Junction spectroscopy techniques and deep-level defects in semiconductors

A. R. Peaker,[1] V. P. Markevich,[1] and J. Coutinho[2]

[1]*School of Electrical and Electronic Engineering and Photon Science Institute, University of Manchester, Manchester M13 9PL, United Kingdom*
[2]*Department of Physics and I3N, University of Aveiro, Campus Santiago, 3810-193 Aveiro, Portugal*

The term junction spectroscopy embraces a wide range of techniques used to explore the properties of semiconductor materials and semiconductor devices. In this tutorial review, we describe the most widely used junction spectroscopy approaches for characterizing deep-level defects in semiconductors and present some of the early work on which the principles of today's methodology are based. We outline *ab-initio* calculations of defect properties and give examples of how density functional theory in conjunction with formation energy and marker methods can be used to guide the interpretation of experimental results. We review recombination, generation, and trapping of charge carriers associated with defects. We consider thermally driven emission and capture and describe the techniques of Deep Level Transient Spectroscopy (DLTS), high resolution Laplace DLTS, admittance spectroscopy, and scanning DLTS. For the study of minority carrier related processes and wide gap materials, we consider Minority Carrier Transient Spectroscopy (MCTS), Optical DLTS, and deep level optical transient spectroscopy together with some of their many variants. Capacitance, current, and conductance measurements enable carrier exchange processes associated with the defects to be detected. We explain how these methods are used in order to understand the behaviour of point defects and the determination of charge states and negative-*U* (Hubbard correlation energy) behaviour. We provide, or reference, examples from a wide range of materials including Si, SiGe, GaAs, GaP, GaN, InGaN, InAlN, and ZnO. *Published by AIP Publishing.* https://doi.org/10.1063/1.5011327

I. INTRODUCTION

Junction spectroscopy is a broad ranging term describing measurements conducted on a semiconductor junction using electrical or electro-optical techniques. Such measurements are widely used to study electrically active defects in semiconductors.

The role of the junction, usually a Schottky diode or p-n structure, is to create a depletion region. In such a region, the population of holes and electrons is low as a result of the repulsion and the drift of these carriers in the depletion field. The use of the depletion region in junction spectroscopy has two important advantages over bulk techniques. The first is that processes, which introduce a small number of carriers or a small change of charge, are easily detectable, whereas in the bulk of the semiconductor, their impact would pass unnoticed. The second is that it is much easier to manipulate the occupancy of defects producing energy levels in the bandgap within the depletion region than in bulk. This enables the properties and behaviour of such defects to be measured.

We use the term defect in its usual sense in crystalline semiconductors. Namely that a defect is any perturbation of the semiconductor lattice, this can be a vacancy, an impurity atom or cluster of atoms, and a self-interstitial or an extended defect such as a dislocation or an atom on the wrong site. In the case of compound semiconductors, an example of the latter would be an antisite, e.g., in GaAs, a Ga atom on an As site. However, some defects are electrically inactive and are not seen by the techniques described here. Usually, these do not impact directly on device performance although they may be important in some aspects of semiconductor technologies.

Surfaces and interfaces with dielectrics are a different class of defects of immense technological importance. However, in this tutorial review, we do not consider the very wide range of techniques, which enable interfaces in MOS and metal-insulator-semiconductor (MIS) structures to be studied. We mention the relevance of specific techniques, which are discussed in relation to studies of bulk defects, to such measurements, but in reality, this topic merits a full review in its own right and so we do not attempt to deal with it in any depth.

This tutorial review is about how defects in the vicinity of a junction are characterised and quantified and how we use calculations of the properties of perturbations of the semiconductor lattice to interpret the measurements. Sometimes, in this way and in combination with other techniques, it is possible to physically identify the defects in terms of their lattice structure and in the case of impurities the chemical species involved. However, what such measurements do normally provide us with are defect parameters, which describe their role in the generation, recombination, and trapping of carriers.

Defects are of crucial importance in practical terms in everyday semiconductor devices. The thermal generation of carriers via deep states causes leakage in reverse biased junctions, which affect DRAM and CMOS. In forward biased junctions, deep states are responsible for recombination processes, which reduce the efficiency of solar cells and light emitting diodes (LEDs). Trapping and release of carriers

from deep states produce noise and instability in some devices. Figure 1 summarises the three main device relevant processes associated with defects, namely, recombination, generation, and trapping. The measured defect parameters can be used to calculate lifetime and generation currents and to estimate the impact the defect has on specific devices via Shockley-Read-Hall (SRH) kinetics.[1]

Some devices use defects to achieve functionality or enhance performance. In wide bandgap semiconductors, impurities are deliberately introduced as luminescence centers, defect states are used in power devices to increase switching speeds, highly defected GaAs is used to generate THz radiation, and defects are used as qubits in quantum information processing.

Historically, deep states were studied in bulk high resistivity materials in which the concentration of deep states exceeded or was comparable with the number of free carriers. In these materials, changes in bulk conductivity could be observed when the temperature was increased due to the release of carriers from traps [i.e., a thermally stimulated current (TSC) technique applied to bulk samples rather than junctions] or when the sample was illuminated by light of different wavelengths (photo-conductivity). Techniques such as these applied to bulk semiconductors are rarely used today although the analogous bulk method of thermo-luminescence still finds application in radiation dosimetry and archaeological dating of pottery. In these applications, the release of carriers from traps in a wide bandgap material filled as a result of radiation exposure produces light which is measured on heating the insulator. This can be related to the radiation the insulator has received since a previous high temperature treatment by a subsequent calibration.

In the 1960s, several groups realised that deep states in semiconductors could be studied with much higher detectivity and with a reliable quantification of concentration using space charge techniques rather than measurements on bulk materials. One of the first publications on junction spectroscopy by Williams describes capture and emission from native defects in GaAs Schottky barriers.[2] This was followed by a sequence of papers looking in detail at the kinetics of photo-excitation in junctions. Of particular importance are papers by Goldstein and Perlman[3] and by Grimmeiss and Olofsson.[4] In parallel, work on thermal excitation of carriers

from defects in space charge regions was being published. A major article by Sah *et al.* reviews the many possible methods of optical and thermal junction spectroscopy and is still an important reference work today.[5]

One of the major problems with the junction spectroscopy methods of this era was that the results needed long hours of routine mathematical treatment to reveal the physical characteristics of the defects. This was dramatically simplified in 1974 by Lang with the invention of Deep Level Transient Spectroscopy (DLTS).[6] Lang's motivation for this work was to look at fast transients,[7] but the widespread adoption of DLTS revolves around the simplicity of the way its results are presented, namely that specific defect states appear as peaks.[6]

The analogue processing of Lang's equipment, referred to as a boxcar system, operates in a part of the noise spectrum, which is particularly quiet (unlike the earlier methods), and so this combined with its inherent signal averaging makes the detectivity high. Combined with the advantage of junction spectroscopy, it is possible to detect incredibly small concentrations of defect states and to characterise them in terms of electronic properties. In a semiconductor with a carrier concentration of 10^{15} cm^{-3}, it is usually possible to characterise electrically active defects with concentrations of 10^{10} cm^{-3}—this is a concentration in the bulk of 1 part in 10^{13}—very few analytical techniques can approach this level of sensitivity. The availability of commercial DLTS instruments has made DLTS easy to adopt and it is widely used.

However, DLTS has a major deficiency, which is that its discrimination between defects with similar emission properties is such that the definitive identification of defects through their emission fingerprint is rarely possible. In addition, the linewidth of the DLTS peak hides the impact of the local environment on the defect such as strain and alloy effects. In fact, the linewidth broadening in DLTS is not an intrinsic limitation of thermal emission techniques, and it results from the instrumental broadening of the simple analogue processing used by Lang. Many alternative schemes have been proposed, but the most satisfactory solution has been to record and average the emission transient at a specific temperature and perform an inverse Laplace transform.[8] Such processing does not give unique solutions in the presence of noise and so the road to Laplace DLTS (LDLTS) has not been easy.[9] However, it is now a mature technique and has proved to be indispensable in many cases but is not as easy to use as conventional DLTS.

An increasingly important aspect of junction spectroscopy is its combination with theory. Identification of defects with junction spectroscopy even with LDLTS is often ambiguous, and theory can provide us with generalisations and in some cases specific information on the physical structure of the defects. A previous tutorial paper in this series has discussed recent developments in combining theory and experiment in relation to defects.[10] In the following tutorial review, we consider the principles and experimental techniques of junction spectroscopy together with specific issues related to the application of theory to the interpretation of junction spectroscopy experiments.
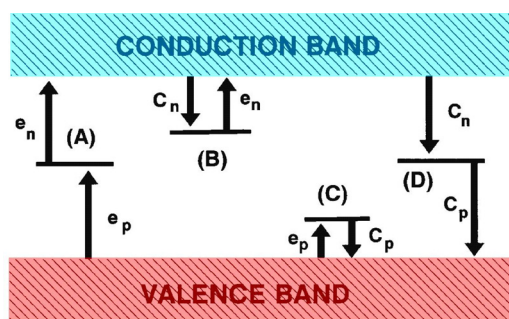

FIG. 1. Schematic band diagram showing (a) generation process via a deep state in a depletion region, (b) electron trapping in n-type, (c) hole trapping in p-type, and (d) recombination of minority carriers in the bulk. The symbols e and C denote emission and capture rates.

II. CONCEPTS

A. Electronic states of defects and transitions

A crystalline defect can be defined as a localized perturbation to the translational symmetry of the many-body crystalline potential. Defects can occur due to deliberate introduction of foreign chemical species into a crystal (electrical, optical, or magnetic dopants), due to unwanted contamination, or produced upon displacement of host atoms (intrinsic defects) due to the impact of implanted ions or irradiation, to mention just a few causes. The perturbation in the potential can be zero-dimensional as for point defects like isolated impurities, one-dimensional for linear defects like dislocations, or two-dimensional for planar defects such as stacking faults and surfaces.

Defects that are attractive to carriers are termed *electrically active*. They introduce electronic states in the forbidden bandgap, meaning that depending on their availability, carriers may become trapped at these states. Defects that are able to become positively charged (after hole capture/electron emission) or negatively charged (after electron capture/hole emission) are termed donors or acceptors, respectively. A further refinement in the classification scheme is related to the binding energy of carriers to the defect—deep acceptors are able to capture free electrons and hold them with a binding energy which is usually a significant fraction of the bandgap. On the other hand, shallow acceptors are more electronegative than the host species, so they attract negative charge from the surrounding atoms and produce empty states just above the valence band top. They are therefore able to emit holes to the valence band with a small activation energy in the meV range. While deep acceptor states are usually localized on a few ligands around the defect, for shallow defects, the acceptor state overlaps a volume which can span many crystalline unit cells, depending mostly on the dielectric properties of the medium. Analogous arguments could be applied for donors, where the shallow ones are more electropositive than the surrounding atoms. We also note that the above classification is not unique. We can, for instance, distinguish shallow from deep acceptors as those that are essentially ionized (negatively charged) at room temperature.

The charge state of an electrically active defect depends on the location of the Fermi energy (E_F) with respect to its transition level ($q/q+1$). The value of E_F for which equal fractions of q- and ($q+1$)-charged defects are found in equilibrium defines the location of the level within the gap. This is a thermodynamic quantity and should not be confused with single-particle states or even optical (vertical) transitions. Defects can have several transition levels. They may be both donors and acceptors, double donors, triple acceptors, etc., and the level ordering is not necessarily sequential as for an isolated atom. Here, the magnitude of the ($n+1$)-th ionization energy (I_{n+1}) is always greater than the n-th one (I_n), simply due to the fact some Coulomb repulsion will be lost after the n-th ionization. The difference $U = I_{n+1} - I_n$, often referred to as Hubbard correlation energy is therefore positive. For some defects, ionization of states in the gap destabilizes the local atomic structure, and that can dramatically change their electronic properties. If the ionization energy of the new

structure becomes lower than the previous one, a second ionization will proceed spontaneously. Defects behaving like that are said to possess a negative-U as they show an inverted order of energy levels.[11] This is possible due to the exchange of some of the Coulomb energy (binding the gap states) by re-bonding (chemical) energy along the ionization sequence. Defects can also be electrically inert or inactive, although they may still interact with carriers, for instance, via scattering events. An example of an inert defect is interstitial oxygen in Si: the strong Si-O bonding states are deep within the valence band, whereas anti-bonding states lie high within the conduction band. This effectively prevents any electronic transition from taking place between free carrier states and defect states.

The exchange of carriers between bound and crystalline states may be measured by several techniques. When photons are not involved in the measured transitions, they are referred to as non-radiative and mostly take place via Auger or multi-phonon emission (MPE) mechanisms. The Auger process involves the conversion of the energy from the captured carrier to excite a nearby electron, whereas in MPE the released energy is dissipated via emission of lattice vibrations, and when these are strongly localized, it can result in the reconfiguration or even migration of the defect.

Figure 2 depicts the capture of a free hole, h^+ (for instance, produced after exciting the sample with above bandgap light, $h\nu \geq E_g$) by a singly negatively charged trap with atomic configuration Q_-. We assume that the defect has a single vibrational degree of freedom. Before capture, the system is found at the upper energy state lying E_g above the ground state (top of the figure). The energy difference

FIG. 2. Configuration coordinate diagram for an electron trap located at $E_c - E_t$ below the conduction band minimum energy. Each parabola represents a vibronic state as a function of a generalized coordinate of atoms. Horizontal segments represent phonon quanta of energy $\hbar\omega$. At the bottom, we show the singly negatively charged state (t^-), which can emit an electron and become neutral (t^0). At the top, we show the t^- state plus an uncorrelated electron-hole pair. Thermally activated rates for electron capture and emission by t^0 and t^- are C_n and e_n, respectively. Hole and electron capture barriers are $\Delta E_{\sigma p}$ and $\Delta E_{\sigma n}$, respectively. Optical (vertical) absorption and luminescence transitions are labelled as E_{abs} and E_{lum}, respectively.

between this excited state and the ground state is equal to the energy needed to create a pair of uncorrelated $e^- + h^+$ carriers (free electron and free hole). For sufficiently high temperatures and large vibrations, the overlap between vibronic states $|t^0 + e^-, n\hbar\omega\rangle$ and $|t^- + e^- + h^+, m\hbar\omega\rangle$ may become large enough to allow level crossing and therefore to end up with a captured hole. After overcoming a capture barrier $\Delta E_{\sigma p}$, atoms within the trap wavefunction will relax to a new equilibrium configuration Q_0, and during this damping step, the localized excess of energy will be dissipated as lattice phonons. Figure 2 also shows how thermally stimulated MPE transitions differ from optical transitions where the Frank-Condon principle holds. The latter are represented by absorption (E_{abs}) and luminescence (E_{lum}) energies involving ground and excited states t^- and $t^0 + e^-$, respectively. For further details on MPE theory, the reader is directed to Refs. 12–14.

Another important effect that involves MPE is *non-radiative recombination*. After carrier trapping, a second capture of a carrier of opposite charge (an electron in the case of Fig. 2) may take place, leading to a recombination event. This is a major detrimental effect in optoelectronic devices such as LEDs and solar cells, whose functioning depends on both carrier types. The kinetics of recombination at deep traps is usually described using Shockley-Read-Hall (SRH) statistics,[1] which ultimately depends on the individual carrier capture rates. The rate at which an electron at a conduction band state is captured by a defect trap is given by

$$C_n = \sigma_n \langle v_n \rangle n, \tag{1}$$

where n is the concentration of free-electrons with thermal velocity $\langle v_n \rangle = (3k_B T/m^*)^{1/2}$ and effective mass m^*, while σ_n is the electron capture cross-section which is usually given in units of cm^2. On the other hand, electron emission from a trap to the conduction band bottom takes place at a rate

$$e_n = \sigma_n \langle v_n \rangle \frac{g_0}{g_1} N_c \exp\left(-\frac{E_c - E_t}{k_B T}\right), \tag{2}$$

with g_0 and g_1 being the degeneracy of the state before and after emission, respectively, N_c is the effective density of states in the conduction band, and $E_c - E_t$ is the depth of the trap with respect to the conduction band minimum. Analogous equations can be written for the case of hole trapping and emission.

The capture cross section depends mostly on the probability of crossing between bound and unbound states and also depends on the re-emission probability.[13] The crossing probability has a thermally activated term which is proportional to $\exp(-\Delta E_\sigma/k_B T)$, where ΔE_σ is usually referred to as the capture barrier, but also depends on the overlap between vibronic states near the crossing point. This quantity is highly dependent on the atomistic and electronic details of the defect in both charge states. The chance that a carrier is re-emitted to a crystalline band mostly depends on the carrier velocity, which is also temperature dependent. It is obvious that a faster carrier is more difficult to be trapped by an attractive potential than a slow one. In the last few years,

some progress has been made in the evaluation of capture cross sections of defect states by first-principles calculations. Most difficulties lie in the calculation of accurate electron-phonon coupling properties near the transition state. We single out the works from Refs. 15 and 16, where both the vibrational and the electron-phonon parts of the carrier-capture problem were completely addressed from first-principles calculations.

B. Defect thermodynamics

Depending on the defect concentration, defect-defect interactions may not be negligible. They are in fact particularly relevant for charged defects in insulators and for those that induce long-range (LR) strain fields. However, there are several factors that often support the validity of a dilute picture. For example, covalent bonding between impurities and their ligands is usually short-ranged (SR), for defects in metals, the screening by free-electrons may quench long-range Coulomb interactions, and, when the defect concentration is low enough, we may assume them to be effectively isolated. In this case, the actual concentration of defects may be evaluated from the general Gibbs free-energy function G of a crystal in contact with its constituents (in their respective standard phases), by imposing the equilibrium condition[17]

$$\delta G = 0 = \sum_i (\partial G/\partial N_i)_{N_j, T, P} \delta N_i, \tag{3}$$

where N_i is the number of constituents of species i with chemical potential given by the derivative $\mu_i = (\partial G/\partial N_i)_{N_j, T, P}$, evaluated at constant temperature, pressure, and number of conjugate elements. We note that the index i may refer to both chemical species and electrons. In semiconductors and insulators, the electronic chemical potential μ_e is the Fermi energy $\mu_e \equiv E_F$. Here, the Fermi energy is limited by $0 \leq E_F - E_v \leq E_g$, where E_g is the bandgap of the material and E_v is the valence band top. As we will see below, chemical potentials in an arbitrary material are usually obtained with the help of *standard chemical potentials* μ_i^0, which stand for the energy per element in its respective standard phase. For instance, for the study of intrinsic defects in ZnO, μ_{Zn} and μ_O depend on μ_{Zn}^0 and μ_O^0, which are obtained from the zinc metal (hexagonal close-packed) and molecular oxygen, respectively.

The free energy of formation needed to introduce a defect in a crystal is given by $G_f = E - \sum_i N_i \mu_i + PV - TS$, where E is the internal energy of the whole defective crystal made of N_i elements of species i. The PV term is negligible, and for the moment, we also neglect the contribution of entropy, S (although the latter assumption may be questionable, particularly at high temperatures). Entropy is often not considered in the calculation of the formation energy of defects. Besides the large effort in obtaining configurational, electronic, vibrational, or elastic entropy terms, one of the reasons for assuming Eq. (4) independent of temperature, particularly for defects in semiconductors and insulators, is the existence of other and more severe sources of error, like the underestimation of the calculated bandgap using local and semi-local approaches to describe the electron-electron interactions. Hence

$$G_{\mathrm{f}} = E(\boldsymbol{R}, q) - \sum_i N_i \mu_i + q E_{\mathrm{F}}, \qquad (4)$$

where the internal energy depends explicitly on the defect structure $\boldsymbol{R}$ and charge state q (which is positive or negative when a defect level within the gap is depleted of or filled with electrons with respect to the neutral state, respectively).

For studying defects in compound crystals, it is particularly useful to define the energy per formula unit (of the crystal) as $\mu_{\mathrm{fu}}^0 = \sum_i n_i \mu_i$, where n_i is the number of atomic species i in the chemical formula. We can also define the heat of formation of the compound (per formula unit) as

$$\Delta H_{\mathrm{fu}}^0 = \mu_{\mathrm{fu}}^0 - \sum_i n_i \mu_i^0, \qquad (5)$$

which is a negative quantity for a compound that is stable against decomposition into its constituent standard phases. From the energy per formula unit and Eq. (5), we obtain

$$\sum_i n_i (\mu_i - \mu_i^0) = \sum_i n_i \Delta \mu_i = \Delta H_{\mathrm{fu}}^0, \qquad (6)$$

implying that μ_i values satisfying Eq. (6) are not unique. They actually depend on the growth conditions (e.g., partial pressure of gas sources), although they are limited by their values in standard phases

$$\mu_i \le \mu_i^0, \qquad (7)$$

with $\mu_i = \mu_i^0$ taking place when there is an excess of i-elements during growth, leading to an instability of the compound against segregation of the standard phase of species i. Combining Eqs. (5) and (7), we arrive at the limits imposed to μ_i under equilibrium conditions

$$\mu_i^0 + \frac{\Delta H_{\mathrm{fu}}^0}{n_i} \le \mu_i \le \mu_i^0, \qquad (8)$$

with the upper and lower bounds representing i-rich and i-poor growth conditions, respectively. The stoichiometric case is obtained midway for $\mu_i = \mu_i^0 + \Delta H_{\mathrm{fu}}^0 / 2 n_i$. We now write a more convenient expression for the formation energy of a defect, obtained by adding ΔN_i atoms of species i to a pristine crystalline sample made of N_{fu} formula units (for atom removal $\Delta N_i < 0$)

$$G_{\mathrm{f}} = E(\boldsymbol{R}, q) - N_{\mathrm{fu}} \mu_{\mathrm{fu}}^0 - \sum_i \Delta N_i \mu_i + q E_{\mathrm{F}}, \qquad (9)$$

with the μ_i values subject to Eqs. (6) and (8). The formation energy depends linearly on $q E_F$, so that the stability of a specific charge state will depend on the Fermi energy. This makes Eq. (9) useful to calculate the position of charge state transitions of defects within the bandgap. Accordingly, the Fermi energy at which the defect formation energies of two consecutive charge states cross, defines a transition energy level. The $E(q/q+1)$ transition level is then found when $G_f(\boldsymbol{R}_q, q, \mu_i, E_F) = G_f(\boldsymbol{R}_{q+1}, q+1, \mu_i, E_F)$, and this takes place at

$$E(q/q+1) = E(\boldsymbol{R}_q, q) - E(\boldsymbol{R}_{q+1}, q+1). \qquad (10)$$

We underline the fact that the atomistic structure $\boldsymbol{R}$ that leads to the internal energies E for charge states q and $q+1$ is not necessarily the same. For defect levels under equilibrium conditions (also referred to as *occupancy levels*), $\boldsymbol{R}$ values correspond to ground-state coordinates. On the other hand, for optical (vertical) transitions, $\boldsymbol{R}_q = \boldsymbol{R}_{q+1}$.

III. MEASUREMENT STRUCTURES

As stated in the introduction, usually Schottky barrier (metal-semiconductor) diodes (SBDs) or p-n structures are used for characterization of deep-level defects with the use of junction spectroscopy methods. Measurements can also be carried out on metal-insulator-semiconductor (MIS) structures and on different types of transistors. Typically, the measurement procedure consists of three stages: (a) the occupancy of a deep state is set to an equilibrium value; (b) the occupancy is perturbed; and (c) the subsequent change in occupancy is measured as a function of time, directly or indirectly, as the state returns to its initial equilibrium condition. This change can be monitored in many ways, particularly by measurements of the current produced by the emitted carriers or AC conductance; however, measurements of depletion capacitance are used most frequently. In this part of the paper, some basic principles of junction measurements and requirements for the measurement structures are briefly discussed.

All junction structures must have contacts for incorporation into a circuit for the measurements of the electrical parameters. The contacts for the p-n structures to be used in junction spectroscopy must be Ohmic, i.e., have linear or quasi-linear current-voltage (I-V) characteristics and a voltage drop on them must be negligible in comparison with the total voltage drop on a junction structure. In the absence of a p-n junction, the sample requires a rectifying Schottky barrier as well as an Ohmic contact to create the depletion region. The Schottky model of metal-semiconductor barriers suggests that Ohmic contacts to a semiconductor can be manufactured by the deposition of metals with a lower work function than that for an n-type semiconductor material, and in the case of a p-type semiconductor, the metal should have a higher work function than the semiconductor.[18] In practice, however, the barrier height is often very dependent on surface states which pin the Fermi level and in some cases make the barrier height independent of the work function of the metal so even for seemingly appropriate combinations of metal-semiconductor pairs, it is difficult to make good quality Ohmic contacts to semiconductor materials with high resistivity. So, for the production of Ohmic contacts, thin layers of highly doped semiconductor are often incorporated into the test structure. Such layers can be manufactured by introducing atoms with shallow doping properties either by ion implantation or by deposition onto the semiconductor surface followed by thermal activation and diffusion treatments. Similarly, it is not easy to predict the best metal to produce a good rectifying contact to a specific semiconductor but empirical information is generally available in the literature.[18]

For characterization of the measurement junction structures, current-voltage (I-V) and capacitance-voltage (C-V) measurements are carried out.[18–22] From such measurements,

a number of parameters of the structures, such as barrier height, doping concentration, and width of the depletion region, can be determined. Some of these parameters are necessary for the determination of characteristics of deep states from further transient measurements.

An analysis of I-V dependencies of p-n or Schottky barrier diodes allows one to estimate the quality of a measurement structure and to determine its parameters such as the barrier height and series resistance. Equations which describe I-V dependencies for p-n structures and Schottky barrier diodes and possible ways of treatments of experimental I-V curves can be found in Refs. 18 and 19 and references therein.

Capacitance-voltage measurements of junction structures are very widely used in various forms to measure the ionized dopant concentration (in the case of shallow dopants, this is equal to the carrier concentration), depletion width, distribution of the electric field, and built-in voltage.[19–22] These parameters are essential prerequisites to the characterization of deep-level defects. The C-V technique relies on the fact that the depletion width changes as a function of applied voltage. It measures the change of capacitance of a junction structure as a function of applied bias.

For illustration of the C-V measurement concepts, a Schottky barrier diode (SBD) on a p-type semiconductor has been chosen since this structure is frequently used as an experimental probe on semiconductor layers. In this case, one contact is Ohmic while the other forms the Schottky barrier. Almost identical arguments can be applied to n^+p junctions, where the n-type layer is very heavily doped compared to the p-type region. The total charge in the depletion region of the SBD can be expressed as

$$Q = qN_AWA, \qquad (11)$$

where q is the elementary charge, N_A is the concentration of ionized acceptors, W is the width of the depletion region, and A is the area of the diode. The differential capacitance of the diodes is

$$C = -\frac{dQ}{dV}, \qquad (12)$$

where V is the junction voltage. Capacitance is measured in the small-signal high-frequency regime, i.e., a small ac sinusoidal voltage is imposed on any quiescent bias, V_R. The measurement frequency is typically 1 MHz, and the ac voltage amplitude is typically $10\,\mathrm{mV} \leq V_{osc} \leq 100\,\mathrm{mV}$. The measuring instrument senses the ac current flow resulting from the oscillator voltage and from the complex impedance extracts the capacitance. It is assumed that the depletion approximation holds good, i.e., the electron density at the edge of the depletion layer rises from zero to its bulk value over a distance which is negligible. The diode is reversed biased (a positive potential is applied to the metal contact on the p-type substrate) so that the total voltage across the structure is

$$V = V_R + V_{bi}, \qquad (13)$$

where V_R is the applied voltage and V_{bi} is the built-in potential. So

$$C = -\frac{dQ}{dV} = \frac{\varepsilon_0\varepsilon_s A}{W} = qAN_A(W)\frac{dW}{dV}. \qquad (14)$$

The only way that Q can vary with applied voltage is by expansion or contraction of the depletion layer. During the incremental reverse bias change due to the sinusoidal ac test voltage, the depletion edge moves out by some increment δW adding to the barrier depletion charge $qN_D\delta W$. During the incremental change of the sinusoid towards forward bias, the depletion edge moves in reducing the depletion charge by the same increment. From Eq. (14), the so called barrier capacitance of the diode can be found as

$$C = A\left[\frac{q\varepsilon_0\varepsilon_s N_A}{2(V_R + V_{bi})}\right]^{1/2}. \qquad (15)$$

Equation (15) can be rewritten as

$$\left(\frac{A}{C}\right)^2 = \frac{2(V_R + V_{bi})}{q\varepsilon_0\varepsilon_s N_A}. \qquad (16)$$

Plotting $(A/C)^2$ against V_R produces a straight line, provided that N_A is uniform. The slope relates to N_A and the intercept on the voltage axis gives V_{bi}. In fact, the value of N_A in the equation is strictly the value of N_A at the depletion edge associated with value of W corresponding to the total potential drop V, and this fact can be used to obtain carrier profiles from C–V measurements.

In a diode with a uniform N_A, the electric field strength, E, decreases linearly with increasing W.[18,20] The maximum value of E_{max} at the metal-semiconductor interface can be determined as[18–20]

$$E_{max} = \left(\frac{2qN_AV}{\varepsilon_0\varepsilon_s}\right)^{1/2} = \frac{2(V_R + V_{bi})}{W}. \qquad (17)$$

Equation (17) shows that E_{max} increases with increasing V_R. It can reach the critical electric field strength, E_{BD}, causing electrical breakdown of the diode. The reverse bias voltage corresponding to the E_{BD} is called breakdown voltage, V_{BD}. According to Eq. (17), the V_{BD} value is a function of semiconductor doping concentration.

Quite obviously, it is possible to profile concentrations of shallow dopants [Eqs. (15) and (16)] and to use junction spectroscopy techniques for studies of majority carrier traps only over the region of the semiconductor through which the depletion region can be swept. In practice, this is limited by a minimum value near the zero bias point and a maximum W at the breakdown voltage. For example, if we consider Si doped to $10^{17}\,\mathrm{cm}^{-3}$, the minimum depth that can be sampled at zero bias is $0.1\,\mu$m. Breakdown occurs at 12 V reverse bias, which is equivalent to a depletion depth of $0.4\,\mu$m. These values increase as the doping level decreases.

It has already been mentioned that Eqs. (13)–(17) are correct within the depletion approximation, i.e., under the assumption that no free carriers exist within the space charge region and that the boundary to the electrically neutral region

is sharp. In reality, there is a small transition region, known as Debye tail, which is only partly depleted of free carriers. The Debye length can be calculated as[20]

$$L_D = \left[\frac{\varepsilon_0 \varepsilon_s k_B T}{q^2 N_A} \right]^{1/2}. \tag{18}$$

The L_D value indicates the width of this transition region. The assumption of the abrupt change in the free carrier concentration between the space charge region and the neutral region is generally satisfied when the width of the space charge region is much longer than the Debye length. The Debye tail related effects can sometimes play a significant role in characterization of deep-level defects by junction spectroscopy techniques and should be taken into consideration.

IV. TECHNIQUES

In this section, we will explain how the concepts described in Sec. II are applied to the structures described in Sec. III to determine the characteristics of defects in semiconductors.

In any measurement of the electrical properties of defects as outlined previously, there are two essential processes: the first is to establish a well defined carrier occupancy of the defect by supplying holes or electrons. This is most frequently done in junction spectroscopy by changing the bias on the diode junction and so shifting the Fermi level or, as we will discuss later, by optical excitation. The second stage is to observe the change of occupancy as the defect returns to equilibrium under different bias conditions or under optical excitation. The electron occupancy f_T at equilibrium is determined by the balance of the electron and hole capture and the carrier emission processes according to

$$f_T = \frac{c_n n + e_p}{e_n + c_p p + c_n n + e_p}. \tag{19}$$

A defect full of electrons has electron occupancy of unity. The symbols e_n and e_p represent the electron and hole emission rates with units of s^{-1}, while c_n and c_p represent the electron and hole capture coefficients with units of $cm^3 \ s^{-1}$. Upper case C_n or C_p is used to denote capture rates with units of s^{-1}, and so, $C_n = n c_n$. This is the conventional notation although in some texts uppercase and lower case are interchanged in this context. The capture cross section σ which we have referred to previously has units of cm^2 and is related to the capture coefficient for the electron case by $\sigma_n = c_n / V_{th}$, where V_{th} is the thermal velocity of electrons.

The change in occupancy is most often quantified in junction spectroscopy by a measurement of capacitance or current transients followed by some form of data processing. Although there are a bewildering number of experimental techniques that all revolve around the concept above, i.e., set the carrier occupancy, change the experimental conditions, and observe the change in occupancy, the techniques are differentiated by different methods of setting the majority or minority carrier occupancy, different methods of relaxing that occupancy, and different signal processing of the resultant transients.

A. Thermally stimulated current and thermally stimulated capacitance (TSCAP) techniques

In the classic junction spectroscopy technique, the defect is filled with majority carriers and then the state is emptied thermally. The simplest experiment to demonstrate this is the technique of thermally stimulated current (TSC) which, although rarely used today, has played an important role in the development of junction spectroscopy. Little equipment is needed to undertake this measurement. The sample is mounted in a cryostat, and although TSC can be used on high resistivity bulk samples (and was originally developed for such materials), we will consider the junction spectroscopy case where we are using an n-type Schottky barrier for the measurement. The process is presented schematically in Fig. 1 as (b). A small forward bias is applied to the diode so as to achieve a near flat band condition with the sample cooled to low temperature.

Under these conditions, the deep states will be filled with majority carriers. The diode voltage is then changed so as to apply a reverse or zero bias. The voltage is limited by what is an acceptable leakage current; usually, <1 pA is the target. The value of bias used also defines the depletion width [Eqs. (14) and (15)] and hence the volume of semiconductor which will be measured. The temperature is then increased at a defined rate, and the thermal release of carriers is monitored.

Figure 3 shows a TSC scan of GaP doped with the shallow donor silicon and nitrogen which is iso-valent with phosphorous.[23] In this case, the temperature at which the occupancy is defined and the scan started is 100 K. On switching to zero bias, a current is seen which represents electron emission from defect states in the depletion region which are sufficiently shallow to be thermally ionised at 100 K. The temperature is then ramped up, in this case at 0.81 K s^{-1}. It can be seen that at $\sim$160 K, a current starts to flow resulting from a defect state at $E_C - E_T = 0.42$ eV. The area under the curve (with the horizontal axis converted to time) gives the charge released. Assuming that each defect releases one electron and knowing the diode area and depletion width, the concentration of defects can be calculated. In this case, $Q_t = 1.36 \times 10^{-10}$ C corresponding to $N_t = 1.52 \times 10^{15}$ cm^{-3}. The peak current of $\sim 10^{-11}$ A is typical of this
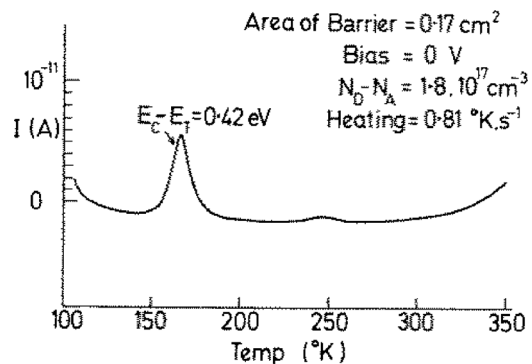

FIG. 3. Thermally stimulated current measurement of n-type GaP:Si,N showing a deep state attributed to a complex containing iso-valent nitrogen. Reproduced with permission from Appl. Phys. Lett. **26**, 122 (1975). Copyright 1975 AIP Publishing.

type of measurement. Using a faster ramp rate increases the amplitude of the peak. However, under these conditions, the peak becomes narrower and the area under the curve remains constant. If deeper traps exist in the sample, they can give rise to peaks at higher temperatures; in this sample, a peak related to a defect with a much lower concentration is seen in the spectrum at about 250 K.

It is conceptually instructive to look at changes in the capacitance equivalent of TSC and referred to as TSCAP (Thermally Stimulated Capacitance). In this technique, the depletion capacitance is measured as a function of temperature after filling the traps with carriers. The physical processes are identical to TSC. In Fig. 4, TSC and TSCAP measurements are compared.[24] In this case, the sample is a n$^+$p Si diode doped with Au. As the n-type region has a very high carrier concentration (hence the designation is n$^+$), so the depletion region spreads predominantly into the p-region and the experiment explores the properties of traps in the p-region which in this case are hole traps.

The TSCAP measurements are undertaken with the Au traps filled with holes and empty of holes as initial conditions. At 77 K, the thermal emission of holes from the Au is negligible, but as the temperature rises, a slow change in capacitance is observed for both cases of the defect being full and empty of holes, which is due to changes of the diffusion voltage and dielectric constant as a function of temperature. In the case where the initial condition was the Au being full of holes, at ∼140 K, a sudden increase in capacitance is

observed due to the emission of holes from the Au. In the TSC plot in the lower part of the diagram, this corresponds to a current peak due to the release of charge. Essentially, as detailed in Sec. II, the carrier emission follows the pattern shown in Fig. 4 and dictated by

$$e_p = A_p \exp(-E_a/kT). \qquad (20)$$

In the TSC and TSCAP methods, it is very difficult to determine E_a although an estimate can be made by varying the ramp rate. This presents experimental problems as at high ramp rates the determination of the sample temperature is fraught with difficulties. In consequence, it became standard practice to determine the activation energy of a defect using a sequence of isothermal measurements at different temperatures. Essentially, this is done by recording the capacitance or current at the selected temperature after filling the trap and then determining the time constant of the exponential change. For a majority carrier trap, the form of the decay is

$$\Delta C(t) = \Delta C_{(total)}(1 - e^{-t/\tau}). \qquad (21)$$

As will be discussed in Sec. IV C, similar methods can be used to study minority carrier traps, i.e., in this case, electrons in p-type. The equation is then of the form

$$\Delta C(t) = \Delta C_{(total)}(e^{-t/\tau}). \qquad (22)$$

The capacitance transients related to majority and minority carrier emission are shown in Fig. 5. The measurement is repeated at different temperatures, and these data are then used to construct an Arrhenius plot from which the activation energy is determined. This is calculated from the slope of a plot of log (e_n/T^2) against $1/T$. The intercept of the plot with the log(e_n/T^2) axis at $1/T = 0$ is used to calculate the apparent or intercept capture cross section. This can be very different to the directly measured capture cross section, and the two should never be used interchangeably.

B. Development of transient capacitance and current techniques (DLTS)

TSC and TSCAP suffer from major deficiencies, in particular, lack of sensitivity, poor discrimination between different defect states, and difficulties in quantifying defect properties. In addition, the measurement of isothermal

FIG. 4. TSCAP and TSC measurements on the same Si n$^+$p junction doped with Au. The horizontal axis plotted as time represents temperature change. The dotted line in the TSCAP result is the difference between the capacitance measurements with different initial conditions representing the Au states full and empty of holes. Reproduced with permission from Appl. Phys. Lett. **20**, 193 (1972). Copyright 1972 AIP Publishing.

FIG. 5. Capacitance–time transients showing majority and minority carrier emission following an appropriate filling pulse.

transients which is necessary for quantification was extremely tedious using the technologies of the 1970s. A consequence of this was that deep level measurements remained the province of specialists, and the techniques were not widely used. However, in 1974, Lang devised an elegant technique,[6] which resulted in the widespread adoption of defect measurements in semiconductor science. This is Deep Level Transient Spectroscopy (DLTS).

Like TSCAP and TSC, it uses a temperature scan to reveal different defect states and is based on filling the defect in a depletion region with carriers and then observing their thermal release. The difference lies in that the defect population is repetitively filled by changing the bias on the diode using an electrical "filling pulse" superimposed on the reverse bias used during the measurement phase. The transient is analysed by sampling at two points t_1 and t_2 as shown in Fig. 6. The DLTS signal is simply the difference in capacitance measured at t_1 and t_2. This difference is averaged over many cycles of the measurement. The transient at the top of the diagram (the highest temperature shown) is so fast that $C(t_1) = C(t_2)$ and hence the DLTS signal is zero. Similarly, at the lowest temperature shown, there is no difference output. However, at some intermediate temperature, the DLTS output reaches a maximum. Assuming that the transient decay is exponential, the DLTS peak will occur when its time constant matches the "rate window." This time constant is given by[6]

$$\tau_{e,\max} = \frac{t_2 - t_1}{\ln(t_2/t_1)}. \tag{23}$$

The t_1 and t_2 times are measured from the end of the filling pulse, and it is crucially important that the sampling times are precisely synchronised with filling pulse.

This method of extracting the time constant of the thermal emission from the defect has become known as the double boxcar technique. It has many virtues, very importantly the output is independent of the baseline, i.e., the background
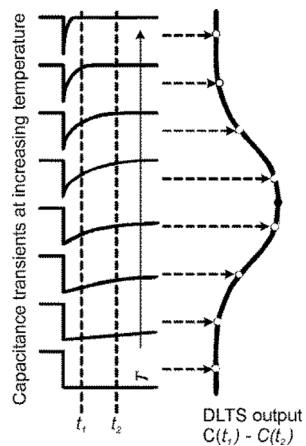
capacitance and the output can be averaged over many filling and emptying cycles so reducing the noise. The values of t_1 and t_2 can be changed, so the temperature at which the peak has a maximum can be determined at different rate windows and the time constant determined as a function of temperature.

The width of the sampling times at t_1 and t_2 is significant in two ways. First, by taking only an instantaneous sample of the transient, some information is lost and so the signal to noise ratio is degraded and information regarding any deviations from the ideal exponential behaviour is not immediately evident. The noise problem can be improved by taking longer samples at t_1 and t_2. However, Eq. (23) is strictly only correct for the case where the sample time is short compared to the overall cycle time and so the actual rate window needs to be calculated in a more complex way depending on the way the signal is averaged during t_1 and t_2. This is referred to later when we consider lock-in methods of DLTS in Sec. IV B 1. The magnitude of the DLTS signal is less than the total capacitance change as only part of the transient is sampled but if t_1/t_2 is maintained constant, while t_1 and t_2 are changed to alter the rate window, the ratio of the DLTS signal to the magnitude of the transient remains constant and a single correction factor can be used to get the magnitude value.

Figure 7 shows two DLTS spectra recorded with different rate windows in a DLTS study of irradiated silicon. It can be seen that the peaks in the spectrum recorded with the highest rate window (corresponding to the higher emission rate) occur at higher temperatures as would be predicted, Eq. (20). The corresponding Arrhenius plots are shown in Fig. 8 derived from the emission data taken at many temperatures (ten in this case).

It is instructive to note that the divacancy in n-type Si can exist in three charge states, from doubly negatively charged to neutral. In the case of one electron occupying the

FIG. 6. Schematic of the rate window principle. The horizontal axis represents the time, and the vertical axis represents the capacitance change at different temperatures (highest at the top). The DLTS output is the difference between the capacitance measured at times t_1 and t_2. Adapted with permission from J. Appl. Phys. **45**, 3023 (1974). Copyright 1974 AIP Publishing.
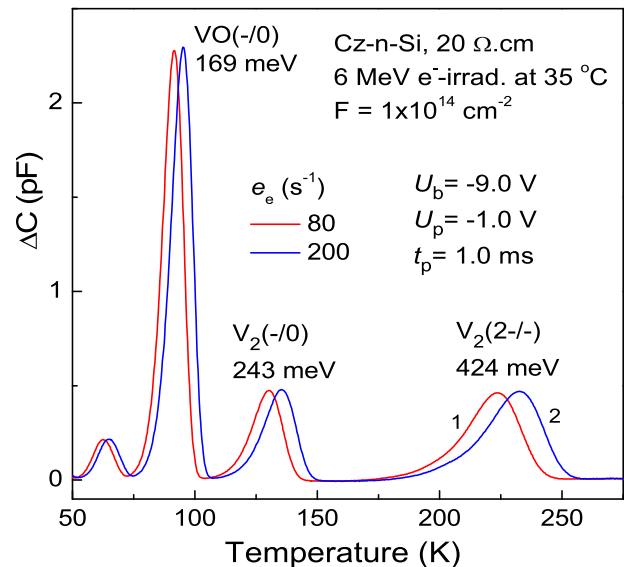

FIG. 7. DLTS spectra taken at two rate windows of electron irradiated n-type Czochralski grown silicon showing the signals from the vacancy-oxygen defect and two charge states of the Si divacancy.
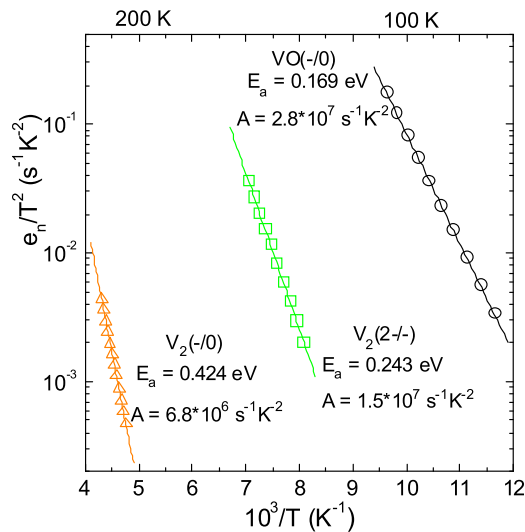
FIG. 8. Arrhenius plots of the thermal emission of electrons from three common defects in implanted or irradiated silicon. These are the vacancy oxygen complex and two charge states of the divacancy. The energy is derived from the slope of the plot and A_n is determined by the intercept with the vertical axis at $T = \infty$.

state where we observe the emission of an electron leaving behind a neutral state (notated $V_2^{-/0}$), the activation energy is higher than that for the release of the second electron giving rise to the reaction $V_2^{2-} \rightarrow V_2^{-} + e^{-}$. This is the usual case of positive Hubbard correlation energy, U. If the energy to release the second electron is higher than that for the release of the first electron, obviously both electrons would be released simultaneously and we would see only one DLTS signal for a defect emitting two electrons. Such a defect property is known as negative U and is discussed briefly in Sec. II of this paper. It should be pointed out that the DLTS signal from a negative-U defect has twice the peak magnitude compared to that for a signal related to traps releasing just one electron. So, care should be taken to prevent a misinterpretation of the defect concentration in cases of the negative-U traps.

The DLTS technique is typically used with a capacitance measurement of the junction, but it can be applied to current measurements. There are important fundamental differences as well as different experimental problems. We will discuss in Sec. IV C junction techniques for characterization of minority carrier traps, but a key difference between capacitance and current transients measurements is that in current measurements, hole and electron transitions are of the same sign, but in capacitance, they can easily be distinguished as they are of opposite sign in Schottky diodes and highly asymmetrical pn junctions ($p^{+}n$ or pn^{+}). In pn junctions with doping levels where the depletion region spreads significantly into both the p and the n regions, it is not possible to decide whether hole or electron emission is being observed even from capacitance measurements. The response time of both current and capacitance measurements can affect the DLTS measurement, but a particular problem of capacitance measurements is that if the diode under test is subjected to a forward bias, the measured capacitance is very high and most capacitance measuring devices overload and do not

give accurate readings for several microseconds. This can cause serious errors in DLTS measurements, and the system must be designed to minimise such overloads, but in addition, the first part of the transient must not be used in analysis of the measurement. These issues and the effect of sampling time have been considered by Day *et al.*[25]

1. Lock-in and correlation methods for time constant extraction

Boxcar equipment is not widely used in semiconductor labs, and this leads to some users developing a simplification of the DLTS technique, namely, to use lock-in based treatments of transients.[26] This equipment normally samples the entire time cycle of a repetitive signal weighting it in two equal parts. The most common weighting function is a square wave with a frequency f chosen to relate to the time constant of the thermal emission exponent so that at the peak response $\tau = 0.42/f$. Because the filling pulse and capacitance meter overload are part of the repetitive cycle, the capacitance signals during the fill pulse and shortly afterwards have to be gated out of the signal to be analysed. This results in complications not only in terms of the derived amplitude of the transient but also in the relationship of the time constant to the lock-in frequency and issues of how to synchronise the lock-in with the sampling pulses in terms of phase. These problems have been considered in depth by Auret.[27]

Conventional DLTS techniques produce an elegant output presenting defects as peaks on a plot of DLTS signal against temperature. Unfortunately, the method does not provide good discrimination between defects of similar energy or more precisely between defects with emission rates close together. This can be understood by looking at the DLTS peak widths. These are very much larger than would be expected from thermal broadening. The origin of the broadening is essentially instrumental, basically because Lang's boxcar technique is extremely crude in signal processing terms. Subsequently, many improvements have been proposed.

Lang was acutely aware of the resolution problem associated with DLTS and so his colleagues at Bell labs[28] and others[29] began working on analogue correlation techniques which might provide a better resolution of emission rates. A key question is the trade-off between detectivity and resolution and although better resolution was obtained, the limitations of analogue techniques resulted in these correlation methods being very difficult to use. Essentially, the community concluded that Lang's method was a good compromise between resolution, noise, and convenience.

The availability of fast 12 bit A to D convertors transformed approaches to DLTS towards the end of the 1980s, and digital systems replicating the box-car and lock-in functions have now almost completely displaced analogue signal processing. Very importantly, digital systems enabled correlation techniques to be implemented in a flexible way including Fourier transform processing. However, fundamentally, the most elegant mathematical solution for analysing

exponents is the inverse Laplace transform to convert the transient signal from the time into the frequency domain.

2. Laplace deep level transient spectroscopy

Extracting exponents with similar time constants from experimental data, on which noise is super-imposed and which has an unknown baseline, is an extremely difficult task. In general, the ability to distinguish closely spaced time constants, the dynamic range of transient amplitudes, and the emission rate range measurable are all dependent on the signal to noise ratio. A comprehensive review of the experimental and analytical issues associated with exponential analysis has been published by Istratov and Vyvenko.[30]

The basic problem is that in the presence of noise, there is no unique solution to de-convoluting a transient into multiple closely spaced exponents with different time constants; in mathematical terms, the problem is ill posed. In transient capacitance or current measurements, this is the normal situation when several transients are superimposed. Lang's technique partially avoids this problem by not separating closely spaced transients. Approximations to the inverse Laplace transform provide a mathematical basis for this, but the ill posed problem remains, so it is always necessary to consider what might constitute a realistic answer to the problem, i.e., does the signal to noise ratio justify a solution with multiple exponential components and if so how many. This process is crucial in avoiding results with no experimental justification and in respect of the Laplace transform is referred to as the regularization process. The Tikhonov regularization method[31] is very effective and can be used with the publicly available numerical calculation CONTIN[32] for the LDLTS case.

Early results using the Laplace transform for transient spectroscopy were not very successful until a realisation that the only way to obtain high resolution of time constants was to design an extremely low noise system. Included in the generic term noise is temperature instability and so "High Resolution" or Laplace DLTS systems as they are now known use isothermal measurements with cryostats of high thermal stability, typically maintaining the temperature to $< \pm 20$ mK for the duration of the measurement.

Improvements in resolution have been achieved often approaching an order of magnitude compared to Lang's technique.[33] The Laplace DLTS method analyzes emission rates at a specific temperature and provides a spectral plot of a processed capacitance signal against the emission rate rather than against temperature. In order to produce an Arrhenius plot, the experiment is repeated at a number of temperatures. This eliminates the line broadening due to the shift of parameters with temperature which are significant when instrumental broadening is eliminated by the use of the inverse Laplace transform. Details of a LDLTS experimental system are available elsewhere.[34]

Good experimental conditions are essential if Laplace DLTS is to provide an order of magnitude higher energy resolution than conventional DLTS techniques. An example of this is shown in Fig. 9, which shows DLTS and LDLTS spectra of hydrogenated silicon containing gold. The
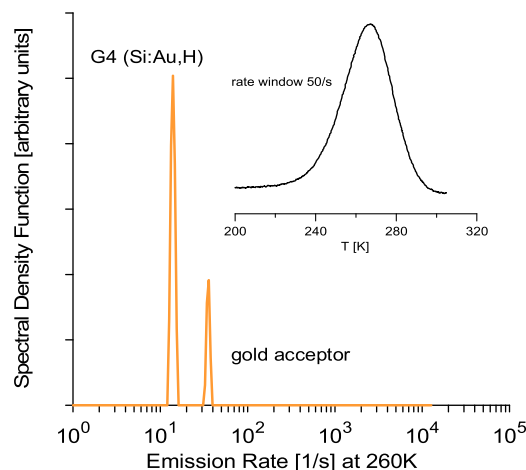
FIG. 9. Comparison of a conventional DLTS plot of the electron emission signal from gold and from a gold-hydrogen complex in silicon with the isothermal LDLTS spectra. The LDLTS measurement clearly separates the two signals which are indistinguishable in conventional DLTS. Adapted with permission from P. Deixler *et al.*, Appl. Phys. Lett. **73**, 3126 (1998). Copyright 1998 IOP Publishing.

conventional DLTS spectrum is shown as an inset at the top right of the figure. The broad peak centered at 260 K is due to electron emission from both the gold acceptor and from the G4 trap which is a gold-hydrogen complex. The linewidth of the DLTS peak showed little deviation from an ideal point defect, but using Laplace DLTS (shown as the main spectrum), the gold acceptor level and the gold-hydrogen level G4 are clearly separated.[35]

LDLTS has had a profound effect on electrical defect spectroscopy enabling the effect of external probes, such as uniaxial stress, and internal perturbations, such as the proximity of atoms isovalent with the host (e.g., the case of SiGe), to be quantified in terms of electronic behavior. Laplace DLTS provides a synergy with other techniques that was difficult or impossible to achieve previously. A comprehensive review of the technique and examples of results obtained are given elsewhere.[8,34]

C. Minority carrier processes and carrier capture measurements

All the methods considered so far only have taken into account majority carriers and have been based on the change of occupancy of a deep level state by shifting the Fermi level and then examining the majority carrier emission. Similar experiments can be done using minority carriers. This is important because if the defect studies are related to recombination, it is essential to know their characteristics in relation to minority carriers. The key issue is how to obtain an initial condition in which the minority carrier occupancy is significant.

Although there are a number of variants on the theme, the two widely used methods are to inject minority carriers in a pn junction using a forward bias or to optically excite minority carriers by irradiation with light. In the latter case, the wavelengths of the exciting light must be slightly above bandgap and the minority carriers must drift or diffuse to the region of the junction that will be measured by transient

spectroscopy. Sub band-gap light can be used to directly change the occupancy of the state according to its optical cross section. The latter technique is referred to as "optical DLTS" (ODLTS) or Deep Level Optical Transient Spectroscopy (DLOS) and is discussed later. If the light generates minority carriers, it is referred to as the minority carrier capture technique (MCC)[36] or minority carrier transient spectroscopy.[37]

For the case of minority carrier injection in a forward biased pn junction, holes will be injected into the n-type region and electrons into the p-type region. The ratio of carrier flux will depend on the relative values of p and n and the absolute value of the sum of fluxes on the current. Observing transients from majority and minority carrier transients in both the p and n regions makes interpretation very difficult, and so, it is usual practice to use p^+n or n^+p junctions. In such samples, the depletion region extends primarily into one polarity of material. Schottky barriers are essentially majority carrier devices and although some minority carrier injection occurs at high current densities in the case of devices with high barrier height, forward biased Schottkys are not used in this context.

We will consider the n^+p junction. Under forward bias, a flux of minority carriers is injected into the p-type region at the measurement temperature. The electron occupancy of any specific defect center is given by Eq. (19). Even if we assume that the emission processes are slow compared to the capture, determining the occupancy under forward bias is still problematic. This is because in the region of the diode, which will be measured by junction spectroscopy, the carrier population consists of both holes and electrons and the minority carrier occupancy of the defect will depend on the ratio of the minority and majority carrier capture cross sections (σ_p/σ_n for the n-type material).

The distribution of minority carriers is spatially dependent, decaying from the junction to a characteristic length dependent on the diffusion length of the minority carriers. As we need to know, the capture rates C_n and C_p to calculate the occupancy, and in general, these are not known and this is an essential prerequisite for determining the concentration of the defect, minority carrier measurements using p-n junctions always have some ambiguity. In the case where the minority carrier cross section is much less than the majority carrier cross section (i.e., in the n^+p case $\sigma_n \ll \sigma_p$ and so $c_p \gg c_n$), it may not be possible to establish a significant minority carrier occupancy of the defect using this technique. However, for the case where the minority carrier cross section is much larger than the majority carrier cross section, the technique works very well. Such a case was described by Lang in the DLTS discovery paper[6] as shown in Fig. 10.

Minority carriers can be generated in p^+n and n^+p junctions by above bandgap light for junction spectroscopy, but the technique is more usually used with Schottky diodes. The first experimental application of the generation of minority carriers by above bandgap light as a technique for manipulating the occupancy of deep states was described by Hamilton *et al.*[36] and referred to as MCC. It was developed into minority carrier transient spectroscopy (MCTS) by Brunwin *et al.*[37] Light can be applied through a semi-
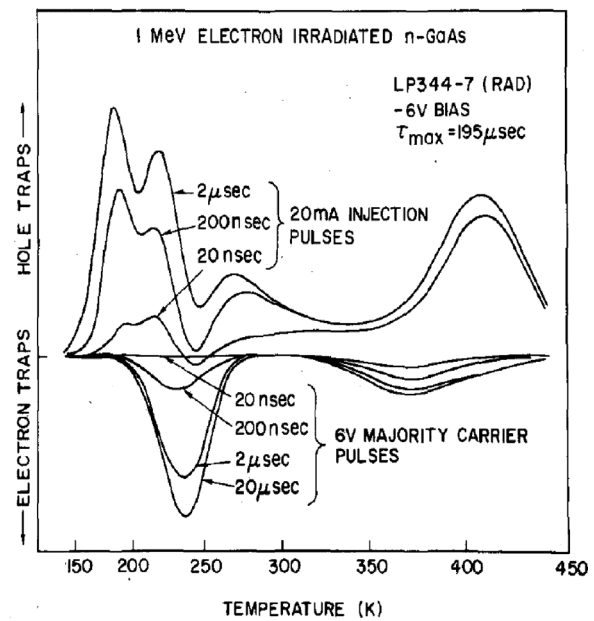

FIG. 10. DLTS scan of a GaAs p^+n diode irradiated with 1 MeV electrons showing minority (hole) and majority carrier traps (electrons). The minority carrier traps (hole traps) are partially filled by a forward bias pulse of different pulse lengths, while the majority carrier traps were revealed by a zero bias pulse showing partial filling according to the pulse width. Note that in this diagram, majority carrier transitions are depicted as negative signals while minority carrier transitions are positive peaks. This convention was used in early publications, but it is now more usual to use positive peaks to show majority carrier emission. Reproduced with permission from J. Appl. Phys. **45**, 3023 (1974). Copyright 1974 AIP Publishing.

transparent Schottky diode, i.e., from the front of the sample or from the backside. If the slice is thick compared to the minority carrier diffusion length, then thinning is necessary as shown in Fig. 11 to allow a sufficient flux of minority carriers to reach the depletion region. However, in most silicon applications, thinning is not necessary.

The disadvantage of excitation from the front is that majority carriers are also created in the depletion region and these play a part in the capture kinetics. The situation is similar to the forward biased p-n junction in that the carrier population is always a mix of minority and majority carriers. In the MCTS case, the light generates equal numbers of holes and electrons which are superimposed on the majority carrier population, and so in general, the majority carrier density is always greater than that of minority carriers. If the minority
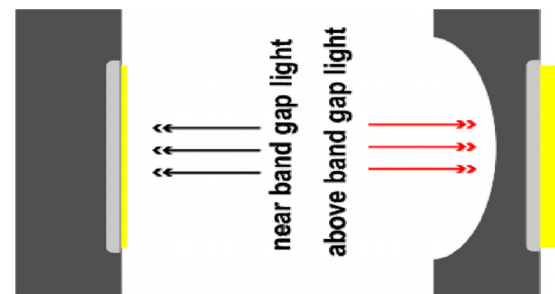

FIG. 11. The concept of minority carrier generation by using a pulse of bandgap light applied to a front face semi-transparent Schottky diode (left) or the back face of a thinned sample (right).

carrier cross section of the defect is greater than the majority carrier cross section ($\sigma_n > \sigma_p$ for p-type), the defect will be occupied by minority carriers and will be seen in MCTS. If excitation is from the back, the problem of both minority and majority carriers being captured is much reduced as the depletion field rejects majority carriers and so if no light is absorbed in the depletion region, the carrier flux is exclusively minority carriers. Thus, defects in the depletion region will capture only minority carriers if the wavelength of light is chosen so that there is sufficient absorption in the slice to ensure that a negligible photon flux reaches the depletion region yet a significant flux of minority carriers can diffuse to the depletion region and then drift through the region. This condition is easier to achieve in indirect gap materials, but the technique has been used successfully in direct gap materials.

Figure 12 compares DLTS and MCTS spectra for an n-type Cz silicon sample with Au Schottky diodes. The MCTS spectrum has been recorded with the use of backside optical excitation pulses from a light emitting diode with a wavelength of 940 nm. MCTS shows a negative peak due to minority carrier emission from powerful recombination centers identified as two C-O-H complexes.[38,39] It appears that there is a small population of majority carriers in the depletion region during the MCTS experiment as can be seen from the majority carrier peak in the low temperature region. This peak is due to the second donor level of oxygen-related thermal double donors having a large electron capture cross section. It is notable that the small flux of majority carriers is not sufficient to provide a detectable majority carrier occupancy of E_1, E_2, and E_3 traps, which have small electron capture cross sections.
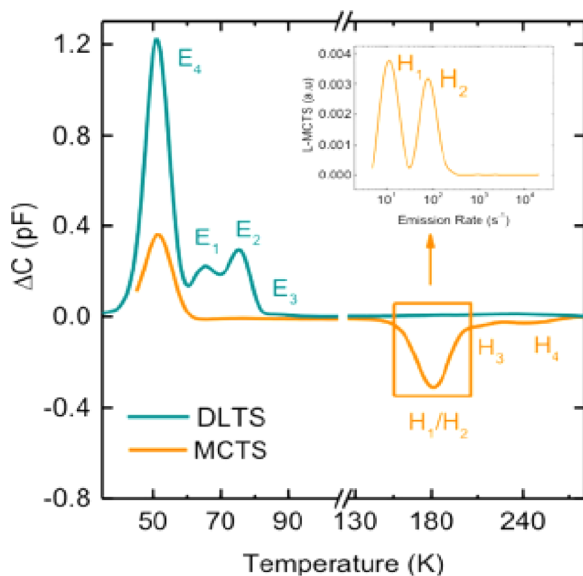
FIG. 12. DLTS and MCTS spectra recorded on a sample from n-type Cz silicon containing oxygen-related thermal double donors (E_4 trap) with a high concentration and C-O-H defects (E_1/H_1 and E_2/H_2 traps).[38,39] The spectra were recorded with a rate window corresponding to $e = 50\,\text{s}^{-1}$, reverse bias $U_b = -2\,\text{V}$, and filling pulse length $t_p = 10\,\text{ms}$. Optical pulses from a 940 nm LED were used in the MCTS measurements. The inset shows the separation of the H_1 and H_2 emission signals in the Laplace MCTS spectrum recorded at 190 K. Reproduced with permission from P. Santos *et al.*, Phys. Status Solidi A **214**, 1700309 (2017). Copyright 2017 Santos *et al.*

D. Depth profiling

In Sec. IV, so far, we have considered rather idealised samples and made some simplifying assumptions about the semiconductor physics we have used. Essentially, we have assumed that we are examining samples in which both the shallow doping level and the deep state population are uniform and in the physical approach have used the depletion approximation. For some samples, this is adequate, but the more general case in which concentrations vary throughout the depth of the sample is quite common. Typical examples are the end of range damage following ion implantation into silicon or carbon contamination in epitaxial growth of GaN.

In junction spectroscopy, it is possible to profile the concentration of defects as a function of the depth from the surface or from a pn junction. The range over which we can profile is limited by the range of depletion widths that can be achieved in the structure as detailed in Sec. III. A working approximation is that it is possible to profile from the zero bias depletion width to the depletion width at the reverse bias where leakage current perturbs the measurement… this is usually significantly less that the breakdown voltage.

Collecting the data for profiling is simple, but the interpretation is far from easy. The amplitude of the signal from the defect is measured as the bias conditions are changed. The profiling methodology can be to change the amplitude of the filling pulse or the amplitude of the reverse bias or both together. The complications arise from the fact that the depletion region contains a tail of majority carriers diffusing from the bulk and being repelled by the depletion field. This Debye tail affects the occupancy of the states and failure to account for it in profiling causes important errors particularly at low biases. Rocket and Peaker[40] were the first to quantify this effect and show how profiling calculations which do not take the Debye tail effect into consideration underestimate the near surface defect concentration by as much as a factor of two. These effects are sometimes referred to as incomplete trap filling although the name implies an oversimplification of the problem. Much has been written on this topic, for example, in Refs. 41–44.

An alternative experimental approach was devised by Lefevre and Schulz[45] referred to as double correlation DLTS (DDLTS) and some commercial systems incorporate this methodology. Essentially, the difference between the DLTS signals resulting from two different filling pulses is recorded, and the difference between the transient signals is analysed. The method is easily implemented in a computer based system compared to the complexity of the original analogue four channel boxcar. The method enables an observation window within the space charge region to be defined which can be well away from the Debye tail. Traps near the Fermi level can be excluded, and so with care, the method eliminates some of the errors of concentration profiling and reduces the variation in electric field that the measured traps experience. This is an issue of importance in cases where the electric-field-induced enhancement of carrier emission is large.

E. High trap concentrations and constant capacitance DLTS

Additional complications arise when the trap concentration is large compared to the free carrier concentration. The limit normally considered for conventional DLTS is that the trap concentration must be less than 10% of the carrier concentration. The problem arises from two main effects: (1) a large change in the depletion width during carrier filling and emission and (2) incomplete filling of the traps. The first issue results in the number of traps being measured changing during the measurement cycle, so producing a non-exponential transient even in the case of an ideal defect. This is particularly significant in Laplace DLTS.

It is possible to eliminate this problem by using a feedback system, which changes the applied bias. The transient measured in such systems is the voltage.[46,47] Designing such a feedback system with unconditional stability is far from trivial and often results in a degradation of response time and noise performance.

F. Optical excitation with below bandgap light

1. Optical DLTS and deep level optical spectroscopy

We have already discussed the generation of carriers with above bandgap light to create carriers, which set the occupancy of traps constituting MCTS and related techniques. However, below bandgap light can be used to excite carriers from traps for photon energies greater than the trap to band energy and the change of occupancy can be determined by observing the capacitance change or the current generated by the emitted carriers. This transient technique is known as optical DLTS (ODLTS) or deep level optical spectroscopy (DLOS).[48] In general, these acronyms are used interchangeably although some authors make a distinction according to the variant of the technique. The method has been built on the steady state techniques established by Grimmeiss.[49] For an electron trap in the upper half of the gap, the optical emission rate, e_n^o, depends on the optical cross section, σ_n^o (which is a function of wavelength) and the photon flux, Φ

$$e_n^o = \sigma_n^o \Phi. \tag{24}$$

If the energy of the photons is sufficient to excite holes from the valence band, then this convolutes with the photo-excited electron emission to the conduction band. Many variants of this technique have been published including the use of two light sources combining with thermal emission. The technique has seen a resurgence in relation to the study of defects in wide bandgap materials where conventional DLTS is complicated by the need for high temperatures. Often, DLOS results are presented as a parameter related to the optical cross section as a function of photon energy as shown in Fig. 13.[50] Fitting the curve to a model of optical cross section yields a threshold energy, which is taken as the energy of the defect from the band. DLOS presents problems compared to conventional and Laplace DLTS. Experimentally, it is necessary to cover quite a wide range of wavelengths so a mono-chromator is normally used. Even with a high intensity
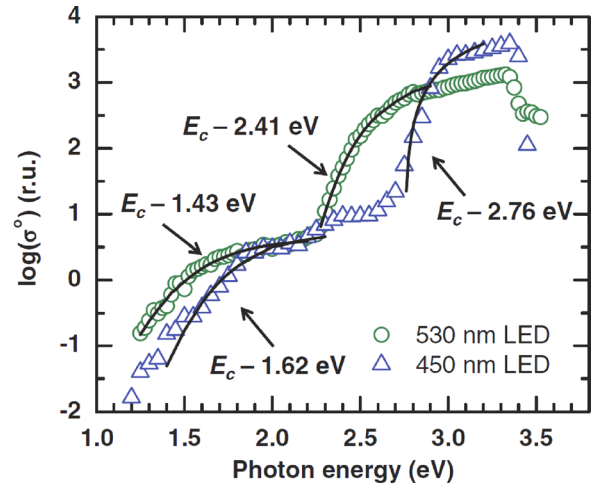

FIG. 13. DLOS spectra of $In_xGa_{1-x}N$ quantum wells in a blue- and a green-emitting LED. Symbols are data and the lines are fits to a model of optical cross section giving the defect energies as marked. Reproduced with permission from A. M. Armstrong *et al.*, Appl. Phys. Express **7**, 032101 (2014). Copyright 2014 The Japan Society of Applied Physics.

source, the photon flux is such that the time constants for emission are slow compared to typical thermal transients. It is difficult to obtain precise concentrations from DLOS if the light excites transitions to both bands. Discrimination between states is not good, and so, analysis is problematic unless the levels are well separated in energy.

2. Excited state spectroscopy

The optical techniques we have described above excite a carrier from the defect into the band referred to as a bound to free transition. As seen in Fig. 13, such transitions exhibit broad spectral features in terms of the energy dependence of the optical cross section. An alternative is to use a smaller photon energy than that required to excite the carrier to the band. Such photon energies can promote transitions of charge carriers from the ground to excited states of a defect. The carriers can then be promoted to the band from the excited states thermally and cause a change in either conductivity or capacitance of a test structure. The technique is known generically as photo-thermal spectroscopy (PTS) or photo-thermal ionisation spectroscopy (PTIS). The excited state sequence carries information on the ground state of the defect although this is not easy to interpret. The advantage of the photo-thermal technique is that the transition from the ground state to an excited state is a bound to bound transition, and in consequence, the photo-thermal spectra would be expected to have a fine line structure. The technique has been used for a large body of work, mostly on bulk samples rather than junctions, done in the 1970s on photo-thermal studies of hydrogenic shallow states. Photo-thermal spectroscopy has been used in studies of deep states (mostly in silicon) by Grimmeiss's group at Lund using several techniques to reveal the high resolution spectra. These include photo-admittance spectroscopy[51] and Fourier photo-admittance spectroscopy.[52] The immense power of these methods is well illustrated by Kleverman *et al.*[53] who used them to study electron-phonon interactions and excited states for the

deep platinum acceptor in silicon. However, as yet, photothermal spectroscopy has not found widespread application in the study of deep states, possibly due to the difficulties of interpretation, particularly when several defect species are present in the sample.

G. Current and conductance measurements

In this section, we draw a distinction between current and conductance DLTS. The former is the case analogous to capacitance DLTS where the charge released from traps is detected by the current that flows between the diode contacts in an external circuit. In the latter case, we consider the drain source conductance in a transistor when the charge in the gate region changes.

There are important differences between capacitance and current techniques. The current produced by holes and by electrons is of the same sense, so current DLTS cannot distinguish between hole and electron emission. So in general, capacitance techniques are favoured because in the case of a one-sided junction (Schottky barriers, p^+n and n^+p junctions), a distinction can be drawn between hole and electron emission simply by observing the sign of the capacitance change.

In current measurements, the integrated current over time represents the total charge released by the traps.[54] Remembering that the transient is faster at higher temperatures, it is evident that faster rate windows will result in higher currents for a shorter time. A consequence of this is that current DLTS shows higher detectivity at higher rate windows than at low emission rates. The same phenomenon also skews the current DLTS peak towards higher temperatures relative to capacitance measurements.[55]

In the case of very high resistivity or semi-insulating materials, the space charge region is not well defined and capacitance measurements are not possible. In these cases, current DLTS combined with optical excitation provides a way of characterising deep states. The samples can be semitransparent Schottky barriers or a bar of material with two Ohmic contacts on the bar. Light is directed between the contacts or through the Schottky barrier. The technique was first proposed by Hurtes *et al.*[56] to study high resistivity GaAs and has since been elaborated by many other groups for use on a range of materials. Usually, above bandgap light is used, but in some cases, the occupancy can be perturbed by sub-bandgap irradiation. It is referred to as Photo-Induced Transient Spectroscopy (PITS) or Photo-Induced Transient Current Spectroscopy (PICTS).

Conductance DLTS has proved to be of immense value in characterising defects in MOS and MESFET structures. In its simplest form, a small potential is applied between drain and source and the drain-source current monitored using conventional DLTS processing. The gate voltage is pulsed so as to fill traps and then relaxed so that the drain source current is modulated by the charge under the gate. In this way, bulk states can be characterised and in depletion mode MOS devices as can the interface state density and their energy distribution. The advantage of undertaking the measurement in this way is that extremely small devices can be

measured, which is not possible in capacitance DLTS because of the very small gate capacitance. The technique is described, and an analysis is developed to enable trap concentrations and parameters to be quantified by Hawkins and Peaker.[57] The technique has been used to characterise III-V FETs, Si MOSFETs, nanowires,[58] and very extensively in recent years GaN HEMTs. In the latter case, conductance DLTS is often combined with optical excitation.[59] An example is shown in Fig. 14 where a FET has been fabricated on a 60 nm ZnO nanowire. In the case of MOSFETs, the effect of individual defect states has been observed using conductance DLTS.

H. Scanning DLTS

All the techniques described so far sample a small lateral area of the sample, and the only way they can be used to explore lateral changes is to compile data from a number of junctions over the surface of the sample. An alternative approach has been implemented by Petroff and Lang[60] and elaborated by several other groups.[61] Essentially, this is to use a large area junction but to fill traps over a small area using an electron beam or light spot which can be scanned across the junction. Suitable junctions are shallow pn structures or Schottky barriers which in the case of optical excitation must be thin enough to be semi-transparent. A modified scanning electron microscope is often used for such measurements requiring a well controlled variable temperature sample stage and beam blanking.

The obvious problem with the technique is that only a small fraction of the diode area is being sampled for deep states at any instant, and so, the method has poor detectivity compared to a normal DLTS measurement. If capacitance DLTS is used, the active area is shunted by a large inactive quiescent capacitance requiring a special capacitance meter;[62] however, if current DLTS is used and fast transients
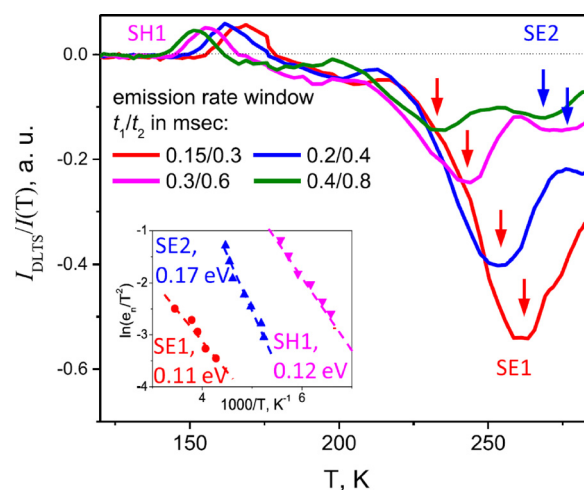

FIG. 14. Conductance DLTS of a ZnO nanowire 60 nm diameter. Source and drain connections have been made to a nanowire $\sim 1\ \mu m$ long using Ti/Au, and backgating was affected using a 120 nm SiO_2 dielectric on Si. Measurement conditions were $V_{ds} = 0.2\ V$ and $\Delta V_g = 10\ V$ for 100 μs. DLTS timings as shown. The inset shows the Arrhenius plots of the observed defects. Reproduced with permission from J. Appl. Phys. **122**, 094305 (2017). Copyright 2017 AIP Publishing.

are examined (very appropriate in a scanning system), the transient current increases compared to using slow rates as discussed in Sec. IV G. It is extremely difficult to obtain absolute values for defect concentration using SDLTS (scanning DLTS) as neither the measurement area nor the initial occupancy is well defined.

The technique works best for high concentrations of defects and is particularly useful in conjunction with other scanning techniques such as lifetime mapping and Electron Beam Induced Current (EBIC).[63] More recent work has focused on the use of DLTS with scanning probe microscopy. In this case, the area of measurement is defined by the probe tip, which limits the measured area. As this is very small, high frequency capacitance measurements are used (~1 GHz) but detectivity is again a real challenge. The technique has been used for point and extended defects and for interface state studies.[64,65]

I. Admittance spectroscopy

Admittance spectroscopy is based on the measurements of the real part of the complex admittance at frequencies and temperatures where the emission rate of carriers from traps is comparable with the measurement frequency. The technique was pioneered by Pautrat *et al.* who applied it to ZnTe and presented a detailed analysis of the methodology.[66] Typically, the complex admittance is measured using an impedance analyser over a frequency range of ~1 kHz to ~10 MHz with the impedance being evaluated as a parallel equivalent circuit. The test signal amplitude is usually ~100 mV, and a DC bias can be applied to profile the defects. The test signal modulates the Fermi level so that the traps fill during a half cycle of the sinusoidal test signal and empty during the other half cycle. If the emission is comparable with the frequency, this shows as an inflection in the capacitance signal and an increase in the real part of the complex impedance. The sample is held in a cryostat as in the case of DLTS. No filling pulse is used, and so, the technique is very simple to set up. The results are analysed by plotting the conductance change divided by the angular frequency of the test signal, $(G - G_{dc})/\omega$, which has units of Ss^{-1}. The technique has been reviewed by Barbolla *et al.* (Ref. 67) and has been used extensively to study thin film solar cells.[68]

Figure 15 from the study by Py *et al.* (Ref. 69) shows the application of admittance spectroscopy to InAlN for the case where the frequency is held at a fixed value and the temperature is scanned. Equally, the temperature could be held constant and the frequency scanned. Admittance spectroscopy provides a simple method to study traps at high emission rates. This is very valuable to study defects with energies only slightly deeper than the shallow donors or acceptors in the sample. It cannot distinguish between hole and electron traps and cannot be used to determine real capture cross sections (only intercept values from the Arrhenius plots). In most experimental implementations, the detectivity is much lower than DLTS. However, in a recent comparison of admittance spectroscopy and DLTS, it is claimed that comparable sensitivities can be achieved.[70]
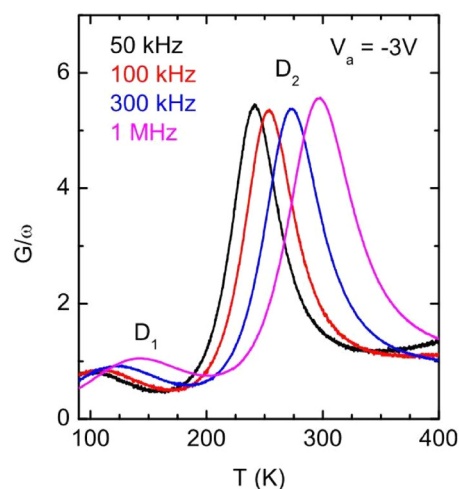

FIG. 15. Temperature dependent admittance measurements of an $In_{0.16}Al_{0.84}N$ unintentionally doped layer showing the frequency dependence of the contribution to the device conductance G/ω of two centers, a shallow donor D1 at 68 meV and a deep acceptor D2 at 290 meV both attributed to oxygen. Reproduced with permission from M. A. Py *et al.*, Phys. Rev. B **90**, 115208 (2014). Copyright 2014 the American Physical Society.

Optical variants of admittance spectroscopy have been developed. Two of these have been discussed in Sec. IV F 2. One involves a conductance measurement while modulating a light source of variable wavelength photo-admittance spectroscopy,[51] while the other used a Fourier transform spectrometer again measuring the sample conductivity. The interferometer is used as the light source, and the transformation of the conductivity signal is undertaken using the spectrometer's routines.[52]

J. Annealing

Some important characteristics of defects can be obtained from studies of changes in their concentration upon heat-treatments. Particularly, from such studies, activation energies for diffusion and solubility of impurity atoms can be found. Further, in some cases, it is possible to determine binding energy of components for a complex defect. It should be pointed out that the above mentioned parameters can differ significantly for different charge states of the same defect. The ability of manipulation of defect charge states in junction structures enables the above parameters to be derived for different charge states of a defect.[71,72]

The junction spectroscopy techniques have been successfully applied for the determination of maximum achievable concentration and activation energies for solubility and diffusion of electrically active impurity atoms in semiconductor materials.[73–75] Impurity atoms are introduced into a semiconductor by either ion implantation or in-diffusion from a surface layer with following heat-treatments at different temperatures terminated by quenching. Conditions of such treatments for solubility measurements (temperature range and durations) are aimed at achieving a uniform distribution of the atoms with a solubility limited concentration at a particular temperature. Then, the concentration of deep level traps related to an impurity is measured, usually in Schottky diodes processed at lower temperatures. For the

determination of impurity diffusion coefficients, concentration depth profiles of deep level traps created by either in- or out-diffusion of the corresponding defects at a particular heat-treatment temperature and duration are measured and analyzed.

Changes in the concentration of a defect in semiconductor materials upon heat-treatments at relatively low and moderate temperatures can be caused by a number of processes including diffusion and interactions with other defects in the lattice, dissociation of a complex defect, transformations between different configurations of the same center, etc. The isochronal and isothermal annealing techniques are usually used to obtain information about origin of the changes and parameters of a corresponding process. Isochronal annealing consists of a sequence of heat-treatments with a constant duration and usually with constant temperature increments, while isothermal annealing is carried out at a constant temperature with accumulated heat-treatment time. From an analysis of concentration changes upon isothermal annealing, $N(t)$, some details of a dominant defect reaction related to the changes can be obtained. In most cases, the reaction rate does not depend on heat-treatment duration, and the $N(t)$ dependence can be described by a mono-exponential decay or growth. Such reactions are called first order reactions. However, in general, the reaction rate can be a function of heat-treatment time, and more complicated $N(t)$ dependencies than the mono-exponential one can occur.[76] We will consider briefly the most frequent first order reactions. From an Arrhenius plot of characteristic time values for such a reaction, the activation energy, E_{ann}, and frequency (pre-exponential) factor, ν_{ann}, of the defect formation/elimination process can be derived. An analysis of frequency factor values gives some information about the origin of the process.

The value of the frequency factor, which is close to the value of the characteristic lattice vibration frequency (ν_0, about 10^{13} s^{-1} in silicon), indicates the occurrence of a dissociation or transformation/reorientation process. For the case of dissociation of a complex defect, E_{ann} is the dissociation energy, which is a sum of a binding energy of components and diffusion energy of a mobile component.

Values of ν_{ann}, which are significantly lower (usually in the range of 10^7–10^{10} s^{-1}) than the ν_0 value, are indicative of reactions consisting of diffusion of some mobile species and their capturing by trapping centers. The ν_{ann} value for such diffusion-trapping related reactions can be presented as[76]

$$\nu_{ann} = 4\pi r D_0 N_T, \tag{25}$$

where r is the so-called capture radius (the distance in the lattice at which the interaction of components occurs), D_0 is the pre-exponential factor in the diffusion coefficient of a mobile component, and N_T is the concentration of traps. The E_{ann} value in such cases is usually equal to the diffusion energy of a mobile species.

K. Identification of deep-level defects—synergy with other techniques

With the use of junction spectroscopy methods, a lot of information about electronic properties of defects can be collected; however, it is not possible to obtain direct information about the chemical identity of atoms incorporated into the defect or details of the defect's atomic configuration from an analysis of the transient spectra recorded. So, an assignment of a defect related signal detected in junction spectroscopy measurements to a particular defect structure is only possible in combinations with other measurement techniques, which are sensitive to the chemical identity of atoms and their position in a semiconductor lattice. Among such techniques are electron spin resonance (ESR) spectroscopy and its derivatives, optical methods [infrared absorption (IR), photoluminescence, and Raman spectroscopy], positron annihilation spectroscopy, secondary ion mass spectrometry, etc. We will briefly describe how these techniques can be complementary to junction spectroscopy methods and consider a few examples of some successful combined studies.

It should be noted that usually the first step in an assignment of an unidentified deep level signal to a particular defect is an analysis of the impurity content in a junction structure studied. Any characterization techniques, which give information about the chemical identity of impurity atoms and their concentration in the host semiconductor lattice, can give some clues for linking the detected defect signal with impurities available in a semiconductor material.

Further, the formation and elimination behaviours of defects detected by junction spectroscopy techniques are compared with those for defects observed and identified in similarly treated semiconductor samples by structure sensitive techniques. It should be noted, however, that such comparisons can sometimes be misleading, especially in the cases when multiple defect reactions occur simultaneously.

ESR related techniques detect electrically active defects having an unpaired electron at a deep or shallow energy level in the gap (the ESR techniques and their applications for defect studies in semiconductors are reviewed in the tutorial paper by Abe *et al.* in this issue of JAP). By both ESR and junction spectroscopy techniques, some parameters related to interactions of an electron with a defect are measured, and so, in some cases, the parameters obtained by these techniques can be directly compared. A combination of ESR and DLTS techniques has been successfully applied for solid identifications of a number of radiation-induced defects with deep levels in silicon by Watkins and co-workers.[77–80] These include a single Si vacancy,[77] an interstitial boron atom,[78] an interstitial carbon atom (C_i),[79] a complex incorporating C_i and a substitutional carbon atom,[80] etc.

IR absorption measurements (the IR absorption techniques and their applications for defect studies in semiconductors are reviewed in the tutorial paper by Stavola *et al.* in this issue of JAP) have been very useful in studies of impurities with atomic masses lighter than those of host atoms in a semiconductor lattice (e.g., oxygen, carbon, nitrogen, and hydrogen in Si).[81,82] Such light impurities in a crystal normally oscillate with frequencies significantly higher than the host phonon frequencies. The modes of the oscillations are denoted as local vibrational modes (LVMs). IR active LVMs give rise to sharp lines at the corresponding frequencies in the absorption spectra. The chemical nature of a defect, which is responsible for an LVM line, can sometimes be

determined from isotope substitution studies. The change of isotope mass for light impurity atoms causes a significant shift of the corresponding LVM line, which can be easily explained as masses of the isotopes and their abundance in a semiconductor material are usually known.[82] Further, LVMs are sensitive to a defect charge state, and some information about a position of a defect energy level can be obtained by recording IR spectra in samples with different positions of the Fermi level. Clues for identification of a number of deep level traps related to light (O, C, and H) impurities in silicon have been found from combined IR absorption and junction spectroscopy studies. These include a complex of Si self-interstitial with an oxygen dimer,[83] a metastable configuration of the vacancy-di-oxygen center,[84] a complex of Si di-interstitial with an interstitial oxygen atom,[85,86] etc.

V. THEORETICAL CALCULATIONS OF ELECTRONIC PROPERTIES OF DEEP LEVEL DEFECTS

A. Electronic structure: All-electron methods

Density functional theory (DFT) is perhaps the most successful quantum mechanical method available, allowing us to address atomistic problems ranging from bulk solids, surfaces, interfaces, molecules, nanostructures, and of course defects in crystals. We refer the reader to Refs. [87] and [88], among the many available reviews and books on the subject. The density functional formalism tells us that the *ground state* of a system of electrons under the effect of an external potential v_{ext} (produced, for instance, from a set of atom nuclei) can be determined from the electron density alone, which is a scalar field that depends on three (space) coordinates only. This is a rather disruptive approach when compared to the Schrödinger equation, whose solution for the ground state of a system with N electrons depends at least on their $3N$ space coordinates, $\Psi_0 \equiv \Psi_0(\mathbf{r}_1, \ldots, \mathbf{r}_N)$. This idea has been applied to the total energy of the ground state E_0, using the variational principle

$$E_0 = \min_{n(\mathbf{r})} E[n], \tag{26}$$

where the electron density n is enclosed within square brackets to emphasize the dependence of E as a functional of n. Hence, according to Hohenberg and Kohn[89]

$$E[n] = \int d\mathbf{r}\, v_{\text{ext}}(\mathbf{r}) n(\mathbf{r}) + F[n], \tag{27}$$

where $F[n]$ was defined as a *universal* (but yet unknown) functional in the sense that it is valid for any external potential, and besides including the electronic kinetic energy, it also describes the complicated many-body electron-electron interactions, $F[n] = T[n] + V_{\text{ee}}[n]$. In order to find an approximation to $F[n]$, Kohn and Sham (Ref. [90]) first decomposed F into

$$F[n] = T_s[n] + \frac{1}{2}\int d\mathbf{r}\, n(\mathbf{r}) v_{\text{H}}(\mathbf{r}) + E_{\text{xc}}[n], \tag{28}$$

where T_s is the kinetic energy that a system of non-interacting electrons with density n would have. The second

term is referred to as the Hartree energy, where $v_H(\mathbf{r}) = \int d\mathbf{r}'\, n(\mathbf{r}')/|\mathbf{r} - \mathbf{r}'|$ is the classical Coulomb potential, and finally, E_{xc} defines the exchange-correlation energy. Besides accounting for the quantum-mechanical electron-electron interactions, the exchange-correlation energy corrects the total energy for the difference $T[n] - T_s[n]$ and for the self-interactions in the Hartree term. All terms in Eq. (28) but the last can be evaluated exactly, and the search for suitable (accurate and trackable) approximations to E_{xc} has been a vivid research topic.

The first approximation to E_{xc} was the local density approximation (LDA)[90] and assumes that the density $n(\mathbf{r})$ is a slowly varying field on the scale of the Fermi wavelength (that corresponds to the radius of the Fermi sphere). Accordingly

$$E_{\text{xc}}^{\text{LDA}}[n] = E_{\text{x}}^{\text{LDA}}[n] + E_{\text{c}}^{\text{LDA}}[n] = \int d\mathbf{r}\, n(\mathbf{r})(\epsilon_{\text{x}}[n] + \epsilon_{\text{c}}[n]), \tag{29}$$

with $\epsilon_x[n]$ and $\epsilon_c[n]$ being the exchange and correlation potentials per particle of a uniform electron gas with density n, respectively. While ϵ_x can be calculated analytically, the correlation potential was obtained by fitting to the data gathered by accurate quantum Monte-Carlo simulations between the low and high density limits.

LDA treats all systems as locally homogeneous. However, real problems are spatially inhomogeneous due to the presence of ionic and electronic potentials. A step beyond the LDA has been taken by including information regarding the first-order variation of density into the exchange-correlation functionals. This is referred to as the generalized gradient approximation (GGA), with

$$E_{\text{xc}}^{\text{GGA}}[n] = \int d\mathbf{r}\, f(n(\mathbf{r}), \nabla n(\mathbf{r})), \tag{30}$$

where the particular choice of f defines a specific flavour of the exchange-correlation functional. Although the above functionals, in particular LDA, consist of very crude approximations for systems that are far from homogeneous, they have proven to be a huge success in many cases. This is in part due to the cancellation of errors from overestimated E_x by the underestimated E_c. The main drawback of LDA/GGA in studying defects in semiconductors and insulators is the severe underestimation of the bandgap. In some extreme cases, such as InN and InAs, materials can be incorrectly predicted as semimetals instead of semiconductors. In recent years, this problem has been greatly mitigated with the emergence of hybrid functionals,[91] which include a Hartree-Fock-like exchange component. The popular HSE06 (Heyd-Scuseria-Ernzerhof)[92] exchange-correlation functional uses an error function screened Coulomb potential to calculate the exchange portion of the energy in order to improve computational efficiency. The exchange energy is then given by

$$E_{\text{x}}^{\text{HSE}} = a E_{\text{x}}^{\text{HF,SR}}(\omega) + (1-a) a E_{\text{x}}^{\text{PBE,SR}}(\omega) + a E_{\text{x}}^{\text{PBE,LR}}(\omega), \tag{31}$$

where $\omega = 0.2$ is a screening parameter setting the separation between the short-range (SR) and long-range (LR)

regions of space, whereas $a = 0.25$ is the exact-exchange mixing parameter. The correlation term is straight from the PBE (Perdew, Burke and Ernzerhof) flavour of GGA.[93]

Hybrid functionals are obviously computationally intense as E_x is now a function of the individual orbitals (and not simply the electron density and its gradient). However, they bring great improvement regarding the quality of the band structure, including the gap widths. For instance, for GaN and InN with experimental bandgaps of 3.30 eV and 0.78 eV, HSE06 predicts 3.03 eV and 0.52 eV, against 1.8 eV and -0.40 eV from LDA, respectively.[94]

A further step towards improving the self-interaction correction can be achieved by many-body perturbation theory within the *GW* approximation.[95] Unlike density functionals, the *GW* method accounts for the many-body electron-electron interactions via screening of the exchange interactions by means of a frequency-dependent dielectric matrix. In practice, quasi-particle (QP) energies can be calculated within the spirit of first-order perturbation theory. The QP energies E_n^{QP} are obtained from

$$\left(-\frac{\nabla^2}{2} + v_H + v_{ext}\right)\psi_n(\mathbf{r}) + \int d\mathbf{r}' \, \Sigma(\mathbf{r},\mathbf{r}',E_n^{QP})\psi_n(\mathbf{r}')$$
$$= E_n^{QP}\psi_n(\mathbf{r}), \tag{32}$$

where the self-energy Σ is obtained to first-order within the GW approximation as

$$\Sigma(\mathbf{r},\mathbf{r}',\omega) = \frac{i}{2\pi}\int d\omega' \, e^{i\eta\omega}G(\mathbf{r},\mathbf{r}',\omega+\omega')W(\mathbf{r},\mathbf{r}',\omega'), \tag{33}$$

with η being infinitesimal, G the single-particle Green's function, and W the dynamically screened Coulomb interaction. Since both G and W depend on ψ_n, Eqs. (32) and (33) can be evaluated self-consistently. This is often avoided by assuming that the single-particle states are those obtained from DFT. This approach is usually referred to as G_0W_0 approximation. While it improves the band structure over that obtained from DFT, a further refined treatment may be obtained beyond the single-shot calculation by performing a sequence of self-consistent calculations of G, W, ψ_n, and E_n^{QP} values.

Periodic boundary conditions provide a natural way for the description of electronic states in a perfect crystal, and therefore, they have been often employed by the solid-state community to solve electronic structure problems. In order to mimic a defect in a crystalline solid, a periodic supercell (constructed by stacking primitive cell replicas) is used, which has to be sufficiently large so that interactions between the periodic images of defects can be neglected. In the case of charged defects, long-ranged Coulomb interactions between charge density fields localized on states bound to the defect converge very slowly with the supercell volume. Consequently, and particularly in crystals with weak electrostatic screening, the desired supercell sizes are usually too large for the available computational power.

A way to circumvent the above problem is to vary the size of the supercell and extrapolate the energy of defects to the infinite host.[96] Since this is not practical in most cases,

several approaches have been developed to obtain the energy of the defective infinite solid, $E(\mathbf{R},q)$, from that of a finite and periodic supercell, $\tilde{E}(\mathbf{R},q)$, with the help of correction schemes (see, for example, Ref. 97 for a recent review). For the sake of example, we describe the method proposed by Freysoldt *et al.*,[98] which has been recently generalized for anisotropic materials.[99] Accordingly, the energy of a defect in an infinite crystal is approximately $E \approx \tilde{E} + E_{corr}$, with E_{corr} being a *charge correction*

$$E_{corr}(\mathbf{R},q) = E_{PC}(q) - q\Delta\overline{\phi}_{PC,ind}(\mathbf{R},q), \tag{34}$$

where $E_{PC}(q)$ is a point-charge correction which for isotropic and cubic materials becomes the Madelung energy of a point-charge immersed in a uniform background charge density of opposite sign, $E_{PC} = \alpha_M q^2 / 2\epsilon L$, which depends on the ratio between the Madelung constant α_M and a characteristic length L (usually a lattice constant), the net charge q, and the dielectric constant ϵ. Also in Eq. (34), the quantity $\Delta\overline{\phi}_{PC,ind}$ is the remote offset between the defect induced average potential $\overline{\phi}_{ind}(\mathbf{R},q) = \overline{\phi}_{def}(\mathbf{R},q) - \overline{\phi}_{bulk}$ and that produced by a point-charge, $\overline{\phi}_{PC}(q)$. The potentials $\overline{\phi}_{def}$ and $\overline{\phi}_{bulk}$ can be evaluated from first-principles from the defective and pristine supercells, respectively, and averaging should be carried out at remote locations from the defect.[98,99] Other defect-defect interactions like dispersion of defect bands (due to wave function overlap across neighbouring cells) and strain can only be mitigated by increasing the supercell size.

B. Calculations of electrical levels

In this section, we will go through the practical details involving the calculations of the binding energy of carriers trapped at defects. We will also demonstrate that the two most popular methodologies, namely, the *formation energy* and *marker methods*, are essentially equivalent.

It was shown in Sec. II that defect electronic transitions can be calculated by finding the crossing-points of the formation energy in different charge states [cf. Eqs. (9) and (10)]. This is usually referred to as the *formation energy method*, according to which the valence band maximum energy is assumed to be the highest occupied state from a band-structure calculation, ϵ_{VBM}, and the position of the level with respect to the valence band top becomes

$$E(q/q+1) - E_v = \left[E(\mathbf{R}_q,q) - E(\mathbf{R}_{q+1},q+1)\right] - \epsilon_{VBM}. \tag{35}$$

We assume that the energies in Eq. (35) are free from the spurious electrostatic effects due to the use of periodic boundary conditions (including charge and potential alignment corrections), i.e., $E(\mathbf{R}_q,q) = \tilde{E}(\mathbf{R}_q,q) + E_{corr}(\mathbf{R}_q,q)$.

We may also use the delta self-consistent field (ΔSCF) method to obtain the levels with respect to the conduction band minimum.[100] This is often convenient when calculating the depth of electron traps. Within the ΔSCF approximation, the bandgap is $E_g = -(A_{bulk} - I_{bulk})$, with $A_{bulk} = \tilde{E}_{bulk}(q=0) - \tilde{E}_{bulk}(q=-1)$ and $I_{bulk} = \tilde{E}_{bulk}(q=+1)$

$-\tilde{E}_{\text{bulk}}(q = 0)$ being calculated straight from the energies of charged bulk supercells. Hence, Eq. (35) becomes

$$E(q/q + 1) - E_{\text{v}} = \left[E(\boldsymbol{R}_q, q) - E(\boldsymbol{R}_{q+1}, q + 1) \right] - I_{\text{bulk}},$$
(36)

or alternatively,

$$E_{\text{c}} - E(q/q + 1) = A_{\text{bulk}} - \left[E(\boldsymbol{R}_q, q) - E(\boldsymbol{R}_{q+1}, q + 1) \right].$$
(37)

Equations (36) and (37) embody the *marker method* of calculating defect levels.[101,102] The term *marker* arrives from the fact that we are comparing the ionization potential (or electron affinity) of the defective supercell with an analogous quantity that is well established (the marker). In this case, the markers are calculated from the bulk solid and stand for the band edge energies. We finally note that when using density-functional methods, the accuracy of the calculated electrical levels depends mostly on the quality of the exchange-correlation description and the size of the supercell or cluster employed. The former approximation impacts the width of the bandgap and the relative location of the transition levels with respect to the band edges. On the other hand, by increasing the size of the supercell (or the cluster) hosting the defect, we may mitigate or even avoid boundary-conditions-related errors like periodic charging effects, strain interactions between defects and their periodic replicas (or the cluster surface), or even defect band dispersion effects which are particularly detrimental when dealing with shallow extended states.

Let us exemplify the use of Eqs. (9), (36), and (37) in the calculation of defect formation energies and electronic transitions. We chose the carbon vacancy (V_{C}) in 4*H*-SiC, which is a technologically important defect due to its presence in the as-grown material and for its minority carrier recombination activity. The V_{C} defect has been connected to a set of deep acceptor traps labelled $Z_{1/2}$ after DLTS measurements.[103,104] These have been ascribed to the superposition of Z_1 and Z_2 signals, each of which arising from a $(=/0)$ negative-*U* emission sequence from V_{C} at different sub-lattice sites of the 4*H* polytype. The binding energy of electrons emitted from $Z_{1/2}^{=}$ and $Z_{1/2}^{-}$ to the conduction band minimum was measured approximately as 0.7 eV and 0.5 eV, respectively.[104]

Hybrid density functional calculations (Ref. 105) of the formation energy of the carbon vacancy at the cubic site, $V_{\text{C}}(k)$, in 4*H*-SiC are summarized in Fig. 16. The results are in line with those reported by Hornos *et al.*[106] The calculated formation energy of the neutral defect ranges from 4.3 eV to 4.9 eV, depending if we consider a carbon chemical potential for C-poor and C-rich conditions. Both limits are separated by the heat of formation of 4*H*-SiC, $\Delta H_{\text{fu}}^0 = -0.6$ eV, which compares to 0.7 eV from experiments. The E_f values also agree very well with formation enthalpies of 4.8–5.0 eV measured from samples grown under C-rich conditions.[107]

For Fermi energies in the upper half of the gap, the lowest energy states are V_{C}^0 and $V_{\text{C}}^{=}$ (thick lines in Fig. 16). The negative charge state is metastable, irrespective of the E_F
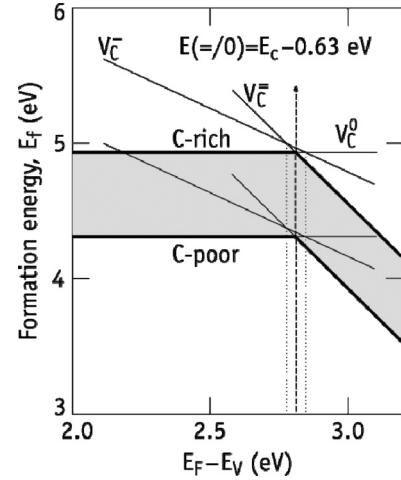

FIG. 16. Formation energy diagram of $V_{\text{C}}(k)$ in 4*H*-SiC as a function of the Fermi energy (only the upper part of the gap is shown). The shaded area is limited by the C-rich and C-poor formation energy lines. Thick lines represent the lowest energy states as a function of the Fermi energy. Neutral, negative, and double negative states correspond to lines with zero, negative, and double negative slope.

value. The crossing point between $E_f(q = 0)$ and $E_f(q = -2)$ is indicated by the dashed line and takes place at $E(=/0) = E_c - 0.63$ eV. Metastable $E(-/0)$ and $E(=/-)$ transitions (vertical dotted lines) are calculated at $E_c - 0.61$ eV and $E_c - 0.64$ eV, respectively. The calculated $(=/-)$ level agrees within 0.1 eV with the experiments, while the $U = E(=/-) - E(-/0) = -0.03$ eV is underestimated by about 0.2 eV.

VI. CODA

In this tutorial review, we have tried to describe and explain all the major techniques used to study defects in semiconductors using junction spectroscopy and provide examples of their use applied to a range of semiconductors. However, there are many variants on these methods based on the principles we have described and it is not possible to cover all of these in the space we have available. However, we do provide an extensive set of references which extend our descriptions and explanations. In the case of MOS measurements, we have written very little in terms of either fundamentals or practical methods. The reasoning behind this is that it is a major subject in its own right and worthy of a separate review or tutorial which we anticipate will be available in the near future.

[1]W. Shockley and W. T. Read, Jr., Phys. Rev. **87**, 835 (1952); R. N. Hall, *ibid*. **87**, 387 (1952); Proc. IEE **106B**, 923 (1960). A simplified description of recombination and generation lifetime with worked tutorial examples appears in D. K. Schroder, *Semiconductor Materials and Device Characterization*, 3rd ed. (John Wiley & Sons, Inc., Hoboken, New Jersey, 2006), Chap. 7.

[2]R. Williams, J. Appl. Phys. **37**, 3411 (1966).

[3]H. B. Goldstein and S. S. Perlman, Phys. Rev. **148**, 715 (1966).

[4]H. G. Grimmeiss and G. Olofsson, J. Appl. Phys. **40**, 2526 (1969).

[5]C. T. Sah, L. Forbes, L. L. Rosier, and A. F. Tasch, Solid-State Electron. **13**, 759 (1970).

[6]D. J. Lang, J. Appl. Phys. **45**, 3023 (1974).

[7]D. J. Lang, Physica B **401–402**, 7 (2007).

[8]L. Dobaczewski, A. R. Peaker, and K. Bonde Nielsen, J. Appl. Phys. **96**, 4689 (2004).

[9]A. R. Peaker, V. P. Markevich, I. D. Hawkins, B. Hamilton, K. Bonde Nielsen, and K. Gościński, Physica B **407**, 3026 (2012).

[10]A. Alkauskas, M. D. McCluskey, and C. G. Van de Walle, J. Appl. Phys. **119**, 181101 (2016).

[11]G. D. Watkins, in *Advances in Solid State Physics*, edited by P. Grosse (Springer, Berlin, Heidelberg, 1984), pp. 163–189.

[12]R. Kubo and I. Toyozawa, Prog. Theor. Phys. **13**, 160 (1955).

[13]C. H. Henry and D. V. Lang, Phys. Rev. B **15**, 989 (1977).

[14]M. G. Burt, J. Phys. C: Solid State Phys. **16**, 4137 (1983).

[15]L. Shi and L.-W. Wang, Phys. Rev. Lett. **109**, 245501 (2012).

[16]A. Alkauskas, Q. Yan, and C. G. Van de Walle, Phys. Rev. B **90**, 075202 (2014).

[17]A. D. Franklin, in *Point Defects in Solids: General and Ionic Crystals*, edited by J. H. Crawford and L. M. Slifkin (Springer, Boston, USA, 1972), pp. 1–101.

[18]E. H. Rhoderick and R. H. Williams, *Metal-Semiconductor Contacts*, 2nd ed. (Clarendon, Oxford, 1988).

[19]S. M. Sze and K. K. Ng, *Physics of Semiconductor Devices*, 3rd ed. (John Wiley & Sons, Inc., Hoboken, New Jersey, 2007).

[20]P. Blood and J. W. Orton, *The Electrical Characterization of Semiconductors: Majority Carriers and Electron States*, Techniques of Physics Series Vol. 14 (Academic Press, London, 1992).

[21]D. K. Schroder, *Semiconductor Materials and Device Characterization*, 3rd ed. (John Wiley & Sons, Inc., Hoboken, New Jersey, 2006).

[22]A. R. Peaker and V. P. Markevich, in *Defects and Impurities in Silicon Materials*, Springer Lecture Notes in Physics Series, edited by Y. Yoshida and G. Langouche (Springer, Japan, 2015), Chap. 3, pp. 129–180.

[23]B. L. Smith, T. J. Hayes, A. R. Peaker, and D. Wight, Appl. Phys. Lett. **26**, 122 (1975).

[24]C. T. Sah, W. W. Chan, H. S. Fu, and J. W. Walker, Appl. Phys. Lett. **20**, 193 (1972).

[25]D. S. Day, M. Y. Tsai, B. G. Streetman, and D. V. Lang, J. Appl. Phys. **50**, 5093 (1979).

[26]L. C. Kimerling, IEEE Trans. Nucl. Sci. **23**, 1497 (1976).

[27]F. D. Auret, Rev. Sci. Instrum. **57**, 1597 (1986).

[28]G. L. Miller, J. V. Ramirez, and D. A. H. Robinson, J. Appl. Phys. **46**, 2638 (1975).

[29]C. R. Crowell and S. Aliphani, Solid State Electron. **24**, 25 (1981).

[30]A. A. Istratov and O. F. Vyvenko, Rev. Sci. Instrum. **70**, 1233 (1999).

[31]C. W. Groetsch, *The Theory of Tikhonov Regularisation for Fredholm Equations of the First Kind* (Pitman, London, 1984).

[32]S. W. Provencher, Comput. Phys. Commun. **27**, 213 (1982).

[33]L. Dobaczewski, P. Kaczor, I. D. Hawkins, and A. R. Peaker, J. Appl. Phys. **76**, 194 (1994).

[34]See www.laplacedlts.eu for information on Laplace DLTS and its experimental implementation.

[35]P. Deixler, J. Terry, I. D. Hawkins, J. H. Evans-Freeman, A. R. Peaker, L. Rubaldo, D. K. Maude, J.-C. Portal, L. Dobaczewski, K. Bonde Nielsen, A. Nylandsted Larsen, and A. Mesli, Appl. Phys. Lett. **73**, 3126 (1998).

[36]B. Hamilton, A. R. Peaker, and D. R. Wight, J. Appl. Phys. **50**, 6373 (1979).

[37]R. Brunwin, B. Hamilton, P. Jordan, and A. R. Peaker, Electron. Lett. **15**, 349 (1979).

[38]M. Vaqueiro-Contreras, V. P. Markevich, M. P. Halsall, A. R. Peaker, P. Santos, J. Coutinho, S. Öberg, L. I. Murin, R. Falster, J. Binns, E. V. Monakhov, and B. G. Svensson, Phys. Status Solidi RRL **11**, 1700133 (2017).

[39]P. Santos, J. Coutinho, S. Öberg, M. Vaqueiro-Contreras, V. P. Markevich, M. P. Halsall, and A. R. Peaker, Phys. Status Solidi A **214**, 1700309 (2017).

[40]P. I. Rocket and A. R. Peaker, Appl. Phys. Lett. **40**, 957 (1982).

[41]S. D. Brotherton, Solid-State Electron. **26**, 987 (1983).

[42]D. Stievenard and D. Vuillaume, J. Appl. Phys. **60**, 973 (1986).

[43]J. H. Zhao, J. C. Lee, Z. Q. Fang, T. E. Schlesinger, and A. G. Milnes, J. Appl. Phys. **61**, 5303 (1987); **61**, 5489 (1987).

[44]D. C. Look, Z. Q. Fang, and J. R. Sizelove, J. Appl. Phys. **77**, 1407 (1995).

[45]H. Lefèvre and M. Schulz, Appl. Phys. **12**, 45 (1977).

[46]G. Goto, S. Yanagisawa, O. Wada, and H. Takanashi, Jpn. J. Appl. Phys. **13**, 1127 (1974).

[47]N. M. Johnson, D. J. Bartelink, R. B. Gold, and J. F. Gibbons, J. Appl. Phys. **50**, 4828 (1979).

[48]A. Chantre, G. Vincent, and D. Bois, Phys. Rev. B **23**, 5335 (1981).

[49]H. G. Grimmeiss, Annu. Rev. Mater. Sci. **7**, 341 (1977).

[50]A. M. Armstrong, M. H. Crawford, and D. D. Koleske, Appl. Phys. Express **7**, 032101 (2014).

[51]M. Kleverman, E. Janzén, and H. G. Grimmeiss, Solid State Commun. **46**, 895 (1983).

[52]E. Janzén, K. Larsson, R. Stedman, and H. G. Grimmeiss, J. Appl. Phys. **53**, 7520 (1982).

[53]M. Kleverman, J. Olajos, and H. G. Grimmeiss, Phys. Rev. B **37**, 2613 (1988).

[54]J. A. Borsuk and R. M. Swanson, IEEE Trans. Electron Devices **ED-27**, 2217 (1980).

[55]Y. Nakano and T. Kachi, Appl. Phys. Lett. **79**, 1631 (2001).

[56]C. H. Hurtes, M. Boulou, A. Mitonneau, and D. Bois, Appl. Phys. Lett. **32**, 821 (1978).

[57]I. D. Hawkins and A. R. Peaker, Appl. Phys. Lett. **48**, 227 (1986).

[58]I. Isakov, M. J. L. Sourribes, and P. A. Warburton, J. Appl. Phys. **122**, 094305 (2017).

[59]A. Sasikumar, A. R. Arehart, G. D. Via, B. Winningham, B. Poling, E. Heller, and S. A. Ringel, Microelectron. Reliab. **55**, 2258 (2015).

[60]P. M. Petroff and D. V. Lang, Appl. Phys. Lett. **31**, 60 (1977).

[61]O. Breitenstein and J. Heydenreich, Scanning **7**, 273 (1985).

[62]O. Breitenstein, Phys. Status Solidi A **71**, 159 (1982).

[63]J. Heydenreich and O. Breitenstein, J. Microsc. **141**, 129 (1986).

[64]L. Dózsa, G. Molnár, V. Raineri, F. Giannazzo, J. Ferencz, and Š. Lányi, Nanoscale Res. Lett. **6**, 140 (2011).

[65]N. Chinone and Y. Cho, J. Appl. Phys. **122**, 105701 (2017).

[66]J. L. Pautrat, B. Katircioglu, N. Magnea, D. Bensahej, J. C. Pfistert, and L. Revoil, Solid State Electron. **23**, 1159 (1980).

[67]J. Barbolla, S. Dueñas, and L. Bailón, Solid State Electron. **35**, 285 (1992).

[68]T. Walter, R. Herberholz, C. Müller, and H. W. Schock, J. Appl. Phys. **80**, 4411 (1996).

[69]M. A. Py, L. Lugani, Y. Taniyasu, J.-F. Carlin, and N. Grandjean, Phys. Rev. B **90**, 115208 (2014).

[70]J. Bollmann and A. Ventner, "Admittance spectroscopy or deep level transient spectroscopy: A contrasting juxtaposition," Physica B (in press).

[71]L. C. Kimerling, H. M. DeAngelis, and J. W. Diebold, Sold State Commun. **16**, 171 (1975).

[72]S. B. Lastovskii, V. E. Gusakov, V. P. Markevich, A. R. Peaker, H. S. Yakushevich, F. P. Korshunov, and L. I. Murin, Phys. Status Solidi A **214**, 1700262 (2017).

[73]H. Nakashima, T. Sadoh, H. Kitagawa, and K. Hashimoto, Mater. Sci. Forum **143–147**, 761 (1994).

[74]K. Graff, *Metal Impurities in Silicon-Device Fabrication*, Springer Series in Materials Science, Vol. 24 (Springer-Verlag, Berlin, 2000).

[75]M. L. Polignano, D. Codegoni, A. Galbiati, S. Grasso, I. Mica, G. Moccia, G. Nardone, and F. Russo, ECS J. Solid State Sci. Technol. **5**, P3048 (2016).

[76]T. R. Waite, J. Chem. Phys. **28**, 103 (1958); **32**, 21 (1960).

[77]G. D. Watkins and J. R. Troxell, Phys. Rev. Lett. **44**, 593 (1980).

[78]J. R. Troxell and G. D. Watkins, Phys. Rev. B **22**, 921 (1980).

[79]L. W. Song and G. D. Watkins, Phys. Rev. B **42**, 5759 (1990).

[80]L. W. Song, X. D. Zhan, B. W. Benson, and G. D. Watkins, Phys. Rev. Lett. **60**, 460 (1988).

[81]R. C. Newman, Mater. Sci. Eng. B **36**, 1 (1996).

[82]B. Pajot and B. Clerjaud, *Optical Absorption of Impurities and Defects in Semiconductor Crystals. II. Electronic Absorption of Deep Centres and*

Vibrational Spectra, Springer Series in Solid-State Science (Springer-Verlag, Berlin, 2013).

[83] V. P. Markevich, L. I. Murin, S. B. Lastovskii, I. F. Medvedeva, J. L. Lindström, A. R. Peaker, J. Coutinho, R. Jones, V. J. B. Torres, S. Öberg, and P. R. Briddon, Solid State Phenom. **108–109**, 273 (2005).

[84] L. I. Murin, J. L. Lindström, V. P. Markevich, I. F. Medvedeva, V. J. B. Torres, J. Coutinho, R. Jones, and P. R. Briddon, Solid State Phenom. **108–109**, 223 (2005).

[85] V. P. Markevich, A. R. Peaker, B. Hamilton, S. B. Lastovskii, L. I. Murin, N. Ganagona, E. V. Monakhov, and B. G. Svensson, Solid State Phenom. **242**, 290 (2016).

[86] V. E. Gusakov, S. B. Lastovskii, L. I. Murin, E. A. Tolkacheva, L. I. Khirunenko, M. G. Sosnin, A. V. Duvanskii, V. P. Markevich, M. P. Halsall, A. R. Peaker, I. Kolevatov, H. M. Ayedh, E. V. Monakhov, and B. G. Svensson, Phys. Status Solidi A **214**, 1700261 (2017).

[87] R. O. Jones, Rev. Mod. Phys. **87**, 897 (2015).

[88] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, 2008).

[89] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[90] W. Kohn and L. J. Sham, Phys. Rev. **140**, A1133 (1965).

[91] A. D. Becke, J. Chem. Phys. **98**, 5648 (1993).

[92] J. Heyd, G. E. Scuseria, and M. Ernzerhof, J. Chem. Phys. **124**, 219906 (2006).

[93] J. P. Perdew, K. Burke, and M. Ernzerhof, Phys. Rev. Lett. **77**, 3865 (1996).

[94] R. R. Pela, M. Marques, and L. K. Teles, J. Phys.: Condens. Matter **27**, 505502 (2015).

[95] L. Hedin, Phys. Rev. **139**, A796 (1965).

[96] C. W. M. Castleton and S. Mirbt, Phys. Rev. B **70**, 195202 (2004).

[97] H.-P. Komsa, T. T. Rantala, and A. Pasquarello, Phys. Rev. B **86**, 045112 (2012).

[98] C. Freysoldt, J. Neugebauer, and C. G. Van de Walle, Phys. Rev. Lett. **102**, 016402 (2009).

[99] Y. Kumagai and F. Oba, Phys. Rev. B **89**, 195205 (2014).

[100] L. Hedin and S. Lundqvist, Solid State Phys. **23**, 1 (1970).

[101] A. Resende, R. Jones, S. Ö. berg, and P. R. Briddon, Phys. Rev. Lett. **82**, 2111 (1999).

[102] J.-W. Jeong and A. Oshiyama, Phys. Rev. B **64**, 235204 (2001).

[103] T. Kimoto, A. Itoh, H. Matsunami, S. Sridhara, L. L. Clemen, R. P. Devaty, W. J. Choyke, T. Dalibor, C. Peppermüller, and G. Pensl, Appl. Phys. Lett. **67**, 2833 (1995).

[104] C. Hemmingsson, N. T. Son, O. Kordina, J. P. Bergman, E. Janzén, J. L. Lindström, S. Savage, and N. Nordell, J. Appl. Phys. **81**, 6155 (1997).

[105] J. Coutinho, V. J. B. Torres, K. Demmouche, and S. Öberg, Phys. Rev. B **96**, 174105 (2017).

[106] T. Hornos, Á. Gali, and B. G. Svensson, Mater. Sci. Forum **679–680**, 261 (2011).

[107] H. M. Ayedh, R. Nipoti, A. Hallén, and B. G. Svensson, Appl. Phys. Lett. **107**, 252102 (2015).