Point Defects in Ion-Implanted Silicon and Silicon Carbide

Paolo Pellegrino

KTH, Royal Institute of Technology
Department of Microelectronics and Information Technology
Stockholm 2001
Table of contents

Abstract 3

List of publications 5

I Introduction 7

II Experimental techniques 10
  2.1 Trapping of electrical carriers in semiconductors
  2.2 Principles of DLTS technique
  2.3 Limitations / range of applicability of the technique
  2.4 Additional features of the DLTS technique

III Principles of ion implantation 28
  3.1 Stopping of energetic particles
  3.2 Ion cascades and defect generation
  3.3 Evolution of the damage in the low dose regime
  3.4 Role of the impurities
  3.5 Applications of defect engineering
  3.6 Future trends and open fields

IV Summary of the results 46

V Appendix 52

VII Acknowledgements 58

VI References 59
ABSTRACT

Ion implantation has been established as a key-process for the introduction of dopants in semiconductor materials. One of the critical secondary effects of such a process is damaging of the crystalline structure caused by collisions with incoming ions, giving rise to atomic displacements. The damage consists to a large extent of point-defects. Point-defects mediate dopant diffusion in semiconductors, at temperatures where thermal diffusion at equilibrium is negligible, resulting in an undesired redistribution of the implanted dopants. They also modify the electrical properties of the crystal, by introducing energy levels in the forbidden band-gap. These levels can act as additional acceptors/donors, charge-carrier traps, recombination centers or/and doping compensating centers, depending on their energy position, capture-rate of the charge carriers and temperature and doping in the substrate. The strong implications of such effects on the performance of modern electronic devices encourage the development of so-called defect-engineering.

The behaviour of semiconductor materials after implantation with different particles is the focus of intensive research. In this thesis the evolution of point-defects during implantation and subsequent annealing of Si and SiC samples has been studied by employing Deep Level Transient Spectroscopy (DLTS). DLTS allows an accurate detection of each deep level introduced in the bandgap. Its application is limited to the low dose regime, where the defect concentration is few percent of the free carrier density.

The increase of prominent vacancy-related point-defects, such as the divacancy (V₂) and vacancy-oxygen (VO) centers, after implantation and thermal treatment at temperatures ≤150 ºC, has been observed in the case of heavy ions (atomic mass ≥72 amu). This kind of ions has been found to be quite effective in generating defect-clusters, which then dissolve during low-temperature annealing.

At higher temperatures the V₂ centers get mobile, and they migrate to traps, with an activation energy of 1.3 eV for the annealing process. The concentration of annihilation centers and the effective diffusion length have been estimated to be at least 5x10¹⁵ cm⁻³ and below 0.1 µm,
respectively. This holds irrespective of the float-zone or Czochralski-
grown silicon wafers.

In contrast to that of V$_2$, the stability of VO is strongly affected by
migrating hydrogen atoms, a common residual impurity in silicon. New
complexes, such as VOH, with a deep level at 0.32 eV below the
conduction band edge, and VOH$_2$ centers are formed by diffusion of
hydrogen from the surface. This process is dominant in the first annealing
stage of VO, due to the high diffusivity of hydrogen. Up to 70% of the
original VO concentration is hydrogenated. A second contribution to the
loss of VO centers is migration of VO, with an estimated activation
energy of 1.8 eV, and formation of vacancy-dioxygen pairs. This process
dominates once the hydrogen supply has been exhausted.

A new experimental concept of studying shifts between vacancy and
interstitial depth-profiles in ion-implanted silicon samples is presented.
The concept is based on DLTS measurements using the filling pulse
 technique, where the two profiles are represented by that of VO and the
interstitial carbon-substitutional carbon pair, respectively. For MeV boron
ions a shift of ~0.5 µm towards larger depths is revealed for the interstitial
profile.

The generation rate of point-defects has been explored in the case of
very low dose of implanted protons, and for different fluxes (dose-rates).
A reduction in production of vacancy-related defects with increasing flux
has been observed at various implantation temperatures, between 70 and
300 K. The imbalance between interstitial and vacancy diffusivity is
responsible of such an "inverse dose-rate effect". The fast diffusing
interstitial, with an estimated activation energy of 0.065 eV, recombines
with vacancy-related secondary defects in adjacent collision cascades. A
dual implant has been used to verify the trap-limited long-range migration
of self-interstitials.

Implantation of silicon carbide n-type epitaxial layers results in a
dramatic compensation of the shallow nitrogen donors. The defects
introduce deep compensating acceptor levels in the band-gap, which
evolve to a stable configuration with a "final" level at $E_C$-0.70 eV.
However, deactivation of the nitrogen donors by interaction with mobile
point-defects appears as the main cause of the carrier compensation.
Without giving rise to deep levels in the bandgap, these nitrogen-defect
complexes show a quite remarkable thermal stability.
List of publications included in the thesis:

I "Reverse Annealing Effect in Heavy Ion Implanted Silicon"
P. Pellegrino, N. Keskitalo, A. Hallén, and B.G. Svensson

II "Impurity-assisted annealing of point defect complexes in ion implanted silicon"
P. Pellegrino, A. Kutnezov, and B.G. Svensson

III "Annealing kinetics of vacancy-related defects in low-dose MeV self-ion-implanted n-type silicon"
P. Pellegrino, P. Lévêque, J. Lalita, A. Hallén, C. Jagadish and B.G. Svensson
submitted to Phys. Rev. B.

IV "Separation of vacancy and interstitial depth profiles in ion-implanted silicon; experimental observation"
P. Pellegrino, P. Lévêque, J. Wong-Leung, C. Jagadish, and B.G. Svensson

V "Rapid Migration of Defects in Ion Implanted Silicon"
J. Lalita, P. Pellegrino, A. Hallén, and B.G. Svensson

VI "Electrically active point defects in n-type 4H-SiC"
J.P. Doyle, M.K. Linnarsson, P. Pellegrino, N. Keskitalo, and B. G. Svensson

VII "Ion Implantation Induced Defects in Epitaxial 4H SiC"
A. Hallén, A. Henry, P. Pellegrino, B.G. Svensson, and D. Åberg,

VIII "Nitrogen de-activation by implantation induced defects in 4H-SiC epitaxial layers"
D. Åberg, A. Hällen, P. Pellegrino and B.G. Svensson
Publications relevant to this thesis, but not included

- "Anomalous Field Dependence of Deep Level Emission in Proton Irradiated Silicon"
  N. Keskitalo, A. Hallén, P. Pellegrino, B.G. Svensson

- "Hydrogen Related Defects Centers In Float-Zone And Epitaxial N-type Proton Implanted Silicon"
  P. Lévêque, P. Pellegrino, A Hallén, B.G. Svensson, V. Privitera

- "Dose-Rate Influence on the Defect Production in MeV Proton-Implanted Float Zone and Epitaxial N-type Silicon",
  P. Lévêque, A Hallén, P. Pellegrino, B.G. Svensson, V. Privitera,
  In manuscript.
I INTRODUCTION

A look on the progress made in the field of defects in semiconductors is instructive to get some perspective on the subject. It is worth noting the close link between many of the most intriguing problems and the device applications.

The use of silicon and germanium as convenient systems in which to study the nature of radiation effects originated in 1947, as a result of the discovery of the profound sensitivity of their electrical properties to energetic particle bombardment. The technological driving force for studies on radiation effects was the interest in using semiconductor devices near nuclear reactors. Lark-Horovitz gave the first correct interpretation of these effects in terms of the generation of point-defects by the interaction of nuclear particles with the crystal atoms [1].

When an atom is displaced from its site in the silicon crystal, elementary damage, a Frenkel pair (a self-interstitial and a vacancy) is created. In the early 1960's it was recognized, by means of electron paramagnetic resonance, that vacancies in silicon are mobile at room temperature [2].

Since then, not only has the field widely broadened to include new materials, but a variety of radiations has been used as well. However, although the field was increasing in complexity, it also started to gain a certain maturity. In addition there were new technological needs which supplied the driving force and support for more refined studies.

During the 1970s, for instance, a large interest in ion-implantation processes made its appearance. The great possibility to use such a technique for doping of semiconductors was accompanied by the observation of the concomitant, unavoidable damage of the crystalline structure.

In the same decade it was found that particle bombardment produces a large destructive effect on solar cells in the Van Allen belt [3]. Then, it was soon recognized that the nature of the damage introduced by all these phenomena was quite similar to the one observed after nuclear particle irradiation, and a common ground for fundamental studies of defect generation was set. All these new subjects resulted in a large growth of the activity for studying of defects in semiconductors. In the 1960s
detailed spectroscopic studies of vacancies, multi-vacancy clusters, and interstitial impurities were started. This early work provides a solid foundation for the understanding of more recently discovered phenomena such as the presence of voids in as-grown silicon crystals and the transient-enhanced-diffusion of ion-implantated dopants.

Fundamental studies on the interaction of hydrogen with native defects in silicon began in the 1970s. They were followed by the recognition in the 1980s that dopants could be passivated by hydrogen, either intentionally or unintentionally. Unintentional hydrogen passivation turned out to be one source of the difficulty of p-doping in wide bandgap materials [4]. During recent years the combination of experiments and theory has revealed the detailed properties of a number of hydrogen-vacancy, and hydrogen-interstitial complexes. The understanding of these defects has made it possible to propose a mechanism for the "smart-cut" exfoliation process, that provides a novel means to fabricate semiconductor-on insulator structures [5].

In addition to the progress made on classical, still unsolved problems, there is a number of new research topics and new experimental and theoretical capabilities. For example, the vacancy's partner, the native, "illusive" silicon self-interstitial has escaped detection and interstitials remain a subject of active research in a variety of host materials.

As the dimensions of Si-integrated circuits rapidly decreased, with oxide-layer thickness of tens of Ångström, and circuit-design rules of tenths of microns, new defect-related problems like stress-induced leakage in MOSFETs and transient enhanced dopant diffusion demand attention [6]. This latter regards defect-assisted diffusion of dopants during post-implant thermal processing of devices, and it appears as an important limiting factor in scaling devices to dimensions in the deep submicron range.

The re-discovery of silicon-carbide as a promising material for power and high frequency applications has enlarged the need to understand imperfection in multicomponent systems. In this material implantation is ideally suited, as the diffusion of most atomic elements has been found to be extremely low at temperatures below ~1800 °C. However, the damage produced by such implantations turns out to be difficult to anneal out and careful studies of the defect reaction kinetics are required to optimize the conditions for implantation and post-annealing.
A variety of new experimental methods has appeared since 1980 and contributes to the understanding of defects. There are the nuclear methods such as perturbed angular correlation, positron annihilation, and muon spin resonance. The direct imaging of defects and interfaces by high-resolution microscopies continues to advance at a rapid pace. The recent achievement of control of the isotopic content of the host crystal has led to advances in spectroscopy and diffusion studies. Heterostructures are used to probe defect properties, a strategy that is often coupled with the high-resolution imaging methods. The ability to probe defects in device structures, which is a strength of the now classical capacitance spectroscopies, has been extended to highly structure-sensitive methods, with the electrical detection of magnetic resonance being a prominent example [7]. Furthermore, nowadays theoreticians have the ability to calculate electronic structures, total energies, and many experimentally observable properties such as vibrational frequencies, hyperfine parameters, and electronic level positions. The calculations include the effects of host relaxations in large supercells or clusters, and are gradually becoming accurate enough to be used as predictions [8].

In conclusion, the technological potential of semiconductor materials has not been exhausted yet. Thus, there will continue to be new, more demanding applications of semiconductors, which will continue to reveal new defect-problems. At the same time, only through the long-time efforts made by several research groups in complementary studies, it has been possible to find a solid explanation of the observed phenomena and make true progress.

The aim of this thesis is to present some recent experimental studies on point-defects, resulting after ion implantation in silicon and silicon carbide. Main attention has been focussed on the production and thermal stability of the introduced damage, as well as the consequent impact on the electrical properties of the semiconductors.

DLTS, the experimental technique used all through this work, is described in section II. It allow to study point-defects up to very low concentrations, such the ones reached in implantations to low doses. A general overview on the main features of ion implantation and damage production is given in section III, while a summary of results and included papers is presented in section IV.
II. EXPERIMENTAL TECHNIQUES

The purpose of DLTS (Deep Level Transient Spectroscopy) measurements is to study deep levels in the band gap, related to specific defects in the semiconductor, by probing the emission rate of charge-carrier trapped at the defect site. This is achieved by detecting the restoring of the steady-state capacitance of a reverse-biased junction, after that a forward-bias pulse has provided the free carriers to fill the trap centers in the depletion region. It is worth to analyze this process in more details.

2.1 Trapping of electrical carriers in semiconductors

In semiconductors, defects alter the periodic potential of the crystal and can introduce discrete levels (or even narrow sub-bands) in the electronic band-gap. With respect to their position relative to the Fermi level $E_F$ and band edges ($E_C$, $E_V$ for conduction and valence band respectively), they can be classified in deep and shallow levels: the shallow ones are closer than $E_F$ to the edge of the conduction (valence) band in n-(p-) type semiconductors, otherwise they can be considered as deep levels [9]. This classification is somehow relative, being $E_F$ strongly dependent on the temperature and doping concentration of the material. However, in the extrinsic regime all the dopants are ionized and $E_F$ lies at $E_C-kT \ln(N_C/N_D)$ (or $E_V+kT \ln(N_V/N_D)$ in p-type materials), where $N_C$ ($N_V$) is the effective density of states in conduction (valence) band, $N_D$ is the donor density, $k$ the Boltzmann constant and $T$ the absolute temperature. Then the definition of deep and shallow level becomes more precise.

Free charge-carriers (electrons or holes) can be captured and released by these levels according to the Shokley-Read-Hall statistics [10]. In particular, the emission rate (or probability to emit one electron/hole per unit time) can be expressed by the following equation:

$$e_{n,p}(T) = v_{th,n,p} \gamma \sigma_{n,p} N_{C,V} \exp[-(\Delta G)/kT] = v_{th,n,p} \gamma \exp[\Delta S/k] \sigma_{n,p} N_{C,V} \exp[-(\Delta H)/kT] \quad (2.1)$$
where \( v_{\text{th} \ n,p} \) is the electron (hole) average thermal velocity, \( \gamma \) is the degeneracy of the level, \( \sigma \) is the electron (hole) capture-cross section, and \( \Delta G \) is the ionization free energy, which can be expressed as the difference between the enthalpy \( \Delta H \) and entropy \( T \Delta S \) terms. This is one of the key-equation in DLTS studies. It is worth to note that, according to eq. 2.1, a shallow level would have an emission time \( 1/e_n \) (\( 1/e_p \)) extremely short, beyond any instrumental capability. On the contrary, a deep level would behave as a "trapping" center, at least in a large temperature interval, showing an emission-time of the order of some ns-ms, then in an accessible time-scale to be observed.

The electron (hole) capture-rate of one discrete level in the bandgap is strictly related to the corresponding emission rate. Being proportional to the free carrier concentration \( n \) (\( p \)), it can be expressed as:

\[
c_n = n \ \sigma_n \ v_{\text{th},n} ; \quad c_p = p \ \sigma_p \ v_{\text{th},p}
\]

where \( c_n \) (\( c_p \)) is the capture coefficient for electrons (holes). It is worth noting the much stronger temperature-dependence of \( e_n \) (\( e_p \)) than \( c_n \) (\( c_p \)).

### 2.2 Principles of DLTS technique

As already mentioned, DLTS mainly focus its attention on the emission process. By means of eq. (2.1), once that the emission rate of a deep level is known as a function of temperature, one can obtain its electrical parameters, the capture cross section \( \sigma \) and the position in the band gap \( E_T \). These two quantities fully characterize the electrical activity of a deep-level defect and, moreover, can provide additional information on other physical properties of the defect. In particular, computer simulations of the atomic configuration of a defect can estimate its electrical properties, which can be used as a mean of finding the lowest energy defect configuration by a comparison with the measured values of \( E_T \) and \( \sigma \). Other experimental techniques than DLTS are sensitive to \( E_T \), while quite few others, if any, are suitable to study \( \sigma \). The magnitude of \( \sigma \) is indicative of which ionization scheme is observed (single negative-charged state to neutral one, or double charge state to single one, and so
In most of the cases it is possible to establish if the emission is coming from a point-like defect or an extended defect.

In order to promote the emission as an overwhelming process, one fundamental requirement of DLTS technique is the need of a rectifying junction in proximity of the damage region. Then in the investigated volume, completely depleted from free carriers, all the capture processes are suppressed, according to eq. 2.2. Moreover, in order to minimize re-capture phenomena while the traps are discharging, the concentration of levels has to be kept below a few percents of the doping concentration. This puts a strong limit on the applicability of the DLTS as experimental method.

An obvious additional condition is that the deep level of interest is filled with carriers just before the emission is recorded. This is the meaning of the filling voltage pulse, which floods the depletion region with free carriers, and thus populates the trap centers.

Let's consider for instance an n-type semiconductor, doped with donors atoms to a density $N_D$. By means of an electrical junction, a depletion region is build up in the semiconductor, up to a length $W$ which is function of the applied bias $V_r$:

$$W(V_r) = \left[ \frac{2 \varepsilon (V_{bi}+V_r)/qN_D}{2} \right]^{1/2} \quad (2.3)$$

where $\varepsilon$ is the dielectric constant of the semiconductor, $V_{bi}$ is the build-in potential of the junction, $q$ the elementary charge. It is possible to associate a capacitance to the depletion region, same as the simple plane parallel plate capacitor:

$$C_\infty = A \frac{\varepsilon}{W} = A \left[ \frac{\varepsilon q N_D}{2 (V_{bi}+V_r)} \right]^{1/2} \quad (2.4)$$

where $A$ is the junction area. Negligible charge-carrier density is present in this volume, such that a steady state regime is settled: the shallows level are all ionized and a spatial charge $N_D$ is present (the deep levels, with density $N_T$, are considered acceptor-like, if not otherwise noted). During the filling pulse regime of amplitude $V_p$ and duration $\tau_p$, the depletion region is shortened according to eq. 2.3, by replacing $V_r$ with $V_r' = V_r - V_p$. In this regime the capture of available free electrons in the probed region dominates. When the square-like pulse is removed, the
shallow levels follow at once the change in bias, emitting the electrons and building up once more the depletion region. The restoring of the initial conditions is then due to the slower carrier-emitters, which respond with a characteristic time equal to $1/e_n(T)$:

$$N_T \text{ (filled)} = N_T \exp(-e_n(T) t) \quad (2.5)$$

This transient to the steady-state regime is recorded as a variation of the reverse-bias capacitance. Making use of eq. 2.3, 2.4, 2.5, in the approximation $N_T \ll N_D$, and with the condition that at time $t = 0$ all the trap-centers are filled, one easily gets:

$$C(t,T) = C_\infty (1-N_T \exp(-e_n(T) t)/2N_D) = C_\infty - \Delta C \exp[-e_n(T) t] \quad (2.6)$$

Then a proportionality between $\Delta C/C_\infty$ and $N_T/N_D$ holds, and an exponential relaxation to the equilibrium is expected, with a time constant $1/e_n(T)$. In Fig. 1 such a behaviour is shown as a function of temperature, which is due to the temperature-dependence of the emission rate.

![Fig. 1: Capacitance transient recorded at different temperatures](image)
It is apparent how the transient becomes more and more rapid as the temperature increases, until the recovery of the capacitance happens in such a short time that it falls beyond the accessible time-scale. A perfect agreement with an exponential drop is found, while the maximum amplitude of the transient is proportional to the defect concentration in the probed region (relative to the doping concentration).

By fitting the capacitance transient as a function of \( T \), one can derive the function \( e_n(T) \), and then, by means of eq. 2.1, all the information about each deep level in the bandgap of the semiconductor is available. The direct use of the transients requires an overwhelming, surely redundant, amount of data, as to say the capacitance \( C(t,T) \) at each sampled time \( t \) in the transient and for each temperature value \( T \). In addition, it is not an easy way to handle the data. For instance, to make comparisons between different measurements by plots like in Fig. 1 is not very illustrative and difficult to evaluate.

A more useful way to deal with the measurement data is to build a DLTS spectrum. This processing of the data resembles the use of a narrow-bandwith filter: during a temperature scan, every time that the emission rate of one defect matches the resonance frequency of the "DLTS filter", a peak appears in the spectrum. By changing the "resonant frequency", a shift of the DLTS peak position is observed in the plot of DTLS signal \( S \) vs. temperature, because the resonant conditions occur for a different emission rate, and in its turn, at a different temperature. Moreover the height of each peak is proportional to the corresponding defect density. The analogous of the "resonance frequency" in DLTS notation is the inverse of the "time window" \( \tau_m \), which is the duration of the recorded transient.
Lang, who invented the method, got a DLTS spectrum just making the difference of two capacitance values recorded at fixed times during the transient, as shown in Fig. 2 [11]. In the present case a more complicated algorithm has been used, which greatly improves the signal to noise ratio, leaving the basic concepts of the DLTS formalism unchanged. In addition, several spectra, each corresponding to a particular "time window" $\tau_m$, can be collected during the same temperature scan. The level parameters can be easily extracted, by applying the following equations:

$$e_n(\text{at peak maximum}) = F_1 / \tau_m$$  \hspace{1cm} (2.7)

$$S(\text{at the peak maximum}) = F_2 \Delta C / C_\infty$$  \hspace{1cm} (2.8)
The particular DLTS algorithm and the corresponding proportionality factors $F_1$ and $F_2$ used in the present case are discussed in appendix 1. An example is shown in Fig. 3, where for each spectrum the corresponding emission rate at the peak maximum is indicated.

A set of $[T, e_n(T)]$ values is then available, and by reshaping eq. 2.1, an Arrhenius behaviour is found:

$$\ln(e_n / T^2) = \ln(\sigma_\infty Q) - \Delta H / kT$$

(2.9)

where the constant $Q$ takes into account the temperature dependence of $N_{C,V}$ and $v_{th}$. The extracted $\sigma_\infty$ includes the entropy and degeneracy factors, and for this reason it can differ from the real $\sigma$ by about one order of magnitude. Later in this chapter a more appropriate method is outlined, normally used to measure the capture cross section of a deep level. A typical case of Arrhenius extrapolation is shown in Fig. 4. As illustrated in appendix 1, a great resolution in calculating $\Delta H$ and $\sigma_\infty$ can usually be achieved from the Arrhenius plot or, in case of overlapping levels, by fitting routines of the DLTS spectra.
This technique offers at the same time an excellent sensitivity in detecting very low defect concentrations, up to around 1/1000 (or even less) of the doping concentration. For this reason DLTS technique has been established as the main experimental tool for investigating behaviour of defects in the low dose of implantation/irradiation regime.

### 2.3 Limitations / range of applicability of the DLTS technique

DLTS is truly a spectroscopic technique because it is able to distinguish the effect of each deep level on the electrical properties of the semiconductor, namely the junction capacitance. Because of this great feature, it is an invaluable tool for studying the properties of the defects, while most of the others electrical characterization techniques face the limitation that they always see a cumulative effect of all the different defects.

Obviously, DLTS presents some important limitations, which have to be taken into account each time an experiment is planned:
-the requirement of an electrical junction implies processing of the test samples in order to get diodes. High quality of the rectifying characteristics is always recommended. This can be difficult to achieve with some materials, which have very high intrinsic doping concentration, and/or ohmic behaviour with most of common metals. Materials with a doping concentration on the order of $1 \times 10^{16}$ cm$^{-3}$ or below are suitable to be investigated. Moreover, only the defects inside the depletion region are accessible, limiting the spatial sensitivity, especially in the very low doped materials;

- in order to observe the correct behaviour during the transient, the material cannot show macroscopical electrical compensation, or, in other words, the maximum concentration of defects should not exceeds 10% of the doping density. This need is opposite to many experimental techniques, which require a high density of defects to be applicable;

- the only defects which are detectable are the ones which give rise to deep levels. In conventional DLTS majority carriers injection is used. Being usually $n >> p$ ($p >> n$) for levels in the upper (lower) part of the bandgap, then in n- (p-) type materials only the upper (lower) half of the band-gap is accessible. This unavoidable inconvenience implies difficulties in tracking the collective behaviour of the defects, in particular the overall production and mutual interactions of the point defects after irradiation/ implantation;

- the assignment of a deep level to a particular defect is quite cumbersome, and mostly relies on the comparative studies with complementary characterization techniques, such as electronic paramagnetic resonance (EPR), infrared light absorption (IR) and photoluminescence (PL). However, by extensive DLTS investigations of the dependence of a deep level on the irradiation/implantation parameters (energy, dose, dose rate, and temperature), and on the characteristics of the material (doping, impurity content, temperature), it is possible to extract a large amount of informations regarding the probable identity of the defect.
2.4 Additional features of the DLTS technique

- Concentration profile vs. depth

DLTS offers the possibility to monitor the defect concentration vs. the junction depth, in the available depth scale, corresponding to the maximum reverse bias provided by the instrumentation. This is particularly important if the investigation regards the damage produced by ion implantation. In this case the created damage resembles a shape of a damage peak around the ion end-of-range and a tail towards the junction interface. Many properties of the emission rate from deep levels are found to depend on the location of the monitored defects [12, 13]. This is attributed to the induced strain in defects-rich regions, which alters the configuration of the point-defects and reduces their capture efficiency [14].

In addition it becomes feasible to study diffusion phenomena in which the electrically active defects are involved, both after implantation and after thermal treatments. In the first case one obtains informations about the diffusivity of primary defects, interstitials and vacancies, as a function of impurity content in the sample (then the estimated values are always extrinsic diffusivities) [15, 16, 17]. By means of future computer simulations it may be possible to reproduce the observed defect profiles, taking into account not only the collision phase, but also the effect of defect diffusion and recombination.

Another important case, verified by DLTS studies, is the displacement in the peak position of the vacancy-related defects respect to the interstitial ones. Due to a preferential initial momentum in the forward direction, the interstitials which escape recombination are located deeper in than the vacancy-type defects, as shown in Fig. 5 (see Paper IV).
The measurement of a defect concentration vs. depth profile is performed as follows. By choosing a particular rate window and keeping the temperature fixed at the peak maximum of the DLTS signal, the reverse bias is fixed to a constant value corresponding to the deeper edge of the explored region. Then a set of measurements is made, in which the height of the filling pulse $V_p$ is progressively increased. As shown in Fig. 6, this means to collect the emission signal $S$ from a region more and more closer to the junction edge. Due to the non uniformity of the defect distribution, as soon as the main defect peak has been covered by the increasing filling pulse height, a saturation on the collected $S$ is observed.

The correlation between the DLTS signal $S(V_p)$ and the corresponding defect density $N_T(x)$ is obtained differentiating eq. 2.8 respect to $V_p$, while the depth scale $x$ is easily found from eq. 2.3 [15].

Fig. 5: [VO] and [$C_{sC_1}$] vs. depth profile from deconvolution of the level at $E_C\cdot0.17$ eV, compared with primary vacancy profile.
By using a p/n junction, it is possible to inject minority carriers in the depletion region, and then to study the emission of minority carrier from the deep levels. Keeping the junction under reverse bias in steady state conditions, the application of a filling pulse of amplitude high enough to forward bias the junction provides the minority carriers, which are then trapped at the defect sites. In this way it is possible to study the different capture coefficients $c_p$ and $c_n$ for each particular defect. For this purpose special care has to be taken in order to control the injection level, as to say the minority carrier concentration.

By means of such a methods, it is possible to overcome the mentioned limitation of "blindness" of the conventional majority-injection DLTS to the deep levels in the "wrong" part of the band-gap. As shown in Fig. 7, a typical spectrum can be difficult to interpret, being a majority carrier component always present (of positive intensity in the figure). This is even more apparent by comparing two different spectra, corresponding to different emission rates: in Fig. 7 the interference of majority-carrier traps

Fig. 6: DLTS intensity $S$ as function of filling pulse $V_p$ (A) and the corresponding defect concentration versus depth (B)
around 200 K is more pronounced for a low emission rate, or, in other words, for long observation time. Anyway, even a semi-qualitative investigation of such a kind brings useful information. For instance the production of particular defects can be examined for different types of dopants and Fermi level positions (as to say, charge state). In addition, the stability and evolution of many defects can be followed at the same time.

- **Filling pulse**

  Conventional DLTS measurements make use of long filling-pulse durations, on the order of milliseconds, in order to ensure the full charging of the available deep levels, even the ones with small capture cross sections (up to of the order of $10^{-18}$-$10^{-19}$ cm$^2$).

  By playing with the duration of the pulse, one can vary the fraction of filled traps before their emission is recorded, then exploring the capture process of a particular trap. After positioning at the temperature which maximizes the DLTS signal for the level of interest, a sequence of measurements is performed, in which the filling pulse duration $\tau_p$ is
progressively increased, while keeping all the other measurement parameters fixed. A plot like the one in Fig. 8 can be built. In the simplest case, verified by most of the point-like defects, an exponential dependence of the signal $S$ on $\tau_p$ of the following kind is found:

$$S(\tau_p) = S_{\text{sat}} (1 - \exp(n_c \tau_p)) \quad (2.10)$$

where $S_{\text{sat}}$ is the DLTS signal at saturation, when all the trap centers are filled by the pulse.

More refined theories take into account deviations from eq. 2.10, due to the non abrupt edge of the depletion region, in which the injected free carrier concentration is not uniform. In this way a very accurate measurement of the capture rate of a deep level is available, much better than from the Arrhenius extrapolation of eq. 2.9. The knowledge of such characteristic is essential in all the applications which make use of deep levels as recombination centers [18, 19].

As shown in Fig. 8, such kind of investigations can lead to the discovery of overlapping between comparable emission from different levels [in this case the VO (-/0) and the $C_5C_1$ (-/0) levels in silicon], and it
provides a tool to quantify the relative contribution of each deep level to the DLTS signal. In addition, after obtaining such "capture fingerprint" of overlapping levels, it is then easy to mask the fast part from the slower one, just by choosing a proper value of $\tau_p$. A development of such a treatment consists of recording the full spectra for different filling pulses, and study the evolution of the DLTS peak shape as the filling pulse increases. This more time-consuming approach may be necessary in cases where strong overlapping of levels, with rather different intensity and similar energy position, occurs. It is the case of the peak at $E_C - 0.43$ eV in silicon, which has been found to consist in the emission from different defects ($V_2$, VP and CP), all with similar energies. The dominant signal of $V_2$ in CZ silicon can be seen growing alone up to filling pulses of the order of 1-10 µs, when the other levels start to appear on the shoulders (see Fig. 9).

Fig. 9: DLTS spectra for different pulse-length: the asymmetric growth of $S$ around the maximum indicates an overlapping level on the low temperature side, as pointed by the arrow.
- Reverse-bias cooling, carriers injection, metastability

Some defects studied by the DLTS technique show an interesting feature, called metastability. By this term one refers to the possibility of switching the defect between alternative configurational states, each of them corresponding to a different level in the band-gap. The (meta-)stability of a particular configuration is a function of many parameters such as temperature, free carrier concentration, charge state of the level, electric field.

![DLTS spectra of one implanted Si p+/n diode recorded after slow cooling with different bias conditions](image)

**Fig. 10:** DLTS spectra of one implanted Si p+/n diode recorded after slow cooling with different bias conditions

To identify a set of deep levels as metastable by means of DLTS, and study the set of conditions under which the defect population switches from one configuration to another, provides a great deal of information on the identification of the metastable defect. Moreover, this kind of studies is of precious importance, once established the metastable character of a particular level, in order to get the correct relationship between DLTS intensity and defect concentration, in quantitative investigations. In fact, the measurement itself can provide a variation in the observed DLTS spectrum for the metastable defects, depending on the choice of the setting and starting conditions.
A typical example of a bistable defect in silicon is the $C_5C_1$ center, which gives rise to two bistable deep levels, $A$ at $E_C-0.17$ eV and $B$ at $E_C-0.11$ eV, respectively. If the defect state is set in one configuration, the sample is cooled, and then the charge-state is changed, the defect will became bistable. When, in fact, such a defect is in the unfilled state, then it has a stable configuration with the lower ionization energy (See the configuration diagram in Fig.11 from Ref. 20).

Once that it has captured one electron, it has a (temperature-dependent) probability $p_{B \rightarrow A}(T)$ per unit-time to switch to the other configuration (corresponding to $E_C-0.17$ eV) before the emission takes place [20]:

\[
p_{B \rightarrow A}(T) = 1.37 \times 10^{12} \exp(-0.145 \text{ eV}/kT) / \text{s}
\tag{2.11}
\]

Then there is a competition between the electron emission from 0.11 eV level and the switching to the $A$ (0.17 eV) configuration. On the other hand, if the lapse between consecutive pulses (roughly equal to the measurement of the transient) is not long enough, no complete back-conversion to the original 0.11 eV takes place.

![Configuration coordinates diagram for the $C_5C_1$ defect](image)

**Fig. 11**: Configuration coordinates diagram for the $C_5C_1$ defect
In fact an analogous, thermally-driven rate of conversion $q_{A\rightarrow B}(T)$ from 0.17 eV to 0.11 eV exists:

$$q_{A\rightarrow B}(T) = 1.33 \times 10^{12} \exp(-0.174 \text{ eV}/kT) /\text{s}$$  \hspace{1cm} (2.12)$$

Then the choice of the measurement parameters strongly influences the relative intensity of the two DLTS signals (the one at 0.17 eV contains also a non negligible contribution from VO ionization). In Fig. 12 the relative population of the metastable level at 0.17 eV is depicted as a function of the typical filling pulse durations $\tau_p$ and transient measurement times $\tau_m$ in the interesting temperature range. This kind of plots, built from eq. (2.11) and (2.12), are quite useful in order to distinguish the two contributions of the electron emission, from VO and $C_5C_1$ defects, when performing a DLTS measurement.

![Fig. 12: Relative population of $C_5C_1$ defects on the acceptor state at $E_C$-0.17 eV, as function of meas. parameters $\tau_m$ and $\tau_p$.](image_url)
III. PRINCIPLES OF ION IMPLANTATION

Understanding of the mechanisms of defect creation by ion implantation in semiconductors is fundamental for electronic device technology. The damage introduced by ion implantation can be studied by analyzing the various collisions taking place once that the projectile impinges on the target. The ultimate goal of such kind of investigations is the prediction of the final defect products, as a function of the implantation parameters (mass, charge, and energy of the implanted ion, dose and flux of the ion beam) and sample characteristics (crystalline orientation, temperature, doping, impurity content).

3.1 Stopping of energetic particles

The starting point of such a modelization is the interaction between an incident particle and one atom in the crystalline matrix. One physical quantity of interest is the energy loss $dE/dx$, as to say, the energy lost in collisions by the projectile per unit of penetration depth. Due to the random nature of such processes, $dE/dx$ is a statistical variable and has to be computed performing simulations on a stochastic ensemble [21].

A distinction is usually made between two different contributions to the energy loss. The first one consider the interaction of the ion with the electrons of the atoms in the target. The core electrons mainly act a screening the coulombic potential with the nucleus. In a first approximation, strictly valid for low energy implantations, they can be treated as part of a quasi-molecule, incoming ion-target atom, and they dissipate energy by ionization to the different electronic energy states of the "molecule". The conduction electrons in the crystal add an additional coulombic term. A complete theory takes into account also quantum-mechanical effects and multi-particles interactions, the electrons in a crystal having a collective behaviour. Such an electron energy loss is the dominant term for ion energies in the MeV range, being up to 90% of the initial energy. Moreover, it depends on the energy of the impinging particle. Each of these inelastic collisions, involving just a limited amount of energy transfer, of the order of the semiconductor band-gap, quite rarely
give rise to an atomic displacement, so it can be regarded as merely a friction term (proportional to the square root of the kinetic energy). No specific permanent effect due to the high electronic excitation has been observed in the most common semiconductors, up to an electronic stopping power of 50 keV/nm, reached by implanting 10 MeV carbon clusters [22].

As soon as the particle loses energy going deeply in the bulk, the probability of knocking out one target atom from its original lattice position increases. This second term can be visualized as mainly an elastic interaction, and the classical theory of scattering can be applied [23]. The cross section $\sigma_n(T)$ of the nuclear interaction between one atom at rest and the incoming ion (defined as the probability to transfer an energy amount $T \pm dT$) is a strong function of the ion energy. In other words, the energy transfer to the target atoms by nuclear collisions is more favourable as the incoming ion reaches low energies (for small deflection angles $\sigma_n$ scales roughly as $1/E$, where $E$ is the ion energy). A strong dependence on the ion mass is also to be considered, heavy ions being much more effective than light ones in causing nuclear collisions (in this case $\sigma_n$ scales roughly as $M$, the ion mass), as shown in Fig. 13.

![Fig. 13: Energy losses in silicon from TRIM simulations; the two ion energies have been set to give a similar implantation depth](image)

---

29
3.2 Ion cascades and defect generation

At some stage the contribution to the energy loss caused by the nuclear collisions becomes crucial. This process is the one responsible for the creation of unrecoverable damage in the sample. In fact, if the transferred energy exceeds a threshold value $E_d$, around 15 eV for common semiconductors, a displacement of the target atom takes place. Otherwise, the atom just transfers some energy to the lattice in the form of phonon excitation.

If the displaced atom does not return to the initial position, the collision has generated 'elementary damage' in the sample, a Frenkel pair. This pair is constituted by a vacancy in the original lattice position, and an interstitial atom. On its turn, if the displaced atom gains sufficient kinetic energy, it can give rise to its own collision-displacements in the lattice, resulting in a concentrated generation of Frenkel pairs, a collision cascade. In the case of large energy transfer, the lapse between consecutive displacements is large enough to consider them independent, and the cascade is defined as linear cascade. However, the interaction between simple defects generated inside the cascade by closely spaced scattering events can result in a complex scenario. In this latter case, modelling by molecular dynamics is required, and no simple relationship between the elastic energy loss and the defect production exists. This is particularly important in the case of heavy ions, where the elastic energy deposition is substantial, and the resulting damage efficiency per ion is greatly enhanced. This fact implies also that a larger energy is needed to attain the same penetration depth than in the case of light ions. This is shown for instance in Fig. 13, where a comparison is made between electronic and nuclear energy losses of 370 keV proton and 6 MeV silicon implantations.

A simplified approach, which tries to estimate semi-quantitatively the damage efficiency, makes use of empirical formulas which gives the average size of a cascade, the energy deposited into the cascade and the number of primary defects created as a function of the initial energy, mass and charge of the projectile. These data can be used in Monte Carlo simulations, which give the vacancy/interstitial and implanted ion distributions, the electronic and nuclear energy deposition, by following a quite large number of individual ion trajectories. In a more detailed
simulation, each individual recoil is followed. The one used in this work is called TRIM (Transport of Ion in Matter), which make use of the simplified classical model (and the relative equations) described above [24]. In TRIM the following limiting assumptions are made: the sample is regarded as amorphous, neglecting any effect of chanelling. Moreover, each introduced ion is treated individually, in other words it does not consider the previously introduced modification ("zero dose approximation").

3.3 Evolution of the damage in the low dose regime

Several studies have been focussed on the investigation of the next step: from the generation of primary point defects, the Frenkel pairs, to the final stable damage configuration. It is well established that just a few-percentage of the generated vacancies/interstitials escape mutual recombination inside the collision cascade, or so-called recombination volume. The usual time scale of this process is on the order of picoseconds. Point defects generated in a collision cascade are spatially correlated, i.e. the probability of finding a defect in the vicinity of another defect is larger than further away. This results in larger recombination rates than in the case of uncorrelated distributions [25].

Both such primary defects can exist in different charge states and to each of these corresponds a certain diffusivity. During the implantation, due to the high ionization occurring in the sample, athermal migration of vacancy or interstitials can then take place after capture or release of a free charge-carrier [26]. This fact introduces a further complication.

The high diffusivity of both the Frenkel defects causes them to be unstable and prevent their direct observation even when performing the implantation at very low temperatures [2, 27, 28].

The following considerations are limited to the case where the implantation dose is kept well below the amorphization threshold, at which the accumulation of damage is high enough to destroy the lattice symmetry in the crystal. A complex hierarchy of competing reactions involving both interstitials and vacancies has been established, but still a satisfactory understanding of the kinetics of formation of secondary defects has not been reached. Hereafter, by the term "secondary defects", one refers to all the defects which are thermally stable at room
temperature and are observed after implantation, when a steady state has been reached.

A first issue to address is the determination of the relative amount of stable defects created by direct impact, without involving any diffusion process. It is the divacancy ($V_2$) and more complex vacancy clusters on one side, and symmetrical interstitial clusters (both divacancies and di-interstitials are known to be thermally stable at room temperature). By the previous discussion, it can be suggested that this contribution becomes more important along with the mass of the implanted ion. Evidence for this conclusion comes from the fact that the production of VO (vacancy-oxygen) centers, direct monitor of the introduced free vacancies, decreases in samples implanted with heavy ions, relative to that of $V_2$, which stays constant as a function of ion mass [29]. However, in MeV electron-irradiated silicon, instead, the ratio between production of $V_2$ by vacancy-vacancy pairing after diffusion and the one by direct impact has been estimated to be between 3 and 4 [30, 31].

The second important issue is the evaluation of reliable data for the diffusion coefficients $D_I$ and $D_V$. In this respect controversy still holds, and, because of the invaluable impact on many technological processes, it is one of the open fields to be explored by means of several experimental techniques and theoretical studies. Some authors assign a larger diffusivity to vacancies, which are then the first ones to stabilize by interacting with impurities or pairing themselves [32, 33]. The excess of free interstitials, partly giving rise to interstitial-related defects, partly annihilating the vacancy-defects, is then responsible of the long transient in the leakage current observed after implantation in silicon implanted diodes [34].

By comparing the vacancy-related defects after implantation at different dose-rates (or fluxes, rates of incoming ions per unit area), in the low dose regime where the defect production is proportional to the implanted ion dose, a completely opposite conclusion has been reached.

This is the so-called "reverse dose-rate" effect, a decrease in vacancy-related defects concentration as the implantation flux gets higher. It can be explained by assuming a large $D_I$, which allows an interaction between subsequent ion tracks, enhancing the annihilation by overlapping self-interstitials.
This phenomenon becomes quite relevant at high dose-rates, when the time-scale of the self-interstitial diffusion between adjacent ion-tracks and the rate of ion introduction are comparable.

According to this model, a strong dependence on the dose is then predicted. The average distance between tracks is reduced more and more when implanting at high doses ("high" on a relative scale, for instance passing from $1 \times 10^9$ cm$^{-2}$ to $5 \times 10^{10}$ cm$^{-2}$), and, because of that, the Frenkel recombination is favorized and a reduction in production of secondary defects is expected. The interstitial diffusion is a thermally activated process, and by varying the implantation temperature it is possible to estimate an activation energy $E_a$ of the diffusion coefficient $D_i$. In the case of proton implantation $E_a$ has been found equal to $0.065 \pm 0.015$ eV, implying that a large contribution to the interstitial diffusion occurs via the ionization-enhancement [Paper V, 35, 36]. The fact that the dose and flux are so strictly linked has another important implication. Every time that a linear dependence in defect production is found as a function of dose, the linear coefficient is dependent on the chosen flux. This is illustrated in Fig. 14 (from Ref. [37]), where a comparison is made between the VO concentrations in Si as a function of flux for three doses.
different MeV proton doses. More studies of this kind are highly needed, especially in p-type material, for a quantitative understanding of the defect build-up.

**Role of the impurities**

The interstitials and vacancies which survive mutual recombination or annihilation at microscopic sinks, like surfaces, dislocations, or stacking faults, are stabilized by pairing themselves or by interacting with one impurity atom. It is quite important to stress once more that a precise knowledge of the various impurities incorporated in the wafers during the growing process is crucial, because this determines to a large extent the final, thermally stable, damage introduced by ion implantation.

A competition between the different kinds of impurities, present in relevant concentration in the sample, takes place when trapping migrating defects. The major impurities to be considered are:

- **Oxygen**: depending on the characteristics of the sample, its concentration is varying from $\sim 10^{18}$ cm$^{-3}$ in Czochralski silicon, to $\sim 10^{16}$ cm$^{-3}$ in float-zone refined silicon, and below $\sim 10^{15}$ cm$^{-3}$ in epitaxially grown silicon. In virgin silicon it occupies an interstitial site, being electrically inactive. It is known to be an efficient trap for both vacancy- and interstitial-related migrating defects [38, 39].

- **Carbon**: it is the second major impurity in silicon, with a concentration ranging from $\sim 10^{16}$ cm$^{-3}$ to $\sim 10^{15}$ cm$^{-3}$; it occupies a substitutional position, and during implantation, by interacting with interstitials, it can turn into interstitial carbon (C$_{i}$). Because of its significant diffusivity at room temperature ($D_{C_{i}} \sim 1x10^{-15}$ cm$^{2}$s$^{-1}$) [40], C$_{i}$ is unstable, and then gives rise to a series of secondary defects by coupling with other impurities (see the diagram below) [41].

- **Dopant impurities**: both n-type (phosphorous, antimony) and p-type (boron) dopants are involved in the production of secondary defects. Boron atoms have the same behaviour as carbon atoms in trapping the diffusing interstitials and forming an interstitial configuration (exchange mechanism). Again, a similar series of secondary defects is created by migrating boron interstitials, as illustrated in the diagram below. All these reactions are diffusion-limited.
The group V elements can be paired with migrating vacancies, giving rise to the so-called E-center, and in this respect compete with oxygen. Alternatively, it is possible that, by coupling with interstitial carbon, one metastable defect is created \( (C_1P_S, \text{see previous chapter}) \) [42]. Also in this case, the reaction is favored in highly doped \( ([P] > 1 \times 10^{16} \text{ cm}^{-3}) \) material with low-oxygen and low-carbon content, to minimize the competition to form \( C_iO_i \) and \( C_SC_1 \) centers.

- Implanted species: some of the most reactive implanted species can form defects, both during the implantation and by subsequent reactions. For instance, a noticeable increase of boron/carbon related defects has been observed after implantations with such ions. Moreover, among the others, hydrogen has been recognized as one of the most prominent species, due to its high diffusivity and reactivity in silicon.

- Metals: many of the transition and noble metals are known to interact with primary and secondary defects, and their detection can be easily made, giving a reliable indication of their contamination in the sample. A great deal of the scientific literature is dedicated to the identification of the deep levels introduced in the band-gap by such defects, in order to study their beneficial or detrimental effects on the electrical characteristics of the samples [43]. In virgin samples, without accidental or intentional contamination, the concentration of such defects is quite low, usually below the detection limit of most of the experimental techniques.

A scheme of the different products from interstitial-related and vacancy-related reactions is depicted in Fig. 15. As Kimerling and al. pointed out [41], the balance between the relative concentration of ultimate defects lies both on the relative concentration of the related impurity, and, to a less extent, also on the interaction cross-section, as well as the temperature at which the reaction is thermally favored. Once that the density of each impurity \( (O_i, B_S, C_S, \text{etc.}) \) is known, by comparing the relative concentrations of secondary defects (for instance \( B_iO_i \), \( B_iC_1 \), etc.), one can estimate the so-called branching ratio, a qualitative measure of the efficiency of one reaction respect to a competing one, for instance:

\[
\frac{[B_iO_i]}{[B_iC_1]} \left( \frac{[C_S]}{[O_i]} \right)
\] (3.1)
By comparing the different ratios, one can get an insight into the interaction between the impurity and migrating interstitial/vacancy-defects. Usually a small branching ratio (1-2) indicates that short-range interactions dominate, while a stronger influence of long-range interactions is deduced from large branching ratio (>5). Being these latter ones of electrostatic kind, a dependence on the bulk Fermi level is expected.

The thermal evolution of the introduced defects is another, closely related aspect in which the role of impurities is crucial. As illustrated in Fig. 15, the secondary defects are subject to different thermal stability. In general, the interaction of point defects with impurities stabilizes them until the temperature is high enough to activate new processes. The difficulty of such investigations resides on the fact that different defects can anneal out in the same temperature range, complicating the identification of the mechanisms involved and the reaction products. In the low-dose regime explored, a major role is played by impurities, which are trapping the moving defects/impurities, giving rise to a new generation of defects. Under these conditions, an overwhelming
characteristic for the trapping is that the impurities can exceed in concentration the point defects by several orders of magnitude.

The two main thermally activated processes which are responsible for the recovery of secondary defects are migration and dissociation. In the former case, the thermal energy displaces one component of the defect, which partially dissociate from the defect. This means that the displacement is just about a few lattice distances, before the strong attraction force, e.g. coulombic, reforms the original defect, resulting in an effective displacement of the defect. This process involves lower energy supply than total dissociation of the defects (when the component is energetic enough to overcome the barrier that enables the re-capture), but, on the other side, it results in a shorter total displacement relative to the original defect position. It means that, if both processes are active in the investigated temperature range, the rate of dissociation and the rate of migration can be comparable, and difficult to separate.

These distinct annealing mechanisms have been recognized for the vacancy-related defects [44, 45, 46], while a pure dissociation has been established for some of the interstitial ones [41].

3.5 Applications of defect engineering

Since the 1960's ion implantation has been established as a standard technique for introducing impurities in a controlled manner in semiconductors. The creation of point-defects and more complex damage is a side-effect responsible for a series of unavoidable phenomena, which can be detrimental or beneficial for the performances of the semiconductor devices. In order to understand and control them, so called "defect engineering" is utilized.

A main application is indeed the doping, by which the introduced impurities provide the carriers (electrons/holes) for electrical conduction. In all the cases when ion implantation is suitable, the advantages over other common techniques, like in-diffusion for surface, are quite remarkable. For instance, the extreme reproducibility and the high spatial control of the implanted ions are key-issues which have promoted this method in the semiconductor industry. The low thermal budget and independence on the processing flow are also two well appreciated
features of ion implantation, even if recently a new generation of technological limitations has emerged.

After a post-implantation anneal (in the range of 700-900 °C), the depth of a p⁺/n junction, as to say the depth at which the concentration of implanted p-type dopant (normally boron) drops, differs from the as-implanted profile, as illustrated in Fig. 16 (from Ref. [47]). This effect is assigned to the so-called "transient-enhanced diffusion" (TED). As one can easily see, it becomes critical as the lateral and vertical dimensions of the junction is reduced to a length of the order of such effect, which is usually some tenths to hundreds nanometers. It has been observed in silicon under many different conditions and affects both n-type and p-type dopants, even if to a different degree.

The explanation of the broadening of the dopant profile, or, in other words, of the enhancement in the diffusivity of the implanted dopant atoms has been found in a supersaturation of self-interstitials which drive the diffusion of the dopants, according to:

\[ D_A = D_A^* \frac{[Si_I]}{[Si_I]^*} \]  \hspace{1cm} (3.2)
where $D_A^* (D_A)$ is the equilibrium (enhanced) diffusion coefficient of the atomic specie A, and $[Si]^* ([Si])$ is the concentration of self-interstitials at equilibrium (at the supersaturation) [48]. A similar behaviour has been seen in the case of Sb, a typical vacancy diffuser. This observation implies that in the case of high dose implantation, both an interstitial- and a vacancy- supersaturation is present, just displaced in depth [49].

An additional, detrimental effect of the release of interstitials during thermal treatment is the clustering of boron atoms, reducing the electrically active dopant concentration. This has been observed in the proximity of the damage peak for high concentrations of boron. Indeed, one can expect that an high concentration of boron in interstitial configuration, produced by interaction with the supersaturated self-interstitials, can agglomerate and form highly stable clusters, not acting as shallow acceptors [49, 50, 51].

The source of interstitials has been identified as the so called {311} defects, a class of extended interstitial defects which resemble a rod along {311} oriented planes. By a modelling of the TED process, the number of interstitials stored in the {311} defects has been estimated, and the efficiency in generating such complexes can be studied as a function of implanted species. For instance, it has been observed that the clusters generation, and consequently TED, is much more pronounced in the case of heavy ion implants [52, 53].

The conditions under which these defects are generated is an object of intense investigation. The progressive agglomeration of interstitials has been followed, from point-defects formed at room temperature, to clusters of larger dimensions at temperatures higher than ~300-350 °C, up to extended {311} defects at temperatures above 550-600 °C [54, 55]. Even if this aspect is still quite unexplored, it seems established that some configurations of the interstitials clusters are more stable than others and are able to grow on the expense of the less stable ones (Ostwald ripening) [56]. A TED effect can be observed for some particular conditions before the {311} defects have been formed [57]. Further, a critical dose at which point defects from the implantation act as nucleating sites for extended defect formation has been found, both in n-type and p-type Si [58, 59]. This dose is dependent on the impurity content, which can inhibit the cluster formation by capturing the migrating interstitials [60, 61].
Interestingly, a way to reduce the concentration of excess interstitials generated by shallow implants (of the order of few tenths of keV) is by a superposition of excess vacancies by a prior MeV ion implantation. In this way the Frenkel recombination is quite efficient because the two types of defect- distributions overlap, widely reducing the agglomeration of interstitial clusters [62].

While TED is a phenomena regarding implantation to high doses, other applications of defect engineering are instead limited to the low-dose regime. A main one is the use of ion implantation for lifetime control.

The intentionally introduced point-defects are meant to act as recombination centers in power electronics devices. The high resolution achievable by ion implantation in creating a spatially confined 3-dimensional damage region is indeed a quite remarkable feature. The reduction of on-state losses in a bipolar power diode imposes the requirement of such a confinement. In fact, the presence of recombination centers in the active portion of an electronic device increases noticeably the leakage current in the blockage mode and the voltage drop in the conduction mode. A good compromise is then attainable by positioning the recombination centers in the low doped bulk region of the device, far from the junctions. This is the region where it is more difficult to remove the excess carriers after the device has been switched to off-state, and here the point-defects can then be more effective, enhancing the carrier recombination. At the same time, the on-state losses are reduced. A look at the recombination/generation process involved provides support to such a statement.

The recombination in indirect band-gap materials can occur via deep levels or by the so called Auger recombination. In fact, a third possibility, the band-to-band recombination, is rather limited, requiring the coupling with a phonon with large momentum. The probability of such an event is quite small, and this process can be neglected. Also Auger recombination is quite small for density of excess-carriers below $5 \times 10^{16}$ cm$^{-3}$.

If only recombination through deep levels is considered, each of them contributes to the electron (hole) lifetime with a recombination time $\tau_n$ ($\tau_p$) which is inversely proportional to the capture coefficient $c_n$ ($c_p$) and to the density of traps $N_T(x)$, according to [63]:

$$\tau_n = 1 / c_n N_T(x), \ \tau_p = 1 / c_p N_T(x)$$

(3.3)
The low level injection lifetime $\tau_L$ and the high level one $\tau_H$ are functions of the partial recombination rates, and in n-type material they can be expressed by the following equations (similar ones hold for p-type material) [10]:

$$\tau_L^{-1} = n_0 \sum \left[ \tau_{nj} (n_0 + n_j) + \tau_{pj} p_j \right]^{-1}, \quad \tau_H^{-1} = \sum (\tau_{nj} + \tau_{pj})^{-1} \tag{3.4}$$

where $n_0 = [N_C N_V \exp(-E_G/kT)]^{1/2}$ is the intrinsic electron concentration ($E_G$ is the energy band-gap), $n_j$ ($p_j$) is the electron (hole) concentration at level $E_{Tj}$, and $j$ is the summation index over the deep levels. A simplification in (3.4) is found for $\tau_L$ in the case of only one dominant deep level, around the middle of the band-gap. Then the carrier lifetime became equal to $\tau_p$, or, in other words, the minority carrier lifetime.

In the case of silicon, more than one deep level is present after ion implantation, and the carrier lifetime is dominated by the contribution of a particular level at low or high injection. The main recombination centers produced by ion implantation are the vacancy-oxygen (VO) and the divacancy (V$_2$) defects. Following eqs. (3.3), (3.4) and the definition of capture coefficient, one can understand that V$_2$ centers, introducing a deep level close to the mid-gap, will dominate the low-injection lifetime. Instead, the main influence by VO defects is at high injection levels, due to its large capture cross-section.

One of the aims achieved by defect engineering is the determination of the set of implantation conditions which optimizes the performance of the power devices under the required excess carrier regime. For instance, a proper choice of the implanted ion results in a well defined ratio of $[V_2]/[VO]$ (which ranges from $\sim 0.5$ to $\sim 5$ passing from MeV proton implantation to heavy ions implants) [29]. On its turn, this is reflected in a suitable ratio $\tau_H / \tau_L$.

The introduced defects act also as carrier generation centers, increasing the leakage current. A generation rate $\nu_g$ can be defined, according to:

$$\nu_g = \sum \{ \tau_{nj} \exp[(E_F - E_{Tj})/kT] + \tau_{pj} \exp[(E_{Tj} - E_F)/kT] \}^{-1} \tag{3.5}$$
The most effective generation centers, as in the case of the low-injection recombination, will be the ones very close to the middle of the band-gap, i.e. near the Fermi level position in a reverse-biased junction. It is worth noting once more the implicit spatial dependence of the partial recombination/generation times $\tau_p$ and $\tau_o$ (and consequently of $\tau_i$, $\tau_i$ and $\nu_g$), in the case of ion implantation. This degree of freedom allows to place the recombination/generation region far away from the electrical junctions, where the electric field is enhancing the generation process.

An additional, and non negligible, effect of ion implantation on the electrical characteristics of an electronic device is the doping compensation. The damage created by energetic ions can interact directly with the dopant atoms, creating impurity-defect pairs which decrease the free-carrier density. The electrically active defects have in general an even more drastic effect, introducing electronic levels in the band-gap which act as trapping centers. According to the emissivity of each of these deep levels and to their character (donor-like or acceptor-like), the effect on the carrier compensation will strongly depend on the doping level $N_D$ ($N_A$ for p-type doping), and the temperature of the substrate. According to the balance equation, one can estimate the reduction in carrier density, in n-type material with several acceptor-like traps, each one with a density $N_{Tj}$ and an ionization energy $E_{Tj}$:

$$n \approx N_D\sum N_{Tj} \exp[(E_F-E_{Tj})/kT]$$

(3.6)

The reasonable approximation of $kT << E_F-E_{Tj}$ has been made. Remembering that in the absence of deep levels, the extrinsic Fermi level $E_F$ lies at $E_C-kT \ln(N_C/N_D)$, it is quite clear from equation (3.6) that the more efficient compensating levels at room temperature are the ones located close to the midgap. A dependence on the electric field can be introduced by replacing $E_F$ with the quasi-Fermi energy level, which contains an additional term involving the electric potential. Then the compensation will depend on the spatial location of the traps, being more effective close to the electrical junctions, where the potential is higher.

The decrease in free-carrier density, or, in other words, the increase in resistivity, results in a corresponding increase of the voltage drop in power devices when they operate in the on-state, driving large currents.
The on-state power losses are then another factor to take into account when the performance of a power device have to be optimized.

A pronounced detrimental effect by carrier compensation is observed in particle detectors used in satellites and high-energy physics experiments. These fields regard mostly irradiation by energetic electron and nuclear particles to high doses, and not ion implantation. However, the damage introduced in the form of point-defects is quite similar in both cases, and the general knowledge acquired in defect engineering can be useful for this multifold range of applications. For instance, the accumulation of point-defects in detectors, progressively increasing the compensation and the resistivity of the device, degrades the reverse-bias characteristics of the detectors. Thus, a concomitant reduction occurs of the active volume of the device, where the particle detection takes place. The lifetime of such devices can be noticeably increased by a proper understanding of the defect build-up, and of the properties of the defects which act as compensation/recombination centers [19].

3.6 Future trends and open fields

Studies of ion-implantation-induced defects, and, in more general, the influence of defects on the electrical properties of semiconductors has been a subject which has grown in size over the last decades. Many aspects remain to be investigated and, in particular, the ones regarding ion implantation. In general, an accurate and quantitative prediction of the final damage evolution as a function of the different implantation parameters is still lacking.

The transient enhanced diffusion is a typical case. The approach followed in order to understand this phenomenon has been somehow driven going 'backwards', from the final effect to the intrinsic causes, i.e. from the dopant diffusion at relatively low temperatures, where the dopants are immobile in equilibrium conditions, to the recognition that a large amount of point-defects survive at these temperatures, promoting the diffusion. The thermal evolution of the defects created by ion implantation is then the key issue, which can ultimately suggest a method to reduce or even suppress the undesired TED. This is, indeed, of a
tremendous impact, once that TED seems to occur also at low temperature budgets.

Moreover, up to now the attention has mostly been focused on the extended defects of interstitial type, which are responsible for the enhancement of the diffusion of boron and phosphorous, the most common dopants in silicon technology. As previously mentioned, a similar enhancement in antimony diffusion has been a key point in proving that a class of vacancy-related defects is present in an amount comparable to that of the interstitial clusters. The identity and the conditions under which these vacancy-defects are produced are still rather unknown.

A new class of defects has been recently found in ion-implanted silicon after high temperature annealing (800-900 °C). The interesting feature is that they are located around half of the projected ion range, as revealed by preferable trapping of metals in this depth range. The identity and properties of such defects are object of intense investigation, and still it results unclear if they consist of interstitial agglomerates [64, 65], or vacancy-like defects [66].

A new application of ion implantation, and strictly related to a spatial confinement of damage, is the one regarding gettering of metal impurities. The formation of microcavities, which are extremely effective in collecting metallic elements in the semiconductors is a rather recent discovery, which appears to be promising for several application. The role of defects (particularly Si interstitials) in both the formation and dissolution has been investigated. In fact, metal transport and trapping at nanocavities are defect mediated processes and subsequent dissolution can also be defect-controlled [67]. A closely related subject is employment of ion implantation in order to produce nanocrystals, amorphous layers, and insulators.

Also the thermal donors, a class of defect-impurity complexes, which are formed after annealing in different conditions at temperatures around 400-500 °C, have been the subject of detailed investigations. Such a denomination comes from the fact that they introduce very shallow donor-like levels, then they change the conductivity of the material. While the presence of oxygen and other species, like hydrogen or nitrogen, has been established, a clear identification and modellization is still lacking, although substantial progress has been made in the 1990's [68].
Many of the previously discussed limitations, which electronic devices suffer due to defects introduced by ion implantation, become more and more stringent as the size of the active device regions shrink, and surface/interface effects play larger roles.

New materials appear for specific commercial applications, and the feasibility and drawbacks of ion implantation in such materials is a key-issue, strictly linked to the defect generation. For instance, an enhancement of the doping compensation by about a factor of ten in 4H n-SiC relative to n-Si, after an equal dose of MeV proton implantation (see Paper VII and VIII). The future application of this compound material to power and high frequency electronics relies heavily on a proper control of the implantation conditions/parameters [69].
IV SUMMARY OF THE RESULTS

A first issue which has been addressed is the thermal stability of the main vacancy-related point defects in silicon, the divacancy and vacancy-oxygen centers. The annealing of such defects provides useful information on the thermally activated reactions involving defects and impurities in the substrate. This argument has been developed in the first three papers, allowing an exhaustive understanding of the annealing mechanisms involved.

A second main argument under investigation has been the production of secondary defects by ion implantation: a first experimental observation of the displacement between vacancy-related and interstitial-related secondary defects in silicon is reported in paper IV, while the influence of dynamic annealing on the production of secondary defects has been addressed in paper V, where the so-called "reverse dose rate effect" has been revealed.

The study of point defects in silicon-carbide is still in its prime: a characterization of deep levels, related to point defects in electron irradiated SiC, has been made and reported in paper VI, while paper VII and VIII discuss the dramatic doping deactivation in SiC due to ion implantation.

4.1 Paper I

By means of DLTS technique, the effect of thermal treatments on heavy ion-induced defects in silicon has been investigated. Wafers with different doping (n- and p-type) and purity (Czochralski and float-zone refined Si) have been implanted with low doses of MeV iodine ions.

In n-type samples, a substantial amount (about 1/3) of defects, with a deep level at about \( E_C - 0.43 \) eV, disappears after a thermal treatment at low temperatures (150 \(-200 \) °C). At the same time, the concentration of divacancy and vacancy-oxygen centers shows an increase, more pronounced in high-purity float-zone material. The vacancy-phosphorus defects are known to anneal out in this temperature range, but their estimated concentration is too low to account for these phenomena.

The dissolution of vacancy-related clusters, with low thermal stability, is supposed to cause this "reverse annealing" effect. Heavy ions are known to efficiently create vacancy-clusters, and their production well correlates with the increasing integrated intensity of the DLTS peak at \( E_C - \)
0.43 eV with increasing the mass of implanted ion. Then a release of vacancies from unstable clusters occurs during the annealing. Other recent studies support this interpretation, observing a similar increase of vacancy-related point-defects in the same temperature range, even if the implantation conditions were different [70, 71].

4.2 Paper II

The stability of ion-implantation induced defects in n-type silicon has been investigated as a function of ion mass. In phosphorous-doped silicon, implantations with different types of ions have been performed, normalizing the implantation dose in order to create a similar amount of damage.

No substantial difference has been detected in thermal stability of divacancy (V₂) and vacancy-oxygen VO centers, the main vacancy-related point-defects measured by DLTS, as a function of implanted ion mass, as well as for electron irradiation. However, in the case of very heavy ions (germanium and iodine), an inverse annealing effect appears after thermal treatment at low temperatures (T_{\text{ann}}< 200°C). This observation is in accordance with the interpretation put forward in paper I: heavy ions are more effective than light particles in the generation of clusters of vacancies. Being the stability of some clusters lower than VO and V₂, a release of vacancies from such clusters favors the additional formation of vacancy-related point-defects. This effect is more pronounced in high purity Float zone material than in Czochralski silicon, where both VO and V₂ centers show a lower thermal stability. Moreover, it is spatially located at the implantation end-of-range, where the defect accumulation is more prominent, and the cluster generation is supposed to occur.

When the annealing is performed at higher temperatures, the signal from V₂ centers, and eventually from VO, shows a decrease in intensity. The annealing of both defects presents an interesting feature, a lack of broadening of the concentration profile versus depth. This is particularly surprising in the case of V₂ defects. For such a defects, the main annealing mechanisms proposed is the migration to trapping centers (possibly impurities), which would involve a redistribution of the V₂ centers by thermal activated diffusion. The reasons of the absence of such a broadening are addressed in Paper III, where the kinetics of the annealing reactions have been investigated.
The recovery of divacancy and vacancy-oxygen centers, introduced by MeV ion implantation in silicon has been studied in MeV ion-implanted n-type silicon. Two sets of samples have been investigated: Czochralski Schottky (CZ) diodes and float-zone-refined p+/n (FZ) diodes. A decrease in concentration for V₂ defects has been detected in the temperature range 200-250°C, independently on the characteristics of the implanted samples. Moreover, the recovery of V₂ follows a first-order reaction kinetics, indicative that a "simple" process is active.

By observing that the VO centers are not affected by the recovery of V₂, some possible reactions, like recombination with mobile interstitials or dissociation, are ruled out. Instead, migration of V₂ to trapping centers is in excellent accordance with the experimental data. The activation energy (\(E_a \approx 1.25-1.3\) eV) and frequency factor (\(\nu_0 \approx 0.5-1 \times 10^9\) s⁻¹) of this process are well in the range of the predicted values, obtained independently by other experiments with complementary techniques. By combining these data with the experimental observation of a diffusion length \(L_D\) below 0.1 µm for migration of V₂ (see previous paper), a lower limit of the concentration of trapping centers is estimated to be \(\approx 5 \times 10^{15}\) cm⁻³. The small dependence on the oxygen content in the samples for the annealing of V₂ excludes oxygen as a candidate for the trapping center.

The recovery of VO, followed by monitoring the intensity of its acceptor level at \(E_C-0.17\) eV, appears to be quite different in the two set of samples. A substantial amount of the VO centers is progressively passivated by highly mobile hydrogen, which initially is incorporated at the surface during the cleaning procedure for the diode manufacturing. By pairing with one migrating hydrogen atom, a VOH center forms and a new deep level appears in the band-gap at \(E_C-0.32\) eV. A strong correlation between the loss of the \(E_C-0.17\) eV level and the corresponding evolution of the \(E_C-0.32\) eV level has been observed after thermal treatment at different temperatures (200-270°C in CZ samples, and 270-375°C in FZ samples). The higher temperature range for the p+/n diodes is attributed to an efficient retardation of hydrogen injection into the implanted region by the p+ layer. The highly B-doped layer acts as an efficient barrier for the hydrogen diffusion from the surface and increases the thermal energy needed to activate the process of hydrogen injection.

While the hydrogen source gets exhausted, and about 60-70% of the original VO centers are converted into VOH₂ complexes, an additional process is activated: the formation of VO₂ pairs by reaction of mobile VO
centers with interstitial oxygen atoms. This competing reaction has a relatively high activation energy ($\approx 1.8$ eV, in accordance with previous IR experiments using electron-irradiated samples). Hence, this process occurs mainly after long annealing times or at high temperatures.

### 4.4 Paper IV

The paper reports on the first experimental observation of a shift in depth between vacancy- and interstitial-related defect distributions, generated by ion implantation. Two secondary defects, vacancy-oxygen (VO) and substitutional carbon- interstitial carbon (CSCI) centers, have been chosen as monitors for the depth distribution of the original Frenkel pairs (vacancies and self-interstitials). A large amount of free vacancies which escape recombination are effectively captured by immobile interstitial oxygen atoms. Interstitial silicon atoms follow a similar evolution, being converted first into carbon interstitials, which, on their turn, are stabilized by pairing with other impurities (mainly oxygen and carbon in low doped, n-type silicon).

Phosphorous doped (65 ohmcm) floating-zone Si material has been used since its impurity content is such that, after an implantation with 6 MeV B ions to a low dose, the two kinds of defects are generated to similar concentrations. Moreover, both the VO and CSCI centers give rise to a deep level at about $E_{C}-0.17$ eV. By means of DLTS measurements in which the filling pulse technique is used, it is possible to distinguish between the two contributions with a large degree of accuracy.

Moreover, the concentration versus depth profile is obtained for each defect at the same sample temperature, by varying the duration of the filling pulse. Thus the two profiles can be recorded with a high relative depth resolution. The peak of the interstitial profile is shown to be displaced by $\sim 0.5 \mu m$ towards larger depths relative to that of the vacancy profile. The shift is primarily attributed to the preferential forward momentum of recoiling Si atoms, in accordance with predictions from simulations.

### 4.5 Paper V

Silicon n-type float-zone samples have been implanted with 1.3 MeV protons at temperatures between 70 and 300 K. The dose-rate was varied from $1 \times 10^7$ to $2 \times 10^{10}$ cm$^{-2}$s$^{-1}$, while keeping the ion dose constant at
5x10^9 cm^-2. The low implantation dose allows the detection of the so-called reverse dose rate effect, a strong reduction in the concentration of VO centers as the dose rate increases. This effect occurs at lower dose-rates as the implantation temperature is lowered, and an activation energy of this process is estimated to be around 0.065 eV.

This effect is attributed to the fast diffusion of Si self-interstitial from one ion-track to adjacent ones, recombining with the less mobile free vacancies. In the low-dose regime, this process is favored when the average lapse between two incoming ions matches the time scale of the diffusion, which is a thermally activated process. For a fixed dose-rate, an increase in temperature allows a higher diffusivity, making the interstitial dilution to occur in a shorter time-scale. Then the threshold in fluxes needed for the onset of the recombination enhancement shifts to higher values. The extracted activation-energy, attributed to ionization-enhanced migration of Si self-interstitials, is in accordance with the reported detection of mobile silicon interstitials even at temperatures around 4 K.

The interstitial diffusion is a trap-limited process, and thus estimated values of the diffusion coefficient have always to be considered as lower limits. By means of experiments using dual implantations, it has been possible to observe long distance migration of self-interstitials at room temperature, in accordance with the dose-rate effect. By performing implantations with 5.6 MeV silicon, a vacancy-related defect distribution has been introduced with a peak at around 3.5 µm, and used as monitor for injection of Si interstitials. A subsequent low-energy implantation provided the injection source of fast migrating defects. A strong reduction in the deeply located vacancy concentration indicates indeed an interstitial diffusion into the sample. The high dose (~10^{13} cm^{-2}), needed to efficiently inject self-interstitials, is related to the concentration of trapping centers, mainly substitutional carbon atoms, which are saturated at high doses. Impurity-trapping has an even larger influence on the diffusion of free vacancies, because of high concentration of the interstitial oxygen atoms.

4.5 Paper VI

The thermal evolution of defects in electron irradiated n-type SiC samples has been investigated. In as-grown epitaxial layers, a uniform concentration (1.5x10^{13} cm^{-3}) of defects, which display a deep level at \( E_C - 0.70 \) eV, has been found. The very low concentration of metallic impurities, such as V, Cr or Ti, as measured by SIMS analysis, rules out
any correlation with the presence of such atoms. Instead, the intrinsic nature of this defect has been indicated by subsequent MeV electron irradiation at low doses ($1 \times 10^{15}$ cm$^{-3}$). Immediately after irradiation, deep levels at $E_C-0.62$, $0.32$ and $0.68$ eV dominate the spectrum. The latter two show a quite low thermal stability, annealing with the same rate and increasing the concentration of $E_C-0.62$ eV defect. They are probably discrete levels originating from the same defect. Then, a gradual shift to a higher ionization energy can be observed, and the final stable configuration of the so-called $Z_1$ center is reached at around 200 °C, with a level at $E_C-0.70$ eV. A possible mechanism which leads to such a transformation can involve dissociation and/or migration of defects and interaction with impurities. Alternatively, a reorientation and structural modification can be invoked.

### 4.6 Papers VII and VIII

Ion implantation has different impact on the electronic properties of semiconductors. A comparison between silicon and silicon-carbide elucidates the mechanisms of such transformation. Both sets of materials, after being implanted with boron or helium ions to a low dose ($2 \times 10^8$ - $1 \times 10^{10}$ cm$^{-2}$), show a reduction in free carrier-concentration at room temperature. In this low dose regime, electrical activation of the implanted species is expected to be irrelevant. Moreover, the electrical compensation in the implanted region can be only partly ascribed to the introduction of acceptor-like defects. The concentration of $Z_1$ defects with a level at $E_C-0.70$ eV is too low to account for the observed phenomena.

Instead, the main process resulting from implantation is the deactivation of the dopants via reaction with mobile implantation-induced defects. This is supported by the observation that such a compensation increases when implanting the same dose of ions in samples with increasing doping level. The compensation effect is very much pronounced in silicon carbide, relative to silicon. In silicon, the fraction of original Frenkel pairs, which give rise to secondary defects stable at room temperature, is below 5%. In silicon carbide, the corresponding fraction is at least five times higher than in silicon. By implanting the silicon-carbide samples at elevated temperatures, up to 600 °C, no significant increase or reduction of the compensation was observed, indicating that defect mobility is not limiting the process.
VI APPENDIX:
SIGNAL PROCESSING FOR DLTS MEASUREMENTS

Hereafter a short description is made of the procedure used to analyze the measurement data and to obtain DLTS spectra. Emphasis is put on the different measurement parameters and signal processes, in order to clarify the interpretation of the resulting spectra.

The following treatment of the measured transient resembles the one of a simulated lock-in amplifier. In general, each measurement can be reduced to the one of a capacitance transient of a total duration $\tau_m$, sampled at regular steps of length $\tau$, at a fixed temperature $T$. By considering just a limited number of measurement points of the total transient, and restricting the calculation to such an interval, called time-window, it is possible to built a DLTS spectrum $S_i$. In the setting used the length of the time-window with index $i$ is equal to $\tau_i = 2^i \tau$. An exponential decay is assumed, with amplitude $\Delta C$ and time constant $(e_n(T))^{-1}$, where $e_n(T)$ is the emission rate of the deep level under investigation.

![Capacitance transient and weighting function for a time window $\tau_i = 320$ ms.](image)

**Fig. A1:** Capacitance transient and weighting function for a time window $\tau_i = 320$ ms.
Then, by considering a step-like weighting function $W(t)$, as depicted in Fig. A1, the DLTS signal $S_i$ can be built as follows:

$$S_i = \frac{1}{2^i} \sum C(t) W(t) \approx (1/\tau_i) \int C(t) W(t) \, dt$$

(A1)

where the summation is taken on the first $2^i$ points, and the integration is performed in the $[\pi/2 \leftrightarrow \pi+\pi/2]$ interval. $W(t) = -1$ for $\pi/2 \leq t \leq (\pi+\pi)/2$, and $W(t) = 1$ for $(\pi+\pi)/2 < t \leq \pi+\pi/2$.

The approximation by the integral in eq. (A1) is quite satisfactory for long $\tau$, while the sum has to be considered for short $\tau$, where the discrete nature of the recorded transient cannot be neglected. A complication arises from the fact that the first measurement point is at $t=\tau$, and not at $t=0$, which is not accessible by the instrumentation.

A simple proportionality holds between $\tau$ and $(e_n)^{-1}$ at the peak maximum, and the proportionality factor $F_1(i)$ is dependent on the index $i$, or, in other words, the number of measurement points taken in the corresponding time-window. In fact, let’s consider the first two measurement points, leading to the spectrum $S_1$. According to eq. (A1):

$$S_1(T) = \frac{1}{2} [C(2\tau, T) - C(\tau, T)] =$$

$$= \frac{1}{2} \Delta C \{\exp[-2e_n(\tau)\tau] - \exp[-e_n(T)\tau]\}$$

(A2)

The maximum in $S_1$, which appears at $T_{\text{max}}$ in a plot $S_1(T)$ vs. $T$, is found deriving eq. (A2) respect to $x = e_n(T)\tau$:

$$\partial S_1[\tau, e_n(T_{\text{max}})] / \partial x = \Delta C \{\exp(-x) - 2\exp(-2x)\} = 0$$

(A3)

Then $x = \log 2 = F_1(1)$, similar to that obtained by Lang [11].

The next spectra are built with a similar algorithm, taking in eq. (A1) the first four points for $S_2$, the first eight ones for $S_3$ and so on. The $x$ value which maximizes $S_i[\tau, e_n(T_{\text{max}})]$ is easily found by solving the associate polynomial in $y = e^x$, and then a set of values $x_i = F_1(i)$ is available. $F_1(i)$ is independent on $\tau$, and depends only on the index $i$. In other words, $e_n(T_{\text{max}})$ can be easily calculated for each time-window, just knowing the sampling step $\tau$ of the transient, and the proper rescaling factor $F_1(i)$ from Table T1.
In analogous manner, by inserting the calculated $x$ value in eq. (A2) (and in the corresponding ones, for the other time-windows), the intensity $S_i$ is found as a function of $\Delta C$, which is proportional to $N_T/N_D$, according to $S_i(T_{\text{max}}) = \Delta C F_2(i)$. For the first time window, from eq. (A2) one gets $F_2(1) = 0.25$, and the values for the other cases are listed in Table T1. The $F_2(i)$ factors allows us to rescale properly the intensity of each DLTS signal $S_i$, in order to have the value $\Delta C$ on the ordinate axis, independently on the window choice.

<table>
<thead>
<tr>
<th>index $i$</th>
<th>$F_1(i) = e_{n}(T_{\text{max}}) \tau$</th>
<th>$e_{n}(T_{\text{max}}) \tau$</th>
<th>$F_2(i) = S_i(T_{\text{max}})/\Delta C$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.69315</td>
<td>1.38629</td>
<td>0.12500</td>
</tr>
<tr>
<td>2</td>
<td>0.44568</td>
<td>1.78272</td>
<td>0.15492</td>
</tr>
<tr>
<td>3</td>
<td>0.26046</td>
<td>2.08367</td>
<td>0.17597</td>
</tr>
<tr>
<td>4</td>
<td>0.14234</td>
<td>2.27745</td>
<td>0.18880</td>
</tr>
<tr>
<td>5</td>
<td>0.074660</td>
<td>2.38913</td>
<td>0.19594</td>
</tr>
<tr>
<td>6</td>
<td>0.038271</td>
<td>2.44937</td>
<td>0.19971</td>
</tr>
<tr>
<td>7</td>
<td>0.0193804</td>
<td>2.48069</td>
<td>0.20165</td>
</tr>
<tr>
<td>8</td>
<td>0.0097519</td>
<td>2.49649</td>
<td>0.20264</td>
</tr>
</tbody>
</table>

Table T1: Values of the rescaling factors $F_1$ and $F_2$ for different time-windows.

It can be found sometimes that the intensity of a DLTS peak differs when the time-window is changed. This can be attributed to a considerable variation in the width of the depletion region with increasing temperature, which leads to an increase/decrease of the amount of trap-centers covered by the applied voltages. This case can be verified by checking the steady capacitance at reverse voltage and during the filling regime vs. temperature. Another possible reason can be the interference between signals from different levels.

In the case of simultaneous emission from different traps, the linear relationship between the total capacitance and the partial contributions during the transient ensures a linearity also in DLTS signal. Great care has to be used in extracting the trap parameters ($N_T$, $E_T$, $\sigma$), because the overlapping of two DLTS peaks can lead to an apparent shift in $T_{\text{max}}$, the
temperature at which the maximum intensity of a DLTS peak occurs. Then the use of a fitting routine is suggested. By means of the simulation of a DLTS spectrum with fitting parameteres \((N_T, E_T, \sigma)\) for each level, the overlap between two or more trap emissions can be resolved.

A quite simple algorithm has been found to reconstruct a DLTS peak from theoretical values. By means of eq. (2-1), the emission rate \(e_n\) as a function of \(T\) is known, from input values \((E_T, \sigma)\); then the first time-window signal \(S_1(T)\) is built according to eq. (A2), while the next ones are obtained in a recursive manner \((\alpha>0)\):

\[
S_{n+1}(T) = 0.5 \ S_n(T) \ \{1+\exp(-e_n(T) \ \tau)\}^\theta \ (A5)
\]

where \(\theta = 2^{\alpha-1}\). The advantage of using such a fitting subroutine is that the full measured peak-shape is analyzed, and the simulated one has to agree at each temperature value, not just at the peak maximum. Hence, the uncertainty of the extracted value is significantly reduced, since the set of simulated DLTS spectra is quite sensitive to small changes in the input parameters.
An effective method to analyze the measured spectra in the case of superposition of DLTS peaks is the use of a more sophisticated weighting function $W(t)$, in eq. (A1). The analogy of a DLTS spectrum with a band-pass filter, transmitting signal only when the emission rate falls inside the "instrumental" frequency band, is illustrative to understand the improvement in using a different $W(t)$. The one which has been used in the present work to resolve closely spaced levels presents several advantages:

- It does not introduce distortions in the DLTS spectra, like negative overshoots, and the peak-shape resembles the "original" one.
- The intensity at the peak maximum is unchanged (representing the defect density), while the broadening is reduced by a factor of 3, both as width at half maximum, and as area under the curve.
- The precision by which $T_{\text{max}}$ is estimated increases, even in the case of widely spaced peaks.
- It does not need to be implemented in the measurement setting, which stores directly the DLTS spectra, but can subsequently be applied off-

Fig. A3: Reduction in broadening of DLTS peaks by using a "sophisticated" algorithm.
line, also on previously measured data, not originally intended for high-resolution studies.

\[
\begin{array}{|c|c|c|c|}
\hline
\text{index } i & F_1(i) = e_n(T_{\text{max}}) \tau_n & e_n(T_{\text{max}}) \tau_i & F_2^*(i) = S_i^*(T_{\text{max}}) / \Delta C \\
\hline
2 & 0.540796 & 2.163184 & 0.20518 \\
3 & 0.279185 & 2.23348 & 0.23926 \\
4 & 0.141765 & 2.26824 & 0.25850 \\
5 & 0.071419 & 2.28541 & 0.26871 \\
6 & 0.035843 & 2.29395 & 0.27396 \\
7 & 0.017954 & 2.29811 & 0.27662 \\
8 & 0.008985 & 2.30016 & 0.27680 \\
\hline
\end{array}
\]

Table T2: Values of the rescaling factors \( F_1^* \) and \( F_2^* \) for different time-windows.

A recursive algorithm generates the new spectra \( S_i^*(T) \) from the original \( S_i(T) \) according to:

\[
S_i^*(T) = \{4 S_i(T) - 1.3 S_{i-1}(T) - 1.8 S_{i+1}(T)\}
\]  \hspace{1cm} (A6)

Being slightly shifted to higher temperatures, the maximum in DLTS intensity corresponds to a different emission rate than the one given in Table T1. By applying the same procedure, and solving once again the polynomials in \( y = \exp[e_n(T_{\text{max}}) \tau] \), a new set of modified \( F_1^*(i) \) and \( F_2^*(i) \) has been obtained, and listed in Table T2. It can be seen by a comparison with Table T1 that the signal to noise ratio for the high resolution spectra is not largely altered.

As shown in Fig. A3, a much clearer interpretation of the measurement is provided by the high resolution spectrum. For instance, it is advantageous to start a fitting procedure using the new modified spectra \( S_i^*(T) \), where most of the overlapping signals have already been resolved. For example, this can be crucial in the case of annealing studies, where a precise evaluation of the intensity of the deep levels is required. Moreover, it is highly feasible for implementation in concentration vs. depth profile measurements.
VI ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to prof. Bengt Svensson, who supervised my work during the whole period of my stay at KTH. He initiated me into the fascinating world of defects in semiconductors, and followed with paternal care the progress of my training. I will always keep as a reference his unexhausted strength in achieving a high scientific level in every kind of research we have gone into. I want to thank him also for the opportunities he has given me to participate in international conferences and to visit other laboratories.

I extend my appreciation to Dr. Anders Hallén, Dr. Nickas Keskitalo, Dr. Andrei Kuznetsov, Dr. Jagadish, who have actively contributed to my research by sharing their unvaluable skills in a wide range of experimental tasks.

I am also grateful to all the colleagues of the Solid State Electronics Department, who have followed the development of my research and studies, bringing useful matter in for discussions and thinking. I would like to acknowledge in particular Dr. J. Lalita and Dr. J. Doyle, who showed me the first rudiments of the experimental techniques. Nicky, Patrick and Denny have helped me to understand many aspects of the experimental work which I was not very conversant with, and I want to thank them as well.

Finally, I would like to acknowledge all the institutions who have financially supported my Ph. D. studies: the Swedish Foundation for Strategic Research, the Swedish Natural Science Research Council, the Swedish Research Council for Engineering Science, and the Swedish Institute.
VII REFERENCES


