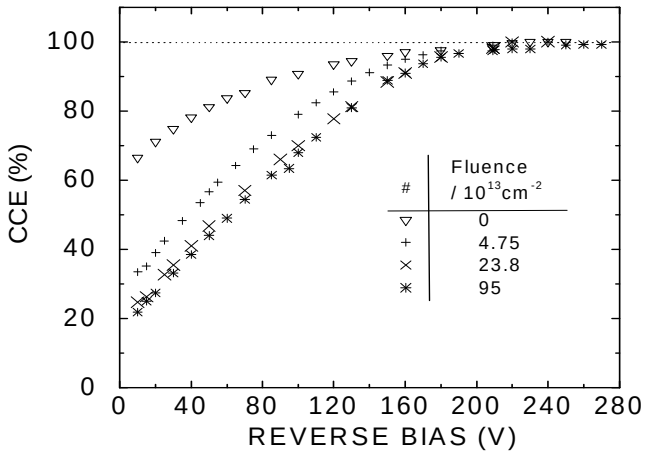


Fig. [7.6] Evolution of CCE characteristics of electron irradiated diodes (measurement performed by F. Nava, University of Modena).

7.2.5. Evolution of C-V characteristics

The compensation of free charge carriers due to the presence in the gap of defect-related deep levels has been monitored in both cases of proton and electron irradiation by means of C-V characteristics. Significant compensation effects were observed mainly in the samples irradiated with the highest doses, i.e. 3.2 x 10^{13} cm^{-2} for proton irradiation and 9.5x 10^{14} cm^{-2} for electron irradiation. In the proton irradiated diodes, C-V measurements were carried out in the temperature interval 300 K – 700 K. Both in as-prepared samples and in samples irradiated with fluences up to 10^{12} cm^{-2}, the free charge carrier concentration remained constant at a value close to n_{CV}=4.5 x 10^{15} cm^{-3} throughout the whole temperature interval. On the contrary, in the sample irradiated with a 3.2 x 10^{13} cm^{-2} proton dose the free charge carrier concentration varied from n_{CV} =3.3 x 10^{15} cm^{-3} at T=300 K to n_{CV} =4.7 x 10^{15} cm^{-3} at T=700 K. The original free charge density is restored at T ~ 700 K even in highly compensated samples. This result disagrees with the compensation model [Åberg01], according to which free charge compensation should be mainly due to nitrogen passivation by defect complexes induced by irradiation. The N atoms are therefore subtracted from their role of active shallow doping centres. If such defects were reasonably stable, they would not dissociate for a relatively low temperature such as T=700K, and there should not be the possibility of a recovery of free charge density to the same value observed before irradiation.

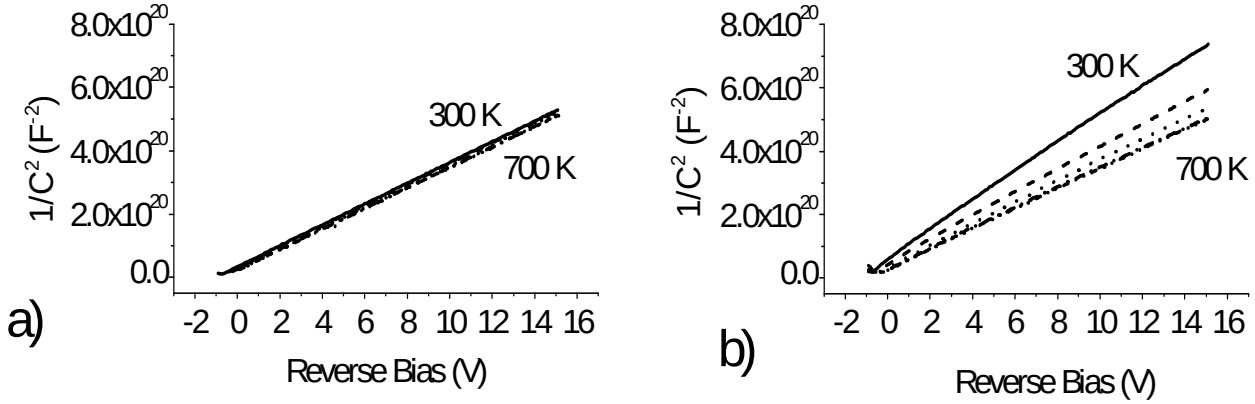


Fig. [7.7]. C^{-2}-V plots of: a) as-prepared sample; b) diode irradiated with 3.2 x 10^{13} cm^{-2} 6.5 MeV protons, showing free charge carrier compensation. The characteristics measured from 300 K to 700 K are shown, with 100 K step.

7.3. Low-temperature annealing of irradiation-induced deep levels

7.3.1. Annealing out of level S2 – DLTS analysis

The annealing out of level S2, together with other minor annealing phenomena, take place over two annealing ranges, as shown in Fig. [7.8]. The first DLTS spectrum was obtained from 80 K up to 360 K (not shown in the picture). Then, the diode was frozen again to 80 K. The second DLTS spectrum (a) was obtained in the temperature range 80 K to 400 K . The first and the second spectrum coincided in the overlapping temperature interval. Spectrum (b), obtained from 400 to 80 K, yields considerably different features: S2A, the dominant peak at T=180 K in spectrum (a), is no longer evident, while the amplitude of S3 increased. Spectrum (c), obtained raising the temperature from 80 K to 470 K, has expectedly the same features of spectrum (b) in the overlapping temperature interval. Spectrum (d), taken from 470 K to 80 K, shows that some further change occurred: the maximum of peak S3 is shifted towards higher temperature, and its amplitude decreased to values similar to those of (a).

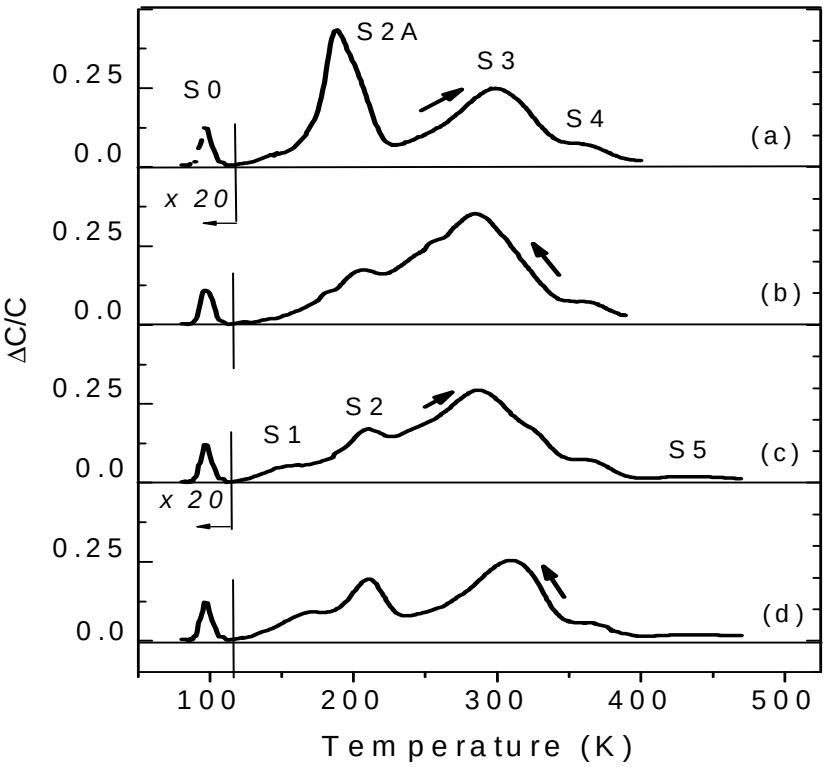


Figure [7.8] DLTS spectra of the 9.5x 10^{14} cm^{-2} irradiated sample. Spectrum (a) from 80 K to 400 K; Spectrum (b) from 400 K to 80 K; Spectrum (c) from 80 K to 470 K; Spectrum (d) from 470 K to 80 K. Emission rate e_n=116.3 s^{-1}. The spectra are shifted in order to display each of them clearly.

Thus, we can identify two separate annealing stages: the former, in which the temperature rises from 360 to 400 K in about 20', is characterized by

a) the dramatic reduction of S2A amplitude.

b) the increase of the amplitude of the peak S3, as well as its lower-temperature shift.

c) the appearance of peaks S1 and S2, which were likely to be previously embedded in the dominant peak S2A.

The latter annealing stage takes place in the temperature interval 400 K – 470 K (in which the sample was kept for about 30'), and is characterized by the amplitude decrease of peak S3, its shape sharpening and its shift towards higher temperature. It is to be noted that after the latter annealing the parameters of the DLTS peak S3 are those of the level known from the literature as Z_1/Z_2 [Dalibor97]. which has been already described in section (7.2.3).

The most remarkable features of the low-temperature annealing are the dramatic decrease of the amplitude of level S2 after the first annealing stage at 360 K- 400 K and the rearrangement of amplitude and enthalpy of peak S3 throughout the two annealing stages in the temperature interval 360 K – 470 K. Correspondingly, the net free carrier density, as shown in the next section, increases after annealing.

A possible explanation of the above results is that the defects associated to level S2 annihilate between 360 K and 400 K. Annealing in this temperature interval has been reported on electron irradiated 3C-SiC investigated by electron spin resonance (stage at T=323 K) [Itoh89] and on Al-implanted 4H SiC by Rutherford backscattering spectroscopy (annealing interval from T=250 K to T=420 K) [Zhang02]. Furthermore, recent theoretical models and numerical simulations [Gao03] predict the recombination of carbon interstitials at T~400 K.

Annealing processes involving peak S3 (Z_1/Z) similar to the one observed in this work have been also reported in other works in slightly different experimental conditions. Doyle et al. [Doyle98]

reported on the changes occurring in the DLTS spectra after a room temperature annealing nine months long in electron irradiated 4H-SiC. The levels at activation energies E_C-E_T equal to 0.32, 0.62 and 0.68 eV show a transition to the Z_1/Z. Nielsen et al. [Nielsen03], [Nielsen05], reported on a bi-stable defect. A correlation might exist between these results at T~400 K and our results. There are, however, substantial differences regarding peak S2:

a) it presents a much higher amplitude change;
b) the decrease of peak S2 and the increase of peak S3 are not related by a 1:1 ratio as according to Nielsen;
c) S3 decreases after raising the temperature up to 470K, significantly lower than 600 K.

7.3.2. Annealing out of level S2 – compensation analysis

The in-depth profile of the free carrier concentration from C-V measurements is illustrated in Fig. [7.9]. The information given by the C-V measurements accounts for the compensation effects due to irradiation and the recovery of free charge carriers occurring upon annealing.

The free carrier concentration *n* decreases by almost a factor 10 after irradiation with Φ=9.5x 10^{14} cm^{-2}, from about 2 x 10^{15} cm^{-3} to 2.5 x 10^{14} cm^{-3}, at the distance x=2 µm from the Schottky junction. After both annealing stages the free carrier concentration increases to ~5 x 10^{14} cm^{-3} at the same depth. The initial decrease of free carriers is a consequence of the introduction of acceptor levels in the semiconductor gap, whereas the subsequent free carrier increase could be related with the disappearance of peak S2A.

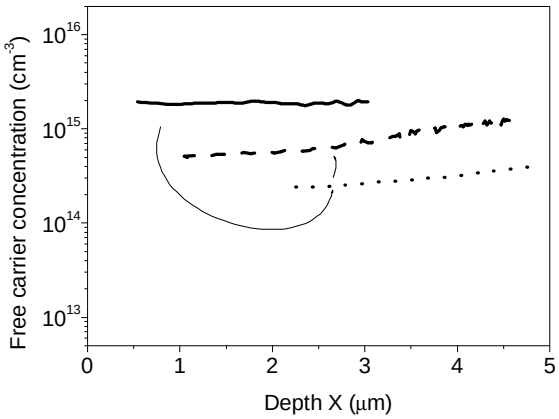


Figure [7.9] Profiles of free carrier concentration versus junction depth from CV characterization. Continuous line refers to the as-prepared diode; dotted line to the diode irradiated with Φ=9.5x 10^{14} cm^{-2}, before the DLTS run up to 400K; dashed line to the same diode, after the DLTS run up to 400K.

To relate C-V and DLTS results Shockley diagrams (Fig. [7.10]) were calculated. The doping concentration N_d=2 x 10^{15} cm^{-3} is set from the free charge concentration measured in the as-prepared diode. The deep level densities and enthalpies were deduced by the DLTS measurements (Table 7.3).

In the irradiated diode the concentration of free charge decreases to n_{free}=(1.2± 0.1) x 10^{15} cm^{-3}, significantly higher than the value measured by C-V characterization n_{CV}=2 x 10^{14} cm^{-3}. Therefore, one has to infer that further levels than those detected by DLTS are introduced by irradiation in the bandgap, accounting for the larger free carrier compensation.

Similarly, in the annealed diode, the free carrier concentration calculated from the Shockley diagrams is n_{free}=(1.5±0.15) x 10^{15} cm^{-3}, while the value from the C-V measurement is n_{CV}=5 x 10^{14}

cm^{-3}. However, the net free charge density increase upon annealing is well reproduced by the Shockley diagram: the difference between the concentration in the irradiated diode and in the annealed diode is roughly equal to 3×10^{14} cm^{-3}, which is also equal to the difference measured by C-V characterization.

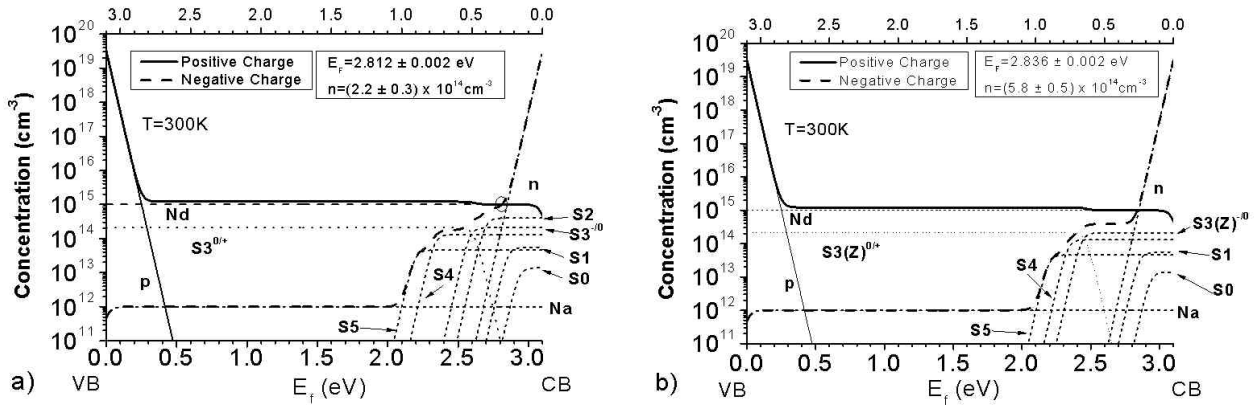


Fig. [7.10]. Shockley diagrams of the annealed diode at T=300 K based on the levels detected by DLTS: a) before and b) after the annealing stage at [360 - 400 K].

7.3.3. Modifications of levels S1 and S5

In Fig. [7.11], the DLTS spectra of samples after annealing at progressively higher temperature are shown. Their features, together with those of the diffusion length, support the hypothesis of a low temperature recovery of irradiation damage. As shown in the inset of Fig. [7.11], the DLTS spectrum is almost flat before irradiation. Only levels S0 (with enthalpy E_C-E_T =0.15eV) and S5 (E_C-E_T =0.89eV), related to a Ti impurity and to a V_C+V_{Si} complex, respectively [LeDonne04], [Dalibor97], [Lebedev00], are here present. After the exposure to the electron beam we observed additional levels, labelled from S1 to S4, with enthalpies E_C-E_T =0.23 eV, E_C-E_T =0.39 eV, E_C-E_T =0.51 eV and E_C-E_T =0.75 eV, respectively, the concentration of which linearly increases with the irradiation fluence. The spectrum of the as-irradiated sample coincides with that of the sample annealed at T=100°C (solid line in fig. [7.11]). After annealing step at T=200°C due to the heating during the DLTS run, the spectrum exhibits the significant variations already described in the previous section.

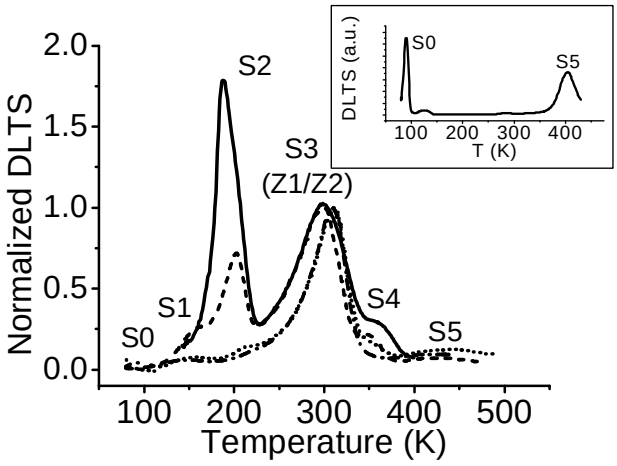


Figure [7.11] DLTS spectra of the irradiated samples after each annealing step, normalized to the height of peak S3. Solid line: annealing at T_{ann}=100°C; dashed line: after annealing at T_{ann} =200°C after DLTS run; dotted line: annealing at T_{ann} =300°C; dashed-dotted line: annealing at T_{ann} =450°C. In the inset, the spectrum of the virgin sample. The emission rate is e_n=46.5 s^{-1}.

After the thermal treatment at T=300°C (dotted line), additional modifications to the spectra occur. The amplitude of peaks S2 and S4 further decreases, whereas peak S3 (Z_1/Z_2) sharpens because of the annealing of the close deep levels previously cited.

The spectrum changes again after the last annealing step, at T=450°C. Peaks S2 and S4 are no longer detected whilst annealing effects are visible also at peaks S1 and S5. Two peaks, S1A (E_C-E_T =0.31eV) and S1B (E_C-E_T =0.22eV) become visible instead of S1 (E_C-E_T =0.27eV) (Fig. [7.12] a), and S5A (E_C-E_T =0.70eV) and S5B (E_C-E_T =1.04eV) appear instead of the broad peak S5 (E_C-E_T =0.96eV) (Fig. [7.12] b). These changes could be justified as follows, dealing only with peaks S1A and S1B as the interpretation for levels S5A and S5B would be the same. Up to annealing temperatures as high as 300°C, the deep level related to peak S1B has a concentration higher than the deep level related to peak S1A. The broad peak S1 results from the emission by both levels. Upon annealing at T=450°C, the concentration of the deep level related to S1B becomes comparable to that related to peak S1A, thus allowing for the single detection of both levels.

Deep level signatures are reported in the Arrhenius plot (Fig. [7.13]). Figure [7.13] not only evidences that levels S0 and S3 maintain their signature after both annealing steps at T=300ºC and at T=450ºC, but also that S1 coincides with S1B and S5 with S5B.

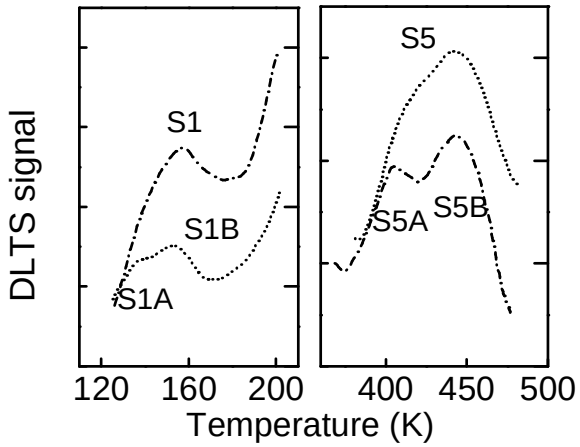


Figure [7.12] Comparison of DLTS spectra after annealing at 300°C and at 450°C in the temperature regions where S1 and S5 are detected. Dotted line: sample annealed at T=300°C; dash-dotted line: sample annealed at T=450°C. The emission rate is e_n=46.5s^{-1}.

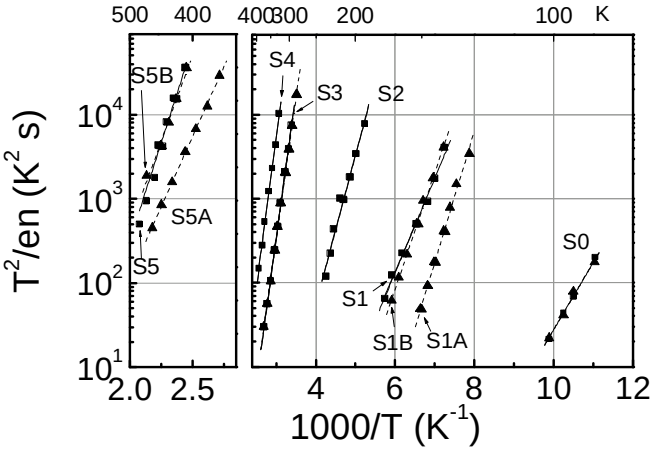


Figure [7.13] Arrhenius plot of samples annealed at T=300°C (squares, solid lines for fits) and at T=450°C (triangles, dashed lines for fits).

7.4. Conclusions

Irradiation with either protons or electrons, at the energies and fluences examined in this work, introduces in the bandgap of 4H-SiC several deep levels, producing free charge carrier compensation, which is dramatic for the highest fluences used. Nevertheless, the material confirms to be highly radiation-hard, as its transport characteristics give evidence: irradiation-induced defects tend indeed to limit the charge transport in diffusion regime, but do not have an equally significant influence on the transport in drift regime. DLTS shows that a few levels (S0 and S5) are already present in detectable concentrations (~10^{13} cm^{-3}) in the as-prepared diodes. By irradiation another set of levels is either introduced or enhanced in concentration (S1 to S5), growing with the irradiation dose. The only level maintaining a constant concentration is S0, likely related to impurity Ti atoms rather than to primary defects. Most deep levels generated by both particles are the same, as deduced by Arrhenius and introduction rate analysis. Generally speaking, protons tend to introduce a density of defects about one order of magnitude (or more) higher than electrons of similar energy, as in agreement with experimental and theoretical results comparing the non-ionising energy loss (NIEL) of either particle type [Lee03] in SiC. On the one hand, one may expect this change in the introduction rates, considering the different particle masses; on the other hand, it is not obvious that such a mass and charge difference yields the same deep levels. Thus, the introduced levels must be elementary intrinsic defects or complexes involving single N or H atoms. Other levels measured for different particle irradiation such as S1 compared with S1* (E_C-E_T ~0.2 eV) and S5 compared with S5* (E_C-E_T ~0.9-1.1 eV) are unlikely ascribable to the same kind of defects, unless to consider an alternative generation kinetics depending upon the irradiation source.

A low temperature annealing effect has been observed for defect-associated energy levels in 8.6 MeV electron irradiated 4H-SiC Schottky diodes. This effect occurs by several annealing stages: in the first annealing stage (360 K – 400 K) the dominant peak S2 dramatically decreases, whereas peak S3, significantly broadened because of a superposition of neighboring levels, shows an amplitude increase; in the second stage (400 K – 470 K) peak S3 exhibits an amplitude decrease, a shape sharpening and a shift towards higher temperature.

The annealing has been interpreted according to the following model: it is known that at T~400 K carbon interstitials recombine [Gao03]. Peak S2 might therefore be related to this defect. This

transformation goes along with configurational rearrangements of other levels related to DLTS peaks overlapping to S3. In the second stage such levels, which are not stable, disappear from the spectrum. Consequently, peak S3 acquires the parameters typical of level Z_1/Z_2.

The annealing out of defect related deep levels expectedly affects also the C-V characteristics, which, successively to a significant irradiation-induced compensation for fluences of ~10^{15} cm^{-2}, evidences a net free carrier increase which correlates well with the amplitude reduction of S2. Despite the free compensation the material undergoes, the device performance as a particle detector is preserved.

The analysis of further annealing stages at slightly higher temperatures led us to the finding that peaks S1 and S5 could be due to the emission of more single levels, which split after annealing at T=450ºC. We observed a coincidence between the amplitude reduction of peaks S1/S1B and S5/S5B. However, this observation was up to now performed over a single annealing step, being insufficient to state whether S1 and S5 are related with each other in a fashion similar to levels S2 and S4.

REFERENCES

[Åberg01] D. Åberg, A. Hallén, P. Pellegrino, and B.G. Svensson: *Appl. Surf. Sci.* **184**, p.263 (2001).

[Achtziger97] N. Achtziger and W. Witthuhn: *Appl. Phys. Lett.* **71**, (1997), p. 110

[Alfieri03] G. Alfieri, E.V. Monakhov and B.G. Svensson, *Proceedings of "International Conference on Silicon Carbide and Related Materials 2003*, Lyon (2003)

[Capano97] M. Capano and R.J. Trew: *MRS Bulletin* vol. **22** No.3, p.19 (1997)

[Castaldini04] A. Castaldini, A. Cavallini, L. Rigutti and F. Nava: *Appl. Phys. Lett.* **85** (2004), p. 3780

[Chantre87] A. Chantre, L. C. Kimerling, *Appl. Phys. Lett.* **48**, 513 (1987)

[Chen02] X.D. Chen, S. Fung, C.C. Ling, C.D. Beling and M. Gong, *J. Appl. Phys.* **94,** 3004 (2002)

[Dalibor97] T. Dalibor, G. Pensl, H. Matsunami, T. Kimoto, W.J. Choyke, A. Schöner, N. Nordell: *Phis. Stat. Sol. (a)* **162**, p. 199 (1997)

[David02] M.L. David, G. Alfieri, E.V. Monakhov, A. Hallén, J.F. Barbot and B.G. Svensson, *Proceedings of "Silicon Carbide and Related Materials 2002"*, Linköping (2002)

[Doyle98] J.P.Doyle M. K. Linnarsson, P. Pellegrino, N. Keskitalo, B. G. Svensson, A. Schöner, N. Nordell, J. L. Lindström: *J. Appl. Phys.* **84**, 1354 (1998) .

[Eberlein03] T. A. G. Eberlein, R. Jones, P. R. Briddon: *Phys. Rev. Lett*. **90**, p. 225502 (2003).

[Fang01] Z-Q. Fang, D.C. Look, A. Saxler, W.C. Mitchell: *Physica B* **308-310** 706-709 (2001)

[Gao02] F. Gao, W.J.Weber and R. Devanathan, *Nucl. Instr. Meth. B* **191**, 487-496 (2002)

[Gao03] Gao F. and Weber W. J.: J. Appl. Phys. **94**, p. 4348 (2003)

[Hemmingsson97] C.G. Hemmingsson, N.T. Son, O. Kordina, J. P. Bergman, E. Janzen, J. L. Lindstrom, S. Savage and N. Nordell: J. Appl. Phys. **81**, p.6155 (1997).

[Hemmingsson98] C. G. Hemmingsson, N. T. Son, A. Ellison, J. Zhang and E. Janzén: Phys. Rev. B **58,** p. R10119 (1998).

[Hemmingsson02] C.G. Hemmingsson, N.T. Son, O. Kordina and E. Janzén, *J. Appl. Phys.* **91** 1324 (2002)

[Itoh89] H. Itoh, N. Hayakawa, I. Nashiyama, E. Sakuma, *J. Appl. Phys.* **66**, 4529 (1989).

[Kawasuso02] A. Kawasuso, M. Weidner, F. Redmann, T. Frank, P. Sperr, R. Krause-Rehberg, W. Triftshäuser and G. Pensl: *Physica B* **308-310**, p. 633 (2002)

[Martin04] D.M. Martin, H. Kortegaard Nielsen, P. Léveque, A. Hallén, G. Alfieri, B.G. Svensson: *Appl. Phys. Lett.* **84,** 10, p.1704 (2004).

[Lebedev99] A.A. Lebedev: *Semiconductors*, **33,** p. 107(1999).

[Lebedev00] A.A. Lebedev, A.I. Veinger, D.V. Davidiv, V.V. Kozlovski, N.S. Savkina, A.M. Strel'chuk: *J. Appl. Phys*. **88**, 11 p.6265 (2000).

[LeDonne04] A. Le Donne, S. Binetti, M. Acciarri, S. Pizzini: *Diamond and Rel. Mat.* **13,** 414 (2004).

[Lee03] K.K. Lee, T. Oshima, A. Saint, T. Kamiya, D.N. Jamieson, H. Itoh: *Nucl. Instr. Meth. B* **210** (2003), p. 489.

[Nava03] F. Nava, E. Vittone, P. Vanni, P.G. Fuochi, C. Lanzieri: *Nucl. Instr. Meth. A* 505, 645 (2003).

[Nielsen03] H. Kortegaard Nielsen, D. M. Martin, P. Lévêque, A. Hallén, B. G. Svensson, *Physica B* **340-342**, 743-747 (2003)

[Nielsen05] H. Kortegaard Nielsen, A. Hallén and B.G. Svensson, *Phys. Rev. B* **72**, 85208 (2005)

[OrtonBlood] P.Blood and J.W. Orton, *The electrical characterization of semiconductors: majority carries and electron states* (Academic Press, London, 1990)

[Pintilie02] I. Pintilie, L. Pintilie, K. Irmscher, and B. Thomas: Appl. Phys. Lett. **81**, p. 4841 (2002)

[Posselt01] M. Posselt, V. Belko and E. Chagarov, *Nucl. Instr. Meth. B* **180**, 17-22 (2001)

[Shur98] M. Shur: EuroConference on Advanced Heterostructure Devices for Micro- and Optoelectronics III, Villa Gualino, TO, Italy, 1998, 22-27 June

[Storasta04] L. Storasta, A. Henry, J.P. Bergman, E. Janzén: *Mater. Sci. For.* **457-460**, , p. 469 (2004)

[Weber03] E. R. Weber: *Physica B* **340-342**, p. 1-14 (2003)

[Zhang02] Y. Zhang, W.J. Weber, W. Jiang, A. Hallén and G. Possnert: *J. Appl. Phys***. 91**, p. 6388 (2002)

APPENDIX

Biographical sketch of the author

Lorenzo Rigutti was born in Porretta Terme (Bo), Italy, on 11 June 1978. He attended the primary school in Udine, Zola Predosa (Bo), and Crespellano (Bo). From 1992 to 1997 he attended the Liceo Classico M. Minghetti in Bologna, reporting the final score 60/60. From 1997 to 2002 he studied Physics at the University of Bologna, Italy, spending one year at the University of Mainz, Germany, thanks to the Erasmus-Socrates program funded by the European Union. He earned his M.S. degree in Semiconductor Physics with the thesis "*Silicon Carbide: a material for harsh environment application*", reporting the score 110/110 cum laude. In 2003 he spent 8 months as an exchange student at the University of Massachusetts Lowell and began his Ph.D. program in the semiconductors group at the Physics department of the University of Bologna led by Prof. Anna Cavallini. His research focused on the study of defects in wide bandgap materials.
He is married to Agnieszka Pocztowska.
He speaks fluently Italian, English, German and Polish.

Lorenzo Rigutti – Scientific publications and conferences

International Journals

2004

A.Cavallini, A.Castaldini, L.Rigutti, F.Nava, P.Vanni, P.G.Fuochi, *Recovery effect of electron irradiated 4H-SiC Schottky diodes*, Proceedings of MRS Fall 2003 Meeting, R 10.6.

A.Cavallini, A.Castaldini, L.Rigutti, F.Nava, *Low temperature annealing of electron irradiation induced defects in 4H-SiC*, Applied Physics Letters 85, 17(2004) p.3780-3782

2005

G. Bertuccio, S. Binetti, S. Caccia, R. Casiraghi, A. Castaldini, A. Cavallini, L.Rigutti, C. Lanzieri, A. Le Donne, F. Nava, S. Pizzini and E. Vittone, *Silicon Carbide for alpha, beta, ion and soft X-ray high performance detectors*, Materials Science Forum 483-485 (2005) p.1015 .

A.Castaldini, A. Cavallini, L.Rigutti, F.Nava, *Electronic levels introduced by irradiation in silicon carbide*, Materials Science Forum 483-485 (2005) p. 359.

A. Cavallini, A.Castaldini, L. Rigutti, F. Nava, C.F. Pirri, S. Ferrero, *Deep Levels by proton- and electron-irradiation in 4H-SiC*, Journal of Applied Physics 98 (2005) p. 53706

A. Castaldini, A. Cavallini, L.Rigutti, M.Meneghini, S. Levada, G. Meneghesso, E. Zanoni, V. Harle, T. Zahner, U. Zehnder, *Short term instabilities of InGaN/GaN light-emitting diodes by capacitance-voltage characteristics and junction spectroscopy*, Physica Status Solidi (c) 2 (2005), p. 2862.

2006

A Castaldini, A. Cavallini, L. Rigutti, S.Pizzini, A. Le Donne,S. Binetti, *Diffusion length and junction spectroscopy analysis of low temperature annealing of electron irradiation-induced deep levels in 4H-SiC*, J. Appl. Phys. (2006) In press

F. Rossi, M. Pavesi, M. Meneghini, G. Salviati, M. Manfredi, G. Meneghesso, E. Zanoni, A. Castaldini, A. Cavallini, L. Rigutti, U. Strass and U. Zehnder, *Influence of short-term low-current DC-aging on the electrical and optical properties of InGaN blue Light-Emitting Diodes,* J. Appl. Phys. (2006) in press.

A. Castaldini, A. Cavallini, L.Rigutti, M.Meneghini, S. Levada, G. Meneghesso, E. Zanoni, V. Harle, T. Zahner, U. Zehnder, *Role of deep levels in DC current aging of GaN/InGaN LEDs studied by capacitance and photocurrent spectroscopy*, Proceedings of MRS Fall 2005 Meeting, in press

A. Castaldini, A. Cavallini and L. Rigutti*, Assessment of the intrinsic nature of defect Z_1/Z_2 by compensation effects in proton-irradiated 4H-SIC*, Semicond. Sci. Tech. (2006) in press

Conferences and Meetings

- MRS Fall 2003 Meeting, 30/11-4/12/2003, Boston, USA

- ECSCRM 2004 European Conference on Silicon Carbide and Related materials, 31/8-4/9/2004, Bologna.

- MRS Fall 2005 Meeting, 28/11-3/12/2005, Boston, USA

Schools

- National school of materials science and nanotribology, Sestri Levante, September 2004.

- 3rd Optoelectronics and Photonics Winter School "Optical Interconnects", Sardagna (TN), February 2005.

Acknowledgements

As far as I know, Ph.D. is a fairly high degree, and if I think about it, it took some years education to get to this goal. Therefore, I think it would be nice to thank here some of the teachers who helped me make it possible. Being chronologically rigorous, my first teachers were two primary school teachers who took care of some specific aspects of my education when I still was in the kindergarden: my grandmother Pia and my grandfather Pino. Then, representing my primary school teachers, Suor Dolores, an absolutely skilled nun. For the Scuola Media, I would like to especially thank Gabriella Strazziari and Amalia Accettella. Among my Liceo Minghetti teachers, special thanks to Fabio Fiorini, who taught me Ancient Greek and Latin, two only apparently dead languages, and Pietro Biancardi, my philosophy's teacher, who, after scaring the class with a three-months long reading of Plato's Symposium, taught us how to analyze texts in depth. Quinto Fattori, my math's teacher, was already thanked in my M.S. thesis. However, his role was too important, so I thank him again here. I also had very good teachers in my University years: among these, I would like to mention here Cesare Parenti. Coming to the Ph.D. years, I would like to thank the Italian Republic, who supported my program, my supervisor, Anna Cavallini, who gave me the chance to work in her group, and the other components of the group: Antonio Castaldini, Beatrice Fraboni, Daniela Cavalcoli and Laura Polenta, as well as my Ph. D. colleagues Marco Rossi, Marica Canino and Filippo Fabbri "Il Laureando". They all provided advice, support, and contributed to the good working environment of this group. Finally, I would like to thank my wife Agusia – who also was one of my teachers, my Polish teacher indeed– , but the reasons I thank her for are much more important and complex, and I wouldn't find adequate words to do it here.