Fakultät für Elektrotechnik und Informationstechnik
der Technischen Universität München,
Lehrstuhl für Technische Elektrophysik

Internal Laser Probing Techniques for Power Devices: Analysis, Modeling, and Simulation

Robert Thalhammer

Vollständiger Abdruck der von der Fakultät für Elektrotechnik und Informationstechnik der Technischen Universität München zur Erlangung des akademischen Grades eines

Doktors der Naturwissenschaften (Dr. rer. nat.)

genehmigten Dissertation.

Vorsitzender:		Univ.–Prof. Dr.–Ing. habil. A. W. Koch
Prüfer der Dissertation:	1.	Univ.–Prof. Dr. rer. nat. G. Wachutka
	2.	Univ.–Prof. F. Koch, Ph.D. Berkeley (USA)

Die Dissertation wurde am 31. 05. 2000 bei der Technischen Universität München eingereicht und durch die Fakultät für Elektrotechnik und Informationstechnik am 05. 10. 2000 angenommen.

Abstract

In this thesis, a physically rigorous model for the simulation of internal laser probing techniques for semiconductor power devices is developed which is employed for a thorough analysis of the physical effects during the measurement process. The theoretical studies provide valuable information for developing and optimizing powerful probing techniques which facilitate space–resolved and time–resolved measurements of carrier concentration and temperature profiles. In the second part of this thesis, it will be demonstrated that these techniques constitute an ideal supplement to the well–established electrical characterization methods and thus enable a significantly improved validation and calibration of electrothermal simulation models.

Internal laser probing techniques detect the modulation of a probing beam which originates from free carrier or temperature induced variations of the complex refractive index. Thus, measurements of the carrier and temperature distribution with a spatial resolution of about $10\,\mu\mathrm{m}$ and a temporal resolution of about 300 ns have become possible. A laterally propagating beam enables a vertical scanning of different profiles, e. g. we can extract the injected carrier concentration from the decrease of the transmitted light intensity, the gradients of carrier density and temperature from the beam deflection, and the temperature evolution from the oscillations of the Fabry–Perot transmission. On the other hand, Backside Laser Probing employs a laser beam which penetrates at the rear surface of the substrate and propagates vertically through the sample, thus providing integral information about the temperature profile. A remarkable feature of this technique, however, is the large measurement range which covers a temperature rise of some 100 K.

Since attendant numerical simulations have proven to be indispensable for the investigation of physical processes, a physical model has been developed which, for the first time, facilitates the simulation of the entire measurement process. It includes the electrothermal device simulation of the investigated structure, the calculation of the beam propagation through the device, the lenses and aperture holes, and the simulation of the detector response. Due to the large computation domain, the calculation of wave propagation in the sample is the most demanding problem. Introducing a suitable set of variables constitutes the decisive step which enables a coarser discretization without loss of accuracy.

Employing this strategy to simulate free carrier absorption measurements reveals that the experimental accuracy is limited by the surface recombination and the lateral extension of the probing beam. Optimum samples are about 20 diffusion lengths in size, thus promising measurements with a negligible error of only a few percent. The decisive effect on the experimental accuracy rather originates from an exact sample alignment and the accurate knowledge of the carrier concentration

dependence of the absorption coefficient.

Under operating conditions with low power dissipation the temperature evolution can be determined with excellent accuracy by laser deflection measurements. Their measurement range, however, is up to now limited to a temperature rise of only some Kelvin. The desired extension requires specific test structures with a shrinked active length and an improved evaluation rule which is deduced from an analytical calculation of the detector signal. Thus, temperature rises up to about 60 K can be detected.

The necessary sample preparation for Backside Laser Probing includes the etching of a window in the collector contact metallization. Current crowding at its edges leads to a local increase of the carrier concentration which gives rise to a tolerable error of about 25 % in the carrier contribution to the phase shift signal. On the other hand, the temperature inhomogeneity amounts to only some percent and can be therefore neglected. Despite the high angular aperture of the probing beam, the lateral beam spreading does not introduce a significant error. Therefore, Backside Laser Probing represents a powerful characterization method for operating conditions with large power dissipation.

Interferometric techniques employing a laterally traversing probing beam basically suffer from a limited measurement range as the superposition of the different rays deteriorates as the internal deflection increases. The detectable temperature rises within specific test structures are comparable to those of the deflection technique. Since interferometric methods do not require additional calibration, their realization would obviously be a useful supplement to the existing techniques.

As measurements of the time–resolved distributions of carriers and temperature are available in addition to the electrical terminal characteristics, a significantly improved validation and calibration of electrothermal simulation models has become possible. Within this thesis, the self–consistent electrothermal extension of the drift–diffusion model is investigated. Its governing equations are derived from the principles of irreversible thermodynamics. The most important material parameters are the recombination rate and the carrier mobility. While the latter is limited by the scattering of phonons, impurities, surface roughness, and the scattering of the carriers among each other, recombination in silicon is dominated by the Shockley–Read–Hall recombination via a trap level in the band gap. The calibration of the corresponding carrier lifetime models is a crucial precondition for predictive simulations and is significantly enhanced by the here introduced methodologies.

The most significant heat generation mechanisms are the Joule heat of electrons and holes as well as the recombination heat in devices with lifetime control. Compared to those, only a small contribution originates from the Peltier effect, which is accurately modeled in the interior of the simulated silicon domain. However, it is neglected at the semiconductor/metal junctions due to a simplified implementation of the thermal boundary conditions. The resulting effects on the temperature distribution are detectable by the developed laser probing techniques.

In order to demonstrate a complete model calibration, two different types of commercially available IGBTs are investigated. The most accurate description of the forward characteristics and the internal carrier distribution is obtained by using cylindrical coordinates within a simulation domain that accurately reflects the volume of a real cell. The simple and computationally efficient Scharfetter–relation can be employed to allow for the doping dependence of the carrier lifetimes if its parameters are transformed to a different set which comprises the minority carrier lifetimes

in the heavily doped regions and the ambipolar lifetime as its independent parameters. While these are extracted from the carrier concentration profile, the channel mobilities are adjusted with reference to the saturation currents of the forward characteristics. Thus, the measured temperature evolution during transient switching is also accurately reproduced by the calibrated model. This example illustrates that founding the calibration procedure on several different probing techniques enables predictive simulations which are valuable tools for a fast and cost effective optimization of semiconductor devices.

Zusammenfassung

Im Rahmen der vorliegenden Arbeit wird ein physikalisch rigoroses Modell zur Simulation interner Lasermessverfahren für Halbleiterleistungsbauelemente entwickelt und für eine umfassende Analyse der physikalischen Vorgänge während des Messprozesses angewandt. Aus den theoretischen Studien ergeben sich wertvolle Ansätze für die Entwicklung und Optimierung leistungsfähiger Messverfahren zur Bestimmung orts– und zeitaufgelöster Ladungsträger– und Temperaturverteilungen. Im zweiten Teil der Arbeit wird gezeigt, dass diese Verfahren eine ideale Ergänzung zu den etablierten elektrischen Charakterisierungsmethoden darstellen und damit eine erheblich verbesserte Validierung und Kalibrierung elektrothermischer Simulationsmodelle ermöglichen.

Interne Lasermessverfahren detektieren die Beeinflussung eines Messstrahls durch träger– oder temperaturinduzierte Änderungen des komplexen Brechungsindexes. Damit erlauben sie erstmalig die experimentelle Bestimmung von Ladungsträgerdichte und Temperatur im Inneren von Leistungsbauelementen mit einer Ortsauflösung von etwa $10\,\mu\mathrm{m}$ und einer Zeitauflösung von etwa 300 ns. Bei lateraler Durchstrahlung der Probe lassen sich verschiedene vertikale Profile abrastern; z. B. können die Trägerdichte aus der Modulation der transmittierten Lichtleistung, die Gradienten von Trägerdichte und Temperatur aus der Strahlablenkung sowie die zeitliche Temperaturentwicklung aus den Fabry–Perot Oszillationen der Transmission extrahiert werden. Dagegen verwendet man beim ”Backside Laser Probing” einen durch die Substratrückseite eintretenden und vertikal propagierenden Laserstrahl, der damit eine integrale Information über das Temperaturprofil liefert. Dieses Verfahren zeichnet sich durch einen großen Messbereich aus, der Temperaturhübe von einigen 100 K umfasst.

Da begleitende numerische Simulationen sich als unverzichtbares Werkzeug zur Untersuchung physikalischer Vorgänge etabliert haben, ist ein physikalisches Modell entwickelt worden, das erstmalig die Simulation des kompletten Messprozesses gestattet. Es umfasst die elektrothermische Devicesimulation des untersuchten Bauelements, die Berechnung der Strahlpropagation durch das Bauelement, die Linsen und die Aperturblenden sowie die Simulation der Detektorantwort. Aufgrund des großen Simulationsgebiets stellt die Berechnung der Wellenausbreitung in der Probe das anspruchsvollste Problem dar. Der entscheidende Schritt besteht in der Einführung eines geeigneten Variablensatzes, der eine gröbere Diskretisierung ohne Genauigkeitsverlust erlaubt.

Die Anwendung dieser Strategie zur Simulation von Absorptionsmessungen zeigt, dass die Messgenauigkeit durch die Oberflächenrekombination und die laterale Ausdehnung des Laserstrahls begrenzt ist. Optimale Teststrukturen weisen daher eine Länge von etwa 20 Diffusionslängen auf, womit sich ein vernachlässigbarer Fehler von wenigen Prozent erzielen lässt. Von entschei-

dender Bedeutung für die Messgenauigkeit ist allerdings eine exakte Ausrichtung der Probe sowie eine genaue Kenntnis der Trägerdichteabhängigkeit des Absorptionskoeffizienten.

Für Betriebszustände mit kleinen Verlustleistungen lässt sich die Temperaturentwicklung mit hoher Genauigkeit durch Laserdeflexionsmessungen bestimmen, deren Messbereich bisher jedoch auf Temperaturhübe von wenigen Kelvin beschränkt ist. Die gewünschte Erweiterung erfordert spezielle Teststrukturen mit einer aktiven Länge von etwa $200\,\mu\mathrm{m}$ sowie eine verbesserte Auswertevorschrift zur Extraktion des Brechungsindexgradienten, die aus einer analytischen Berechnung des Detektorsignals gewonnen wird. Damit können dann Temperaturhübe bis etwa 60 K bestimmt werden.

Für das Backside Laser Probing ist es notwendig, ein Fenster in die Kollektormetallisierung zu ätzen, was an dessen Rändern zu einer Stromkonzentration und damit zu einem lokalen Anstieg der Trägerdichte führt. Dies bewirkt einen tolerierbaren Fehler von etwa 25 % im Ladungsträgerbeitrag zur Phasenverschiebung, während die resultierenden Temperaturinhomogenitäten nur wenige Prozent betragen und daher vernachlässigbar sind. Trotz der hohen Apertur verursacht die laterale Strahlaufweitung keinen nennenswerten Fehler, so dass Backside Laser Probing ein leistungsfähiges Charakterisierungsverfahren bei hohen Verlustleistungen darstellt.

Das Grundproblem interferometrischer Verfahren mit lateraler Durchstrahlung der Probe besteht in der Beschränkung des Messbereichs aufgrund der Tatsache, dass sich die Überlagerung der Einzelstrahlen mit zunehmender interner Strahlablenkung verschlechtert. Mit Hilfe von Teststrukturen können Temperaturhübe gemessen werden, die denen des Deflexionsverfahrens vergleichbar sind. Da interferometrische Verfahren keine Kalibrierung erfordern, bietet sich ihr Aufbau als Ergänzung zu den bestehenden Methoden an.

Mit den zeitaufgelösten Ladungsträger– und Temperaturprofilen stehen zusätzlich zum Klemmenverhalten der Bauelemente experimentelle Ergebnisse zur Verfügung, die nunmehr eine erheblich umfassendere Validierung und Kalibrierung elektrothermischer Simulationsmodelle erlauben. Im Rahmen dieser Arbeit wird die selbstkonsistente elektrothermische Erweiterung des Drift–Diffusionsmodells untersucht, deren Grundgleichungen auf den Prinzipien der irreversiblen Thermodynamik beruhen. Die wichtigsten Materialparameter dieses Modells sind die Rekombinationsrate und die Beweglichkeit. Während letztere durch die Streuung an Phononen, Störstellen, Oberflächenrauhigkeiten und den Trägern untereinander begrenzt wird, ist der wichtigste Rekombinationsmechanismus in Silizium die Shockley–Read–Hall Rekombination über Trapniveaus in der Bandlücke. Die Kalibrierung der zugehörigen Lebensdauermodelle ist eine entscheidende Voraussetzung für prädiktive Simulationen und wird durch die hier entwickelten Methoden erheblich verbessert.

Als die dominanten Wärmegenerationsmechanismen werden die Jouleschen Wärmen von Elektronen und Löchern sowie bei lebensdauereingestellten Bauelementen die Rekombinationswärme identifiziert. Im Vergleich dazu ist nur ein kleiner Beitrag auf den Peltiereffekt zurückzuführen, der im Inneren eines simulierten Siliziumgebiets richtig beschrieben ist. Dagegen wird er an Metall–Halbleiterübergängen aufgrund einer vereinfachten Implementierung der thermischen Randbedingungen vernachlässigt, was mit den entwickelten Lasermessverfahren nachgewiesen werden kann.

Eine vollständige Modellkalibrierung wird am Beispiel zweier kommerziell erhältlicher IGBTs demonstriert. Die beste Beschreibung des Kennlinienfeldes und des internen Ladungsträgerprofils erhält man bei der Verwendung von Zylinderkoordinaten in einem Simulationsgebiet,

welches das Volumen einer realen Zelle richtig wiedergibt. Um die Dotierungsabhängigkeit der Lebensdauer mit dem einfachen und damit numerisch effizienten Scharfetter–Modell beschreiben zu können, wird eine Parametertransformation mit den Minoritätsträgerlebensdauern in den hochdotierten Buffer– und Substratbereichen und der ambipolaren Lebensdauer als den unabhängigen Parametern durchgeführt. Diese werden mit Hilfe der Ladungsträgerprofile bestimmt, während die Kanalbeweglichkeiten aus den Sättigungsströmen des Kennlinienfeldes extrahiert werden. Das so kalibrierte Modell gibt auch die gemessene Temperaturentwicklung während des transienten Schaltens korrekt wieder. Dieses Beispiel verdeutlicht, dass eine auf mehreren experimentellen Charakterisierungsverfahren beruhende Kalibrierung prädiktive Simulationen ermöglicht, die ein wertvolles Hilfsmittel zur schnellen und kostengünstigen Optimierung von Halbleiterbauelementen darstellen.

Contents

Chapter 1

Introduction

Since the birth of modern semiconductor industry — the invention of the transistor by Bardeen, Brattain [1], and Shockley [2] in 1947 — microelectronics has captured an increasing influence on today's life. The rapid technological progress has been enhancing the scale of integration up to some billions of components within a single chip today. During the same time, discrete devices for specific functions have been continuously improved and novel concepts of operating principles, manufacturing techniques, and semiconductor materials have emerged on the market.

Among the huge variety of applications, power electronics constitutes the key technology for the final stages of microelectronic equipment and electric plants. Covering a power range from a few Watts to some Megawatts, power semiconductor devices are employed within modulators for power conversion or controllers of three–phase–current motors in household appliances, industrial machines, or traction applications.

1.1 Power devices and modules

The diversity of power devices [3] which are available on the market today may be divided into diodes and power switches. The former are employed as rectifiers and, in this case, optimized with respect to a low voltage drop in the forward conducting state, whereas a fast switching behavior is the major developmental goal for free wheeling diodes. On the other hand, the field of power switches with a maximum blocking voltage up to 3.5 kV and a maximum chip current up to 50 A is dominated by the transistors. Power MOSFETs are a common choice if blocking voltages up to 600 V or high switching frequencies are desired. In the medium power range, the Insulated Gate Bipolar Transistor (IGBT) is gaining increasing significance and IGBTs with higher and higher blocking voltages are being developed [4]. Nevertheless, the thyristors still represent the devices which offer the highest blocking voltages up to almost 9 kV. Related concepts are the GTO (Gate–Turn–Off) thyristor and the MCT (MOS Controlled Thyristor) which can also be turned off at their maximum load current.

Even the simple structure of a power diode (cf. fig. 1.1) incorporates the typical features which are common to all kind of power devices: The heavily doped emitter regions — in case of a diode the anode and the cathode — are only a few microns in depth while most of the structure consists of a large and weakly doped region. In the forward conducting state, its conductivity is

enhanced by the injection of carriers. While the doping level is about $10^{12}\,\mathrm{cm}^{-3}$ to $10^{14}\,\mathrm{cm}^{-3}$, the injected carrier densities amount to some $10^{15}\,\mathrm{cm}^{-3}$ or $10^{16}\,\mathrm{cm}^{-3}$, i. e. forward conducting power devices are operating under high injection conditions. In this case, charge neutrality enforces an equal electron and hole concentration. This is why we will simply speak of the injected *carrier* concentration.

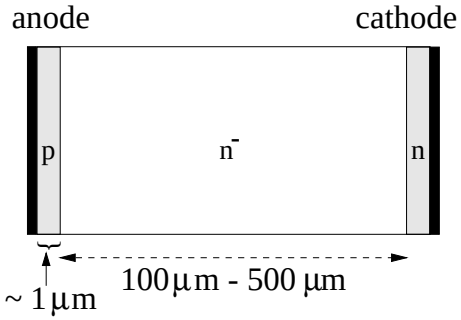


Figure 1.1: Schematic sketch of a pin power diode.

The fundamental problem in designing power devices consists in the necessary trade–off between a low forward voltage drop and a fast turn–off behavior. Whereas a high carrier injection favors the forward conductivity, it also lengthens the extraction of the stored charge and thus decreases the maximum switching frequency. The device performance can be improved by a clever adjustment of the carrier distribution which is managed by a proper choice of the emitter efficiencies and the purposeful incorporation of recombination centers for lifetime control [5]. A precise knowledge of the carrier concentration profile is therefore of great interest for the optimization of power devices.

Among the power switches, the Insulated Gate Bipolar Transistor (IGBT) has become the most popular device for blocking voltages from 600 V to 3.5 kV since it combines the simple gate control of MOSFETs with the low forward voltage drop of bipolar transistors. The entire chip which is able to conduct some Amperes comprises an array of thousands of single cells (cf. fig. 1.2). Each of them consists of a pnp bipolar transistor whose base is driven by a n–channel MOSFET. Hence, a positive gate bias leads to forming a n–channel in the p–well through which electrons are drifting into the n^-–base of the bipolar transistor. As a consequence, holes are injected by the p–emitter at the bottom, thus flooding the n^-–region with carriers. Although the emitter of the *internal* pnp transistor is connected to the bottom contact it has become common practice to call this contact the ”collector” and the top contact the ”emitter”. These terms are taken from transistor operation at the *circuit level*.

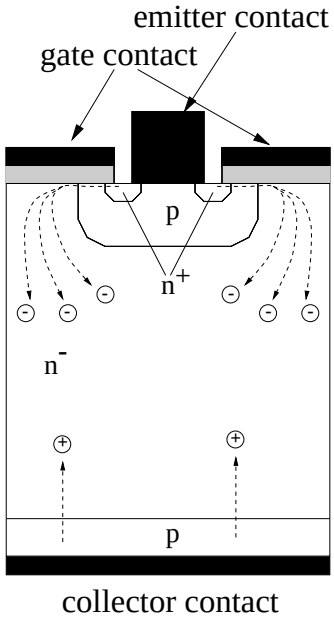


Figure 1.2: Schematic sketch of an Insulated Gate Bipolar Transistor (IGBT) cell.

One of the major goals in the development of IGBTs is to improve the overload ruggedness. At very high operating current densities the hole current gives rise to a voltage drop across the p–well which is sufficient to trigger the parasitic thyristor formed by the n^+–source, the p–well, the n^-–base, and the rear p–emitter. In this case, the device current can not be switched off by decreasing the gate bias. This so–called ”latch–up” condition is suppressed by an additional p^+–doping in the center of the p–well.

For practical applications, customers ask for devices which can be turned off even after a specific period of time under short circuit operation, typically $10\ \mu\mathrm{s}$. Since the entire supply voltage

of some kV drops across the device while it is conducting a short circuit current density of several $100\ \mathrm{A/cm^2}$, the resulting power dissipation of almost $1\ \mathrm{MW/cm^2}$ may easily cause the thermal destruction of the device. To prevent a thermal runaway the devices are designed to exhibit a negative temperature coefficient of the saturation current density. It ensures a homogeneous current distribution across the whole chip area and thus averts current filamentation.

The basic circuit (cf. fig. 1.3) which is incorporated in power converters and pulse width modulators includes a power switch, e. g. an IGBT, and a free wheeling diode to suppress inductive voltage spikes. Complete pulse width modulators comprising three half bridges are available for blocking voltages up to 3.5 kV and switching currents of 1.2 kA. For optimum performance, the IGBT and the diode have to be well–adapted to each other, in particular with respect to their switching behavior. While the IGBT is in its conducting state the diode is reverse biased, with the space charge region covering most of the intrinsic region. During turn–off an increasing fraction of the load current has to flow through the diode. Since the latter is bare of charge carriers, the forward voltage drop across the device can rise up to 60 V resulting in an enormous power dissipation. A similar situation occurs during turning on the power switch: The forward conducting diode which is flooded with carriers has to sustain an increasing blocking voltage. If the latter rises too quickly, the electric field at the pn–junction may dynamically exceed the critical field strength, thus generating additional charge carriers. However, a soft reverse recovery behavior of the diode is desirable [6] since a sudden current decrease will result in great induction voltage spikes. Mastering these excessive operating conditions requires a thorough investigation of the physical processes in the interior of the devices. For that purpose, electrical characterization techniques are employed for determining the terminal behavior under stationary and transient operating conditions. Further information is gained from various optical beam testing and electron beam testing methods [7, 8] which facilitate probing the distributions of different physical quantities, such as the charge carriers, the heat sources, or the electric field.

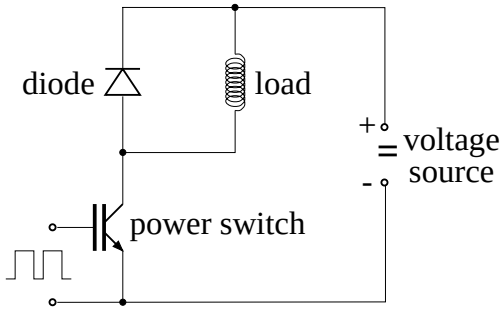


Figure 1.3: Basic circuit of a power switch and a free wheeling diode for the control of inductive loads.

1.2 Numerical simulations

In addition to the experimental characterization, the performance of semiconductor devices is investigated by numerical simulations, i. e. the numerical solution of an appropriate set of partial differential equations which govern the electrical, thermal, optical, and/or mechanical behavior under specific operating conditions. For the simulation of power devices semi–classical models are employed which are derived by a momentum expansion of the Boltzmann transport equation (cf. chapter 6). Two major goals are pursued with numerical simulations:

- First, the theoretical model is able to calculate a variety of physical quantities which are not

directly accessible by currently available experimental methods or cannot even be measured at all. In addition, the spatial and temporal resolution is usually much higher than that of the experimental methods. For instance, the dynamic expansion of the space charge region and the maximum field strength during the turn–off process of power devices are of particular interest for the optimization of the switching behavior. Although some probing techniques (e. g. OBIC [9, 10], EBIC [11]) are sensitive to the electric field distribution, they do not enable a space–resolved and time–resolved detection in a non–invasive manner. Therefore, numerical simulations constitute the more powerful method to visualize the physical processes in the interior of the device, thus significantly enhancing our understanding of the devices' operating principles.

- Second, theoretical studies can replace numerous experimental tests, avoiding the processing of the necessary test structures and their characterization. Provided that calibrated physical models are available, they help to reduce costs and development time. A modification of the channel length in an IGBT, for example, requires the corresponding adaptation of the diffusion process which forms the p–well. As the entire manufacturing process typically takes some weeks and several iterations for subsequent improvements are usually necessary, the whole optimization procedure may consume several months. On the other hand, an optimum structure and its sensitivity to each processing parameter is revealed by numerical studies within a few days. This is especially true as the rapid progress of computer technology produces more and more powerful machines which even facilitate the implementation of costly models for complex physical phenomena.

However, since numerical simulations stick to the principle "garbage in, garbage out" [12] the decisive precondition for predictive calculations is a comprehensive validation of the employed models, both with respect to their analytical structure as well as to the quantitative determination of the model parameters. For that purpose, a continuous comparison with experimental data is indispensable. The use of combined characterization methods helps the calibration procedure to become more profound, yielding reliable models of more general validity.

1.3 Device characterization methods

The first approach to the experimental investigation of semiconductor devices usually consists in the electrical characterization of the terminal behavior. Both static and dynamic measurements at various temperatures are employed to extract information about technological parameters and device performance. Capacitance–voltage (C–V) measurements, for example, are a common technique to determine the doping profile of MOS structures and pin diodes. In addition, pulsed current–voltage (I–V) measurements reveal a variety of decisive device parameters, such as threshold voltages, current or voltage gain, saturation currents, breakdown voltages, holding currents, or snapback voltages.

A lot of effort has been spent to ascertain the carrier lifetime and the injected charge which is stored in the interior of a power device during its forward conducting state. During turn–off, the excess carrier concentration is partly reduced by internal recombination and partly extracted by

a negative terminal current. Suppressing one of each possibilities is the fundamental idea of two special measurement techniques:

First, the stored charge Q_S can be calculated by integrating the reverse recovery current during a rapid turn–off (cf. fig. 1.4). This idea was first proposed by Hoffmann and Schuster [13] who applied a sudden blocking pulse to a diode carrying a forward current I. The calculation of Q_S, however, is based on the assumption that the internal recombination during the switching process is negligible. Satisfying this condition requires an experimental setup with low parasitic wire inductivities which facilitate a steep current slope during turn–off. An effective carrier lifetime $\tau_{eff} = Q_S/I$ can be extracted from these measurements [14] which have become the most frequently employed technique to determine the stored charge [15, 16, 17, 18, 19, 20]. For further details about the reverse recovery process and the correlation to the carrier recombination in the interior of the device the reader is referred to these publications.

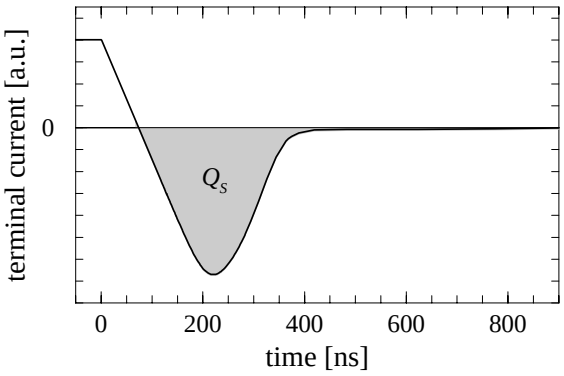


Figure 1.4: Reverse recovery of a pin–diode.

Second, the carrier lifetimes are determined by open–circuit–voltage–decay (OCVD) measurements. For that purpose, the forward current flow through a diode is suddenly interrupted, thus suppressing any reverse recovery current. The internal recombination of the excess carrier concentration leads to continuously decaying diffusion potentials across the junctions. Therefore, the resulting decrease of the terminal voltage is related to the carrier lifetime. Since the invention of this method for the investigation of pn–junctions [21, 22], it represents a common technique for measuring effective carrier lifetimes in junction devices [23, 24, 25, 20, 26].

A thorough analysis of the switching process reveals which physical effects govern each period during turn–off. The corresponding sectors of the current or voltage transients thus reveal specific information about, e. g., the carrier lifetime under high and weak injection or the minority carrier lifetimes in the heavily doped regions. However, pure measurements of the terminal behavior in principle do not allow scanning the distribution of a physical quantity, as, e. g., the carrier concentration. However, the knowledge of the latter is of great benefit for the development of power devices since, as mentioned above, a suitable adjustment of the carrier profile improves trade–off between a low forward voltage drop and a fast turn–off behavior.

The development of experimental techniques for the spatially resolved measurement of the carrier distribution in power devices started in 1952. Haynes and Briggs [27] reported that infrared radiation is emitted by silicon and germanium samples into which carriers are injected. Although these materials are indirect semiconductors, there is a small fraction of phonon or exciton assisted radiative recombination which constitutes the physical origin of the detected radiation [28]. This effect can be exploited to determine the distribution of carriers in the interior of forward biased power devices. For that purpose, a specific sector of the sample is observed through an aperture hole. A photo multiplier is employed to amplify the emitted infrared radiation, whose intensity is a measure of the local carrier concentration. The 2D distribution of the carriers is thus obtained

by shifting the device along its vertical and horizontal axis [29, 30, 31, 32, 33, 34, 35]. While one of the most remarkable features of recombination–radiation measurements is the excellent spatial resolution of about $6\,\mu\mathrm{m}$ to $10\,\mu\mathrm{m}$, this technique suffers from several major drawbacks [36, 37, 38]: First, it is sensitive to the product np, i. e. carriers can only be detected if both types are present simultaneously. Consequently, this method is not applicable for the investigation of the minority carrier distribution. Second, the weak radiation intensity requires highly sensitive preamplifiers and an extensive averaging for a sufficient amount of time, thus raising great demands for stable operating conditions during the experiment. Third, since reabsorption attenuates the radiation which is emitted from the inner regions of the sample, the detected intensity primarily originates in those carriers which recombine close to the device surfaces. Therefore, the measurement results are seriously affected by surface recombination unless the samples are prepared properly.

During the same decades, another probing technique for the measurement of carrier distributions has been established, which exploits the dependence of the absorption coefficient on the carrier concentration [39, 40, 33, 41, 42, 43, 44]. While an infrared laser beam traverses the device under test, the latter is subjected to current pulses. Since photon scattering by free carriers causes an enhanced absorption during the duty cycle, the local concentration of the injected carriers can be extracted from the decrease of the transmitted light intensity. A detailed description of this method is given in section (2.2.1). Compared to recombination–radiation measurements, the free carrier absorption technique provides a superior sensitivity and a higher time resolution [37, 38]. It will be demonstrated in section (4.2) that an optimized optical setup facilitates an improved spatial resolution which is then comparable to that of recombination–radiation measurements.

Since the invention of optical probing techniques, the temperature distribution has become accessible to measurements, too. For example, thermal expansion causes surface displacements, which are detected by a laser interferometer [45]. This method has been employed for the investigation of transport phenomena [46], the testing of power devices [47] and opto–electronic devices [48], and the calibration of another temperature–sensitive measurement technique [49].

The fundamental physical principle of another class of characterization methods is the temperature dependence of the refractive index. Since the reflection coefficient is therefore affected by the surface temperature, the latter can be extracted from the intensity modulation of a laser beam which is reflected off the chip surface [50, 47, 51]. Laser Reflectance Thermometry constitutes a useful technique for the investigation of thermal and plasma waves which are excited by a modulated pump laser beam [50, 46].

Pioneering work has been done by Fournier and Boccara who introduced the Photothermal Deflection Spectroscopy, also known as the ”mirage effect” [52, 53]. The basic mechanism is the deflection of a probing beam due to thermally induced *gradients* of the refractive index. As the sample is periodically heated by a pump laser beam, temperature inhomogeneities are built up in the adjacent air around the sample, thus deflecting a second laser beam which is propagating closely above and parallel to the device surface (*external* mirage effect [54, 55]). The same physical principle is the foundation of the *internal* mirage effect. For that purpose, the probing beam penetrates the sample and is deflected in the *interior* of the device [56]. Since the refractive index of semiconductors is affected by both the temperature and the carrier density, this non–destructive and non–invasive technique is well–suited for the investigation of electric and thermal transport phenomena and material properties. A similar setup without a pump laser is employed for the

characterization of power devices to detect the gradients of carrier concentration and temperature which arise from the electrothermal behavior during typical switching conditions [49, 57].

Since parallel surfaces of the investigated devices form a Fabry–Perot resonator cavity whose optical thickness is affected by modulations of the refractive index, the temperature evolution can be extracted from the oscillations of the reflected or transmitted intensity [58]. As an alternative, the phase shift of the probing beam during a single propagation through the sample is detected by interference with a reference beam [59, 60, 61].

Fabry–Perot thermometry, free carrier absorption and laser deflection measurements are representative examples of internal laser probing techniques (cf. chapter 2) which exploit the dependence of the complex refractive index on carrier concentration and lattice temperature. Detecting the absorption, the deflection, or the phase shift of a probing beam, they directly provide space–resolved and time–resolved information about the distribution of carriers and heat in the interior of power devices. Due to their excellent spatial and temporal resolution, a more comprehensive analysis of the electric switching behavior and the thermal power dissipation has become possible, thus promising great benefit for the development of power devices and the validation of numerical simulation models.

1.4 Scope of this thesis

While the above mentioned internal laser probing techniques are still subject to intense research activities, novel concepts of additional characterization methods exploiting the electro–optical and the thermo–optical effect are currently being discussed, as well. The first part of this thesis is therefore dedicated to a theoretical study to support the improvement and development of these techniques. A physically rigorous model for the simulation of the entire measurement process is introduced which includes an electrothermal simulation of the device under test, the calculation of the laser beam propagation through the sample, the lenses and aperture holes, and finally the simulation of the detector response. In addition, analytical models are employed to supplement the numerical analysis of specific physical phenomena and to deduce the evaluation rules for the detector signals. Both the analytical and the numerical calculations are carried out to improve our understanding of the measurement process, in particular to identify and to tackle the most significant sources of error, to investigate parasitic effects such as multiple reflections, and to find a way to cope with practical restrictions as, for example, the minimum sample size which can be handled and prepared. Although attendant simulations represent a valuable methodology to study the measurement process itself, inverse modeling is impracticable for series measurements. The experiments are therefore designed in such a way that that the desired measurand can be directly extracted from the detector signal by applying the corresponding evaluation rule. Major goals of the optimization are a minimum experimental error, a large measurement range, and a minimum sensitivity to undesired effects. Some tendencies are already well–known, but the numerical analysis reveals *quantitative* results for the optimum parameters of the experimental setup (e. g. the beam aperture) and the sample geometry (e. g. the size of the active region) and enables to judge the accuracy of the optimized measurements.

In the second part of this thesis, it will be demonstrated that the availability of experimental

carrier concentration and temperature profiles crucially improves the validation and calibration of electrothermal device simulation models. Special attention is paid to the self–consistent extension of the drift–diffusion model which is derived from the principles of irreversible thermodynamics [62]. Whether it properly reflects the physics of power devices is judged by a thorough comparison of calculated and measured results which include the terminal characteristics as well as the internal carrier and temperature distribution under stationary and dynamic operating conditions. Since, for example, laser deflection measurements are able to detect temperature gradients of only $10\,\mu\mathrm{K}/\mu\mathrm{m}$ [49], a precise analysis of the heat source distribution has become feasible which reveals even minor shortcomings of the thermal model. In addition, as a single characteristic may be affected by several models in the same way, the comprehensive variety of experimental data favors the identification and extraction of the model parameters, thus supporting the endeavor to find reliable and generally valid simulation models.

Chapter 2

Internal Laser Probing Techniques

The introduction of optical characterization methods has opened up promising perspectives for the investigation of semiconductor devices. In this work, we will focus our interest on internal laser probing techniques which exploit the electro–optical and the thermo–optical effect and work in following manner: Under transient switching conditions, variations in temperature and the injection or removal of carriers give rise to modulations of the complex refractive index in the interior of the investigated sample. Detecting the resulting changes of the absorption, the deflection, or the phase shift of an incident probing beam, information about the local carrier concentration and lattice temperature can be extracted. Thus, space–resolved profiles of these quantities are scanned by shifting the device along its vertical or horizontal axis.

This chapter will illustrate the underlying physical principles and give a survey of the existing probing techniques as well as of the novel concepts which are currently being developed.

2.1 Physical background – modulation of the refractive index

In the following section, the interaction of infrared electromagnetic waves with free carriers (plasma–optical effect) and lattice vibrations (thermo–optical effect) is discussed. However, this discussion is restricted to an overview of the common models and a summary of the most significant experimental results. For further details the reader is referred to the cited publications (see also [57, 58] and the references therein).

In this work, n_{Si} denotes the real part of the refractive index in silicon, since it is common practice in semiconductor physics to represent the concentration of free electrons by the symbol n.

2.1.1 The plasma–optical effect

The optical absorption near the band gap of silicon proceeds by two major processes, namely band–to–band absorption (interband transitions) and free carrier absorption (intraband transitions) [63]. On account of the high density of states at the band edges, the optical excitation of an electron from the valence band to the conduction band exhibits only a very weak dependence on the carrier concentration [64]. Therefore, a *change* of the absorption coefficient due to the injection of carriers

originates from an increase of the free carrier absorption. Together with the corresponding effect on the real part of the refractive index this will be discussed in the following paragraphs.

The Drude theory

A classical description of the plasma–optical effect is known as the Drude model (cf. e. g. [65]) which regards the charge carriers as harmonic oscillators with vanishing binding energies. Solving the equation of motion

$$m^*_{n,p}\,\ddot{\vec{r}}_{n,p} + \frac{m^*_{n,p}}{\tau^c_{n,p}}\,\dot{\vec{r}}_{n,p} = \pm q\vec{E} \tag{2.1}$$

in the frequency domain yields the electrical polarization $\vec{P}_n = -qn\vec{r}_n$ and $\vec{P}_p = qp\vec{r}_p$ due to the displacement of the electrons and the holes, respectively. The refractive index is then calculated by expanding the square root of the dielectric constant $\sqrt{\varepsilon_R(\omega)} = n_{Si} + i\alpha/2k_0$ in terms of $1/\omega$. Hence, we obtain

$$n_{Si} = n_r - \frac{q^2\lambda^2}{8\pi^2\varepsilon_0 n_r c^2}\left(\frac{n}{m^*_n} + \frac{p}{m^*_p}\right) \tag{2.2}$$

$$\alpha = \frac{q^3\lambda^2}{4\pi^2\varepsilon_0 n_r c^3}\left(\frac{n}{m^{*2}_n\mu_n} + \frac{p}{m^{*2}_p\mu_p}\right) \tag{2.3}$$

where n_r denotes the static refractive index. Inserting the material parameters of silicon (cf. section 6.2) leads to the following result ($\lambda = 1.3\,\mu\mathrm{m}$):

$$n_{Si} = 3.5 - 8.33\cdot 10^{-22}\,\mathrm{cm}^3\,n - 5.70\cdot 10^{-22}\,\mathrm{cm}^3\,p \tag{2.4}$$

$$\alpha = 2.51\cdot 10^{-19}\,\mathrm{cm}^2\,n + 3.91\cdot 10^{-19}\,\mathrm{cm}^2\,p \tag{2.5}$$

However, an accurate description of the plasma–optical effect needs to consider all scattering processes that are assisted by phonons or impurities since free carrier absorption requires the interaction with a third particle for momentum conservation. Although these effects are implicitly included in the relaxation times $\tau^c_{n,p}$ (or the mobility), they are not treated rigorously by the Drude model. Another major drawback is the assumption of parabolic band edges which fails for optical excitation to higher states in the same band.

Absorption spectra and the Kramers–Kronig relation

Huang et al. [66] carried out a quantum mechanical analysis of the absorption coefficient which is assumed proportional to the transition rate

$$W_\pm = \int_0^\infty W_\pm(k) f(E_k) 2VN(E_k)\,dE_k\ . \tag{2.6}$$

Here, $W_\pm(k)$ denotes the transition rate for phonon emission or absorption, E_k the electron energy, $f(E_k)$ the distribution function, and $N(E_k)$ the density of states. Huang et al. employed the models for acoustic–phonon scattering, optical–phonon scattering, and impurity scattering [67] and

added a description of the nonparabolic band structure. Since the scattering mechanisms are proportional to the free carrier concentration, the absorption coefficient in the near–infrared spectrum also exhibits a linear dependence on the electron and hole concentrations.

Knowing the frequency dependent absorption coefficient, the real part of the refractive index is calculated from the Kramers–Kronig relation (cf. e. g. [68, 66]) which relates the real part n_{Si} of the refractive index to its imaginary part and thus to the absorption coefficient α

$$n_{Si}(\omega) = 1 + \frac{c}{\pi} P \int_0^\infty \frac{\alpha(\omega')}{\omega'^2 - \omega^2}\, d\omega' \ . \tag{2.7}$$

In this equation the symbol P denotes the Cauchy principal value. For a carrier density of $10^{16}\,\mathrm{cm}^{-3}$, which is a typical concentration in the electron–hole plasma of forward biased power devices, Huang et al. reported a concentration dependence proportional to $n^{1.07}$ and $p^{1.03}$.

A similar strategy is followed by Soref and Bennett [64]. They collected experimental data about the absorption spectra of heavily doped samples [69, 70] and calculated the real part of the refractive index from the Kramers–Kronig relation. Their results are proportional to $n^{1.05}$ and $p^{0.805}$ for a wavelength of $1.3\,\mu\mathrm{m}$.

Experimental results

The carrier concentration dependence of the absorption coefficient was ascertained by several authors from absorption measurements on uniformly doped samples [69, 70, 71, 72]. In this case, the major attenuation of the probing beam originates from the interaction with the majority carriers which are thermally generated by the doping atoms. These experiments separately reveal the dependence on both the electron concentration and the hole concentration, i. e. $\partial\alpha/\partial n$ and $\partial\alpha/\partial p$, respectively. However, the probing of the carrier distribution in power devices requires knowledge of the absorption coefficient under high injection conditions.

The optical absorption of an electron–hole plasma was measured on forward conducting pin diodes by Horwitz and Swanson [73]. They employed a vertically propagating probing beam which penetrates through a grid contact at the top and is reflected at the bottom metallization layer. The detected damping is therefore sensitive to the integral of the carrier distribution which, in turn, is calculated from the measured diffusion potential across the junctions. As a result, they reported a non–linear dependence of the absorption coefficient on the injected carrier density.

Laterally irradiating the sample by a monochromatic light source, Schierwater [33] detected the transmitted intensity at various current densities, at each of which the total stored charge was determined by reverse–recovery measurements (cf. section 1.3). For a wavelength of $1.3\,\mu\mathrm{m}$, he obtained the linear dependence $\Delta\alpha = 2.80 \cdot 10^{-18}\,\mathrm{cm}^2 \cdot \Delta n$ on the injected carrier density Δn. While Schierwater's setup is sensitive to an average concentration of the injected carriers, recent experiments employed a focussed laser beam which enables a space–resolved determination of the characteristic carrier concentration profile. Again, calibration was done by measuring the reverse–recovery behavior [74, 75] or the diffusion potential [76].

A direct measurement of the real part of the refractive index was performed by Yu et al. [77]. They detected the transmittance of a Fabry–Perot resonator while its optical thickness is modulated by optically generated charge carriers. However, their results are about one order of magnitude

larger than the theoretical predictions based on the Kramers–Kronig relation. The difference might arise from the evaluation strategy which attributes the decrease of the transmitted light intensity merely to the modulation of the refractive index but neglects the enhanced free carrier absorption. Another approach is reported by Deboy [49, 57] et al. who performed simultaneous laser absorption and deflection measurements (cf. section 2.2.2). Comparing the carrier contribution to the deflection signal with the calibrated absorption signal yields the refractive index modulation $\Delta n_{Si} = -4.58 \cdot 10^{-21}\,\mathrm{cm}^3 \cdot \Delta n$.

Comparison of the available data

If we introduce the definitions

$$\left(\frac{\partial n_{Si}}{\partial C}\right)_T := \left(\frac{\partial n_{Si}}{\partial n}\right)_{p,T} + \left(\frac{\partial n_{Si}}{\partial p}\right)_{n,T} \tag{2.8}$$

$$\left(\frac{\partial \alpha}{\partial C}\right)_T := \left(\frac{\partial \alpha}{\partial n}\right)_{p,T} + \left(\frac{\partial \alpha}{\partial p}\right)_{n,T} \tag{2.9}$$

the isothermal modulations of the complex refractive index under high injection conditions ($n = p$) can be written as

$$\Delta n_{Si} = \left(\frac{\partial n_{Si}}{\partial C}\right)_T \Delta n\ , \tag{2.10}$$

$$\Delta \alpha = \left(\frac{\partial \alpha}{\partial C}\right)_T \Delta n\ . \tag{2.11}$$

A comparison of reported data is shown in figure (2.1) and table (2.1). As far as the carrier concentration dependence of the absorption coefficient is concerned, we realize a satisfying agreement between the data reported by Huang et al., Soref et al., and Horwitz et al., while Schierwater's and

	$(\partial n_{Si}/\partial C)_T$	$(\partial \alpha/\partial C)_T$
Drude theory	$-1.40 \cdot 10^{-21}\,\mathrm{cm}^3$	$6.42 \cdot 10^{-19}\,\mathrm{cm}^2$
Huang et al. [66]	$-4.36 \cdot 10^{-22}\,\mathrm{cm}^3$	$5.11 \cdot 10^{-18}\,\mathrm{cm}^2$
Soref, Bennett [64]	$-3.57 \cdot 10^{-21}\,\mathrm{cm}^3$	$5.92 \cdot 10^{-18}\,\mathrm{cm}^2$
Schierwater [33]		$2.80 \cdot 10^{-18}\,\mathrm{cm}^2$
Horwitz, Swanson [73]		$4.85 \cdot 10^{-18}\,\mathrm{cm}^2$
Deboy [49]	$-4.58 \cdot 10^{-21}\,\mathrm{cm}^3$	
Schlögl [76]		$8.1 \cdot 10^{-18}\,\mathrm{cm}^2$

Table 2.1: Carrier concentration dependence of the refractive index and the absorption coefficient for $n = p = 1 \cdot 10^{16}\,\mathrm{cm}^{-3}$, $T = 300\,\mathrm{K}$, $\lambda = 1.3\,\mu\mathrm{m}$. (Most of the data is published graphically. In this case, the fit functions are adopted from [57].)

Schlögl's experimental results differ by a factor of 2. On the other hand, the published data about the carrier concentration dependence of the refractive index varies about one order of magnitude, which is not very surprising if we regard the difficulties in measuring this coefficient. Fortunately, this uncertainty does not constitute a serious problem since, as will be shown later, the typical operating conditions of power devices modulate the real part of the refractive index predominantly by the thermo–optical effect.

In this thesis, the following dependence on the carrier concentration is assumed [78]:

$$\left(\frac{\partial n_{Si}}{\partial C}\right)_T = -1.81 \cdot 10^{-21}\,\mathrm{cm}^3 \tag{2.12}$$

$$\left(\frac{\partial \alpha}{\partial C}\right)_T = 5.11 \cdot 10^{-18}\,\mathrm{cm}^2 \tag{2.13}$$

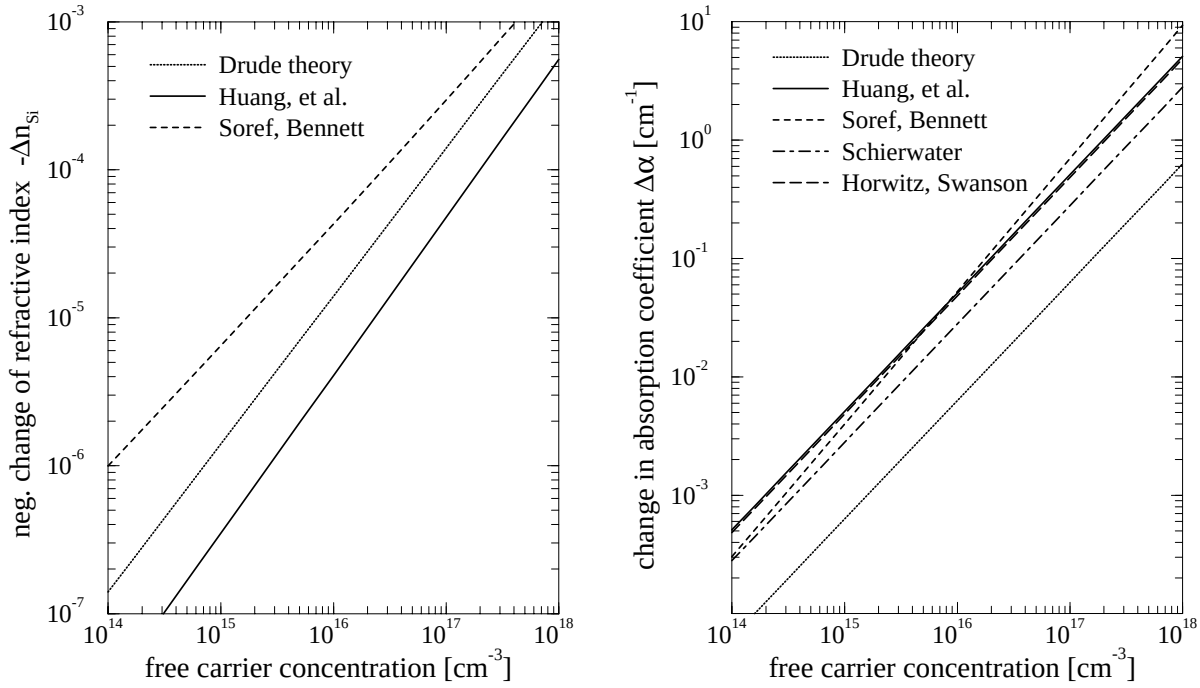


Figure 2.1: Silicon refractive index (left) and absorption coefficient (right) for equal carrier concentrations ($T = 300\,\mathrm{K}$, $\lambda = 1.3\,\mu\mathrm{m}$).

2.1.2 The thermo–optical effect

Thermally induced modulations of the complex refractive index originate from three major effects, namely the changes in the distribution functions of carriers and phonons, the temperature induced shrinkage of the band gap, and the thermal crystal expansion. As temperature rises, the latter mechanism decreases the optical density and, consequently, the refractive index. Since an increase is observed in reality, the two former effects are obviously decisive.

Temperature dependence of the absorption coefficient

Considering interband transitions with the absorption or the emission of an acoustic or an optical phonon, Macfarlane et al. [79] developed a model for the temperature dependence of the band–to–band absorption. Their model reproduces accurately the measured absorption coefficient at elevated temperatures where interband absorption becomes the most significant mechanism at a wavelength of $1.3\,\mu\mathrm{m}$ [63, 80]. A semi–empirical relation accounting for the temperature dependence of free carrier absorption processes was gained from spectral emissivity measurements [81].

Temperature dependent measurements of the absorption coefficient are reported by Schierwater [33] and Schlögl [76]. Either of them carried out his above mentioned experiment at various temperatures ranging from 150 K to 400 K. From their graphically published data, the temperature dependence of the absorption coefficient can be extracted and amounts to about $(\partial\alpha/\partial T)_{n,p} \approx 10^{-4}\,\mathrm{cm}^{-1}\mathrm{K}^{-1}$ at a carrier concentration of $1 \cdot 10^{16}\,\mathrm{cm}^{-3}$.

However, as temperature rises in the interior of power devices subjected to the typical operating conditions during absorption measurements, the crucial effect on the absorption coefficient originates from an enhanced injection of carriers by the emitter regions, which, in turn, increases the free carrier absorption. In other words, among the two contributions

$$\frac{d\alpha\,(n,p,T)}{dT} \overset{n=p}{=} \left(\frac{\partial\alpha}{\partial T}\right)_{n,p} + \left(\frac{\partial\alpha}{\partial C}\right)_{T} \cdot \frac{dn}{dT} \tag{2.14}$$

the latter significantly exceeds the former. The temperature dependence $(\partial\alpha/\partial T)_{n,p}$ of the absorption coefficient may therefore be neglected at all.

Temperature dependence of the refractive index

The crucial parameter for temperature measurements by means of internal laser probing is the temperature coefficient of the refractive index. This coefficient can be directly measured by various optical techniques. Their basic idea is to detect the thermally induced modulation of the optical sample thickness by ellipsometry [82], by interferometry exploiting the Fabry–Perot effect [83, 84, 85, 78], or by interferometry with respect to a reference beam of constant phase [86]. If the measurements are carried out on electrical devices, the temperature rise originates from self–heating and has to be ascertained from the electrical power dissipation [78], from the saturation currents of a MOSFET [86], or from the thermal expansion [49], for example. Another method is reported by Bertolotti et al. [87] whose samples were shaped as prisms and illuminated by a monochromatic light ray. The refractive index of the prism can thus be determined from the minimum refraction angle.

Table (2.2) presents a comparison of the various experimental results. In this thesis, the following temperature coefficient at $1.3\,\mu\mathrm{m}$ is used [78]:

$$\left(\frac{\partial n_{Si}}{\partial T}\right)_{n,p} = 1.60 \cdot 10^{-4}\,\mathrm{K}^{-1} \tag{2.15}$$

	$\partial n_{Si}/\partial T$	temperature range
Magunow [84]	$2.0 \cdot 10^{-4}\,\mathrm{K}^{-1}$	room temperature
Bertolotti et al. [87]	$1.8 \cdot 10^{-4}\,\mathrm{K}^{-1}$	$15^0\,\mathrm{C} \ldots 35^0\,\mathrm{C}$
Jellsion et al. [82]	$2.4 \cdot 10^{-4}\,\mathrm{K}^{-1}$	$25^0\,\mathrm{C} \ldots 75^0\,\mathrm{C}$
Icenogle et al. [83]	$1.5 \cdot 10^{-4}\,\mathrm{K}^{-1}$	room temperature
Seliger et al. [86]	$1.6 \cdot 10^{-4}\,\mathrm{K}^{-1}$	room temperature
Hille et al. [78]	$1.6 \cdot 10^{-4}\,\mathrm{K}^{-1}$	$25^0\,\mathrm{C} \ldots 125^0\,\mathrm{C}$

Table 2.2: Temperature dependence of the refractive index for a wavelength of $\lambda = 1.3\,\mu\mathrm{m}$ (data adapted from [78]).

2.2 Measurement techniques

Even several decades ago, laser probing was employed for investigating the carrier distribution in semiconductor samples [39, 40]. In recent years, additional methods have been developed which exploit the electro–optical and the thermo–optical effect and thus facilitate either a vertical or a lateral scanning of carrier concentration and temperature profiles. The following section provides an overview of the internal laser probing techniques which are already available or will be realized in the near future.

2.2.1 Free carrier absorption measurements

The carrier distribution of an electron–hole plasma in the intrinsic region of power devices can be determined by free carrier absorption measurements. A basic sketch of the experimental setup and a typical measurement signal are shown in fig. (2.2) and (2.3), respectively.

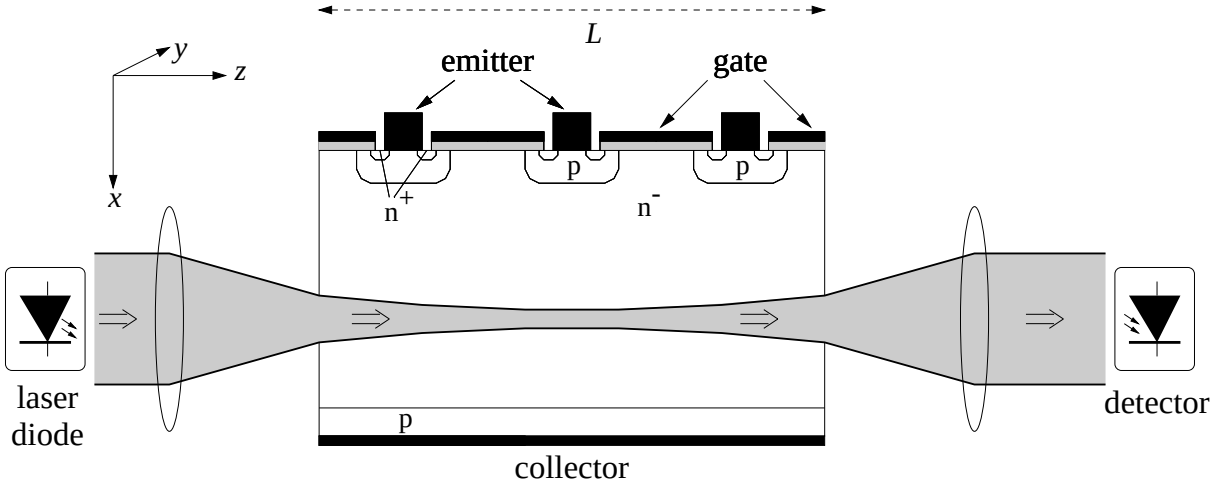


Figure 2.2: Experimental setup for free carrier absorption measurements.

The cw (continuous wave) laser beam is focussed onto the device under test which is subjected to periodic current pulses. Since the injection of carriers enhances the absorption, a decrease of the transmitted light intensity is detected by the photo diode. To improve the signal–to–noise ratio, the detector signal is averaged for several hundreds of current pulses. The excess carrier concentration $\Delta n(x_0, t)$ at the current beam position x_0 is then extracted from the absorption law

$$I_{on}(x_0, t) = I_{off}(x_0) \exp\left(-L\frac{\partial\alpha}{\partial C}\Delta n(x_0, t)\right) \tag{2.16}$$

where I_{on} and I_{off} denote the transmitted intensities during on–state and off–state of the device, respectively, and L is the interaction length. To eliminate the unknown incident intensity which may be subject to a thermal drift during the measurement, the AC component of the photo current is scaled to the DC component. This quantity shall be called the absorption signal (cf. fig. 2.3). In case of small duty cycles, it is equal to $I_{on}/I_{off} - 1$. Repeating the measurements at different beam positions by shifting the devices along its vertical axis enables a vertical scanning of the carrier concentration profile. The spatial resolution is about $10\,\mu\mathrm{m} - 20\,\mu\mathrm{m}$, while the time resolution is limited by the detector capacity to typically some hundred nanoseconds.

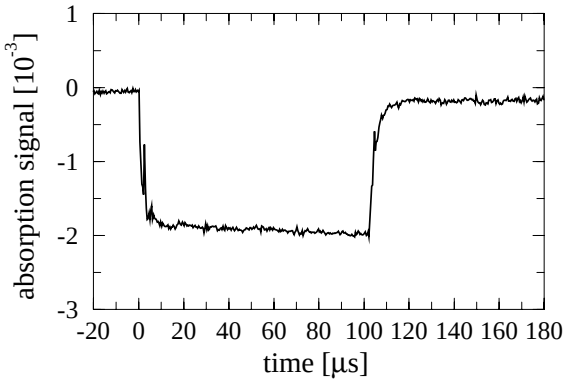


Figure 2.3: Measurement signal for free carrier absorption measurements (pulse duration $100\,\mu\mathrm{s}$).

Over the years, various modifications of this probing technique have been developed. Commonly employed light sources are HeNe lasers with a wavelength of $\lambda = 3.39\,\mu\mathrm{m}$ [42, 43, 44, 88], InGaAs laser diodes [49, 89] ($\lambda = 1.3\,\mu\mathrm{m}$) or an incoherent ASE (Amplified Spontaneous Emission) light source with a wavelength of $1.55\,\mu\mathrm{m}$ [74, 75, 90]. Special periods during the current pulse can be investigated by pulsing the laser beam or by using a mechanical chopper [43]. The sample preparation includes the polishing of the surfaces and the deposition of an antireflective coating. The latter step is not required if multiple reflections at the parallel surfaces are suppressed by the use of an incoherent light source or by an appropriate sample alignment exploiting the vanishing reflection coefficient at the Brewster angle [43, 88, 89].

It should be pointed out that the evaluation of the detector signal according to equation (2.16) is based on the following assumptions:

- The observed intensity decrease is due to an enhanced free carrier absorption. This assumption is violated, for example, if self–heating modulates the optical thickness $n_{Si}(T)L(T)$ of the sample, thus changing the transmittance of the Fabry–Perot cavity which is formed by the parallel device surfaces.

- The lateral extension of the probing beam is negligible.

- The carrier distribution along the beam path is constant. Otherwise, the product $n(x_0)L$ has to be replaced by $\int_0^L n(x_0, z)\, dz$. In this case, the equation (2.16) cannot be solved for the unknown carrier distribution.

- The optical axis is perpendicular to the device surfaces. If the probing beam traverses the sample at an angle Φ_{Si} in silicon, the intensity decrease at the beam position x_0 is given by

$$I_{on}(x_0) = I_{off}(x_0)\, \exp\left(-\frac{\partial\alpha}{\partial C}\int_0^L \Delta n(x_0 + z\tan\Phi_{Si}, z)\, dz\right) \quad . \tag{2.17}$$

- The deflection of the probing beam due to internal gradients of the refractive index is negligible. This condition is not satisfied during operating conditions with high power dissipation. The resulting temperature gradients give rise to a beam displacement which can, for example, easily amount up to $20\,\mu\mathrm{m}$ in the interior of a sample which is only $800\,\mu\mathrm{m}$ in length (cf. fig. 4.31).

2.2.2 Internal Laser Deflection measurements

Inhomogeneous distributions of any state variable that affects the refractive index give rise to a corresponding gradient in the refractive index. Detecting the resulting deflection of an incident laser beam constitutes the fundamental idea of the photothermal deflection spectroscopy. This mechanism is also known as the mirage–effect and was exploited to study electrical and thermal transport phenomena [56]. The same principle underlies the Internal Laser Deflection method [49] which has emerged as a sensitive technique to ascertain temperature profiles in the interior of power devices. Except for the detector, the setup (cf. fig. 2.4) is similar to the laser absorption technique described in the previous section.

Under transient switching conditions, both the injection (or removal) of carriers and a local change in temperature cause a gradient of the refractive index which deflects the focussed laser

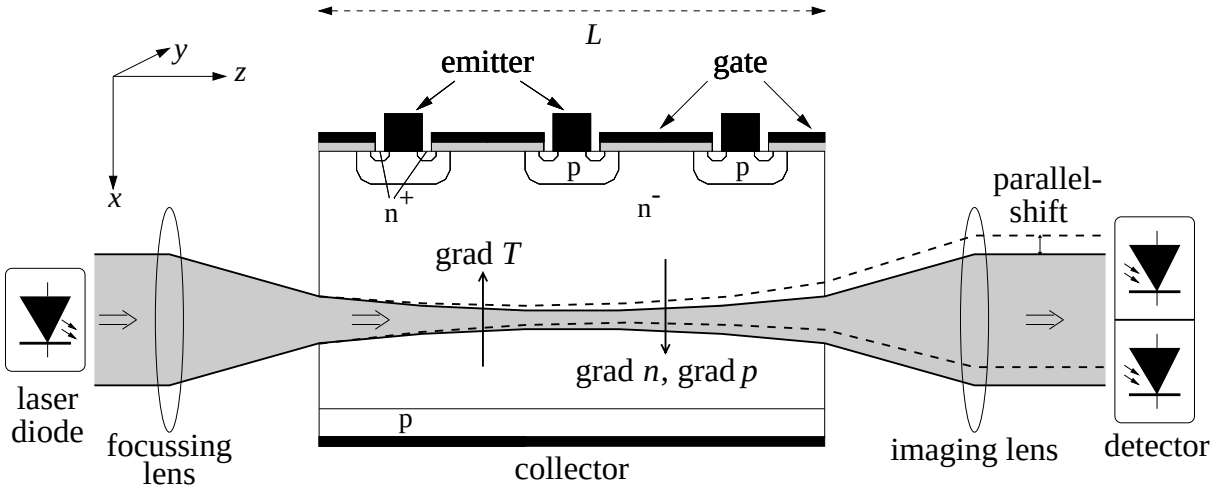


Figure 2.4: Experimental setup for Internal Laser Deflection measurements.

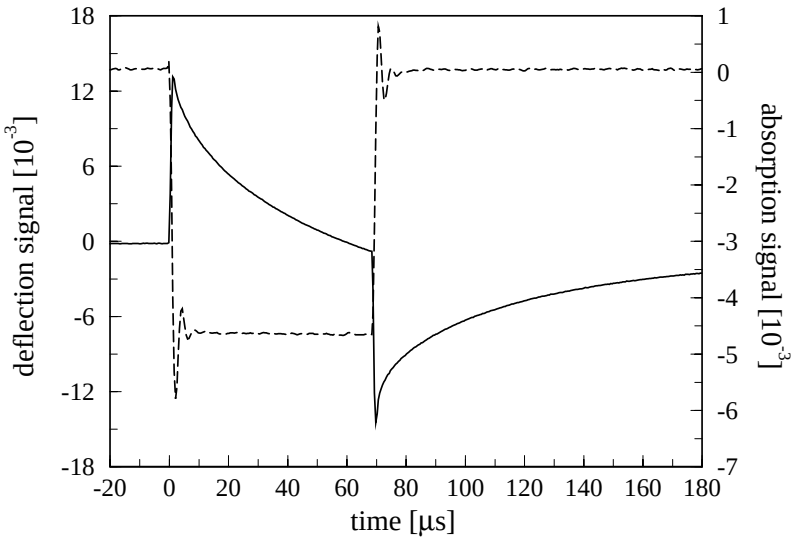


Figure 2.5: Deflection signal (solid line) and absorption signal (broken line) for Internal Laser Deflection measurements on a power diode subjected to a current pulse of $70\ \mu\mathrm{s}$ *in duration.*

beam. A long distance objective transforms the internal deflection approximately into a parallel shift which is detected by a four–quadrant photo diode. While the difference between the photo currents of two opposite segments is related to the vertical or horizontal deflection, the sum of all photo currents is proportional to the total intensity. Thus, laser deflection and laser absorption measurements can be performed simultaneously.

Like the absorption signal, the deflection signal is also scaled to the total transmitted intensity in order to eliminate externally induced disturbances. An example of typical measurement signals is shown in fig. (2.5). The rapid change of the deflection signal immediately after turn–on and turn–off ($t = 0$ and $t = 70\ \mu\mathrm{s}$, respectively) originates from the injection or removal of carriers in the intrinsic region. The thermal contribution to the deflection signal is visible on the time scale of several tens of microseconds.

Assuming the paraxial approximation, a quantitative relation of the refractive index gradient and the scaled deflection signal has been deduced [57]:

$$\begin{aligned} M(x,t) &:= \frac{I_1(x,t) - I_3(x,t)}{I_1(x,t) + I_2(x,t) + I_3(x,t) + I_4(x,t)} \\ &= \frac{Q}{\eta_A}\left(\int_0^L \frac{\partial n_{Si}}{\partial x}(x,z,t)\,dz + \frac{1}{n_{Si}F}\int_0^L \frac{\partial n_{Si}}{\partial x}(x,z,t)\,z\,dz \right) \end{aligned} \tag{2.18}$$

In this equation, η_A and Q denote the angular aperture of the probing beam and the slope of the detector response function (cf. section 4.3.5), respectively, I_1, I_2, I_3, and I_4 are the photo currents of the four detector segments, and F the distance between the device and the imaging lens. It should be mentioned, however, that Q depends on the laser spot size on the detector [78, ch. 5.4.2]. Changing the position of the lenses therefore demands a recalibration of the response function.

The desired information about the internal carrier concentration and temperature is extracted in the following way: In case $n_{Si}(x,z)$ is constant along the beam path, i. e. $n_{Si}(x,z) = n_{Si}(x)$, equation (2.18) can be solved for the refractive index gradient:

$$\frac{M(x,t)\cdot\eta_A}{QL(1+\frac{L}{2n_{Si}F})} = \frac{\partial n_{Si}}{\partial x}(x,t) = \left(\frac{\partial n_{Si}}{\partial C}\right)_T \frac{\partial n}{\partial x}(x,t) + \left(\frac{\partial n_{Si}}{\partial T}\right)_{n,p} \frac{\partial T}{\partial x}(x,t) \tag{2.19}$$

If the carrier concentration profile is in its stationary state before a significant temperature gradient is built up at the current beam position, the two contributions to the deflection signal can be separated due to their widely differing time constants (cf. fig. 2.5). This condition is violated if heat sources are closely located near the beam path. In this case, the carrier contribution has to be determined from the absorption signal or an electrothermal simulation and subtracted from the deflection signal.

To prevent disturbances by a thermal drift of the experimental setup, the detector is mounted on a piezo x–y–translation unit, which is controlled by low pass filters. Thus, the average deflection signals are kept at zero and the detector position is dynamically adjusted in the center of the laser beam.

Laser deflection and free carrier absorption measurements require the same sample preparation steps: After polishing the surfaces, an antireflective coating is deposited, unless the interference of repeatedly reflected rays in the interior of the device is suppressed by employing an incoherent light source or aligning the sample such that the laser beam propagates at the Brewster angle.

While the Internal Laser Deflection technique facilitates an excellent sensitivity of $25\,\mu\mathrm{K}/\mu\mathrm{m}$ for temperature measurements [57], it suffers from its limited measurement range due to the saturation of the detector response. During typical transient switching conditions, the maximum detectable temperature gradient with a typical interaction length of a few mm is exceeded if the peak temperature rises to only a few Kelvin (cf. [57, ch.7.4.2]).

2.2.3 Backside Laser Probing and Differential Backside Laser Probing

Like the Internal Laser Deflection method, Backside Laser Probing (cf. fig. 2.6 and 2.7) also exploits the dependence of the refractive index on carrier concentration and lattice temperature. The detected quantity, however, is the phase shift $\Delta\varphi(t)$ of a vertically propagating probing beam due to modulations of the optical sample thickness [91]:

$$\Delta\varphi(t) = 2 \cdot \frac{2\pi}{\lambda} \int_0^L \left[\left(\frac{\partial n_{Si}}{\partial T} \right)_{n,p} \Delta T(x,t) + \left(\frac{\partial n_{Si}}{\partial C} \right)_{T} \Delta n(x,t) \right] dx \tag{2.20}$$

This phase shift is detected by interference with a reference beam which is reflected off the rear metallization layer. While the laser diode is operating in cw (continuous wave) mode [61, 92, 93] or in pulsed mode [61], the device under test is subjected to periodic current pulses. A schematic view of the experimental setup is shown in fig. (2.6) and (2.7).

Heinrich et al. [59] proposed an interferometer including a Wollaston prism to split the original laser beam. His setup has been modified by Goldstein et al. [60] who exploited Bragg diffraction by the acoustic waves in the interior of an acousto–optic modulator (AOM). The diffraction angles are controlled by the driving frequencies ω_1 and ω_2, thus enabling an arbitrary positioning of the probing beam and the reference beam. After reflection and passing the AOM cell again, the two beams interfere on the detector, yielding an intensity signal of the form $\sin(2\Delta\omega t + \Delta\varphi(t))$.

Two different ways of signal processing are available: Including a local oscillator with frequency $2\Delta\omega$ (cf. fig. 2.7) [61, 92, 58] facilitates a highly sensitive detection of either the intensity modulation or the phase modulation of the probing beam. The phase of the local oscillator determines which quantity is actually observed. For the sake of an increased bandwidth, however, the

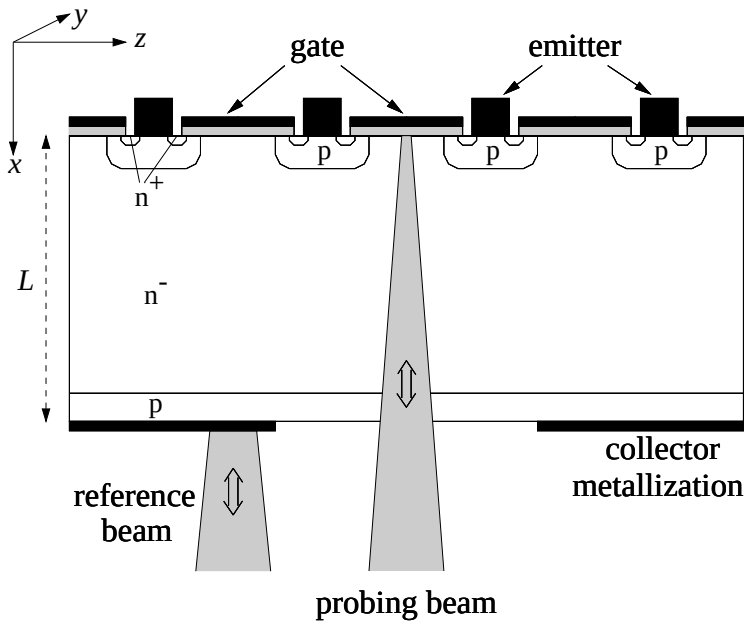


Figure 2.6: Backside Laser Probing of an IGBT sample: While the probing beam penetrates the active area and is reflected at the top metallization layer, the reference beam is reflected at the collector contact metallization layer.

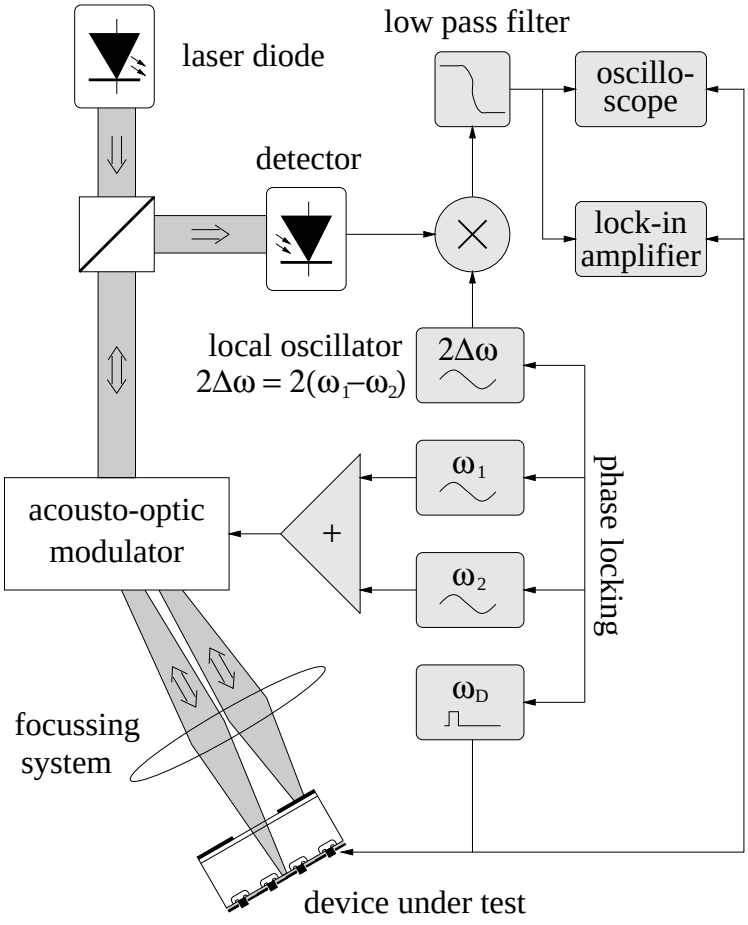


Figure 2.7: Experimental setup for Backside Laser Probing.

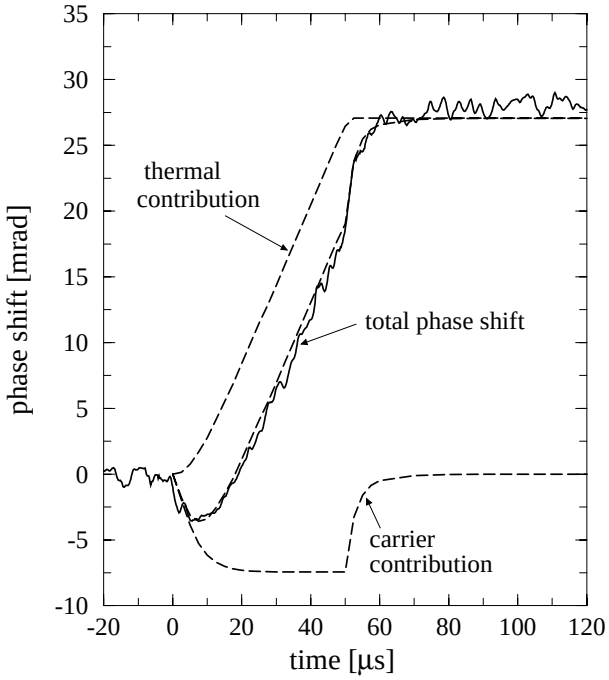


Figure 2.8: Phase shift signal under operating conditions with low power dissipation (pulse duration $50\,\mu\mathrm{s}$). To illustrate each contribution simulation results are included (broken lines).

local oscillator is omitted, enabling measurements of the phase modulations during fast transient switching conditions [58, 94].

A typical phase shift signal is shown in fig. (2.8). The two contributions originating from the injection/removal of carriers and the temperature rise during the current pulse can be easily identified due to their differing time constants. Though opposite in sign, they are of the same order of magnitude if the power dissipation in the interior of the sample is small. However, for a large temperature rise, as it typically occurs during short circuit operation of power devices, for example, the phase modulation by the carriers can be neglected in view of the thermal contribution.

The Backside Laser Probing technique is a valuable method for investigating lateral carrier concentration and temperature inhomogeneities. Its spatial resolution is limited by the laser spot size to approximately $2\,\mu\mathrm{m}$. Although only integral information on the vertical profiles is provided, this technique constitutes a useful characterization method for vertical power devices since it is capable of detecting large temperature rises of 100 K or more.

As a slight modification of the optical setup, *Differential Backside Laser Probing* has been employed for the investigation of IGBTs by Fürböck et al. [95]: Directing both beams into the active area of the sample enables a direct measurement of the difference between the temperatures at the two beam positions [96].

The Backside Laser Probing techniques require a special sample preparation: To provide access through the rear side a window of approximately $70\,\mu\mathrm{m} \times 70\,\mu\mathrm{m}$ in size has to be opened in the metallization layer. This is done by photolithographic structured etching with two subsequent etching steps using $\mathrm{HNO_3}$ and HF acids. Thus, the metallization can be removed without etching the silicon. The effect of the contact window on the carrier and temperature distribution is discussed in section (5.1.1). Finally, an antireflective coating is deposited to suppress multiple reflections within the substrate.

2.2.4 Other interferometric techniques

The above described measurement techniques have already been successfully employed for device characterization. However, additional internal laser probing techniques could be imagined which exploit the carrier concentration and temperature dependence of the refractive index. Regarding the typical operating conditions of power devices, the major interest is focussed on the capability of detecting a large temperature rise of 100 K or more. Some novel ideas have been already proposed and preliminary setups are realized by now. Their basic concepts shall be briefly outlined in this section.

Fabry–Perot reflectivity measurements

The reflectivity of a Fabry–Perot resonator is modulated by thermally induced variations of its optical thickness $n_{Si}(T)L(T)$. Hence, the temperature evolution within the sample can be detected by monitoring the reflected intensity [93, 97, 98]. The experimental setup which was employed to investigate lateral SOI power devices [58] is shown in fig. (2.9).

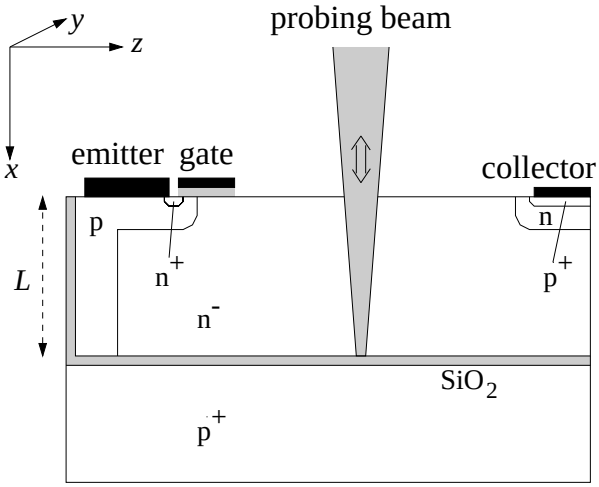


Figure 2.9: Determination of the temperature distribution in a SOI IGBT by Fabry–Perot reflectivity measurements. The probing beam is reflected back and forth at the top surface of the wafer and predominantly at the upper silicon/oxide interface.

Representing the reflection coefficients at the front surface and at the rear surface by r_f and r_r, respectively, the reflectivity R_{FP} of a lossy Fabry–Perot resonator of thickness L becomes [58]

$$R_{FP} = \frac{(r_f + r_r e^{-\alpha L})^2 - 4 r_f r_r e^{-\alpha L} \sin^2(\frac{2\pi}{\lambda} n_{Si} L)}{(1 + r_f r_r e^{-\alpha L})^2 - 4 r_f r_r e^{-\alpha L} \sin^2(\frac{2\pi}{\lambda} n_{Si} L)} \quad . \tag{2.21}$$

While two adjacent maxima of the reflectivity are observed, the optical thickness $n_{Si}(T) \cdot L(T)$ is changed by $\lambda/2$. The corresponding temperature difference ΔT is therefore equal to

$$\Delta T = \frac{\lambda}{2} \left(n_{Si} \frac{\partial L}{\partial T} + L \frac{\partial n_{Si}}{\partial T} \right)^{-1} \quad . \tag{2.22}$$

The temperature distribution in lateral SOI IGBTs has been determined by Fabry–Perot reflectivity measurements [99] by means of a vertically propagating laser beam. On the other hand, scanning temperature profiles in *vertical* devices requires a *lateral* incidence of the probing beam.

In this case, however, a limited measurement range is to be expected as the increasing internal deflection decreases the amplitude of the intensity oscillations. This problem is inherent to all interferometric techniques with a lateral beam propagation and will be addressed in greater detail in sections (5.2) and (5.3). Nevertheless, at least an integral information could be obtained by an experimental setup similar to fig. (2.6), including a laser beam impinging at the rear surface of the device. Such measurements, however, have not been reported, yet.

Fabry–Perot transmission measurements

A related concept is founded on the temperature dependence of the transmittance T_{FP} of a Fabry–Perot resonator:

$$T_{FP} = \frac{(1 - r_f r_r)^2 e^{-\alpha L}}{(1 + r_f r_r e^{-\alpha L})^2 - 4 r_f r_r e^{-\alpha L} \sin^2(\frac{2\pi}{\lambda} n_{Si} L)} \tag{2.23}$$

Though thermally induced Fabry–Perot oscillations of the transmitted intensity are an undesired effect during laser absorption measurements, they can be exploited for extracting the temperature rise in the interior of the sample. Choosing a coherent light source and omitting the deposition of an antireflective coating, the same experimental setup (cf. fig. 2.2) as for the absorption measurements can be employed. The temperature difference between two adjacent maxima is then calculated according to equation (2.22).

Mach–Zehnder interferometry

Mach–Zehnder interferometry (cf. fig. 2.10) constitutes another experimental method to scan vertical carrier concentration and temperature profiles [57]. The modulations of the refractive

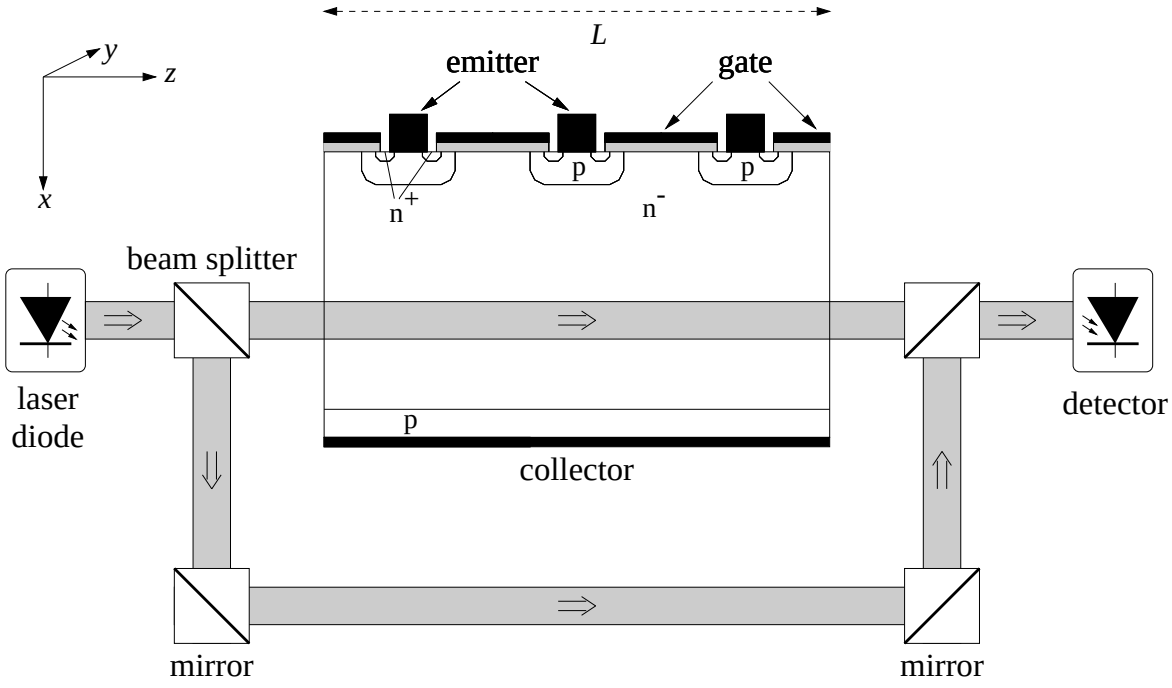


Figure 2.10: Experimental setup of the Mach–Zehnder interferometer.

index give rise to a phase shift of the laterally penetrating probing beam. This phase shift is detected by interference with a reference beam which is guided around the sample.

The Mach–Zehnder interferometry requires the use of a coherent light source. Multiple reflections between the parallel surfaces have to be suppressed by the deposition of an antireflective coating.

2.3 Summary

Internal laser probing techniques exploit the electro–optical and the thermo–optical effect and thus facilitate time–resolved measurements of the carrier and temperature distribution in the interior of power devices.

Table (2.3) compares the presented methods. Although they only reveal integral information along the beam path, the sensitivity and the measurement range is expressed in terms of the local carrier concentration and temperature values, assuming typical interaction lengths for the conversion. As far as the probing techniques have not been realized, yet, expected data is presented.

	free carrier absorption measurements	laser deflection measurements	backside laser probing	FP reflectivity measurements	FP transmittance measurements	Mach–Zehnder–interferometry
beam propagation	lateral	lateral	vertical	vertical	lateral	lateral
scan direction	vertical	vertical	lateral	lateral	vertical	vertical
detected quantities	$\int n\,dz$	$\int \partial_x n\,dz$, $\int \partial_x T\,dz$	$\int n\,dx$, $\int \Delta T\,dx$	$\int n\,dx$, $\int \Delta T\,dx$	$\int n\,dz$, $\int \Delta T\,dz$	$\int n\,dz$, $\int \Delta T\,dz$
sensitivity	$10^{14}\,\mathrm{cm}^{-3}$	$10^{11}\,\mathrm{cm}^{-3}/\mu\mathrm{m}$ $25\,\mu\mathrm{K}/\mu\mathrm{m}$	$10^{14}\,\mathrm{cm}^{-3}$ $5\,\mathrm{K}$	$10^{17}\,\mathrm{cm}^{-3}$ $10\,\mathrm{K}$	$10^{16}\,\mathrm{cm}^{-3}$ $1\,\mathrm{K}$	$10^{16}\,\mathrm{cm}^{-3}$ $1\,\mathrm{K}$
measurement range up to	$10^{18}\,\mathrm{cm}^{-3}$	$10^{14}\,\mathrm{cm}^{-3}/\mu\mathrm{m}$ $50\,\mathrm{mK}/\mu\mathrm{m}$	$10^{19}\,\mathrm{cm}^{-3}$ $> 500\,\mathrm{K}$	$10^{19}\,\mathrm{cm}^{-3}$ $> 500\,\mathrm{K}$	?	?
spatial resolution	$15\,\mu\mathrm{m}$	$15\,\mu\mathrm{m}$	$5\,\mu\mathrm{m}$	$5\,\mu\mathrm{m}$	$15\,\mu\mathrm{m}$	$15\,\mu\mathrm{m}$
calibration	–	detector response	–	–	–	–
light source	incoherent	incoherent	coherent	coherent	coherent	coherent
preparation	polishing of surfaces	polishing of surfaces	etching of a window deposition of an anti-reflective coating	–	polishing of surfaces	polishing of surfaces, deposition of an anti-reflective coating

Table 2.3: Comparison of internal laser probing techniques.

Chapter 3

Modeling of Optical Probing Techniques

The internal laser probing techniques described in the previous chapter provide valuable information that crucially enhances our insight into the internal behavior of semiconductor devices. An accurate evaluation of the measurement signals, however, requires a consistent analysis which exceeds the concepts of geometrical ray tracing or simple one–dimensional models. Within the scope of this work, a physically rigorous numerical model of the entire measurement process has been developed [100]. Parasitic effects which may be introduced by the sample preparation and inherent phenomena arising from the wave propagation of the probing beam can thus be investigated thoroughly. The major goal of the presented model is to support the design of the experiments and to improve the interpretation of the measurement results.

The following sections are dedicated to a detailed discussion of the basic ideas for modeling optical probing techniques. Throughout this chapter, the coordinate system is chosen in such a way that the laser beam propagates along the positive z–axis.

3.1 Simulation steps – an overview

Since the optically absorbed power is by several orders of magnitude smaller than the electric power dissipation, the electrothermal behavior of the investigated device is not affected by the penetrating probing beam. Hence, we can follow the simulation strategy sketched in fig. (3.1).

The first step is a transient electrothermal device simulation of the operating conditions the structure is subjected to during the measurement. The calculation takes into account specific properties of the prepared sample as, e. g., a window in the metallization layer (cf. section 2.2.3) or an enhanced surface recombination rate which originates from the sawing and polishing of the surfaces. Employing the general purpose device simulator $\mathrm{DESSIS_{ISE}}$ [101] which is founded on the electrothermal model described in chapter (6), the carrier concentration and temperature profiles are obtained as functions of space and time.

From these the space and time dependent modulations of the complex refractive index are calculated according to the dependencies summarized in section (3.2). The corresponding functional relation is incorporated into the input parser of the optical simulator, which reads the output files of the electrothermal simulator $\mathrm{DESSIS_{ISE}}$.

The most demanding problem is the calculation of the beam propagation through the device

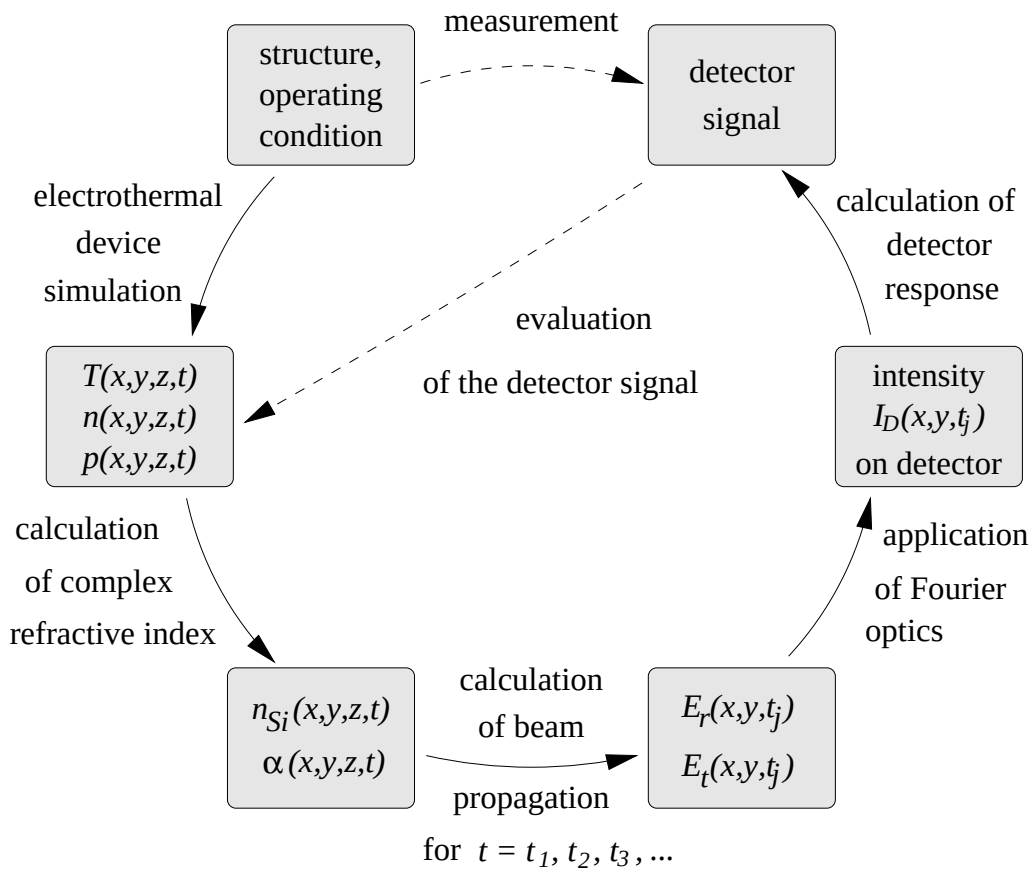


Figure 3.1: Simulation strategy for the physically rigorous modeling of internal laser probing techniques.

under test. Since the oscillations of the laser field are much faster than the modulations of any electrical or thermal quantity, the optical problem is solved for an appropriately chosen set of points of time t_1, t_2, t_3, ... at each of which the refractive index distribution is treated stationary. An efficient algorithm is derived from Maxwell's equations in the frequency domain (cf. section 3.3). As a result, we obtain the field distribution of the reflected and the transmitted waves at every point of time under consideration.

The image of the emerging wave on the detector is formed by an arrangement of lenses, aperture holes, etc. A suitable model for the beam propagation through this optical elements is gained from Fourier optics (cf. section 3.4).

Finally, integrating the field distribution over the active areas of the (possibly dynamically translated) detector (cf. section 2.2.2) reveals the photo currents of each detector segment. From these, the time–dependent measurement signal is constructed.

The following sections provide a more detailed discussion of the algorithms which are employed for each simulation step.

3.2 Calculation of the refractive index modulations

Principally, any of the models in section (2.1) can be chosen to describe the dependence of the complex refractive index on the carrier concentrations and the temperature. As mentioned above, however, the published data exhibits a rather large discrepancy. For that reason, we take another point of view: Since any of the proposed models exhibits an exactly or approximately linear dependence on carrier concentration and temperature, we use a first order Taylor expansion of the complex refractive index to evaluate the resulting effect on the laser beam:

$$\begin{aligned} n_{Si}(n,p,T) &= n_{Si,0} + \left(\frac{\partial n_{Si}}{\partial n}\right)_{p,T}\Delta n + \left(\frac{\partial n_{Si}}{\partial p}\right)_{n,T}\Delta p + \left(\frac{\partial n_{Si}}{\partial T}\right)_{n,p}\Delta T \\ &\overset{n=p}{=} n_{Si,0} + \left(\frac{\partial n_{Si}}{\partial C}\right)_{T}\Delta n + \left(\frac{\partial n_{Si}}{\partial T}\right)_{n,p}\Delta T \end{aligned} \tag{3.1}$$

$$\begin{aligned} \alpha(n,p,T) &= \alpha_0 + \left(\frac{\partial \alpha}{\partial n}\right)_{p,T}\Delta n + \left(\frac{\partial \alpha}{\partial p}\right)_{n,T}\Delta p + \left(\frac{\partial \alpha}{\partial T}\right)_{n,p}\Delta T \\ &\overset{n=p}{=} \alpha_0 + \left(\frac{\partial \alpha}{\partial C}\right)_{T}\Delta n + \left(\frac{\partial \alpha}{\partial T}\right)_{n,p}\Delta T \end{aligned} \tag{3.2}$$

Hence, it is sufficient to know the expansion coefficients. The constant terms may be determined from special parasitic effects, e. g. from the Fabry–Perot oscillations of the absorption signal (cf. section 4.2.4) whose amplitude decreases as α_0 increases. However, the desired detector signals do not depend on $n_{Si,0}$ and α_0 since the probing techniques actually detect the difference between off–state and on–state of the device under test. Therefore, the linear expansion coefficients represent the information we need to know. In this thesis, the following data is assumed (cf. section 2.1):

$$\begin{aligned} \left(\frac{\partial n_{Si}}{\partial C}\right)_{T} &= -1.81\cdot 10^{-21}\,\mathrm{cm}^3 & \left(\frac{\partial n_{Si}}{\partial T}\right)_{n,p} &= 1.60\cdot 10^{-4}\,\mathrm{K}^{-1} \\ \left(\frac{\partial \alpha}{\partial C}\right)_{T} &= 5.11\cdot 10^{-18}\,\mathrm{cm}^2 & \left(\frac{\partial \alpha}{\partial T}\right)_{n,p} &\approx 10^{-4}\,\mathrm{cm}^{-1}\mathrm{K}^{-1} \end{aligned} \tag{3.3}$$

Under typical operating conditions of power devices the most important effects are due to the carrier concentration dependence of the absorption coefficient $(\partial\alpha/\partial C)_T$ and the temperature dependence of the refractive index $(\partial n_{Si}/\partial T)_{n,p}$. It should be mentioned that in general all the expansion coefficients depend on the state variables and therefore have to be measured at a representative set of operating points if widely differing operating conditions are investigated.

3.3 Wave propagation in inhomogeneous media

In order to perform a physically rigorous simulation of the beam propagation through the device under test we have to calculate wave propagation in a medium with an inhomogeneous and in general arbitrary distribution of the complex refractive index. Its spatial variations arise for two major reasons: First, different kind of materials within the structure, e. g. silicon, oxide, or metallization layers, form regions with rather widely differing refractive indices. Second, additional variations

during transient switching of the device are due to the electro–optical and the thermo–optical effect. These modulations are therefore small but time–dependent.

3.3.1 Algorithms reported in the literature

Various numerical methods have been proposed in the literature (cf. e. g. the summary in [102]). Without claiming to present a complete overview, some of the most important shall be mentioned briefly. Since we have to calculate the optical field distribution in the entire sample at numerous points of time, the suitability of the algorithms is judged by their computational economy within a large simulation domain.

The fundamental idea of surface integral techniques, e. g. the method of moments (MOM) or the boundary element method (BEM), consists in expanding the unknown field distribution in terms of a set of basis functions representing an exact solution of the partial differential equation. The unknown variables are the expansion coefficients, which are determined such that the boundary conditions are satisfied. In particular, they have to minimize the scalar products of the residuals of the governing equations and a set of weighting functions. Common choices for the latter are the basis functions themselves (Garlerkin method) or Dirac's delta distributions at the grid nodes (point matching techniques).

A special case is the generalized multipole technique [103, 104]. The electromagnetic vector fields in homogeneous regions are expressed in terms of the electromagnetic multipole functions at different origins. Each of them represents an exact solution of the Helmholtz equation. The boundary conditions at a finite number of matching points constitute an overdetermined system of equations which is solved for the expansion coefficients in the least square sense. Although the rather small number of unknowns promises an efficient calculation this technique suffers from its restricted applicability to homogeneous domains and its difficult model setup [105].

A straightforward and generally applicable technique is gained from the finite element or finite difference discretization of the wave equation or Maxwell's equations [106, 107]. In order to get an acceptable discretization error, however, the mesh has to be refined up to 10 nodes per wavelength. Since the latter is about 400 nm in silicon, the typical dimensions of power devices would require a huge amount of grid nodes.

Considering a sequence of parallel planes perpendicular to the optical axis, beam propagation methods [108] calculate the field distribution on the following plane from the distribution on the preceding plane. The relation is derived from Maxwell's equations [109] or the scalar wave equation [110, 111, 112] by means of the paraxial approximation which enables an analytical solution for the unknown field values. Each propagation step therefore just comprises a matrix multiplication, thus constituting a very fast algorithm. However, the applicability of beam propagation techniques is restricted to paraxial wave propagation in one direction. Although longitudinal reflections can be included by calculating the propagation back and forth, numerical stability may be seriously affected by the inherent presence of parasitic modes with an exponentially increasing field distribution.

Multiple reflections can be calculated with a single computation step by multiplying matrices which allow for the propagation of both forward and backward travelling waves within a layer between two adjacent planes. The waveguide model [113, 114] assumes a longitudinally piece-

wise constant refractive index and matches the Fourier expansions of the electric and magnetic field distribution in each layer to ensure the continuity at the interfaces. The field distribution at the emergence plane is thus related to that on the entrance plane which is finally calculated from the radiation boundary conditions in propagation direction. The same strategy is the foundation of the differential method [115, 116, 117] which, however, includes a different derivation of the propagator matrices: A lateral Fourier transformation of Maxwell's equations leads to a first order differential equation in z–direction. Employing an explicit Euler discretization of the latter yields the desired linear relation of the field distributions on two adjacent planes. With these methods, the field distribution on the entrance plane is considered a superposition of forward and backward propagating waves, whereas it is equal to the incident wave if the above mentioned beam propagation techniques are employed.

The algorithm which is proposed here is similar to the waveguide model and the differential method. As the major improvement, a different set of variables enables a much coarser discretization without loss of accuracy. Furthermore, the propagator matrix is derived by a longitudinal integration of Maxwell's equations which is more accurate and stable than the above described techniques and can be combined with different lateral discretization methods. A detailed description is given in the following sections.

3.3.2 The propagator matrix

Within this section, a linear relation of the field distributions on the entrance plane and the emergence plane is derived which rigorously takes into account an arbitrary superposition of forward and backward propagating waves. Choosing the computation variables appropriately leads to an algorithm which works accurately on coarse meshes, too.

Basic assumptions

The presented derivation of the propagator matrix is based on the following assumptions where the positive z–axis is chosen as the direction of the laser beam propagation:

- Three dimensional effects are negligible, i. e. all quantities are assumed to be independent of the y–coordinate. In particular, derivatives with respect to y vanish.
- The field distribution is TE polarized, i. e.

$$\vec{E}(\vec{x}) = E_y(\vec{x})\,\vec{e}_y \ . \tag{3.4}$$

  As a consequence, the magnetic field writes $\vec{B}(\vec{x}) = B_x(\vec{x})\vec{e}_x + B_z(\vec{x})\vec{e}_z$.
- The dielectric constant $\varepsilon_R(\vec{x},t) = \left[n_{Si}(\vec{x},t) + i\frac{c}{2\omega}\alpha(\vec{x},t)\right]^2$ varies slowly in view of the oscillations of the electromagnetic field. At any point of time t_j, the beam propagation can therefore be calculated for a stationary dielectric constant $\varepsilon_R(\vec{x},t_j)$. Hence, the argument t_j is omitted in the following discussion.
- The magnetic permeability is equal to μ_0.

As no polarization dependent effects have been observed in the measurements, the second constraint does not limit generality. Whereas the two latter assumptions are satisfied naturally, the first one is the major restriction of the model. However, regarding the feature sizes and the aspect ratios of power devices, we do not have to expect significant effects demanding a 3D calculation. Therefore, the 2D model also facilitates meaningful investigations.

Computational variables and their governing equations

Based on these assumptions, Maxwell's equations in the frequency domain with the electric field $\vec{E}$ and the magnetic field $\vec{B}$ as their variables are equivalent to the following set of equations:

$$\frac{\partial}{\partial z}E_y(x,z) = -i\omega B_x(x,z) \tag{3.5}$$

$$\frac{\partial}{\partial z}B_x(x,z) = -i\frac{\omega}{c^2}\varepsilon_R(x,z)E_y(x,z) - \frac{i}{\omega}\frac{\partial^2}{\partial x^2}E_y(x,z) \tag{3.6}$$

The two degrees of freedom are represented by E_y and B_x. The remaining component B_z can be calculated from $B_z(x,z) = -\frac{i}{\omega}\frac{\partial}{\partial x}E_y(x,z)$.

The key idea for enhancing computational economy consists in replacing the variables E_y and B_x by the electric fields of the forward and backward propagating waves and analytically splitting off their rapid oscillations. Formally speaking, we choose a parameter κ and employ the following variable transformation:

$$\boxed{\begin{aligned} E^F(x,z) &:= \tfrac{1}{2}e^{-i\kappa z}\left[E_y(x,z) - \tfrac{\omega}{\kappa}B_x(x,z)\right] \\ E^B(x,z) &:= \tfrac{1}{2}e^{i\kappa z}\left[E_y(x,z) + \tfrac{\omega}{\kappa}B_x(x,z)\right] \end{aligned}} \tag{3.7}$$

The electromagnetic fields are reconstructed from the inverse transformation

$$\begin{aligned} E_y(x,z) &= e^{i\kappa z}E^F(x,z) + e^{-i\kappa z}E^B(x,z) \\ B_x(x,z) &= \tfrac{\kappa}{\omega}\left[-e^{i\kappa z}E^F(x,z) + e^{-i\kappa z}E^B(x,z)\right] \end{aligned} \tag{3.8}$$

From the mathematical point of view, κ is just an arbitrary global parameter for the transformation formulas. If it is chosen as the propagation constant along the z–axis, the variables E^F and E^B represent the envelopes of the waves travelling in positive and negative z–direction, respectively. E^F and E^B vary the more slowly the closer κ matches the propagation constant in z–direction. Though the latter is a priori unknown, the value $n_{Si}\frac{\omega}{c}$ is a good guess for paraxially propagating waves. In practical applications, a grid refinement of several microns has proven sufficient, which is about two orders of magnitude coarser than the required discretization width for calculations in terms of E_y and B_x.

Inserting (3.8) into (3.5) and (3.6) reveals the differential equations that E^F and E^B have to satisfy:

$$\begin{aligned}\frac{\partial}{\partial z}E^F &= \frac{i\omega^2}{2c^2\kappa}\varepsilon_R\,E^F + \frac{i}{2\kappa}\frac{\partial^2}{\partial x^2}E^F - \frac{i\kappa}{2}E^F \\ &\quad + \frac{i\omega^2}{2c^2\kappa}\varepsilon_R\,e^{-2i\kappa z}E^B + \frac{i}{2\kappa}e^{-2i\kappa z}\frac{\partial^2}{\partial x^2}E^B - \frac{i\kappa}{2}e^{-2i\kappa z}E^B \qquad (3.9)\\ \frac{\partial}{\partial z}E^B &= -\frac{i\omega^2}{2c^2\kappa}\varepsilon_R\,E^B - \frac{i}{2\kappa}\frac{\partial^2}{\partial x^2}E^B + \frac{i\kappa}{2}E^B \\ &\quad - \frac{i\omega^2}{2c^2\kappa}\varepsilon_R\,e^{2i\kappa z}E^F - \frac{i}{2\kappa}e^{2i\kappa z}\frac{\partial^2}{\partial x^2}E^F + \frac{i\kappa}{2}e^{2i\kappa z}E^F \qquad (3.10)\end{aligned}$$

The propagator matrix

Since the numerical solution of the equations (3.9) and (3.10) is calculated on a rectangular tensor grid, the lateral discretization (x–axis) and the longitudinal discretization (z–axis) can be treated independently, thus enabling flexible combinations of different methods.

Several approaches can be pursued for the lateral discretization. The most accurate is the transformation to k_x-space (e. g. [112, 114, 117]). In this case, the operator $\frac{\partial^2}{\partial x^2}$ is represented by the factor $-k_x^2$, while the product $\varepsilon_R(x,z)E^{F,B}(x,z)$ becomes a convolution integral. The discretized algebraic system of equations therefore includes fully occupied matrices, which is the major drawback of this approach. In addition, the expansion in terms of periodic functions involves considerable difficulties if Dirichlet boundary conditions to model reflecting surfaces at $x = x_1$ or $x = x_{N_x}$ have to be implemented.

Within this work, finite difference discretization is employed in the lateral direction. The resulting restriction that the largest grid spacing along the x–axis must not exceed 1/10 of the smallest lateral wavelength $\lambda_x = 2\pi/k_x$ is easily satisfied for paraxially propagating waves. Its greatest advantage is that finite difference discretization facilitates a very quick evaluation of the discretized system of equations since the operator $\frac{\partial^2}{\partial x^2}$ is represented by a band structured matrix with only the first super– and subdiagonals occupied.

Special attention has to be paid to the longitudinal discretization to accurately allow for the rapidly oscillating terms $e^{\pm 2i\kappa z}$ in the equations (3.9) and (3.10). For that purpose, these equations are integrated analytically over the interval $[z_n, z_{n+1}]$, assuming a linear interpolation of the computation variables $E^{F,B}$ and the dielectric constant ε_R. For a compact notation, the unknown fields at the plane $z = z_n$ are summarized in a vector

$$\vec{u}(z_n) = \left[E^F(x_1,z_n), E^B(x_1,z_n), \ldots, E^F(x_{N_x},z_n), E^B(x_{N_x},z_n)\right]^T \quad . \qquad (3.11)$$

Then, the discretized equations (3.9) and (3.10) can be written as

$$M_2\vec{u}(z_{n+1}) = M_1\vec{u}(z_n) \qquad (3.12)$$

In case of finite difference discretization in lateral direction, the matrices M_1 and M_2 are band structured matrices with complex 2×2 matrices as their coefficients. A more detailed description is given in the appendix (A). Thus, we can define a matrix P_n which describes the relation between the field distributions at two adjacent planes $z = z_n$ and $z = z_{n+1}$.

$$\vec{u}(z_{n+1}) = P_n\vec{u}(z_n) \quad \text{with} \quad P_n = M_2^{-1}M_1 \qquad (3.13)$$

Multiplying all these matrices, we finally obtain a propagator matrix P which relates the field distribution at the front surface $z = z_1$ to the field distribution at the rear surface $z = z_{N_z}$:

$$\vec{u}(z_{N_z}) = P\vec{u}(z_1) \quad \text{with} \quad P = \prod_{n=1}^{N_z-1} P_n \tag{3.14}$$

It should be pointed out that P accounts for wave propagation in *both* directions in a physically rigorous manner since it represents the discretized Maxwell equations (3.9) and (3.10). The derivation merely assumes that the fields $E^{F,B}$ and the dielectric constant can be linearly interpolated in every interval $[z_n, z_{n+1}]$.

3.3.3 Boundary conditions in propagation direction

The first order partial differential equations (3.9) and (3.10) demand the specification of the field distributions E^F and E^B at one of the device surfaces. Actually, the field distributions of the incident waves are known. The initial value problem is therefore transformed to a boundary value problem with one half of the boundary conditions at each of the surfaces $z = z_1$ and $z = z_{N_z}$, respectively. Deriving these so called radiation boundary conditions [118] therefore requires a decomposition of the computation variables in terms of forward and backward propagating waves, which are related to the incident waves by the refraction law at the interfaces.

Boundary condition in the k_x–z–space

Throughout this section, the representation of any quantity in k_x–z–space will be denoted by a ˜ above the corresponding symbol.

At any given k_x, the amplitudes of waves propagating in positive and negative z–direction (cf. fig. 3.2) are related by the refraction law [119] at the interface $z = 0$. The simplest representation is gained in k_x–z–space:

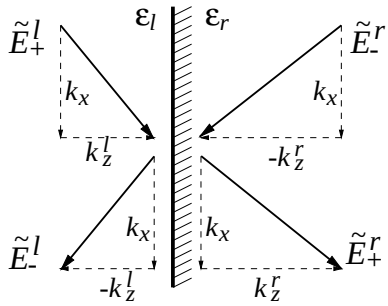


Figure 3.2: Refraction at an interface.

$$\tilde{E}^l_+(k_x, z=0) = \frac{k^l_z + k^r_z}{2k^l_z}\,\tilde{E}^r_+(k_x, z=0) + \frac{k^l_z - k^r_z}{2k^l_z}\,\tilde{E}^r_-(k_x, z=0) \tag{3.15}$$

$$\tilde{E}^l_-(k_x, z=0) = \frac{k^l_z - k^r_z}{2k^l_z}\,\tilde{E}^r_+(k_x, z=0) + \frac{k^l_z + k^r_z}{2k^l_z}\,\tilde{E}^r_-(k_x, z=0) \tag{3.16}$$

where $k^{l,r}_z = \sqrt{\varepsilon_{l,r}k_0^2 - k_x^2}$. The travelling waves are expressed in terms of the electric and magnetic fields

$$\tilde{E}^{l,r}_\pm(k_x, z) = \frac{1}{2}\left(\tilde{E}^{l,r}_y(k_x, z) \mp \frac{\omega}{k^{l,r}_z}\tilde{B}^{l,r}_x(k_x, z)\right) \tag{3.17}$$

and finally in terms of the computation variables $\tilde{E}^F$ and $\tilde{E}^B$:

$$\tilde{E}^{l,r}_\pm(k_x, z) = \frac{1}{2}\left(1 \pm \frac{\kappa}{k^{l,r}_z}\right)e^{i\kappa z}\tilde{E}^F(k_x, z) + \frac{1}{2}\left(1 \mp \frac{\kappa}{k^{l,r}_z}\right)e^{-i\kappa z}\tilde{E}^B(k_x, z) \tag{3.18}$$

Note that this equation also identifies $\tilde{E}^{F,B}$ as the envelopes of the forward and backward propagating waves if the parameter κ matches the propagation constant k_z^l or k_z^r, respectively.

At the entrance plane $z = z_1$, the field $\tilde{E}_+^l$ is equal to the incident laser field $\tilde{E}_i$, while $\tilde{E}_-^l$ represents the unknown reflected wave. Besides, $\tilde{E}_\pm^r$ refer to the fields in the interior of the sample. Inserting (3.18) into (3.15) and (3.16) reveals the boundary condition at the left hand boundary of the simulation domain

$$\boxed{\tilde{E}_i(k_x) = \frac{1}{2}\left(1 + \frac{\kappa}{k_z^a}\right) e^{i\kappa z_1} \tilde{E}^F(k_x, z_1) + \frac{1}{2}\left(1 - \frac{\kappa}{k_z^a}\right) e^{-i\kappa z_1} \tilde{E}^B(k_x, z_1)} \tag{3.19}$$

and the conditional equation of the reflected wave

$$\tilde{E}_r(k_x) = \frac{1}{2}\left(1 - \frac{\kappa}{k_z^a}\right) e^{i\kappa z_1} \tilde{E}^F(k_x, z_1) + \frac{1}{2}\left(1 + \frac{\kappa}{k_z^a}\right) e^{-i\kappa z_1} \tilde{E}^B(k_x, z_1) \ . \tag{3.20}$$

In these relations, ε_a denotes the dielectric constant in front of the entrance plane (i. e., for $z < z_1$) and k_z^a is given by $k_z^a = \sqrt{\varepsilon_a k_0^2 - k_x^2}$.

If no wave impinges from the back, a similar condition applies to the right hand boundary of the simulation domain at $z = z_{N_z}$

$$0 = \frac{1}{2}\left(1 - \frac{\kappa}{k_z^b}\right) e^{i\kappa z_{N_z}} \tilde{E}^F(k_x, z_{N_z}) + \frac{1}{2}\left(1 + \frac{\kappa}{k_z^b}\right) e^{-i\kappa z_{N_z}} \tilde{E}^B(k_x, z_{N_z}) \tag{3.21}$$

where $k_z^b = \sqrt{\varepsilon_b k_0^2 - k_x^2}$ and ε_b is the dielectric constant behind the rear surface (i. e., for $z > z_{N_z}$). The transmitted wave is calculated from

$$\tilde{E}_t(k_x) = \frac{1}{2}\left(1 + \frac{\kappa}{k_z^b}\right) e^{i\kappa z_{N_z}} \tilde{E}^F(k_x, z_{N_z}) + \frac{1}{2}\left(1 - \frac{\kappa}{k_z^b}\right) e^{-i\kappa z_{N_z}} \tilde{E}^B(k_x, z_{N_z}) \ . \tag{3.22}$$

For the sake of a compact notation, the components $\tilde{E}_i(k_x)$, $\tilde{E}_r(k_x)$, and $\tilde{E}_t(k_x)$ are summarized in the vectors $\tilde{\vec{E}}_i$, $\tilde{\vec{E}}_r$, and $\tilde{\vec{E}}_t$, respectively. Defining $N_x \times 2N_x$ matrices $\tilde{B}_\pm^{a,b}$ by

$$\begin{aligned} \left[\tilde{B}_\pm^{a,b}\right]_{i,2i} &:= \frac{1}{2}\left(1 \pm \frac{\kappa}{k_z^{a,b}}\right); \quad \left[\tilde{B}_\pm^{a,b}\right]_{i,2i+1} := \frac{1}{2}\left(1 \mp \frac{\kappa}{k_z^{a,b}}\right) \quad \text{for } i = 1, \ldots, N_x \\ \left[\tilde{B}_\pm^{a,b}\right]_{i,j} &:= 0 \quad \text{otherwise} \end{aligned} \tag{3.23}$$

makes the boundary conditions read

$$\boxed{\tilde{\vec{E}}_i = \tilde{B}_+^a \, \tilde{\vec{u}}(z_1) \quad \text{and} \quad 0 = \tilde{B}_-^b \, \tilde{\vec{u}}(z_{N_z})} \ . \tag{3.24}$$

Each of these matrix equations (3.24) comprises N_x scalar equations, thus representing the required $2N_x$ boundary conditions for the $2N_x$ unknown field components in $\tilde{\vec{u}}$ (cf. eq. 3.11). Similarly, the conditional equations of the emerging waves become

$$\tilde{\vec{E}}_r = \tilde{B}_-^a \, \tilde{\vec{u}}(z_1) \quad \text{and} \quad \tilde{\vec{E}}_t = \tilde{B}_+^b \, \tilde{\vec{u}}(z_{N_z}) \ . \tag{3.25}$$

Matrix representation of the boundary conditions in real space

In order to derive a matrix representation $B_{\pm}^{a,b}$ of the boundary conditions in real space two different strategies can be followed:

First, the refraction law and the relations in the previous section are regarded as operator equations. Since k_x writes as $-i\frac{\partial}{\partial x}$ in real space, the operator $k_z^{a,b}$ is represented by $\sqrt{\varepsilon_{a,b}k_0^2 + \frac{\partial^2}{\partial x^2}}$. The simplest evaluation of this square root is to neglect k_x with respect to $\sqrt{\varepsilon_{a,b}}k_0$, thus approximating $k_z^{a,b}$ by $\sqrt{\varepsilon_{a,b}}k_0$ to define the matrices $B_{\pm}^{a,b}$ (cf. eq. 3.23). This assumption is only valid for small angles of incidence.

As this approach can be considered to be the constant term of a Taylor expansion of $1/k_z$ in terms of k_x, we will obtain a more accurate expression if the linear term is also included [120, 118]:

$$\frac{1}{k_z^{a,b}} = \frac{1}{\sqrt{\varepsilon_{a,b}}k_0}\left[1 - \frac{1}{2\varepsilon_{a,b}k_0^2}\frac{\partial^2}{\partial x^2} + O\left(\frac{\partial^4}{\partial x^4}\right)\right] \tag{3.26}$$

Inserting this approximation into (3.18) yields relations which can be easily discretized in real space. The resulting matrices $B_{\pm}^{a,b}$ exhibit a band structure and constitute a sufficiently accurate approximation for angles of incidence up to 15°.

The second strategy for gaining the matrix representation $B_{\pm}^{a,b}$ exploits the linearity of the Fourier transformation which can be written as a matrix product:

$$f(x_n) = \sum_m f(k_{x,m})e^{ik_{x,m}x_n} = \sum_m T_{nm}f(k_{x,m}) \quad \text{with} \quad T_{nm} := e^{ik_{x,m}x_n} \tag{3.27}$$

Multiplying (3.24) by T, we obtain the boundary conditions

$$\vec{E}_i = B_+^a\,\vec{u}(z_1) \quad \text{and} \quad 0 = B_-^b\,\vec{u}(z_{N_z}) \tag{3.28}$$

in which the desired matrix representation $B_{\pm}^{a,b}$ in real space is calculated from

$$B_{\pm}^{a,b} = T\,\tilde{B}_{\pm}^{a,b}\,T^{-1}\ . \tag{3.29}$$

The accuracy of these approaches is compared in appendix (B).

3.3.4 Summary of the algorithm

In summary, beam propagation through the device under test is calculated by the following steps:

First, the propagator matrix P which relates the unknowns $\vec{u}(z_{N_z})$ at the rear surface to the unknowns $\vec{u}(z_1)$ at the entrance plane is calculated from the equations (3.13) and (3.14). This consumes about 75 % of the total CPU–time. Second, the boundary conditions (3.28) are expressed in terms of the following system of equations

$$\begin{pmatrix} B_+^a \\ B_-^b P \end{pmatrix}\vec{u}(z_1) = \begin{pmatrix} \vec{E}_i \\ 0 \end{pmatrix} \tag{3.30}$$

which is solved for the unknown fields at the entrance plane. Finally, the reflected and transmitted waves are extracted:

$$\begin{pmatrix} \vec{E}_r \\ \vec{E}_t \end{pmatrix} = \begin{pmatrix} B_-^a \\ B_+^b P \end{pmatrix}\vec{u}(z_1) \tag{3.31}$$

In order to obtain the field distribution in the interior of the sample, the field distributions $\vec{u}(z_n)$ ($1 < n < N_z$) have to be calculated from $\vec{u}(z_1)$ by iteratively using equation (3.13).

This algorithm incorporates an important advantage: Since only the right hand side of equation (3.30) depends on the incident beam, the field distribution for several different incident waves can be calculated simultaneously with almost no additional effort:

$$\begin{pmatrix} B_+^a \\ B_-^b P \end{pmatrix} \left(\vec{u}_1(z_1),\, \vec{u}_2(z_1),\, \vec{u}_3(z_1),\, \ldots \right) = \begin{pmatrix} \vec{E}_{i,1} & \vec{E}_{i,2} & \vec{E}_{i,3} & \cdots \\ 0 & 0 & 0 & \end{pmatrix} \tag{3.32}$$

3.4 Fourier optics

The reflected or transmitted waves are projected onto the detector by focussing lenses and aperture holes. An appropriate model is derived from Fourier optics [121]. Regarding a sequence of planes perpendicular to the optical axis (cf. fig. 3.3), the field distribution $\vec{E}(x, y, z_n)$ on the following plane is calculated from the field distribution $\vec{E}(x, y, z_{n-1})$ on the preceding plane. Contrary to the algorithm derived in the previous section, this model only allows for waves propagating in the positive z–direction.

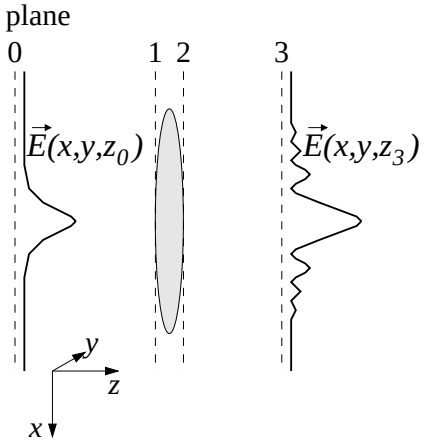


Figure 3.3: Modeling the image formation by optical elements.

3.4.1 Image formation by aperture holes

In the limit $z_{n-1} = z_n$, the imaging by an aperture hole of radius R with the center located at the optical axis $(0, 0)$ is described by an amplitude modulation in real space:

$$\vec{E}(x, y, z_n) = \Theta\left(R^2 - x^2 - y^2\right) \vec{E}(x, y, z_{n-1}) \tag{3.33}$$

where $\Theta(x)$ is the unit step function, which is equal to 1 for $x > 0$ and equal to 0 for $x < 0$.

3.4.2 Image formation by thin lenses

A lens is said to be a thin lens if the translation of the ray within the lens can be neglected. Thus, only the phase shift $\varphi(x, y)$ of the wavefront has to be taken into account. It is proportional to the optical thickness at the position (x, y):

$$\varphi(x, y) = k_0 n_L \Delta(x, y) + k_0[\Delta_0 - \Delta(x, y)] \tag{3.34}$$

In this relation, $k_0 n_L \Delta(x, y)$ represents the phase delay within the lens of refractive index n_L while $k_0[\Delta_0 - \Delta(x, y)]$ is the phase delay along the remaining distance in the surrounding air (cf. fig. 3.4). Deriving $\varphi(x, y)$ by geometrical considerations, the phase delay on the paraxial approximation can be finally expressed as [121]

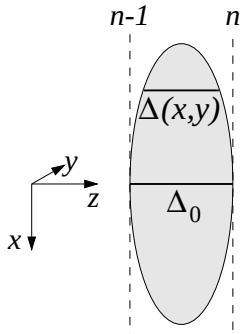


Figure 3.4: Phase delay within a thin lense.

$$\varphi(x,y) = k_0 n_L \Delta_0 + \frac{k_0}{2f}(x^2+y^2) \tag{3.35}$$

where $f = \left((n_L-1)\left(\frac{1}{R_1}-\frac{1}{R_2}\right)\right)^{-1}$ denotes the focal length of the lens. The image formation by a thin lens therefore introduces the following phase modulation in real space:

$$\vec{E}(x,y,z_n) = \exp(ik_0 n_L \Delta_0)\, \exp\left[i\frac{k_0}{2f}(x^2+y^2)\right]\, \vec{E}(x,y,z_{n-1}) \tag{3.36}$$

3.4.3 Propagation in free space

Since only waves with a positive propagation constant $k_z = +\sqrt{k_0^2 - k_x^2 - k_y^2}$ along the optical axis are taken into account, the field distribution between the planes $z = z_{n-1}$ and $z = z_n$ can be expressed in terms of the Fourier integral

$$\vec{E}(x,y,z) = \int \tilde{\vec{E}}(k_x,k_y)\, e^{ik_x x + ik_y y + ik_z(k_x,k_y)z}\, dk_x\, dk_y \ . \tag{3.37}$$

At $z = z_{n-1}$ and $z = z_n$, this relation becomes

$$\vec{E}(x,y,z=z_{n-1}) = \int \tilde{\vec{E}}(k_x,k_y)\, e^{ik_z z_{n-1}}\, e^{ik_x x + ik_y y}\, dk_x\, dk_y \tag{3.38}$$

$$\vec{E}(x,y,z=z_n) = \int \tilde{\vec{E}}(k_x,k_y)\, e^{ik_z z_n}\, e^{ik_x x + ik_y y}\, dk_x\, dk_y \tag{3.39}$$

These equations can be considered the two–dimensional Fourier transforms of the field distributions $\vec{E}(x,y,z_{n-1})$ and $\vec{E}(x,y,z_n)$, respectively. Therefore, propagation in free space is modeled by a phase change in k_x–k_y–space:

$$\tilde{\vec{E}}(k_x,k_y,z_n) = e^{i\sqrt{k_0^2-k_x^2-k_y^2}\,(z_n - z_{n-1})}\, \tilde{\vec{E}}(k_x,k_y,z_{n-1}) \tag{3.40}$$

3.5 Detector response

The illumination of the detector generates photo currents which are electronically amplified and converted. For the sake of simplicity, we assume the switching times of the device under test to be large enough so that the photo diode capacitance and any delay in the signal processing can be neglected. Since except for the laser deflection technique the detector comprises a simple pin diode, the measurement signal is proportional to the integrated intensity distribution, with the integration boundaries matching the active area of the detector.

To perform laser deflection experiments, the detector contains a four quadrant pin diode which is dynamically translated to keep the DC component of the deflection signal vanishing (cf. section 2.2.2). Due to the very small duty cycle of the periodic current pulses, this DC component is equal to the deflection signal during the off–state of the device under test.

Representing the boundaries of the detector segments by x_1 to x_4 (cf. fig. 3.5), the photo currents I_1 and I_2 of two opposite segments are proportional to the integrated intensities

$$I_1(t) = const \int_{x_1}^{x_2} \|\vec{E}(x,t)\|^2 dx \,, \quad (3.41)$$

$$I_2(t) = const \int_{x_3}^{x_4} \|\vec{E}(x,t)\|^2 dx \,. \quad (3.42)$$

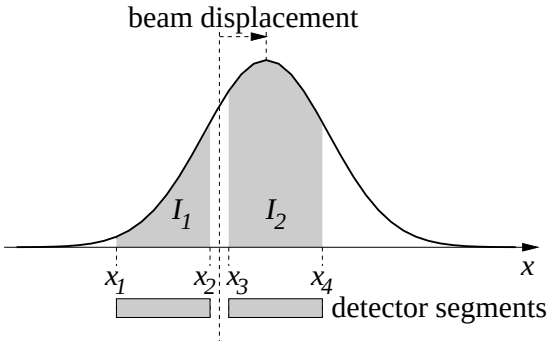


Figure 3.5: Intensity distribution on the detector.

While the widths $x_2 - x_1$ and $x_4 - x_3$ and the gap $x_3 - x_2$ between the segments are fixed by the detector geometry, the absolute position x_1 is calculated to satisfy the condition $I_1^{off} = I_2^{off}$ during the off–state. Then, the deflection signal M is obtained by

$$M(t) = \frac{I_1(t) - I_2(t)}{I_1(t) + I_2(t)} \quad (3.43)$$

while the simultaneously detected absorption signal is calculated from

$$A(t) = \frac{I_1(t) + I_2(t)}{I_1^{off} + I_2^{off}} - 1 \quad (3.44)$$

which actually represents the AC component of the detected time–dependent intensity distribution scaled to its DC component (cf. section 2.2.1). It should be noted that these ratios merely depend on the detector geometry whereas they are not affected by its (probably unknown) sensitivity and the amplification factors.

Chapter 4

Analysis of Free Carrier Absorption and Internal Laser Deflection Measurements

In the previous chapter, a physically rigorous simulation strategy has been introduced which facilitates a thorough numerical study of optical laser probing techniques. Supplemented by analytical estimations, it is now employed to carry out a comprehensive analysis of the Free Carrier Absorption and the Internal Laser Deflection method. The major goals of these investigations are to enhance our understanding of the measurement process itself, to support the design and the optimization of the experiments with respect to a minimum experimental error, a large measurement range, and a minimum sensitivity to parasitic effects, and to provide quantitative results for the optimum setup and the optimum sample geometry.

4.1 Virtual experiments and the optimization strategy

Although the engineers' interest is focussed on the carrier and temperature distribution in the interior of product devices, internal laser probing can only be performed on specific samples which are prepared for the measurements (cf. chapter 2 and table 2.3). They differ from the corresponding product devices, for example, with respect to their size, an enhanced surface recombination velocity due to sawing and polishing, or a window in the metallization layer which is required for Backside Laser Probing. For that reason, the strategy illustrated in fig. (4.1) is applied to optimize the experiments: The desired carrier concentration n_{ref} and temperature T_{ref} within the product device are gained from electrothermal device simulation. This data is kept as a fixed reference during the optimization study. On the other hand, the simulation of the measurement process which takes into account the specific properties of the prepared samples yields a calculated detector signal. From this signal, a carrier concentration n_{extr} and a temperature T_{extr} are extracted according to the evaluation rule which is also applied for the real measurements (for example, the absorption law 2.16 or equation 2.19). While these "virtual experiments" are performed with different setups and sample geometries, the goal of an optimized experiment is achieved if the extracted profiles n_{extr} and T_{extr} match the reference quantities n_{ref} and T_{ref} as closely as possible. In addition, the remaining deviation is a measure of the experimental error.

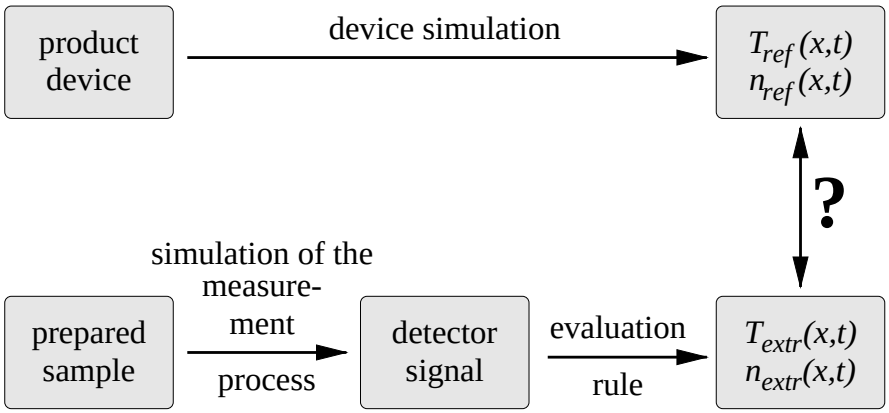


Figure 4.1: Basic strategy for optimizing the optical experiment.

4.2 Free carrier absorption measurements

As a first example, we investigate free carrier absorption measurements which are employed to determine the charge carrier distribution in the interior of a power device during its forward conducting state. Many of the findings, in particular those results concerning wave propagation effects or the lateral extension of the probing beam, can be transferred to other techniques with a similar setup, for example the laser deflection method.

4.2.1 Optical field distribution

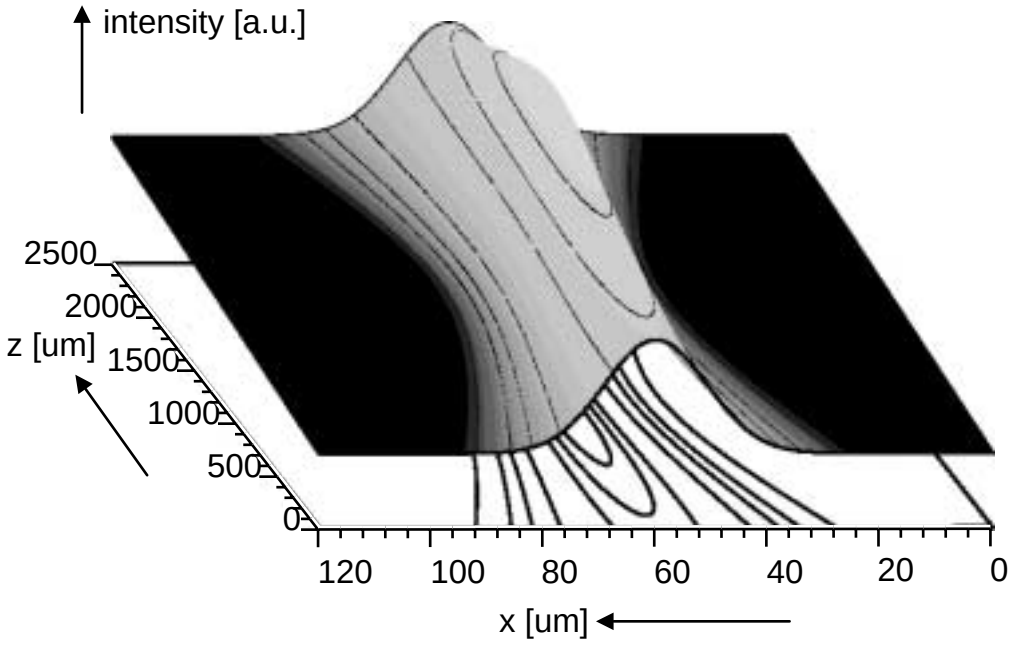


Figure 4.2: Optical field distribution of a Gaussian beam ($\eta_A = 0.1$) incident at a depth $x = 60\ \mu$m.

The field distribution in the interior of the device is shown in fig. (4.2). For the sake of clarity, the number of grid nodes in the figure is reduced so that standing waves due to multiple reflections at the surfaces are not visible.

Assuming an incident Gaussian wave with an angular aperture of $\eta_A = 0.1$, the beam profile in the interior of the sample is also nearly Gaussian. The focal plane is located in the center of the sample where the minimum spot diameter is about $15\,\mu\mathrm{m}$. The spreading of the laser beam increases the spot diameter to about $25\,\mu\mathrm{m}$ at the surfaces of a sample which is 2.5 mm in length.

4.2.2 Propagation in the boundary regions

If the probing beam impinges close to the top or bottom surfaces of the device under test, undesirable distortions of the beam profile are observed (cf. fig. 4.3). Reflections at the metallization layer give rise to an interference pattern with several additional intensity peaks [122].

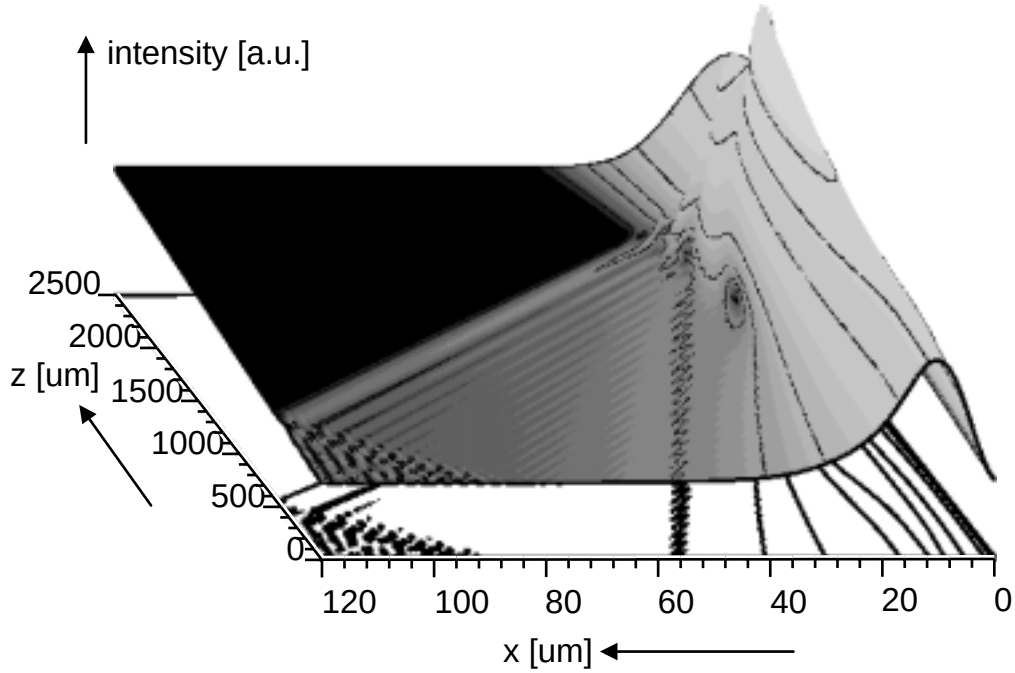


Figure 4.3: Optical field distribution of a Gaussian beam ($\eta_A = 0.1$) incident at a depth $x = 10\,\mu\mathrm{m}$.

These simulation results are confirmed by the infrared images which are taken by the camera included in the optical setup. While the beam profile in the bulk region is approximately Gaussian (right hand image of fig. 4.4), a diffraction and interference pattern is observed for probing beams propagating close to the metallizations layers (left hand image of fig. 4.4).

4.2.3 Longitudinal averaging/samples with cell structure

As mentioned above, the extraction of the carrier concentration from the measured decrease of the transmitted intensity according to the absorption law (2.16) assumes a constant carrier distribution along the path of the probing beam. However, this assumption does not hold for samples

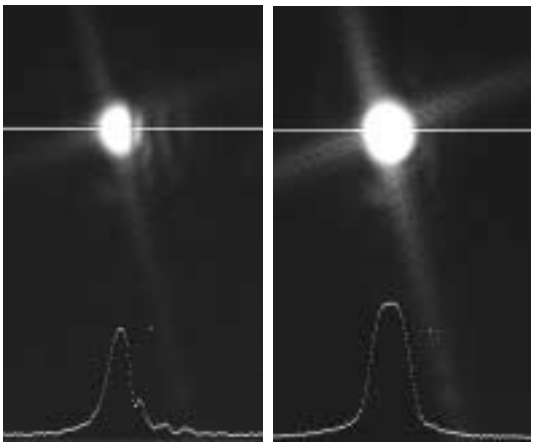


Figure 4.4: Experimental field distribution at the rear surface for a beam position near the anode boundary (left) and in the center (right) of the sample. The images are taken by an IR camera.

comprising a cell structure, as e. g. IGBTs (cf. fig. 8.1). To investigate the resulting effect on the measurement signal, the periodic carrier concentration profile is expanded into a Fourier series

$$n(x,z) = n_0(x) + \sum_{k=1}^{\infty} \Delta n_k(x) \cos(k \cdot 2\pi z/D) \tag{4.1}$$

where D denotes the cell width. Thus, within a sample of length L, the integrated carrier density becomes

$$\frac{1}{L}\int_0^L n(x,z)\,dz = n_0(x) + \sum_{k=1}^{\infty} \Delta n_k(x) \frac{D}{2\pi k L} \sin(k \cdot 2\pi L/D) = n_0(x)\ . \tag{4.2}$$

Since the sample length is an integer multiple of the cell width, the absorption signal is thus expected to reflect the average concentration $n_0(x)$.

However, this simple consideration neglects the deflection of the probing beam due to refractive index gradients and possible reflections at the metallization layers. To carry out a consistent analysis, the carrier concentration profiles are extracted by simulating the measurement process on samples of different lengths and compared to the carrier distribution obtained by an electrothermal device simulation. As it can be seen from fig. (4.5) an excellent accuracy is achieved for depths

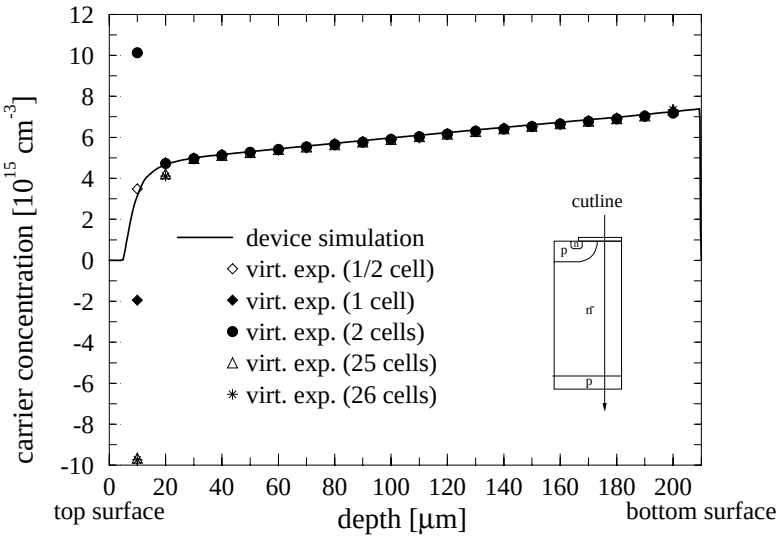


Figure 4.5: Vertical carrier distribution in an IGBT sample. The dots refer to the carrier concentrations which are obtained by virtual experiments on samples of different lengths (i. e. different number of cells).

greater than $20\,\mu\mathrm{m}$. On the other hand, significant deviations are observed if the probing beam impinges close to the top surface. The space charge regions around the reversed biased p–wells form areas with a higher refractive index, thus acting as focussing lenses (cf. fig. 4.7). The probing beam is therefore deflected towards the upper boundary of the device where it traverses regions of lower carrier concentration. As a consequence, the transmitted light intensity may even exceed the transmission during the off-state (cf. fig. 4.6, depth $10\,\mu\mathrm{m}$) which would be interpreted in terms of a negative injected carrier concentration. Keeping in mind the distortion of the beam profile due to reflections at the metallization layers (cf. section 4.2.2), the observed variation of the extracted concentration values close to the top surface (cf. fig. 4.5) is not very surprising.

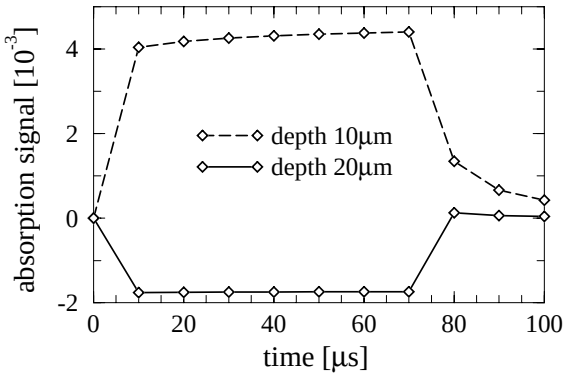


Figure 4.6: Calculated absorption signal for a sample comprising 25 IGBT cells (pulse duration $70\,\mu\mathrm{s}$*).*

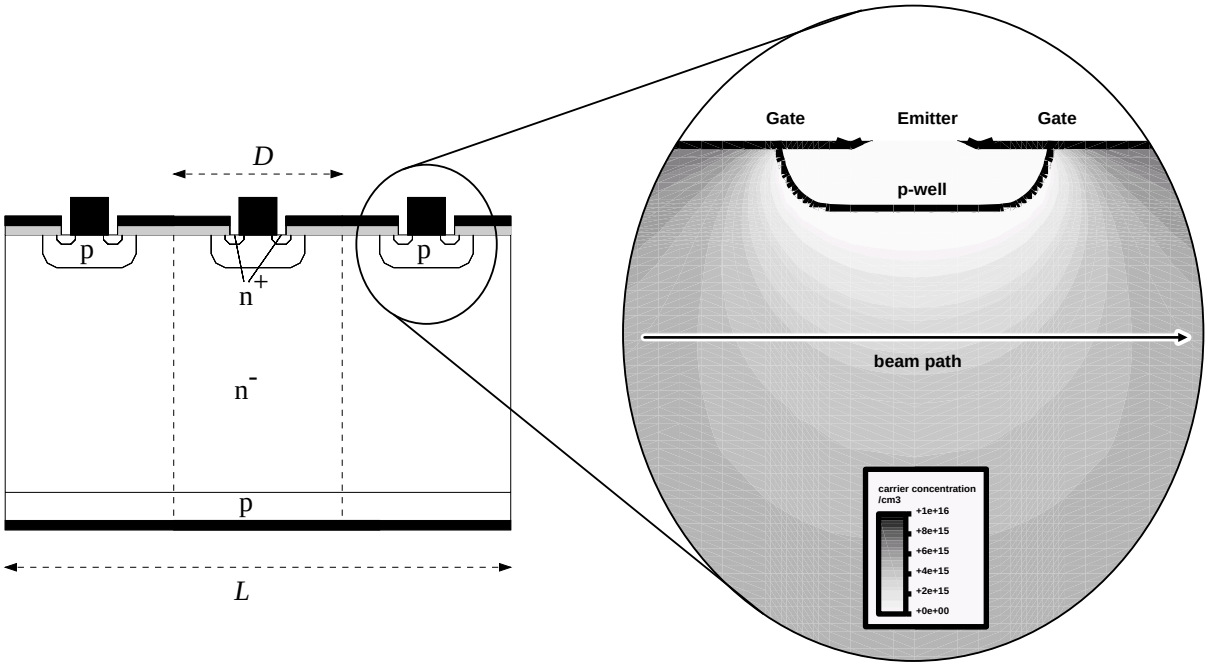


Figure 4.7: Carrier distribution at a current density of $145\,\mathrm{A/cm^2}$ *in an IGBT sample near the top p–well.*

In summary, we conclude that the absorption signal reflects the average concentration of the periodic carrier distribution in the interior of samples with a cell structure. However, if the probing beam impinges close to the top surface, deflection and reflection lead to a corruption of the measured concentration which sensitively depends on the sample size. At an interaction length of about 1 mm, for example, we therefore have to keep a minimum distance of about $20\,\mu\mathrm{m}$ to the top surface to get reliable results.

4.2.4 The Fabry–Perot effect

Multiple reflections at the polished surfaces of the device give rise to the Fabry–Perot effect which modulates the transmitted intensity as the optical thickness $n_{Si} \cdot L$ of the sample changes [119]. To increase the latter by one wavelength within a device of 2.4 mm in size, a temperature rise of 1.7 K is sufficient which may easily originate from self–heating during the typical operating conditions of power devices. For example, a pin diode subjected to a current pulse of $150\,\mathrm{A/cm^2}$ exhibits an average power dissipation of $110\,\mathrm{W/cm^3}$ which results in a temperature increase of approximately $0.01\,K/\mu s$. Therefore, oscillations with a period length of $160\ \mu\mathrm{s}$ are observed on the measurement signal (cf. fig. 4.8, left).

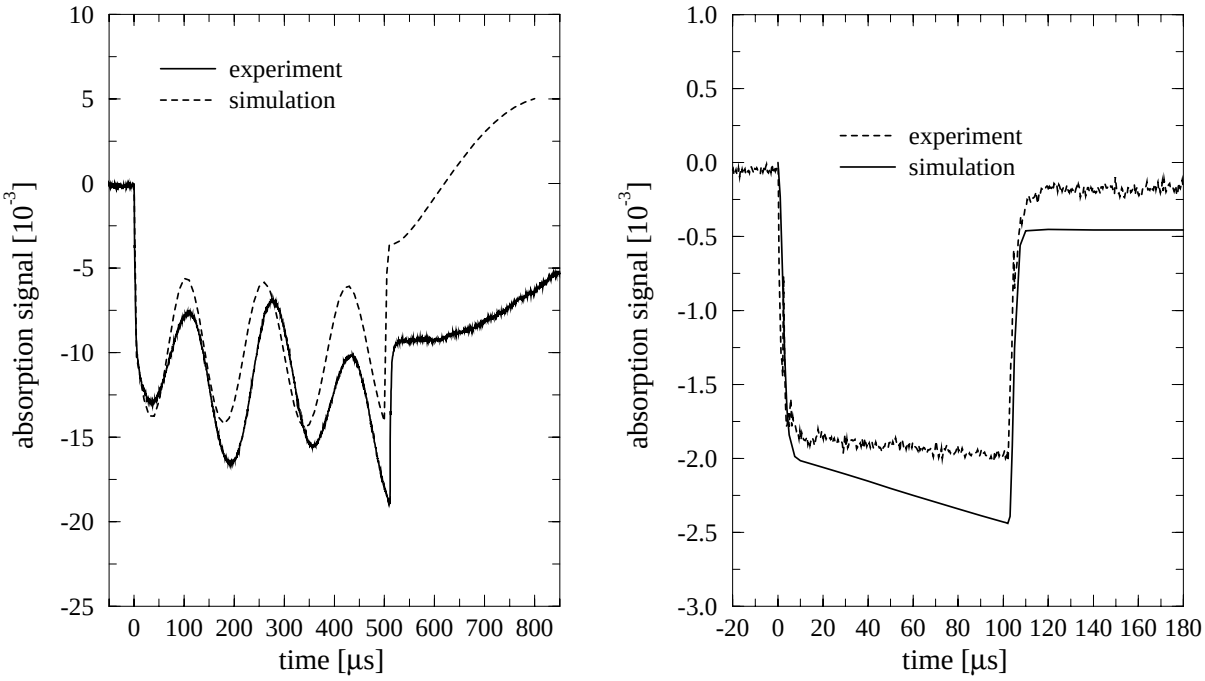


Figure 4.8: Absorption signal measured on a 1700 V pin–diode. During operation with high power dissipation (current density $150\,\mathrm{A/cm^2}$, pulse duration $500\ \mu\mathrm{s}$), Fabry–Perot oscillations are observed (left), whereas the absorption signal exhibits a linear drift in case of a low power dissipation (right hand figure; current density $15\,\mathrm{A/cm^2}$, pulse duration $100\ \mu\mathrm{s}$).

However, with a low power dissipation during a short current pulse, the temperature rise may increase the optical sample thickness only by a small fraction of a wavelength. In this case, a drift on the signal is observed (cf. fig. 4.8, right) whose slope is governed by the linear Taylor expansion of the Fabry–Perot transmittivity at the initial optical thickness. Hence, the drift on the absorption signal may be either positive of negative in sign or may even vanish, if the optical thickness is by chance a half integer multiple of the wavelength.

To suppress these undesired effects, various precautions can be taken, e. g. depositing an antireflective coating, polishing the surfaces at a slightly tilted angle of 1° to prevent interference of the multiply reflected rays, employing an incoherent laser source [74], or aligning the sample appropriately to exploit the vanishing reflection coefficient at the Brewster angle [43].

4.2.5 Spatial resolution

The spatial resolution of free carrier absorption measurements is limited by the lateral extension of the probing beam. To estimate this effect, the carrier concentration is expanded into a Taylor series at the beam position x_0 [78]:

$$n(x) = n(x_0 + \Delta x) = n(x_0) + n'(x_0)\Delta x + \frac{1}{2}n''(x_0)\Delta x^2 + \dots \tag{4.3}$$

Then, the convolution with a Gaussian beam profile $I(x,y,z) = \frac{2}{w^2(z)\pi}\exp(-\frac{(x-x_0)^2+(y-y_0)^2}{w^2(z)/2})$ centered around (x_0, y_0) with a spot diameter $2w(z)$ becomes:

$$\iint_{-\infty}^{\infty} n(x)I(x,y,z)\,dx\,dy = n(x_0) + \frac{w^2(z)}{8}n''(x_0) + \dots = n(x_0)\left(1 + \frac{w^2(z)}{8L_D^2}\right) + \dots \tag{4.4}$$

The totally absorbed power of a Gaussian beam with the z–dependent spot radius

$$w(z) = w_0\sqrt{1 + \left(\frac{2z}{n_{Si}k_0 w_0^2}\right)^2} \tag{4.5}$$

is calculated by integrating the convolution integral (4.4) along the beam path. Hence, the carrier concentration extracted from the absorption signal is expected to be

$$\begin{aligned} n_{extr}(x_0) &= n(x_0)\left(1 + \frac{1}{8L_D^2}\frac{1}{L}\int_{-L/2}^{L/2} w^2(z)\,dz\right) = n(x_0)\left(1 + \frac{w_0^2}{8L_D^2} + \frac{L^2}{24L_D^2 n_{Si}^2 k_0^2 w_0^2}\right) \\ &= n(x_0)\left(1 + \frac{1}{2L_D^2 k_0^2 \eta_A^2} + \frac{L^2\eta_A^2}{96L_D^2 n_{Si}^2}\right) \quad . \end{aligned} \tag{4.6}$$

Though this estimation neglects the deflection of the probing beam and a possible deviation from the Gaussian profile, it reveals some important results: First, the carrier density extracted from the absorption signal appears always larger than the real concentration. Second, the relative error amounts to $\frac{1}{2L_D^2 k_0^2 \eta_A^2} + \frac{L^2\eta_A^2}{96L_D^2 n_{Si}^2}$ which is only a few percent in the case of typical sample lengths L and diffusion lengths L_D. Third, the "apparent" diffusion length $\sqrt{n_{extr}/n''_{extr}}$, which is extracted from a measured carrier concentration profile n_{extr}, matches the "real" diffusion length $L_D = \sqrt{n/n''}$ as long as the factor of equation (4.6) is constant, i. e. the diffusion length L_D is constant within the volume covered by the probing beam.

Defining the spatial resolution Δ by the condition that the measured concentration $n_{extr}(x_0)$ is equal to the carrier density $n(x_0 + \Delta)$ at the position $x_0 + \Delta$, we obtain from the equations (4.6) and (4.3) at the minimum of the profile ($n'(x_0) = 0$):

$$\Delta = \sqrt{\frac{w_0^2}{4} + \frac{L^2}{12n_{Si}^2 k_0^2 w_0^2}} = \sqrt{\frac{1}{k_0^2\eta_A^2} + \frac{L^2\eta_A^2}{48n_{Si}^2}} \tag{4.7}$$

Inserting a typical sample length of 2 mm and an angular aperture of 0.04 yields a spatial resolution $\Delta = 7\ \mu\mathrm{m}$ which is comparable to that of the recombination radiation measurements (cf. section 1.3).

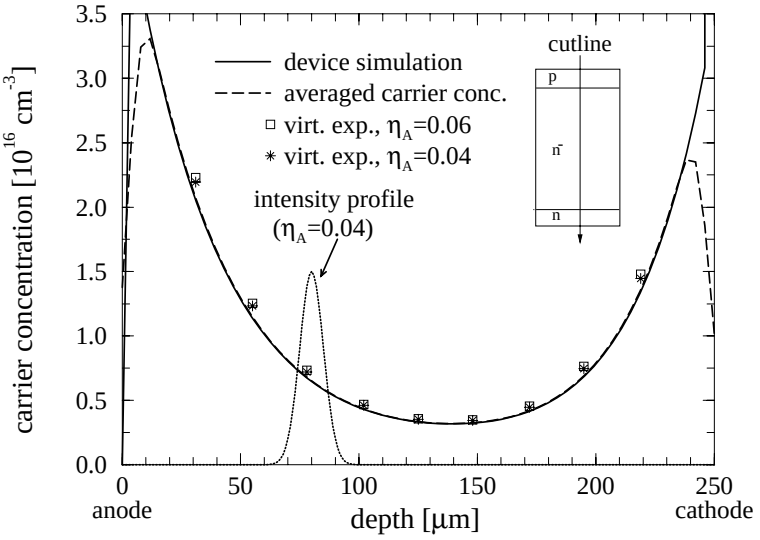


Figure 4.9: Vertical carrier distribution in a pin diode with an ambipolar carrier lifetime of $0.8\,\mu$s and a sample length of 2 mm. The averaged carrier density is calculated by the convolution of the beam profile and the concentration profile obtained by the device simulation.

For further illustration, the convolution of the beam profile and the carrier distribution obtained by a device simulation is plotted in fig. (4.9). Additionally, the carrier concentrations extracted from virtual experiments with different angular apertures of the probing beam are included. We can clearly see from the figure that the lateral extension of the probing beam does not introduce a significant error.

4.2.6 Optimizing the optical setup

The discussion in the previous section is based on two important parameters of the experiment, namely the sample length L and the angular aperture η_A of the probing beam. Within this section, the angular aperture shall be optimized at a given sample length. As the major goals, we strive for a minimum experimental error, i. e. a small deviation of the extracted concentration from the real carrier concentration, and a large measurement range, i. e. a small minimum distance to the top and bottom device boundaries which is required to prevent disturbances of the measurement signal due to diffraction phenomena (cf. section 4.2.2).

For a quantitative analysis, the following quantities are regarded: First, we extract the carrier concentration $n_{extr}(x)$ from the detector signal of a virtual experiment and calculate the relative deviation E_{rel} from the reference concentration $n_{ref}(x)$ obtained by an electrothermal device simulation

$$E_{rel} := \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{n_{extr}(x_i)}{n_{ref}(x_i)} - 1\right)^2} \quad . \tag{4.8}$$

The second optimization criterion is the necessary minimum distance to the metallization layers which is defined by the position x_i, where the relative deviation $|\frac{n_{extr}(x_i)}{n_{ref}(x_i)} - 1|$ exceeds 10 %.

Employing small angular apertures, the most significant error originates in the large spot diameter. On the other hand, increasing the angular aperture decreases the minimum spot size but increases the spreading of the probing beam and with that the spot diameter at the device surfaces. Hence, at a given sample length there is an optimum angular aperture which increases as the device extension decreases.

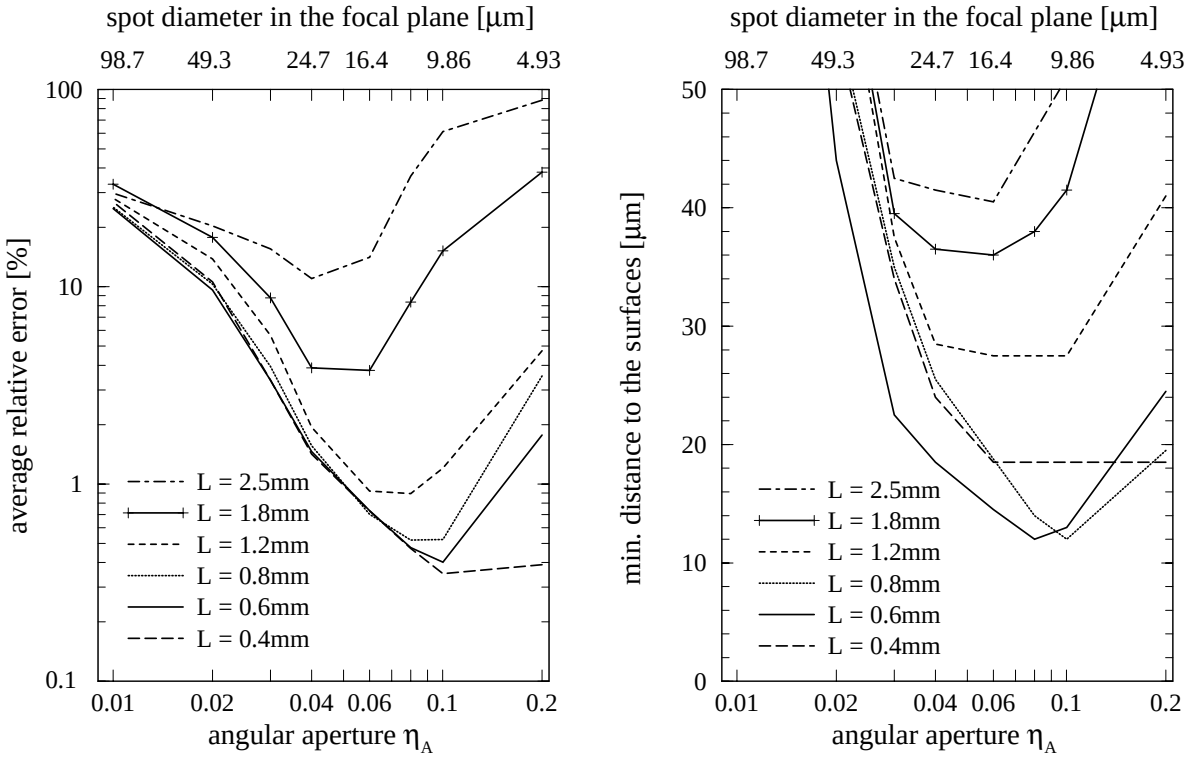


Figure 4.10: Optimizing the angular aperture of the probing beam. Left: Average deviation of the carrier concentration profile gained by a virtual experiment from the reference profile obtained by device simulation [122]; Right: Necessary minimum distance to the boundary for which the local deviation is below 10 %. The carrier lifetime in the sample is $0.8\,\mu s$.

The analytical estimation (cf. eq. 4.6) presented in the previous section predicts an error of $\frac{1}{2L_D^2 k_0^2 \eta_A^2} + \frac{L^2 \eta_A^2}{96 L_D^2 n_{Si}^2}$ which attains its minimum for an angular aperture of

$$\eta_{A,opt} = \sqrt{\frac{4\sqrt{3}\, n_{Si}}{k_0 L}} \ . \tag{4.9}$$

Consequently, the following optimum values for the spot radius $w_{0,opt}$ in the focal plane, the averaged spot radius W_{opt}, and the minimum relative error $E_{rel,opt}$ are obtained:

$$\begin{aligned} w_{0,opt} &= \frac{2}{k_0 \eta_{A,opt}} = \sqrt{\frac{L}{\sqrt{3}\, n_{Si} k_0}} \quad ; \quad W_{opt} = \sqrt{\frac{1}{L} \int_{-L/2}^{L/2} w_{opt}^2(z)\, dz} = \sqrt{\frac{2L}{\sqrt{3}\, n_{Si} k_0}} \ ; \\ E_{rel,opt} &= \frac{L}{4\sqrt{3}\, n_{Si} k_0 L_D^2} \ . \end{aligned} \tag{4.10}$$

Inserting a sample length of $L = 800\,\mu\mathrm{m}$ and a diffusion length of $38\,\mu\mathrm{m}$, we obtain $\eta_{A,opt} = 0.085$, $w_{0,opt} = 5.8\,\mu\mathrm{m}$, $W_{opt} = 8.2\,\mu\mathrm{m}$, and $E_{rel,opt} = 0.6\,\%$ at a wavelength of $1.55\,\mu\mathrm{m}$.

The results of the rigorous numerical study are shown in fig. (4.10). The calculated optimum angular apertures and the relative errors agree quite well with the predictions of the simple estimation. For a sample length of $800\ \mu\mathrm{m}$, the required minimum distance to the device boundaries amounts to $12\ \mu\mathrm{m}$ which is a bit larger than the average spot radius $W_{opt} = 8.2\ \mu\mathrm{m}$. This originates from the deflection of the probing beam and from the diffraction of the tails of the beam profile.

However, some important issues should be pointed out: First, this analysis is valid in case of a negligible surface recombination rate. For that reason, the attainable minimum error continuously decreases as the sample length decreases. A corresponding study which also includes surface recombination effects is presented below. Second, preparation and handling of the sample is only manageable if its length is larger than 0.5 mm. A reduction of the interaction length below this limit requires the design of specific test structures which will be discussed in section (4.3.7). And third, an acceptable signal–to–noise ratio demands a minimum interaction length of about several $100\ \mu\mathrm{m}$.

Focus position

The previous discussion assumes the focal plane of the field distribution to be located in the center of the sample. This is expected to be the optimum setup since it enables the smallest average spot diameter in the sample. Figure (4.11) shows a comparison of the carrier concentration profiles from virtual experiments with the focal planes located at the front surface, in the center, and at the

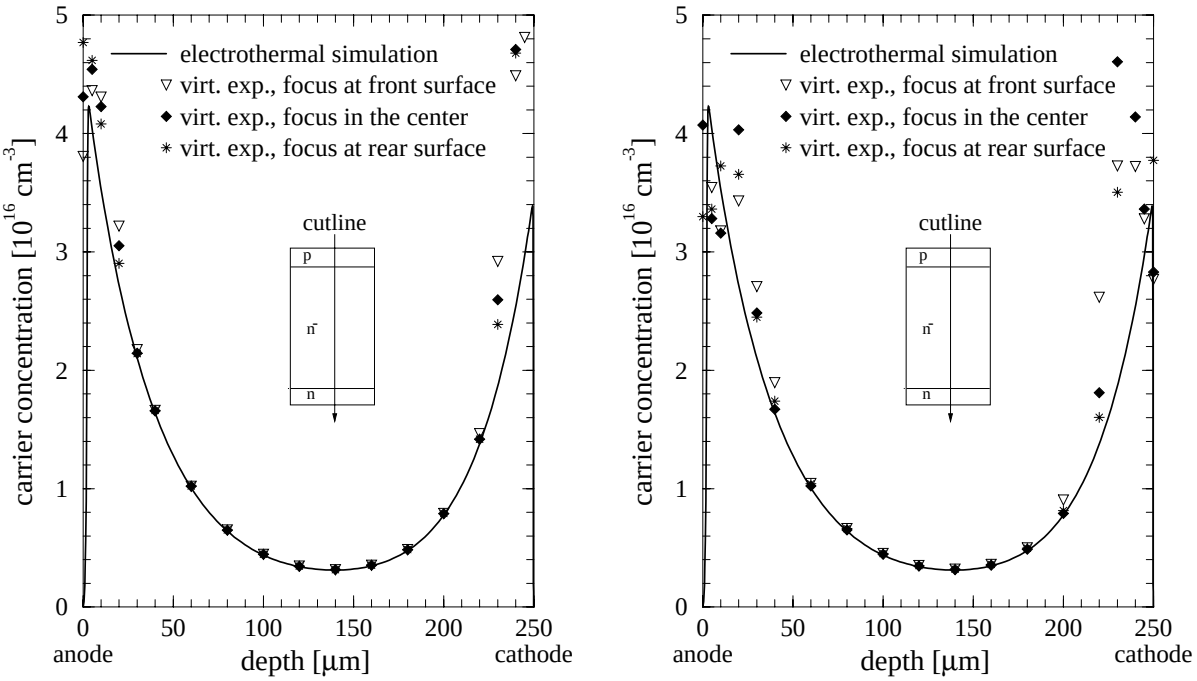


Figure 4.11: Vertical carrier distribution in a pin diode (ambipolar carrier lifetime $0.8\ \mu\mathrm{s}$). The dots mark the concentrations extracted from virtual experiments with different focus positions (left hand figure: sample length $800\ \mu\mathrm{m}$, $\eta_A = 0.08$; right hand figure: sample length 2.5 mm, $\eta_A = 0.04$).

rear surface, respectively. While the relative error for $L = 800\,\mu\mathrm{m}$, $\eta_A = 0.08$ is about 0.6 % in case of a symmetric field distribution, it increases up to 1.2 % if the focal plane is located at the front surface. Similarly, the required distance to the metallization layers increases from $14\,\mu\mathrm{m}$ to $16\,\mu\mathrm{m}$. The effect becomes more pronounced for $L = 2.5\,\mathrm{mm}$, $\eta_A = 0.04$ (cf. fig. 4.11, right), in which case the accuracy deteriorates from 11 % to 29 % and the minimum distance from $40\,\mu\mathrm{m}$ to $60\,\mu\mathrm{m}$. Regarding the absolute values, however, we conclude that a slight shift out of the focal plane does not introduce a significant experimental error, in particular in case of samples with a small interaction length.

Angle of incidence

The above discussion assumes that the sample surfaces are perpendicular to the optical axis. In the following, however, we address the effects of an obliquely impinging probing beam. If the latter propagates at an angle Φ_{Si} in the interior of the sample, the interaction with the carrier distribution is governed by the integral

$$n_{extr}(x_0) = \frac{1}{L} \int_{-L/2}^{L/2} n(x_0 + z \tan \Phi_{Si})\, dz \ . \tag{4.11}$$

Making use of the refraction law at the entrance plane, the angle Φ_{Si} is related to the angle Φ of incidence. Equation (4.11) enables an analytical estimation of the relative error which is plotted as

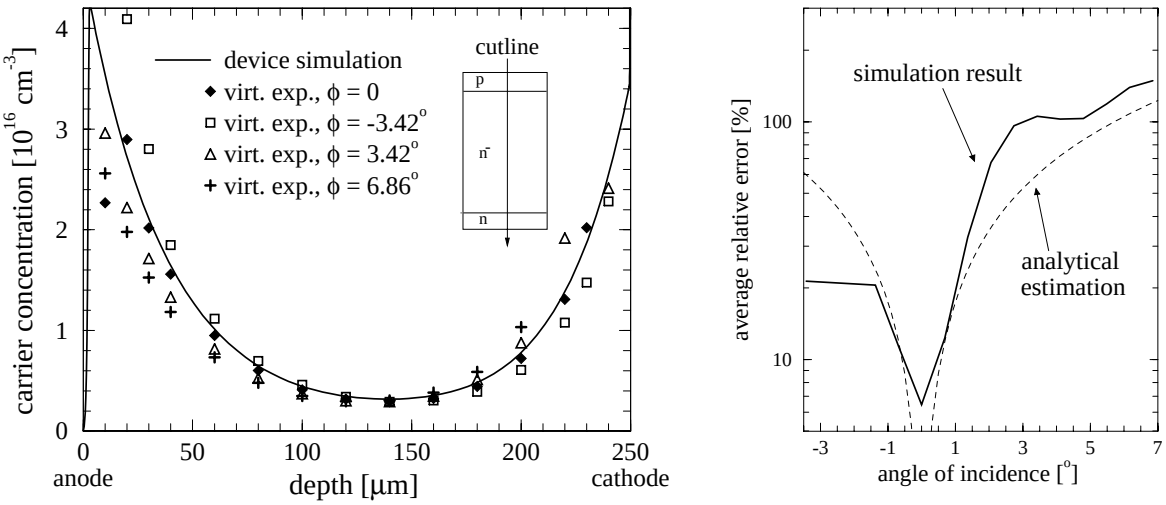


Figure 4.12: Effect of an oblique incidence of the probing beam. Left: Vertical carrier distribution in a pin diode (ambipolar carrier lifetime $0.8\,\mu\mathrm{s}$). The dots mark the concentrations extracted from virtual experiments with different angles of incidence (sample length 0.8 mm, $\eta_A = 0.08$). Right: Relative error as a function of the angle of incidence. The solid line represents the numerical dependence while the broken line is gained by an analytical integration along the tilted beam path neglecting the deflection due to the refractive index gradient.

a dashed line in fig. (4.12). The solid line indicates the result of the rigorous numerical calculation which also takes into account the beam deflection in the interior of the sample. For that reason, the numerical result is not symmetric with respect to $\Phi = 0$.

It is important to note that a possible misalignment of the device under test results in an error which may easily amount to several 10 % and thus constitutes a significant source of error. For example, an angle of incidence $\Phi = 1^{\circ}$ (which corresponds to $\Phi_{Si} = 0.3^{\circ}$) gives rise to a deviation of 20 %. Hence, a precise sample alignment is a crucial precondition of accurate measurement results.

4.2.7 Effects of surface recombination

In this section, we will discuss the effects of an enhanced surface recombination rate which originates from the sample preparation, in particular the sawing and polishing of the surfaces. As a

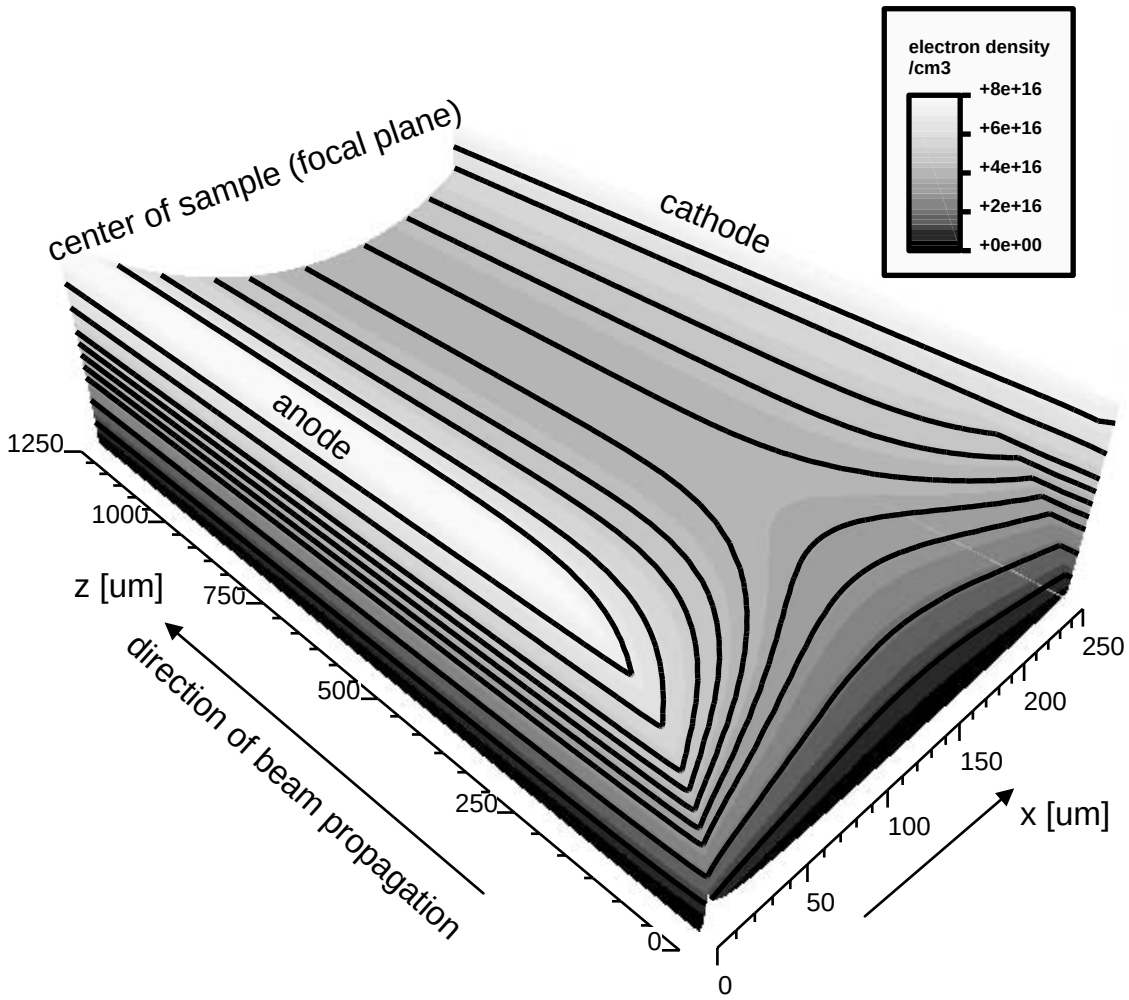


Figure 4.13: 2D carrier distribution in a pin diode (carrier lifetime $8\,\mu\mathrm{s}$, surface recombination velocity $10^5\,\mathrm{cm/s}$). Due to symmetry only one half of the sample is shown.

consequence, the 2D carrier distribution exhibits a saddle area (cf. fig. 4.13). While the vertical profile in the bulk is formed by the typical $\cosh(x)$–like function, the assumed surface recombination velocity of $10^5\,\mathrm{cm/s}$ is sufficient to decrease the carrier concentrations at the surfaces to their equilibrium values.

For that reason, it is no longer desirable to prepare samples with an interaction length as small as possible. We rather expect an optimum length which facilitates a minimum overall error due to both the carrier concentration decrease at the surfaces and the spreading of the probing beam within large devices. Therefore, two parameters are subject of the optimization procedure, namely the sample length L and the angular beam aperture η_A. As in the previous section, the figures of merit are a small deviation of the measured distribution from the reference profile and a minimum distance from the top and bottom surfaces which has to be kept to prevent a corruption of the detector signal due to reflections of the probing beam. Since the product devices are usually much larger in size and therefore not affected by an enhanced surface recombination velocity, the reference profile is calculated without including the surface recombination effects.

Devices with small carrier lifetimes

In case of a carrier lifetime of $0.8\,\mu\mathrm{s}$, surface recombination effects do not reach far into a sample of 2.5 mm in length. As a consequence, the averaged concentration $N(x) := \frac{1}{L}\int_{-L/2}^{L/2} n(x,z)\,dz$, which is gained by integrating the two–dimensional carrier distribution *along* the beam path, matches the carrier concentration obtained by a one–dimensional calculation without surface recombination effects. (cf. fig. 4.14). Hence, in case of a large sample length, the most significant error is introduced by the limited resolution due to the *lateral* extension of the probing beam.

On the other hand, the decrease of the carrier concentration at the surfaces is no longer negligi-

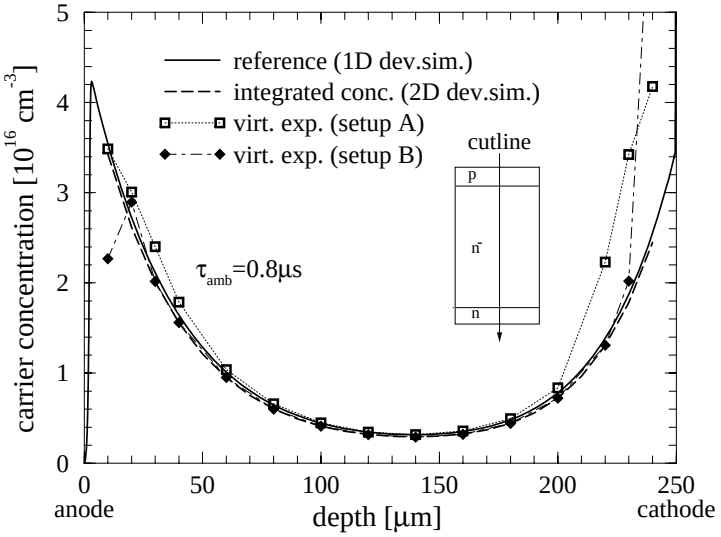


Figure 4.14: Carrier distribution in a pin diode (ambipolar carrier lifetime $0.8\,\mu\mathrm{s}$). The dots indicate the concentration values extracted from virtual experiments including surface recombination effects for two different setups A and B ($L = 2.5\,\mathrm{mm}$, $\eta_A = 0.02$ for setup A and $L = 0.8\,\mathrm{mm}$, $\eta_A = 0.08$ for setup B).

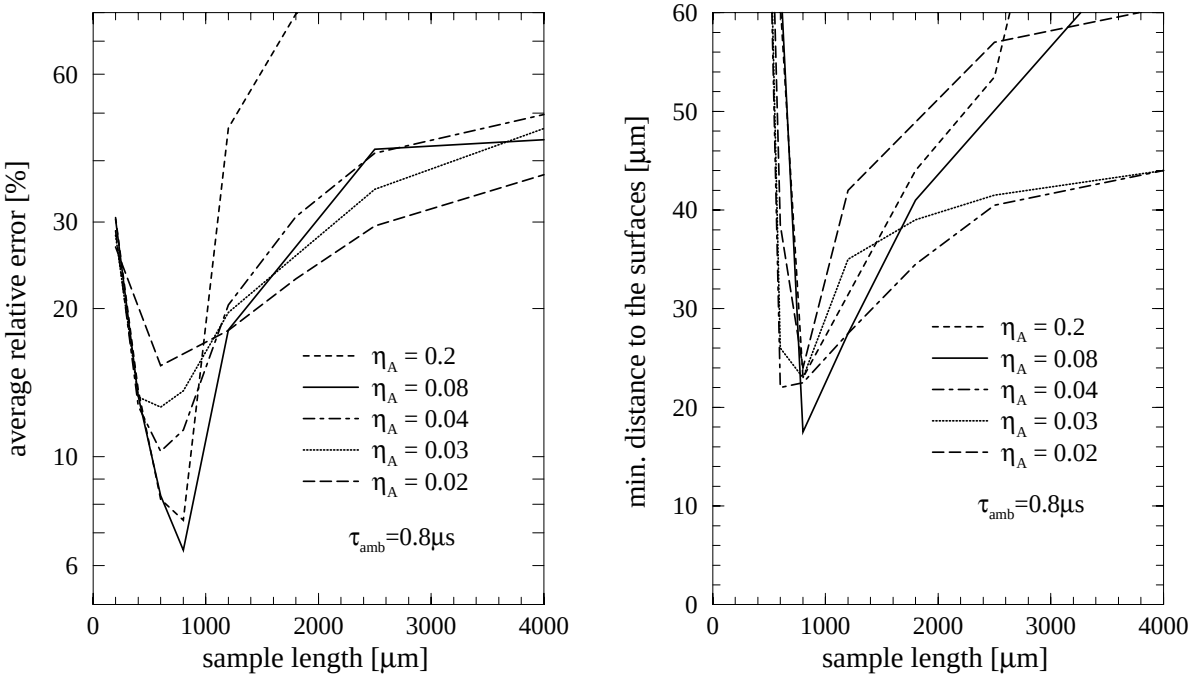


Figure 4.15: Optimizing the sample geometry and the angular aperture of the probing beam. Left: Average deviation of the carrier concentration profile gained by virtual experiments from the reference profile obtained by a 1D device simulation. Right: Required minimum distance to the boundary for which the local deviation is below 10 %. The ambipolar carrier lifetime in the sample is $0.8\ \mu s$.

ble, if the device length is reduced below 1 mm. Simulating the measurement process for different geometries reveals an optimum sample length of $800\ \mu\mathrm{m}$ which is about 20 times the diffusion length. The corresponding optimal angular aperture is about 0.08 (cf. fig. 4.14 and 4.15) which agrees well with the result of the simple estimation (4.9) from the previous section. The numerical study predicts a lower error limit of 6 % and a minimum distance to the metallization layers of about $16\ \mu\mathrm{m}$.

Devices with large carrier lifetimes

In case of large diffusion lengths, the curvature in the vertical direction is small. Consequently, the convolution of the carrier distribution with the lateral beam profile does not introduce a significant error. On the other hand, the effects of the surface recombination reach far into the sample. The absorption signal is therefore sensitive to the averaged carrier density $N(x) := \frac{1}{L}\int_{-L/2}^{L/2} n(x,z)\,dz$, which essentially differs from the 1D reference concentration (cf. fig. 4.16).

The numerical study reveals the optimum arrangement which reproduces the reference profile as closely as possible (cf. fig. 4.17). For that purpose, the sample has to be 2.5 mm in length, which is again 20 times the diffusion length. The corresponding optimum angular aperture is about 0.02 which is lower than the value 0.04 estimated from equation (4.9). Thus, measurements on devices with large carrier lifetimes can also attain an average error down to 0.06 % and can be carried out up to a minimum distance from the boundaries of only $16\ \mu\mathrm{m}$.

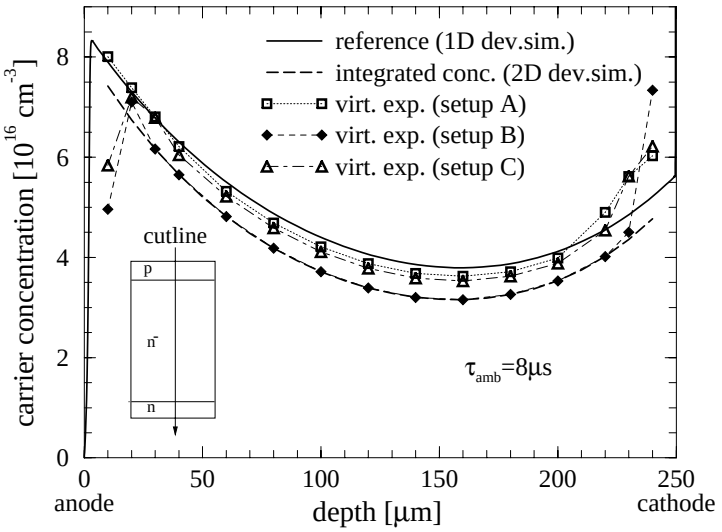


Figure 4.16: Carrier distribution in a pin diode (ambipolar carrier lifetime $8\,\mu$s). The dots indicate the concentration values extracted from virtual experiments including surface recombination effects for three different setups A, B, and C ($L = 2.5\,$mm, $\eta_A = 0.02$ for setup A, $L = 0.8\,$mm, $\eta_A = 0.08$ for setup B, and $L = 1.8\,$mm, $\eta_A = 0.03$ for setup C).

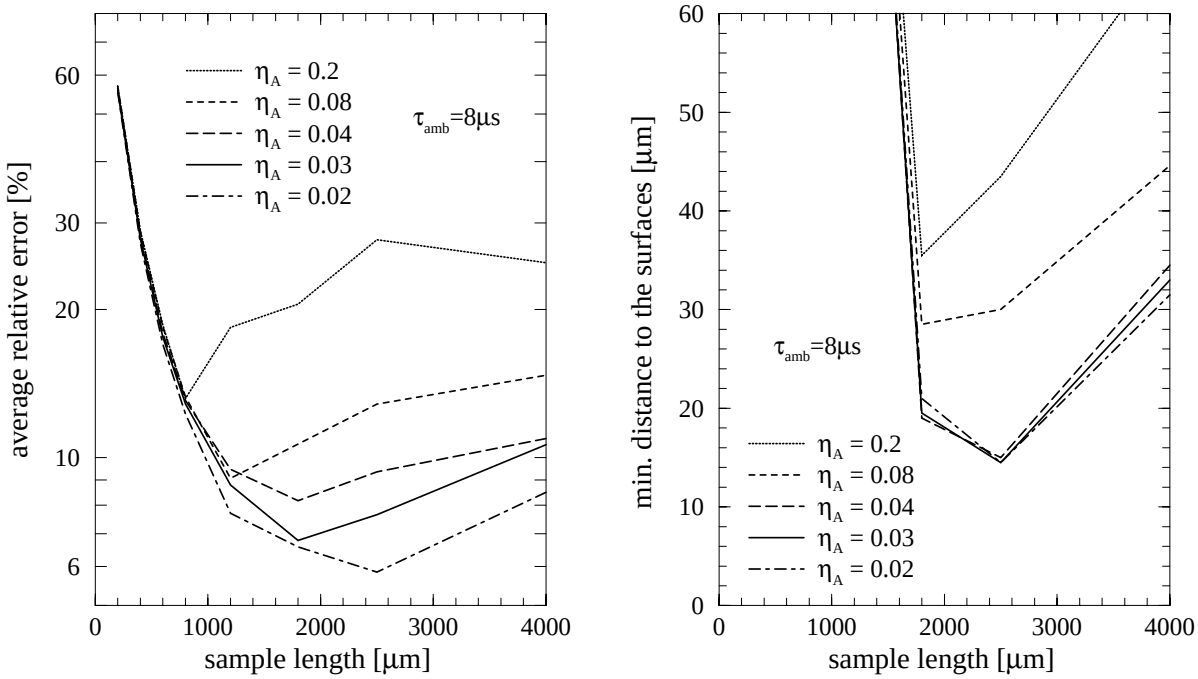


Figure 4.17: Optimizing the sample geometry and the angular aperture of the probing beam. Left: Average deviation of the carrier concentration profile gained by virtual experiments from the reference profile obtained by a 1D device simulation. Right: Required minimum distance to the boundary for which the local deviation is below 10 %. The ambipolar carrier lifetime in the sample is $8\,\mu s$.

However, it should be mentioned that the curvature of the integrated profile $N(x)$ is different from that of the reference profile. Consequently, the diffusion length can not be directly extracted from the measured carrier distribution. An analytical first order correction is gained by the following consideration which is based on an expansion of the 2D carrier distribution in terms of $\cos(\gamma_j \frac{z}{L/2})$–functions (cf. [31, appendix])

$$n(x,z) = \int_0^\infty n(x,\gamma)\, \cos\left(\gamma \frac{z}{L/2}\right)\, d\gamma \ . \tag{4.12}$$

Each component $n(x,\gamma)$ has to satisfy the ambipolar diffusion equation

$$\left(\frac{\partial^2}{\partial x^2} - \gamma^2 \frac{4}{L^2}\right) n(x,\gamma) = \frac{1}{L_D^2}\, n(x,\gamma) \ . \tag{4.13}$$

As the boundary conditions

$$D_{amb}\, \frac{\partial n(x,z)}{\partial z} = \mp v_{sr}\, n(x,z) \qquad \text{for } z = \pm\frac{L}{2} \tag{4.14}$$

result in the relation

$$\gamma_j\, \tan(\gamma_j) = \frac{v_{sr} L}{2 D_{amb}} \tag{4.15}$$

with an infinite set of discrete solutions γ_j, we end up with

$$n(x,z) = \sum_{j=1}^{\infty} A_j \cosh\left(\frac{x - x_0}{L_j}\right)\, \cos\left(\gamma_j \cdot \frac{2z}{L}\right) \quad \text{where} \quad \frac{1}{L_j} := \sqrt{\frac{1}{L_D^2} + \gamma_j^2 \frac{4}{L^2}} \ . \tag{4.16}$$

The integrated carrier concentration finally becomes

$$N(x) = \sum_{j=1}^{\infty} A_j \cosh\left(\frac{x - x_0}{L_j}\right)\, \frac{\sin(\gamma_j)}{\gamma_j} \ . \tag{4.17}$$

Assuming an infinite surface recombination velocity (thus, $\gamma_j = (2j-1)\pi/2$) and approximating this expansion by the first moment, the desired first order correction for the unknown diffusion length L_D is obtained as

$$\frac{1}{L_D} = \sqrt{\frac{1}{L_1^2} - \frac{\pi^2}{L^2}} \ , \tag{4.18}$$

where L_1 is extracted from the curvature of the measured distribution $N(x)$. In the limit of small carrier lifetimes ($L_D \ll L$), this equation simplifies to $L_D \approx L_1$, i. e. the diffusion length can be directly extracted from the curvature of the measured profile. This result is already expected in view of the findings from the previous section.

4.2.8 Summary

The theoretical study of the free carrier absorption measurement technique leads to the following conclusions:

- Reflection and diffraction at the top or bottom sample surfaces give rise to unpredictable distortions of the beam profile, thus corrupting the measurement signal. To avoid these effects, a minimum distance from the device boundaries is required.

- Self–heating can easily change the optical thickness of the sample by several wavelengths. To suppress the resulting Fabry–Perot oscillations of the measurement signal special precautions have to be taken, e. g. depositing an antireflective coating or employing an incoherent light source.

- The space charge regions around the reverse biased p–wells in IGBT samples act as lenses deflecting the probing beam. For sample lengths below 1 mm, however, keeping a distance of approx. $20\,\mu\mathrm{m}$ to the top surface facilitates reliable measurements which are sensitive to the average carrier concentration in the cell structure along the beam path.

- Our strategy in optimizing the experiment strives for minimizing two criteria: First, the deviation of the carrier concentration gained by a virtual experiment from the reference concentration obtained by an electrothermal simulation, and second, the minimum distance to the device boundaries which is required to prevent a corruption of the measurement signal.

- At any given sample length L, the optimum angular aperture of the probing beam can be calculated from the condition of minimizing the average beam cross section. The resulting analytical formula reproduces the optimum, which was found numerically by means of virtual experiments with different probing beams.

- A surface recombination velocity of $10^5\,\mathrm{cm/s}$ decreases the carrier concentrations at the sample surfaces almost to their equilibrium values. Consequently, an optimum sample length is about 20 diffusion lengths. The corresponding optimum angular aperture can be approximated by the equation minimizing the average beam cross section. Thus, the measurements can be carried out at an average error of about 6 %, if a minimum distance of $16\,\mu\mathrm{m}$ to the top and bottom device surfaces is kept.

- The exact position of the focal plane of the probing beam is not a critical parameter. Even if it is located at the sample surfaces, instead of in the center, the average error merely increases by a factor of 2.

- A possible misalignment of the sample constitutes the most significant source of error. Tilting the sample at an angle of a few degrees gives rise to an error of several 10 %. In addition, the necessary minimum distance to the boundaries increases rapidly.

- The accuracy of the real measurement results sensitively depends on the precise knowledge of the carrier concentration dependence of the absorption coefficient. This source of error is not apparent in the numerical studies as they rely on the same value to calculate the refractive index modulations (cf. section 3.2) and to evaluate the detector signal. However, since this coefficient enters as a factor in the evaluation rule, the *relative* profiles are accurate in any case.

4.3 Laser deflection measurements

Internal laser deflection measurements are sensitive to the gradient of the refractive index and therefore facilitate the experimental determination of carrier concentration *gradients* and temperature *gradients*. Both contributions can be usually separated due to the widely differing electric and thermal time constants. Alternatively, the carrier contribution is calculated from the absorption measurements and subtracted from the deflection signal.

Many of the findings of the previous chapter are expected to be valid for deflection measurements as well, especially as far as the longitudinal averaging in samples comprising a periodic cell structure (cf. section 4.2.3) and the lateral convolution of the beam profile and the refractive index distribution (cf. section 4.2.5) are concerned. It is also quite evident that a crucial source of error originates from a possible misalignment of the sample.

Due to the saturation of the detector response function, the current setup suffers from a limited measurement range which is already exceeded if the maximum temperature rise within the sample is greater than $5 - 10\,\mathrm{K}$ [57]. Therefore, the major challenge is to facilitate the detection of high temperature gradients up to $1\,\mathrm{K}/\mu\mathrm{m}$ which typically occur during the short circuit operation of power devices. For that purpose, two strategies are investigated, namely the optimization of the optical setup and an appropriate design of specific samples or test structures.

4.3.1 State of the art: Deflection measurements in case of low power dissipation

Measurements of the temperature profiles in power devices operating with low power dissipation

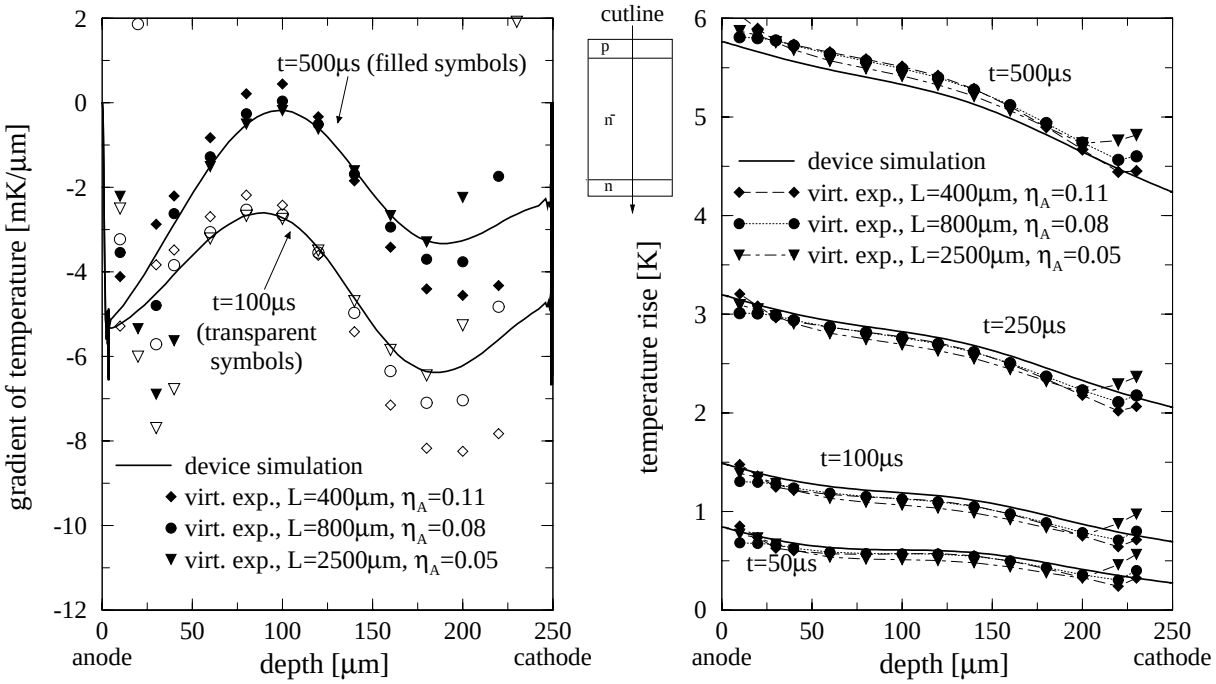


Figure 4.18: Temperature distribution within a pin diode subjected to a current pulse of $150\,\mathrm{A}/\mathrm{cm}^2$.

have been performed successfully [49, 78, 123, 124, 125, 126]. In this case, the angular deflection and the internal shift of the probing beam are sufficiently small so that the beam displacement on the detector remains within the linear range of the response function.

To judge the accuracy of these measurements, virtual experiments according to the strategy sketched in fig. (4.1) are carried out, using different angular apertures and sample lengths. The carrier contributions are calculated from the simulated absorption signals (cf. section 4.2.7) and subtracted from the deflection signals. Thus, the remaining contributions reveal the temperature gradients which are compared to their reference values obtained by an electrothermal device simulation (cf. fig. 4.18). Finally, an integration yields the absolute temperature profiles if the integration constants are calculated from the total electric power dissipation assuming adiabatic boundary conditions.

As it can be seen from fig. (4.18), the extracted temperature gradients deviate from the reference gradients by about 10 % to 20 %. The variations predominantly originate from the uncertainty of the carrier contributions. However, as the integration procedure significantly smoothes the profiles we merely observe a negligible error in the absolute temperature distributions, regardless of the sample length and the beam aperture. This clearly demonstrates the technique's capability to detect small temperature variations at high sensitivity and excellent accuracy.

4.3.2 Simultaneous free carrier absorption measurements

The experimental setup for laser deflection measurements offers a simple way to include simultaneous absorption measurements since the transmitted intensity can be calculated by summing up

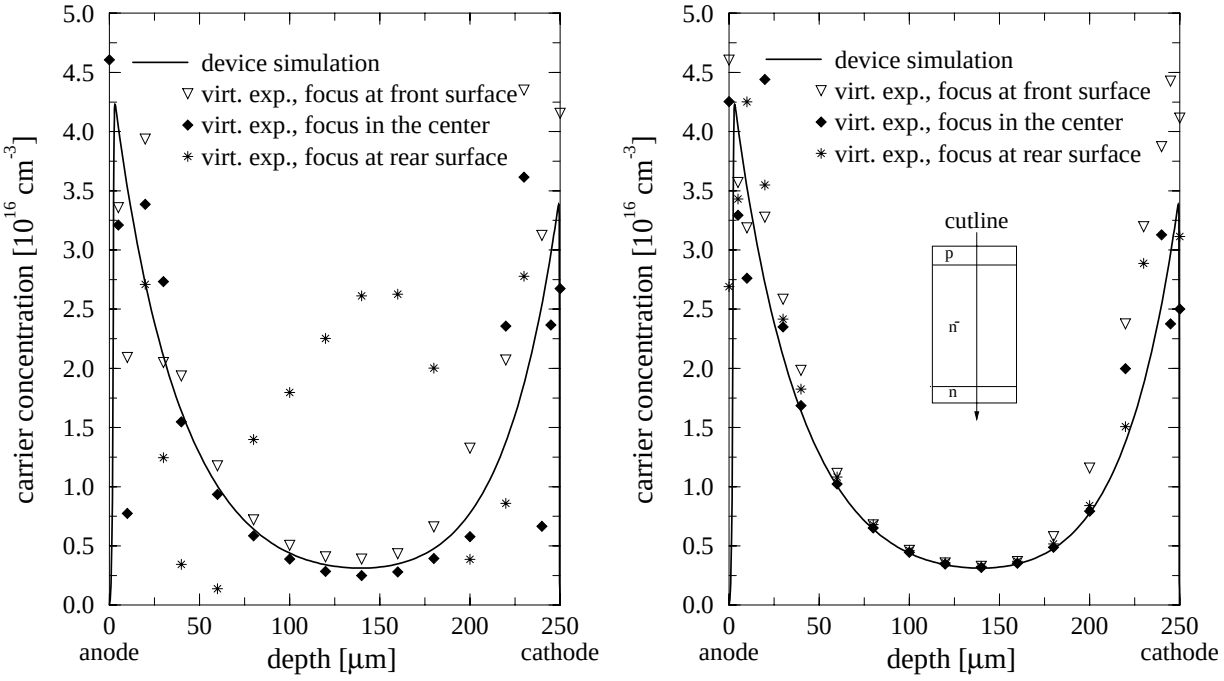


Figure 4.19: Extracting the carrier distribution from the absorption signal measured by a four quadrant photo diode (left) or a separate pin diode (right) (interaction length $L = 2.5\,\mathrm{mm}$, $\eta_A = 0.06$).

all the photo currents of the four–quadrant detector [57]. Even though this method is feasible with almost no additional experimental effort, several difficulties arise: First, accurate absorption measurements require a large active area of the photo diode to detect the total incident IR power even if the intensity profile on the detector is shifted due to the internal beam deflection. As will be shown below, however, a smaller detector is preferable for laser deflection measurements as it facilitates a larger measurement range. Second, the field distribution on the detector is at its maximum in the center of the four–quadrant diode, i. e. in the inactive area between the segments. Thus, a deflection of the laser beam due to temperature gradients, for instance, increases the detected IR power which would be misinterpreted in terms of a decreasing carrier concentration.

Fig. (4.19) compares the carrier distributions which are extracted from the absorption signal detected by a four–quadrant photo diode and a simple pin diode, respectively. Since the latter is sensitive to the total intensity profile without cutting any segments or the tails of the intensity distribution, the undesired coupling of the absorption signal and the internal beam deflection is suppressed. Consequently, employing the pin detector improves the accuracy of the measurement results and decreases the sensitivity to a possible delocation of the focal plane within the device (cf. fig. 4.19). Despite the more complicated experimental setup, an additional beam splitter and a separate pin diode detector for simultaneous absorption measurements are therefore preferable.

4.3.3 The Fabry–Perot effect

Due to multiple reflections in the interior of the sample, thermally induced changes of its optical thickness give rise to Fabry–Perot oscillations on the absorption signal (cf. section 4.2.4). The standardized deflection signal, however, is not affected by pure *amplitude* modulations since any constant factor is canceled (cf. eq. 2.18). On the other hand, in case of plane–parallel facets each of the multiply reflected rays of different order is shifted by a distance which is the larger the more times the ray has been reflected back and forth. Consequently, the detected displacement of the total intensity distribution depends on whether the contributions interfere constructively or destructively (cf. fig. 4.20). As with the free carrier absorption measurements, the thermally induced Fabry–Perot effect is observed as a small drift in case of low power dissipations (cf. fig. 4.21, left) or as complete oscillations if the optical thickness is increased by several wavelengths (cf. fig. 4.21, right).

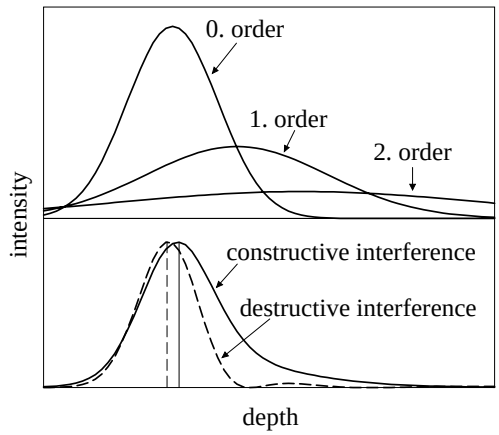


Figure 4.20: Fabry–Perot effect of the deflection signal. Above: Intensity distribution of the multiply reflected rays. Below: Total intensity distributions in case of constructive or destructive superposition (for the sake of clarity, each profile is scaled to 1).

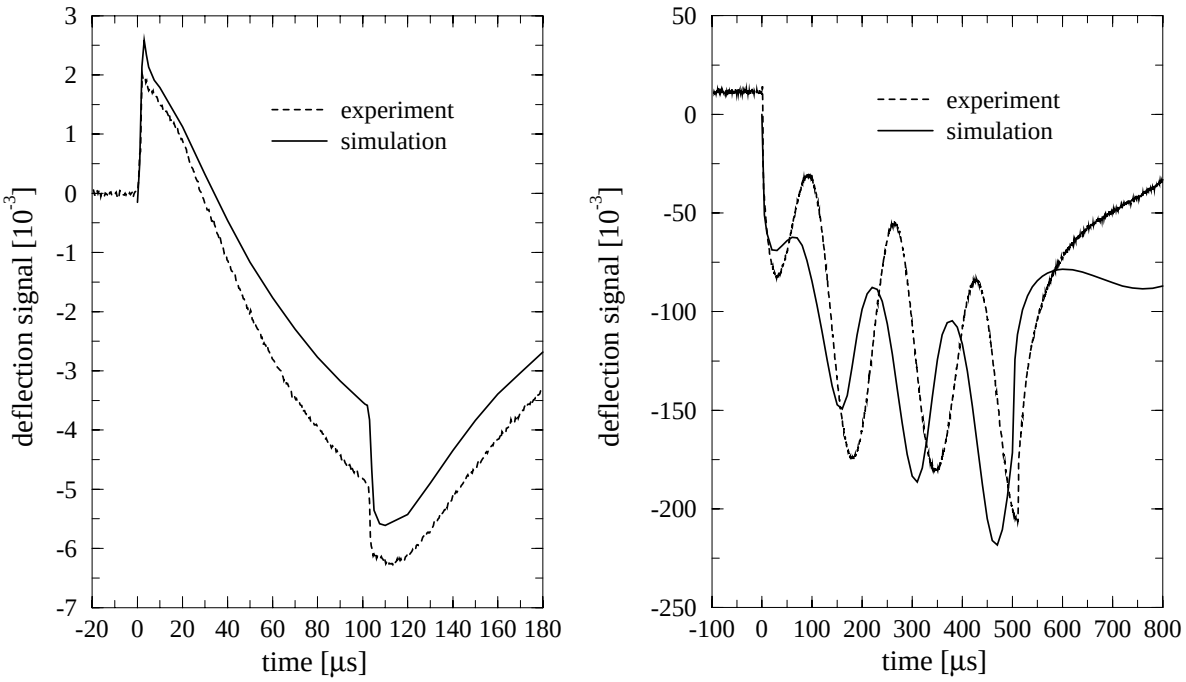


Figure 4.21: Deflection signals in case of low power dissipation (left, current density $15\,\mathrm{A/cm^2}$) and high power dissipation (right, current density $150\,\mathrm{A/cm^2}$) if multiple reflections within the sample are not suppressed.

4.3.4 Analytical calculation of the measurement signal

After it has been demonstrated that the Internal Laser Deflection technique constitutes a sensitive and accurate method to determine small temperature gradients, we will now address the question of how the limited measurement range can be increased. As the final goal, we strive for the capability of detecting peak temperatures up to 100 K to 200 K, which typically occur within power devices if they are, e. g., operating under short circuit conditions. The following discussion, however, will be restricted to the currently feasible modifications of the optical setup and the sample preparation. Thus, proposals will be elaborated which are intended to facilitate a measurement range of at least 50 K to 100 K.

Within this section, an *analytical* expression of the deflection signal is derived for two major reasons: First, it clearly reveals the parameters of the optical setup and the sample geometry which affect the detector signal and therefore shall be subject of optimization. And second, extracting the desired temperature distribution from the measurement signal requires an evaluation rule which is also valid in case of a large beam deflection. We therefore present a consistent derivation which is compared to the model reported in [57].

The fundamental idea is to regard the probing beam as a Gaussian beam whose optical axis is deflected in the interior of the sample according to the principles of geometrical optics. The internal deflection is represented by two parameters, namely $\Delta\tilde{x}$ and α_2, since the field distribution at the rear surface can be imagined as intensity distribution of a *non–deflected* Gaussian beam which emerges from the focal plane of the focussing lens at an angle α_2 and a shift $\Delta\tilde{x}$ (cf. fig. 4.22). Finally, Fourier optics is employed to calculate the projection of this Gaussian beam onto the detector plane.

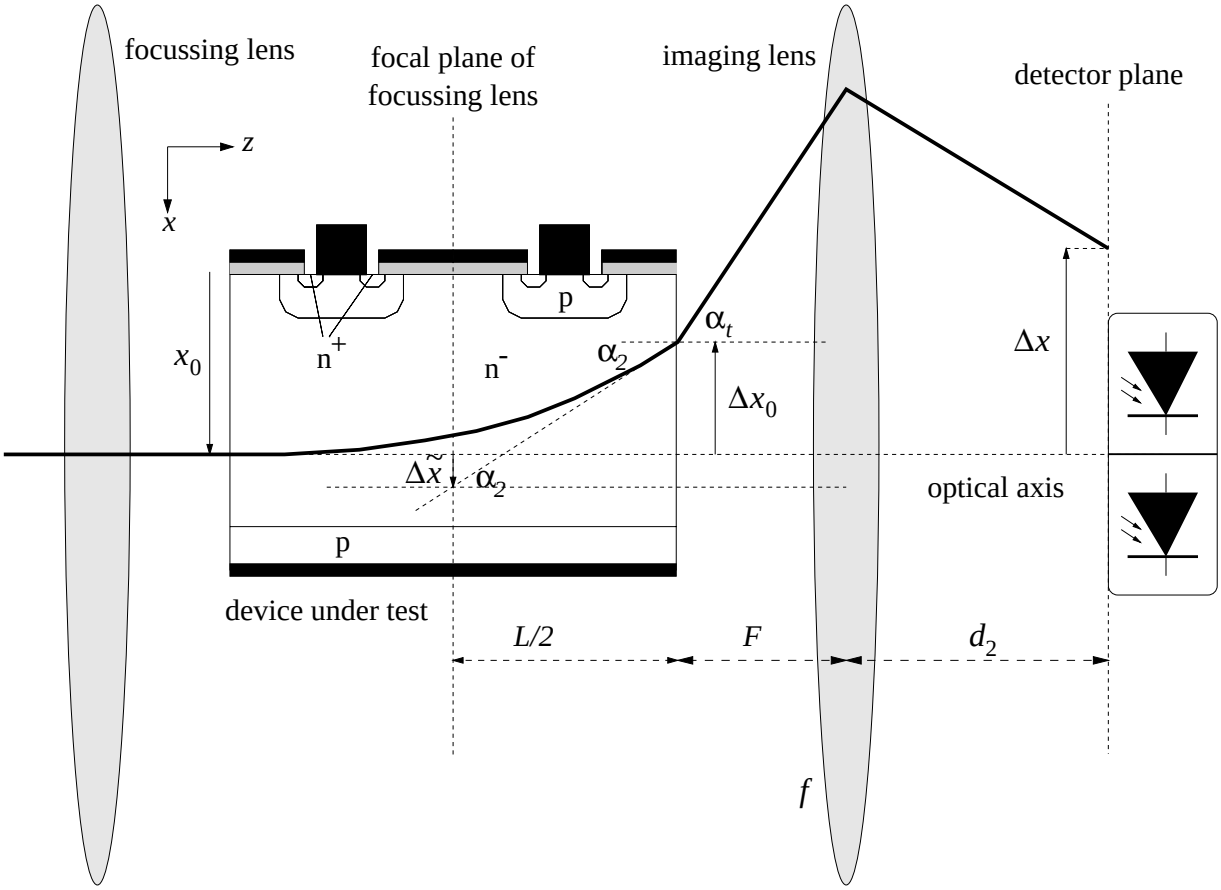


Figure 4.22: Internal deflection and projection of the probing beam onto the detector plane.

Internal beam deflection

In appendix (C), the internal beam path is calculated from Fermat's principle. Assuming a vanishing angle of incidence and a constant gradient of the refractive index within the area covered by the probing beam, we obtain the following result

$$\Delta x_0 = l \cdot [\cosh(L/l) - 1] = \frac{L^2}{2l} + O\left(\frac{L}{l}\right)^4 = \frac{dn_{Si}}{dx} \cdot \frac{L^2}{2n_{Si}(x_0)} + O\left(\frac{L}{l}\right)^4 \quad (4.19)$$

$$\tan(\alpha_2) = \frac{dx}{dz}|_{z=L} = \sinh(L/l) = \frac{L}{l} + O\left(\frac{L}{l}\right)^3 = \frac{dn_{Si}}{dx} \cdot \frac{L}{n_{Si}(x_0)} + O\left(\frac{L}{l}\right)^3 \quad (4.20)$$

where we have introduced the definition $l := \frac{n_{Si}(x_0)}{dn_{Si}/dx}$ (cf. eq. C.8). The leading terms of these Taylor expansions exactly match the expressions reported in [57].

The same field distribution at the rear surface of the sample is produced by a non–deflected, rotated Gaussian beam whose focus is located in the focal plane of the focussing lens and shifted by a distance $\Delta\tilde{x}$ (cf. fig. 4.22). Geometrical considerations reveal

$$\Delta\tilde{x} = \Delta x_0 - \frac{L}{2}\tan\alpha_2 = l \cdot \cosh\left(\frac{L}{l}\right) - l - \frac{L}{2} \cdot \sinh\left(\frac{L}{l}\right) = -\left(\frac{L}{l}\right)^4 \cdot \frac{l}{24} + O\left(\frac{L}{l}\right)^6 . \quad (4.21)$$

Projection by the imaging lens

The field distribution in the focal plane of the focussing lens is projected onto the detector plane by the imaging lens. In order to take into account the refraction at the rear surface of the sample, the angle α_2 and the distance $L/2$ have to be replaced by α_t and and $L/2n_{Si}$, respectively. Therefore, we consider a Gaussian beam which is rotated by an angle α_t and shifted by $\Delta\tilde{x}$ from the optical axis. Neglecting the curvature of the phase fronts near the focus, its field distribution can be approximated by

$$\vec{E}(x) = \vec{E}_0 \, \exp\left[-\frac{(x-\Delta\tilde{x})^2}{w_0^2/\cos^2(\alpha_t)}\right] \exp\left[ik_0 \sin(\alpha_t)\cdot x\right] \ . \tag{4.22}$$

With an optical distance d_1 to the imaging lens equal to $d_1 = F + L/2n_{Si}$ (cf. fig. C.3), this field distribution is transformed to a Gaussian intensity profile on the detector (cf. appendix C):

$$|\vec{E}_D(x)| = \mathit{const}\cdot \exp\left[-\frac{(x-\Delta x)^2}{w^2}\right] \ . \tag{4.23}$$

The displacement Δx and the spot radius w on the detector are given by the relations

$$\begin{aligned} \Delta x &= \frac{a}{2V_t}\sin(\alpha_t) - \frac{\Delta\tilde{x}}{V_t} = \\ &= \frac{a\, n_{Si}}{2V_t}\cdot\frac{L}{l} + O\left(\frac{L}{l}\right)^3 = \frac{aL}{2V_t}\cdot\frac{dn_{Si}}{dx} + O\left(\frac{L}{n_{Si}}\frac{dn_{Si}}{dx}\right)^3 \end{aligned} \tag{4.24}$$

$$\begin{aligned} w &= \frac{2}{V_t k_0 \eta_A \cos(\alpha_t)}\sqrt{1+\frac{k_0^2\eta_A^4\, a^2\cos^4(\alpha_t)}{16}} = \\ &= \frac{\sqrt{16+k_0^2\eta_A^4 a^2}}{2V_t k_0\eta_A} + \frac{n_{Si}^2}{4V_t}\cdot\frac{16-k_0^2a^2\eta_A^4}{k_0\eta_A\sqrt{16+k_0^2a^2\eta_A^4}}\left(\frac{L}{l}\right)^2 + O\left(\frac{L}{l}\right)^4 \end{aligned} \tag{4.25}$$

in which we have included the definitions (cf. eq. C.19)

$$V_t := \frac{1}{d_2/f-1} \qquad a := 2\left(\frac{L}{2n_{Si}} + F - V_t d_2\right) \ . \tag{4.26}$$

As it is explained in appendix (C), the parameter V_t constitutes the inverse of the transverse magnification, while a represents twice the distance between the focal plane of the focussing lens and the image of the detector plane which is formed by the imaging lens.

Detector signal

The intensity profile on the four–quadrant photo diode gives rise to a deflection signal M which is related to the standardized beam displacement $\Delta x/w$ by the detector response function. In case of a small displacement, we obtain the linear dependence

$$\begin{aligned} M = Q(w)\cdot\frac{\Delta x}{w} &= Q(w)\cdot\frac{a\,L\,k_0\eta_A}{\sqrt{16+k_0^2\eta_A^4a^2}}\cdot\frac{dn_{Si}}{dx} + O\left(\frac{L}{n_{Si}}\frac{dn_{Si}}{dx}\right)^3 = \\ &\approx Q(w)\cdot\frac{L}{\eta_A}\left(1+\frac{8}{k_0^2\eta_A^4a^2}\right)\cdot\frac{dn_{Si}}{dx} \ . \end{aligned} \tag{4.27}$$

As the major drawback, however, it turned out that the slope Q of the response function depends on the spot size w of the probing beam on the detector. Experimental evidence is provided in [78]. Consequently, any modification of the optical setup, in particular the variation of the beam aperture or a shift of the imaging lens, demands a recalibration of the response function. For that purpose, the detector is shifted by the piezo translators during the off–state of the device under test, recording the corresponding modulation of the deflection signal. This preprocessing step can be included into the program control of the experimental setup. However, since the spot size w on the detector is unknown, it is more convenient to consider the detector response with respect to the absolute displacement, i. e. without scaling it to the spot radius w. Rewriting (4.27) in terms of $\tilde{Q}(w) := Q(w)/w$ results in

$$\begin{aligned} M = \tilde{Q}(w) \cdot \Delta x &= \tilde{Q}(w) \cdot \left[\frac{a}{2V_t} \sin(\alpha_t) - \frac{\Delta \tilde{x}}{V_t} \right] \\ &= \frac{\tilde{Q}(w)}{V_t} \cdot \left[\frac{L^2}{2n_{Si}} + L\left(F - V_t d_2\right) \right] \frac{dn_{Si}}{dx} + O\left(\frac{L}{n_{Si}} \frac{dn_{Si}}{dx} \right)^3 . \end{aligned} \quad (4.28)$$

Solving for the refractive index gradient results in the following evaluation rule:

$$\boxed{ \frac{dn_{Si}}{dx} = \left(\frac{\partial n_{Si}}{\partial T} \right)_{n,p} \frac{dT}{dx} + \left(\frac{\partial n_{Si}}{\partial C} \right)_{T} \frac{dn}{dx} = \frac{M \cdot V_t / \tilde{Q}(w)}{\frac{L^2}{2n_{Si}} + L\left(F - V_t d_2\right)} + O(M^3) } \quad (4.29)$$

Parameters of the experiment

Equation (4.28) clearly reveals the degrees of freedom which can be used to optimize the experiment with respect to a higher measurement range. Actually, there a three factors that have to be taken into consideration:

First, the interaction length L. A reduction of the sample length obviously decreases the internal deflection at a given gradient of the refractive index, which, in turn, facilitates the detection of higher temperature gradients. On the other hand, the shrinking of the sample size is limited by practical constraints. We will therefore discuss the design of specific test structures in section (4.3.7).

Second, the detector response function. As its saturation constitutes the limiting effect of the measurement range, a lower sensitivity (i. e. a smaller slope of the response function) favors the maximum detectable beam displacement. Details are addressed in section (4.3.5).

Third, the position of the image of the detector plane, i. e. the parameter $b := F - V_t d_2 = F - f d_2/(d_2 - f)$. As mentioned above, b can be interpreted as the distance between the rear surface of the sample and the image of the detector plane which is formed by the imaging lens. Note that the transverse magnification $1/V_t$ does not contribute an independent parameter since scaling the intensity distribution on the detector has the same effect as the inverse scaling of the detector itself (except for the bandwidth which is related to the detector area). Formally speaking, the parameter V_t can be combined with the slope of the response function $\tilde{Q}(w)$. It is interesting that the image formation by the experimental setup is therefore characterized by one single degree of freedom! In other words, if the parameters of the optical configuration, which are obviously chosen as F, f, d_2,

are replaced by a different set, namely $V_t = f/(d_2 - f)$, $F + d_2$, and $b = F - fd_2/(d_2 - f)$, only the latter affects the deflection signal. This is also true in case the image formation is performed by means of several lenses, as they can be replaced by a single thick lens if all distances are measured to the corresponding principal planes. Modifying the parameter b is discussed in section (4.3.6).

4.3.5 Detector response

The limitation of the measurement range originates from the saturation of the detector response function, the typical shape of which is depicted in fig. (4.24). Within a two–dimensional consideration, the standardized deflection signal writes

$$M = \frac{I_1 - I_2}{I_1 + I_2} = \frac{I_1/I_2 - 1}{I_1/I_2 + 1} \tag{4.30}$$

where I_1 and I_2 denote the photo currents of the two detector segments. If one of the photo currents is much greater than the other one, the deflection signal saturates at ± 1 and any further increase of the beam displacement does not increase the detector signal. Therefore, the crucial condition is that within the desired measurement range, the photo currents are of equal order of magnitude.

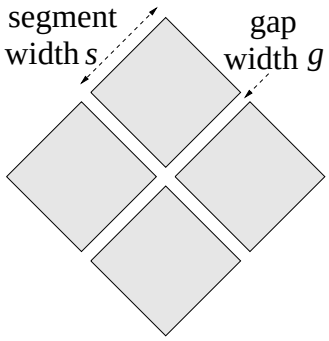


Figure 4.23: Geometry of a four–quadrant photo diode detector.

The geometry of the detector (cf. fig. 4.23) is characterized by the segment size s and the width g of the gap between the segments. Only the ratios of these quantities and the laser spot size on the detector plane are decisive, since scaling the field distribution and the detector geometry by the same factor does not alter the photo currents.

Dependence on the gap size

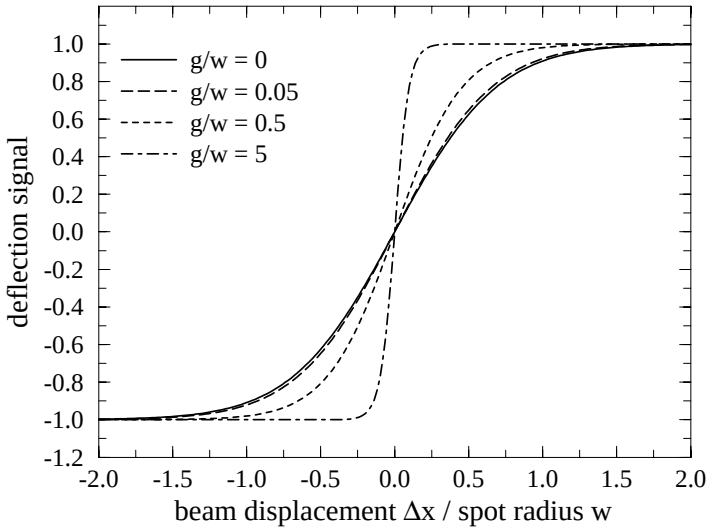


Figure 4.24: Detector response functions for different gap widths g (standardized detector segment size $s/w = 1$).

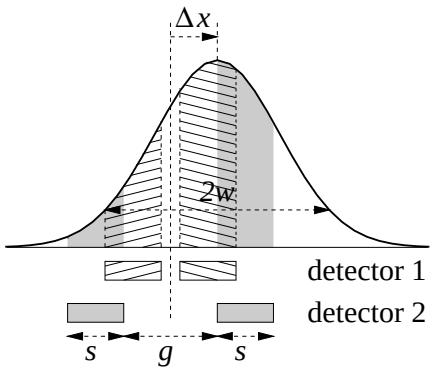


Figure 4.25: Four–quadrant diodes with a small gap (detector 1) and a large gap (detector 2). The photo currents are proportional to the highlighted areas.

Fig. (4.24) compares the response functions of detectors with different gap widths. The depicted detector signals are calculated analytically by integrating a shifted Gaussian intensity profile over the intervals corresponding to the active areas of the detector. Since isotropic scaling does not alter the deflection signal, all geometric lengths are standardized to the spot radius w.

With a given displacement of the laser beam, the ratio of the photo currents gets closer to unity if the gap width between the segments is reduced (cf. fig. 4.25). As a consequence, the slope of the detector response function decreases which, in turn, leads to a larger measurement range. As it can be seen from fig. (4.24), a gap width $g/w = 0.05$ and the optimum geometry of a vanishing gap size result in nearly the same response function. Hence, we can employ a commercially available four–quadrant photo diode whose gap width is 1/20 of the segment size or below.

Dependence on the detector segment size

A similar tendency is observed for different segment sizes (cf. fig. 4.26). Since shrinking the detector segments improves the balance of the photo currents (cf. fig. 4.27), a smaller detector facilitates a larger measurement range.

Although reducing the detector size also favors the time resolution of the detector system, there are some practical limitations: First, if the displacement is larger than the detector size, the photo currents are induced by the tails of the field distribution. Their intensity decreases exponentially with increasing beam displacement, thus significantly raising the noise level. And second, the response of many available photo detectors is disturbed if the borders of the active area are significantly illuminated. Therefore, the spot radius and the beam displacement should not exceed the detector size. On the other hand, it can be seen from fig. (4.26) that the slope of the response function decreases by just a factor of 3 if the detector segment size is reduced from $0.5w$ to $0.2w$. For that reason, we must conclude that a feasible shrinkage of the detector geometry enlarges the

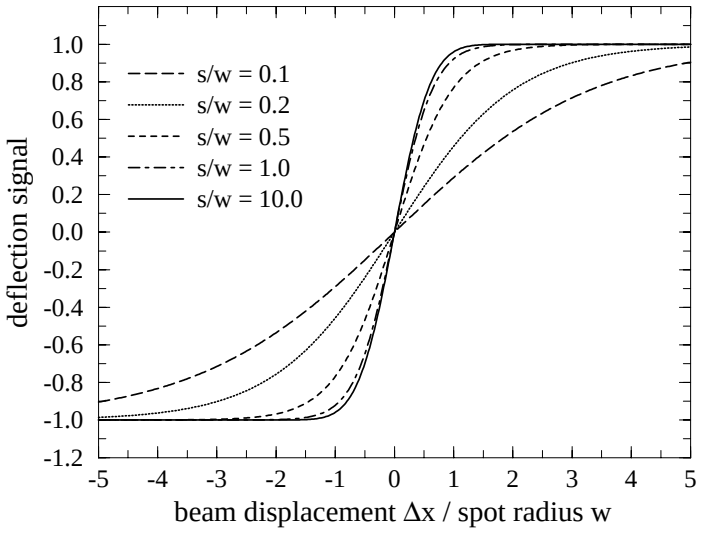


Figure 4.26: Detector response for different detector segment sizes s (standardized gap width $g/w = 0.05$).

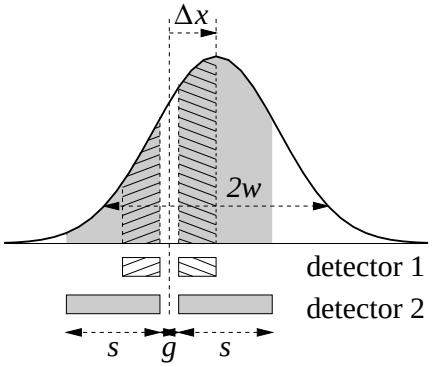


Figure 4.27: Four–quadrant diodes with a small segment size (detector 1) and a large segment size (detector 2). The photo currents are proportional to the highlighted areas.

measurement range by approximately a factor of 2 but does not facilitate the desired extension by at least one order of magnitude.

Evaluation of the differential signal

As another attempt to increase the measurement range, one might think of different ways of signal processing. For example, we can omit the scaling of the differential signal to the total detected intensity and simply regard the difference of the photo currents $I_1 - I_2$ of two opposite detector segments. The corresponding detector response functions are plotted in fig. (4.28).

Moving the laser beam across the detector, the photo current of one segment increases and decreases, as the spot enters and leaves the segment area, respectively. In case of small detectors, the ascending and descending wings overlap, while a sufficiently large segment size leads to a period of constant photo current (cf. fig. 4.28). If the gap width exceeds the spot diameter, the response function exhibits two peaks which are opposite in sign and arise as the laser beam sweeps across one of the detector segments.

It should be noted that contrary to the standardized deflection signal, the slope of this response function is independent of the detector geometry (cf. fig. 4.28). Therefore, the range of linearity cannot be extended. It would rather be necessary to employ a position sensor comprising an array of single segments. Although it would thus be possible to detect large beam displacements, the spatial resolution will be seriously affected by the large internal beam deflection (cf. section 4.3.7).

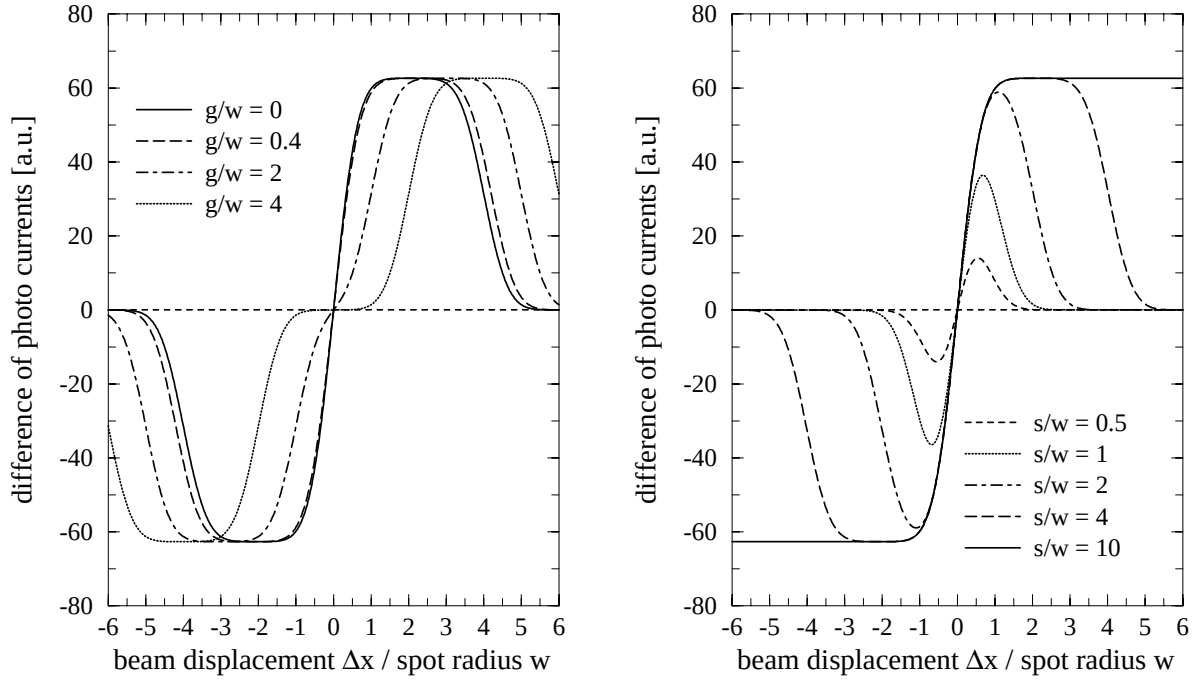


Figure 4.28: Detector reponse with respect to the difference signal (i. e. without scaling to the sum of the photo currents) for different gap widths (left; $s/w = 4$) and for different segment sizes (right; $g/w = 0.05$).

4.3.6 Modification of the image formation

As mentioned above (cf. last paragraph of section 4.3.4), the image formation by the optical setup is characterized by the parameter $b = F - fd_2/(d_2 - f)$. With a given sample length L, it can be replaced by $a = 2b + L/n_{Si}$. As it is shown in appendix (C), increasing a will increase the total displacement as well as the spot radius on the detector. In case of a large a, however, the standardized displacement does become independent of a but the detector sensitivity decreases as the spot radius increases (cf. section 4.3.5). As a consequence, enhancing the parameter $a = 2b + L/n_{Si}$ reduces the deflection signal (cf. fig. 4.29).

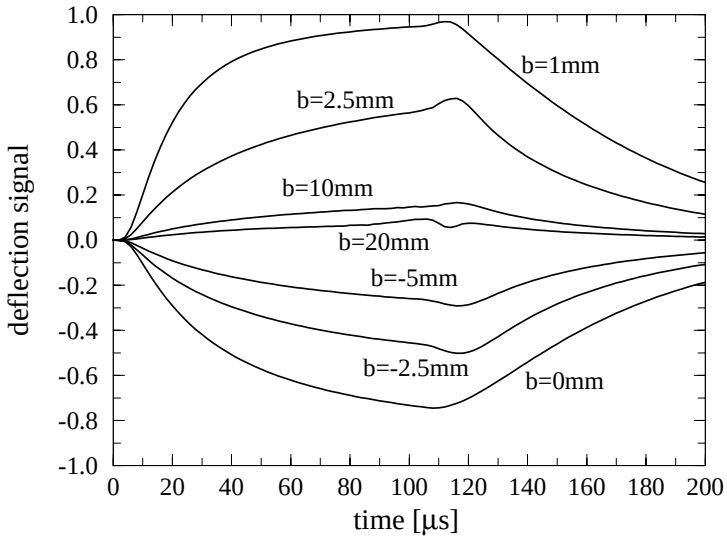


Figure 4.29: Deflection signal in case of different image formation properties. The optical setup is characterized by the parameter $b = F - fd_2/(d_2 - f)$.

Although this effect might be exploited to extend the measurement range, we should remember that increasing b enlarges the spot diameter on the detector. This is equivalent to a reduction of the detector size and thus introduces the same difficulties mentioned in section (4.3.5).

4.3.7 Deflection measurements in case of large power dissipation

As it has been demonstrated above, a modification of the optical setup facilitates a reduction of the sensitivity and thus an extension of the measurement range by at most a factor of 2. Therefore, a suitable preparation of specific samples is the only way to enable the detection of large temperature gradients. Within this section we will discuss the design of test structures to overcome the difficulties originating from a large power dissipation in the interior of the investigated device.

As a representative model problem, we consider the following situation: A block of silicon with a wafer thickness of $200\ \mu\mathrm{m}$ is heated by a constant heat flow of $j_Q = 2 \cdot 10^4\ \frac{\mathrm{W}}{\mathrm{cm}^2}$ through a thermal contact at its top surface during the period $0 < t < 100\,\mu\mathrm{s}$. The bottom surface is connected to a heat sink of 300 K via a thermal resistor of $1\,\mathrm{K\,cm^2/W}$. The evolution of the temperature distribution (cf. fig. 4.30) is calculated by a numerical solution of the heat flow equation. Though quite simple, this model reflects the typical situation of power devices within which most of the heat is dissipated in a small layer close to the top surface. For example, short circuit operation of IGBTs with a voltage drop of 200 V at a current density of $100\ \mathrm{A/cm^2}$ results in a heat generation

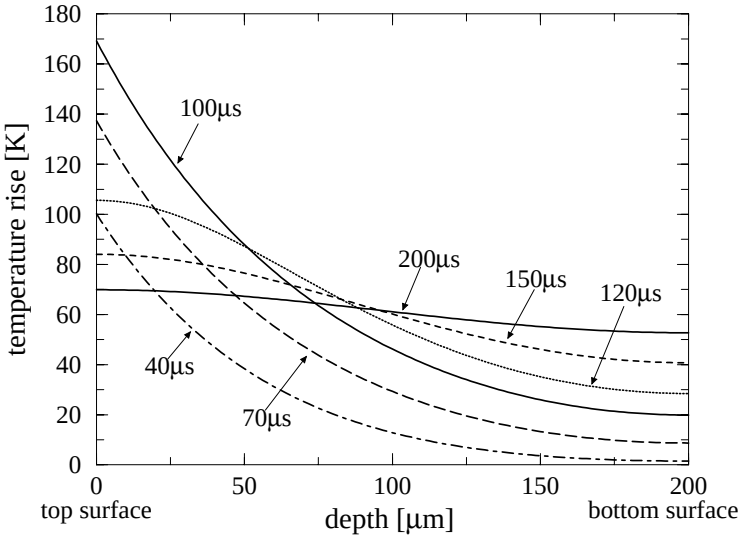


Figure 4.30: Temperature evolution in a silicon block heated by a constant heat flow through the top surface ($j_Q = 2 \cdot 10^4 \frac{\mathrm{W}}{\mathrm{cm}^2}$, pulse duration $100\ \mu\mathrm{s}$).

rate of $2 \cdot 10^4 \frac{\mathrm{W}}{\mathrm{cm}^2}$ per square area, which is almost completely dissipated in the MOS channels and the space charge region at the top of the device.

The crucial problem is evident from fig. (4.31). Even within a sample of only $800\ \mu\mathrm{m}$ in length, the temperature gradients give rise to a beam displacement of some tens of microns. For instance, a laser beam entering at a depth of $40\ \mu\mathrm{m}$ is shifted by $20\ \mu\mathrm{m}$ after $100\ \mu\mathrm{s}$. Apart from the difficulties in detecting this deflection, the spatial resolution will seriously suffer from such a large displacement since the measurement signal will be sensitive to the average temperature gradient between $x = 20\ \mu\mathrm{m}$ and $x = 40\ \mu\mathrm{m}$.

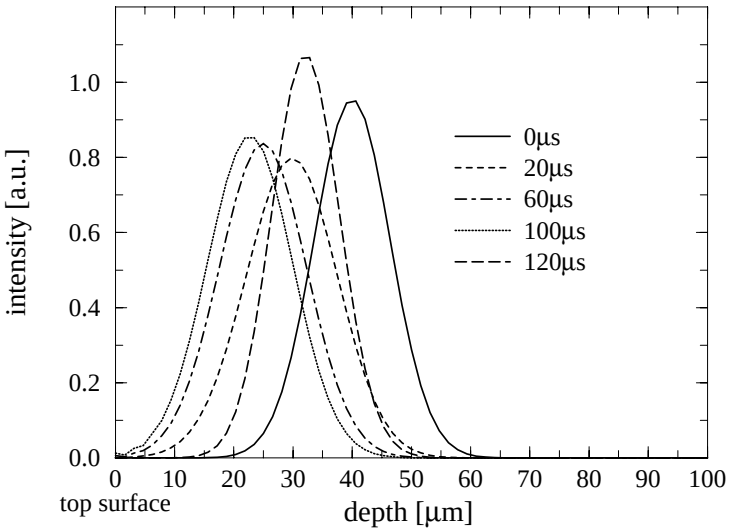


Figure 4.31: Intensity distribution at the rear surface of the device (heat flow of $2 \cdot 10^4\,\mathrm{W/cm^2}$ into the top contact, pulse duration $100\ \mu\mathrm{s}$). The probing beam enters at a depth of $40\ \mu\mathrm{m}$.

Test structures

Reducing the interaction length L will not only improve the spatial resolution but also decrease the deflection signal and thus facilitate a larger measurement range. However, the shrinking of the sample is limited by practical constraints since the handling of the device is only manageable for dies larger than ca. $500\ \mu\mathrm{m}$. In addition, the blocking capability will suffer from removing the edge

termination. For that reason, we will investigate test structures comprising an active area of only some $100\,\mu\mathrm{m}$ in size which is embedded in a larger silicon substrate (cf. fig. 4.32). Compared to product devices ($L_a \approx L_s$), the total deflection of the probing beam is smaller due to the smaller length L_a of the active area, but a significant error is expected due to the current and heat spreading, resulting in a two–dimensional carrier and temperature distribution (cf. fig. 4.32).

Internal Laser Deflection measurements on test structures with different lengths of the active area are investigated by performing virtual experiments as described in section (4.1). Assuming

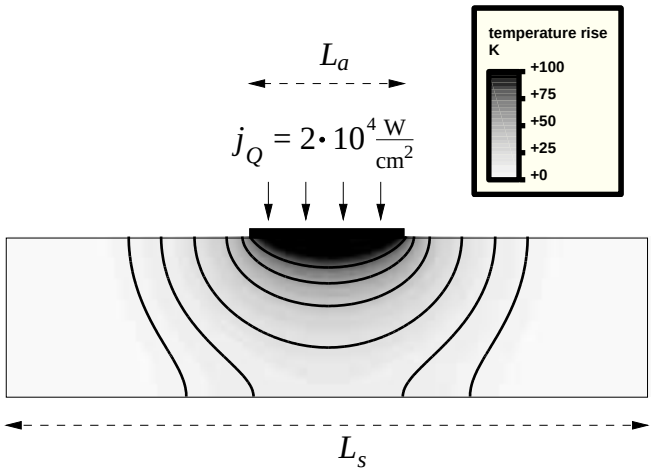


Figure 4.32: Temperature distribution in the test structure at $t = 100\,\mu\mathrm{s}$ (sample length $L_s = 800\,\mu\mathrm{m}$, length of the active area $L_a = 100\,\mu\mathrm{m}$). The sample is heated by a constant heat flow j_Q through the thermal contact at the top surface.

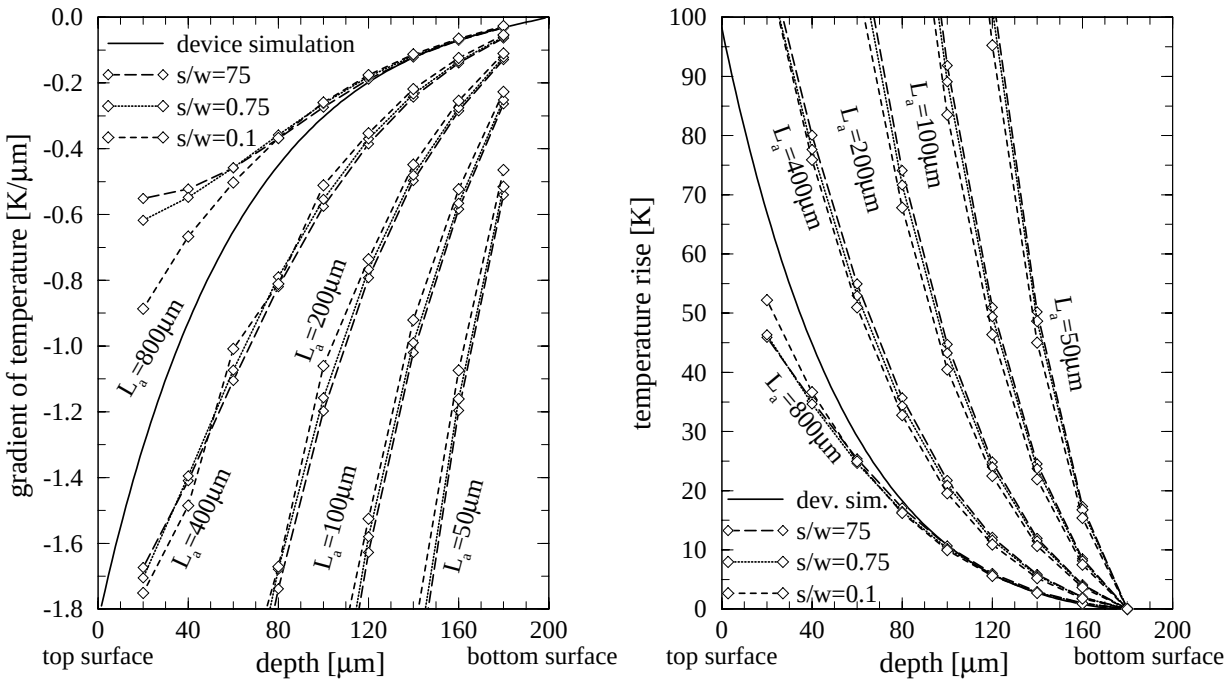


Figure 4.33: Temperature distribution at $t = 100\,\mu\mathrm{s}$ extracted from the simulated deflection signal on test structures (sample length $L_s = 800\,\mu\mathrm{m}$, length of active area L_a). Different detector geometries are employed (gap width $g = 0$, segment size/spot radius s/w as indicated in the figure).

an interaction length $L = L_a$, the temperature gradients are extracted according to the evaluation rule (4.29) and compared to their reference values (cf. fig. 4.33). The latter are obtained by a *one–dimensional* solution of the heat flow equation since product devices are much larger in size and therefore do not exhibit the mentioned heat spreading effects. For further illustration, the absolute temperature profiles are also included in the figure. Each integration constant is set to zero and, for the sake of comparability, the reference temperature profile is shifted so that the temperature rise at the bottom vanishes.

Two tendencies are clearly visible: First, the saturation of the response function results in a maximum detectable temperature gradient which is larger in case of a smaller detector geometry (cf. section 4.3.5) or a smaller length of the active area. Second, simply inserting the length L_a of the active area as the interaction length L leads to a significant overestimation of the temperature gradient. This is especially due to the first term in the denominator of eq. (4.29), representing the quadratic dependence on L. Hence, the deviations are larger as the absolute value of $F - V_t d_2$ is smaller in view of $L/2n_{Si}$, i. e. the image of the device rear surface is located closer to the detector plane. Note that $L_a = 800\,\mu\mathrm{m} = \mathrm{L_s}$ represents a one–dimensional structure which therefore does not exhibit two–dimensional heat diffusion.

Effective interaction lengths and the modification of the evaluation rule

Extracting the temperature gradients from the deflection signals measured on test structures obviously demands the introduction of an effective interaction length L_{eff} which reflects the effects of the two–dimensional heat spreading. To end up with a generally applicable evaluation strategy, a simple formula of the effective interaction length is desired which is independent of the regarded operating condition. Since the deflection signal exhibits a linear and a quadratic dependence on L (cf. eq. 4.29), the following *definitions* are motivated:

$$L_{\mathit{eff},lin}(x) \quad := \quad = \left(\frac{dT_{ref}}{dx}\right)^{-1} \cdot \int_0^L \frac{dT}{dx}(x,z)\,dz \tag{4.31}$$

$$L^2_{\mathit{eff},qu}(x) \quad := \quad = \left(\frac{dT_{ref}}{dx}\right)^{-1} \cdot \int_0^L \frac{dT}{dx}(x,z)\,z\,dz \tag{4.32}$$

Integrating the simulated temperature distributions in the interior of the test structures, these expressions are calculated and plotted in fig. (4.34). The diffusion of heat broadens the temperature profile along the beam path, but reduces the maximum temperature rise. These two effects nearly balance each other in case of the linear effective interaction length, which therefore almost fits the length L_a of the active area. On the other hand, the quadratic effective interaction length can be approximated by the empirical expression $\sqrt{L_s L_a}$. For that reason, the following evaluation rule is proposed

$$\boxed{\frac{dn_{Si}}{dx} = \left(\frac{\partial n_{Si}}{\partial T}\right)_{n,p} \frac{dT}{dx} + \left(\frac{\partial n_{Si}}{\partial C}\right)_{T} \frac{dn}{dx} = \frac{M \cdot V_t/\tilde{Q}(w)}{L_a\left(\frac{L_s}{2n_{Si}} + F - V_t d_2\right)} + O(M^3)} \tag{4.33}$$

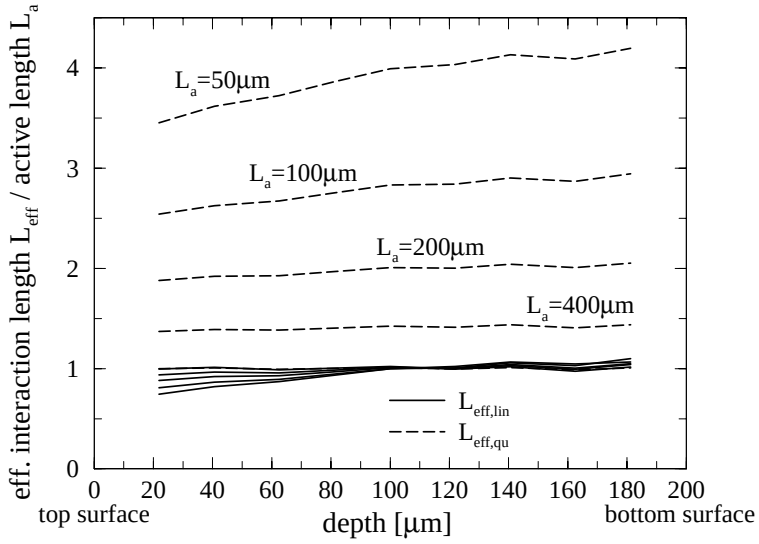


Figure 4.34: Linear and quadratic effective interaction lengths $L_{eff,lin}$ and $L_{eff,qu}$ (definitions in the text) evaluated at $t = 100\ \mu$s.

where $V_t = f/(d_2 - f)$ denotes the inverse of the transverse magnification (cf. appendix C). Usually, $L_s/2n_{Si}$ amounts to a few 100 microns while the other terms in the parenthesis are in the order of some cm. Therefore, the most significant effect arises from the linear dependence on L_a, unless the image of the device rear surface is located on the detector plane. In this case, also the quadratic dependence on $L_a L_s$ is pronounced.

Evaluating the virtual experiments according to equation (4.33) yields much more accurate results (cf. fig. 4.35). As it can be seen from the figure, the extracted temperature distributions match

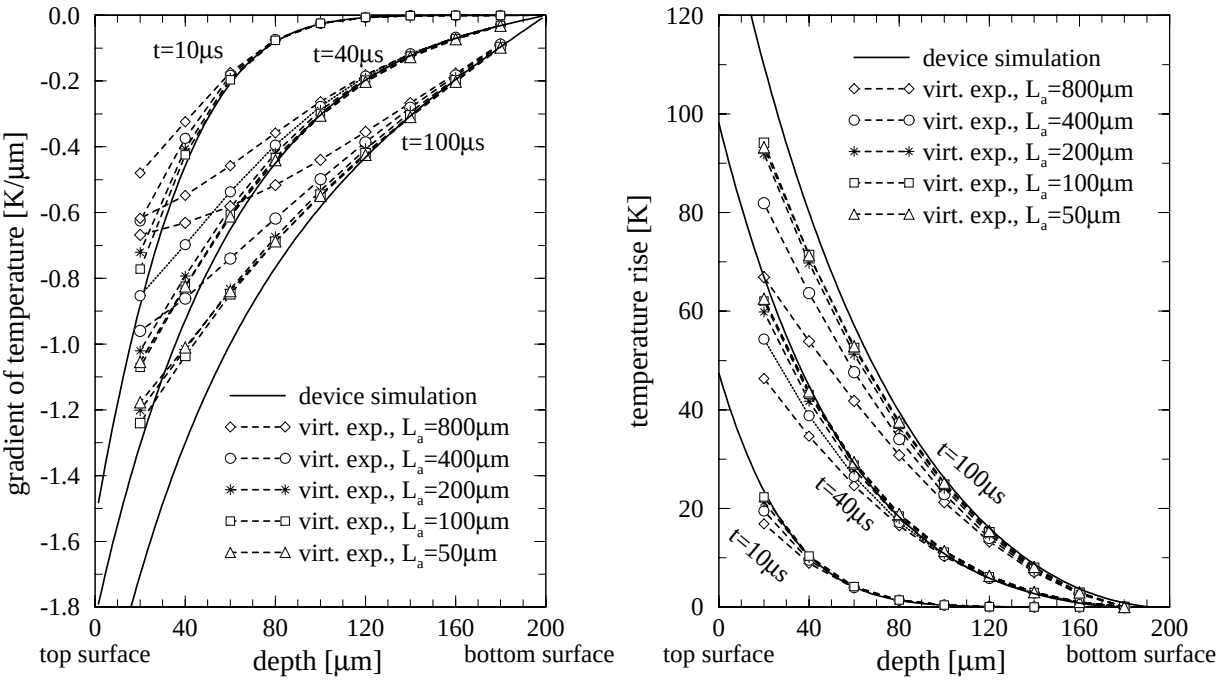


Figure 4.35: Temperature distribution extracted from the simulated deflection signal on test structures according to eq. (4.33) (sample length $L_s = 800\ \mu$m, length of active area L_a, detector segment size/spot radius $s/w = 0.75$, detector gap width $g = 0$).

the reference temperature profiles with satisfying accuracy up to a specific gradient. Decreasing the active length increases the maximum detectable temperature gradient. $L_a \approx 200\,\mu\mathrm{m}$ turns out to be an optimum geometry, since no additional improvement is attainable by a further reduction of L_a. In this case, the measurement range extends up to approximately $0.6\,\mathrm{K}/\mu\mathrm{m}$. Since the deviations in the absolute temperature profiles are even less significant, typical temperature rises up to about 60 K can be detected.

However, it is important to remember the assumptions the presented formula is based on:

First, the *total* internal displacement and the angular deflection of the probing beam have to be small enough so that the paraxial approximation remains valid. In case of the test structures with a small active area, this restriction even holds during operations with large power dissipation.

Second, the *gradient* of the refractive index is assumed to be constant within the area covered by the probing beam. On account of the rather small lateral extension of the probing beam (cf. section 4.2.1), this precondition is satisfied with reasonable accuracy.

Third, the proper effective interaction lengths have to be inserted. In particular, L_a constitutes the length of the active area, i. e. that region of the structure within which the electric power is dissipated. In addition, L_s represents the total substrate length. However, in the interior of very large silicon dies, the outer regions are not affected by the heat dissipation within the active area. It is obvious that in this case, the sample length L_s has to replaced by the width which is covered by the heat diffusion during the duty cycle.

Fourth, since this study has been carried out with a fixed heat dissipation rate, we implicitly assume that during the device operation the thermal feedback on the electrical behavior is negligible. However, this restriction is violated, for example, during short circuit operations of IGBT samples comprising a very small number of cells. Since the saturation currents significantly depend on the channel temperature, the lower peak temperature within the smaller test structures results in a higher short circuit current which, in turn, increases the heat dissipation rate.

4.3.8 Summary

The most important results of this section are summarized by the following statements:

- The temperature distribution within power devices operating with low power dissipation can be detected by Internal Laser Deflection measurements with high sensitivity and excellent accuracy.

- Adding the photo currents of all segments of the four–quadrant detector yields the total transmitted intensity. Although laser deflection and free carrier absorption measurements can thus be performed simultaneously, a separate pin diode detector essentially improves the accuracy and the reliability of the absorption measurements.

- Thermal modulations of the optical sample thickness result in Fabry–Perot oscillations of the deflection signal, which have to be suppressed by depositing an antireflective coating or using an incoherent light source.

- To gain a quantitative evaluation rule of the deflection signal, the probing beam is considered a Gaussian beam whose optical axis is deflected in the sample according to the laws of geo-

metrical ray tracing. The field distribution at the rear side is projected by the imaging lens, resulting in a Gaussian field distribution on the detector. The deflection signal is affected by three parameters, namely the interaction length within the sample, the image formation properties of the imaging lens, and the detector response function.

- The saturation of the detector response function constitutes the limiting effect of the measurement range, which can be extended by reducing the detector size. However, an increase by at most a factor of 2 is attainable on account of practical restrictions.

- The image formation by the imaging lens is characterized by two parameters, namely the transverse magnification and the distance of the device rear surface to the image of the detector plane. While the former affects the deflection signal in the same way as a corresponding scaling of the detector geometry, the increase of the latter enlarges the spot diameter on the detector. Thus, an extension of the measurement range by a factor of 2 can be achieved, unless the ratio of the spot diameter and the detector size must not be increased due to the same practical constraint as in the previous item.

- The most effective way to increase the measurement range is to reduce the interaction length L. Probing test structures which comprise a small active area embedded in a larger silicon die requires the definition of effective interaction lengths which reflect the effects of the additional lateral heat diffusion. Thus, laser deflection measurements facilitate the detection of temperature gradients up to about $0.6\,\mathrm{K}/\mu\mathrm{m}$, which typically correspond to temperature rises of about 60 K if the heat sources are located near the top surface.

Chapter 5

Numerical Simulation of Interferometric Techniques

In this chapter, we will investigate laser probing techniques which are based on interferometry and therefore promise several important advantages: First, the periodically oscillating detector signal lacks of saturation effects and thus enables a large measurement range, which is merely limited by parasitic effects. Second, the desired information is extracted by exploiting the knowledge that the optical path length of the probing beam is modulated by one wavelength as two adjacent intensity maxima are observed on the detector. Consequently, there is no need for calibrating the optical setup, such as the magnification of the imaging system or the sensitivity of the detector, for example. Third, as interferometric techniques are sensitive to the optical sample thickness they do not detect the gradients but directly reveal the absolute value of the refractive index. However, the oscillating signal reflects the temperature *evolution* and thus requires an integration in time space, but the initial temperature distribution during transient switching is usually known. In particular, it is homogeneous and equal to the ambient temperature in case of the operating conditions investigated below.

On the other hand, a major drawback is common to all interferometric techniques: The oscillating signal does not allow to discriminate whether it originates from an ascending or a descending temperature evolution. Consequently, it could be difficult to identify the temperature maximum during a transient process.

5.1 Backside Laser Probing and Differential Backside Laser Probing

As a first example, we will discuss Backside Laser Probing, which has been successfully employed for investigating a variety of devices, e. g. power MOSFETs [86], smart power devices [99], MEMS [127], and ESD protection devices [94].

Contrary to the measurement techniques discussed in the previous chapter, Backside Laser Probing operates by means of a vertically propagating laser beam. Thus, only integral information on the carrier concentration and the temperature profile in vertical power devices is provided.

However, since the beam path is *parallel* to the gradient of the refractive index, the probing beam is not subject to a lateral deflection, which will be shown to limit the measurement range of scanning techniques employing a laterally propagating laser beam.

To provide access for the probing beam penetrating the device from the rear side a window has to be etched in the bottom metallization layer. Its effect on the device performance is one of the most interesting questions and will be addressed in the following section.

5.1.1 Effects of the sample preparation

To investigate the effects of etching a window in the collector contact simulations of multiple IGBT cells have been performed. As it can be seen from fig. (5.1), current crowding at the contact edges results in an enhanced carrier concentration at the bottom of the device. This phenomenon is typical of potential driven particle currents and is also observed, for example, as the water level rises in front of the piers of a bridge in the river. For a thorough quantitative analysis, lateral cuts in various depths of the structure are plotted in fig. (5.2). Compared with an unprepared device, the carrier density at the edges of the window is almost twice as large. This excess concentration

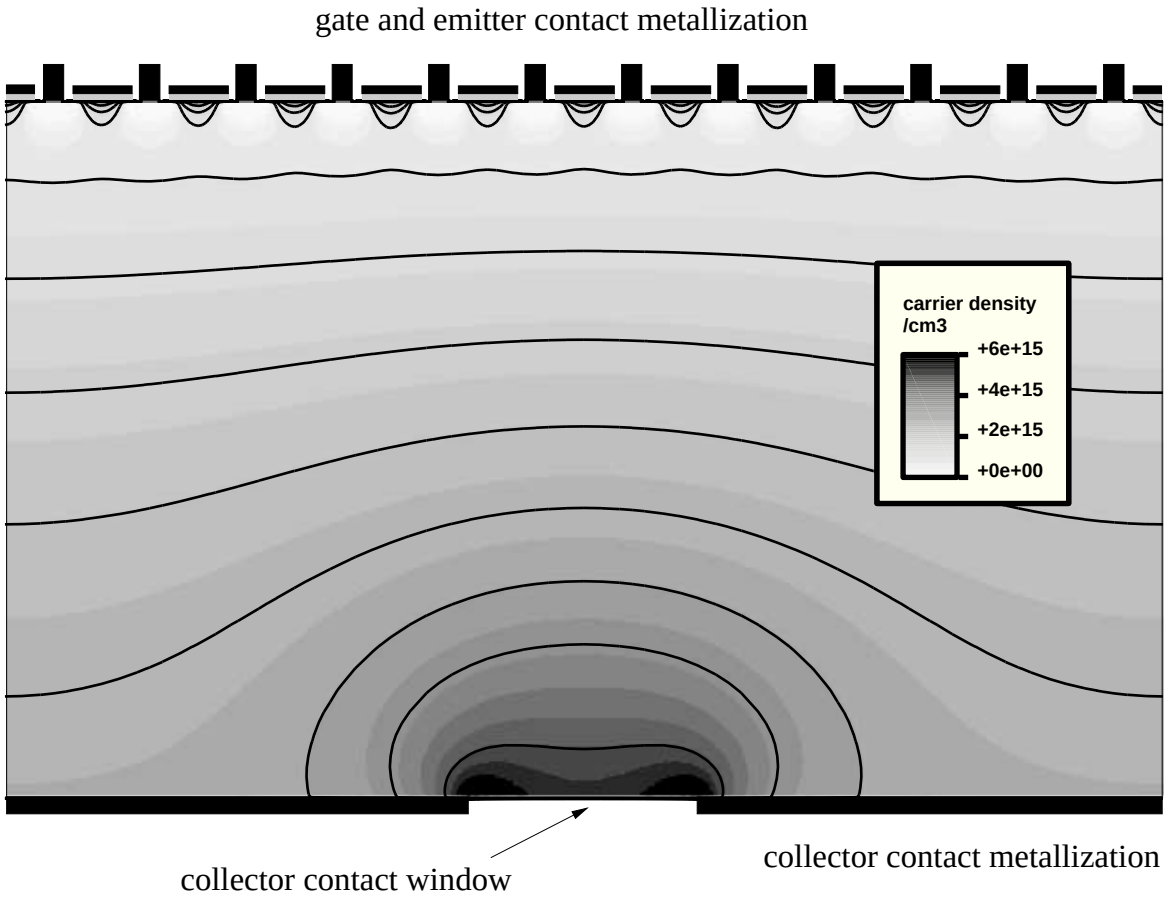
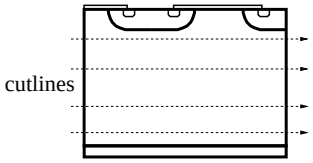


Figure 5.1: Carrier distribution in the interior of an IGBT sample with a window in the collector contact metallization layer (current density $25\ \mathrm{A/cm^2}$).



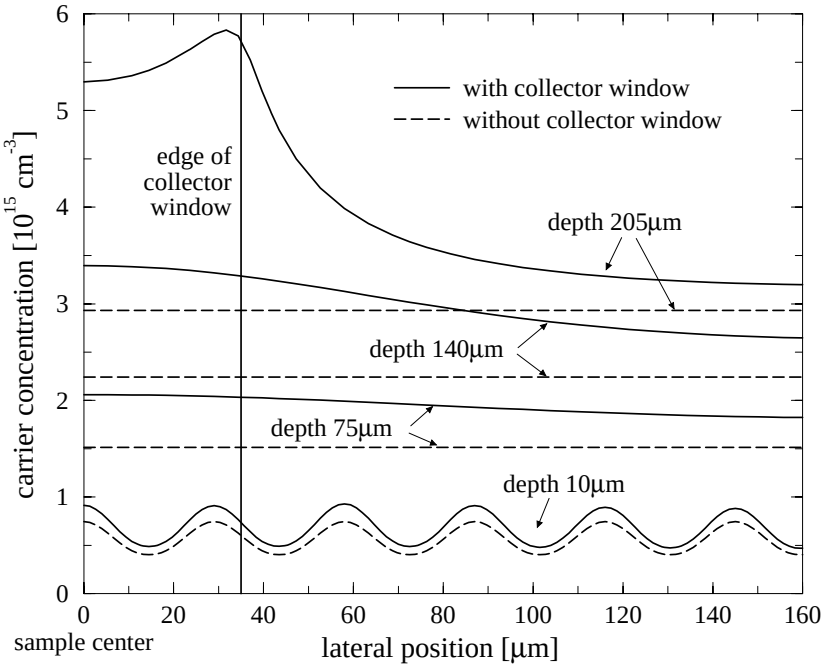


Figure 5.2: Carrier distribution in various depths within an IGBT sample prepared for Backside Laser Probing (current density $25\,\mathrm{A/cm^2}$). Due to symmetry, only the right hand half of the sample is shown.

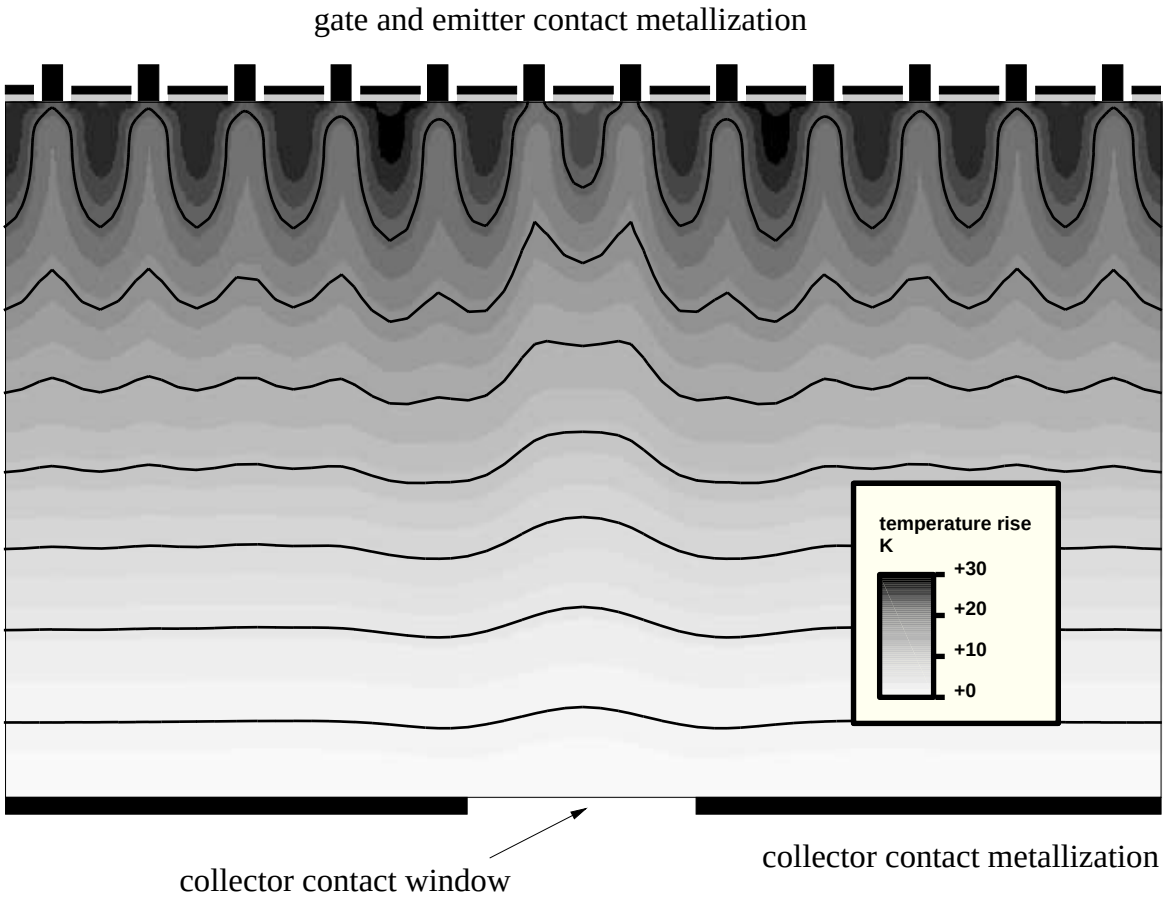
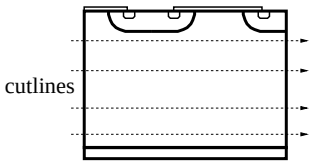


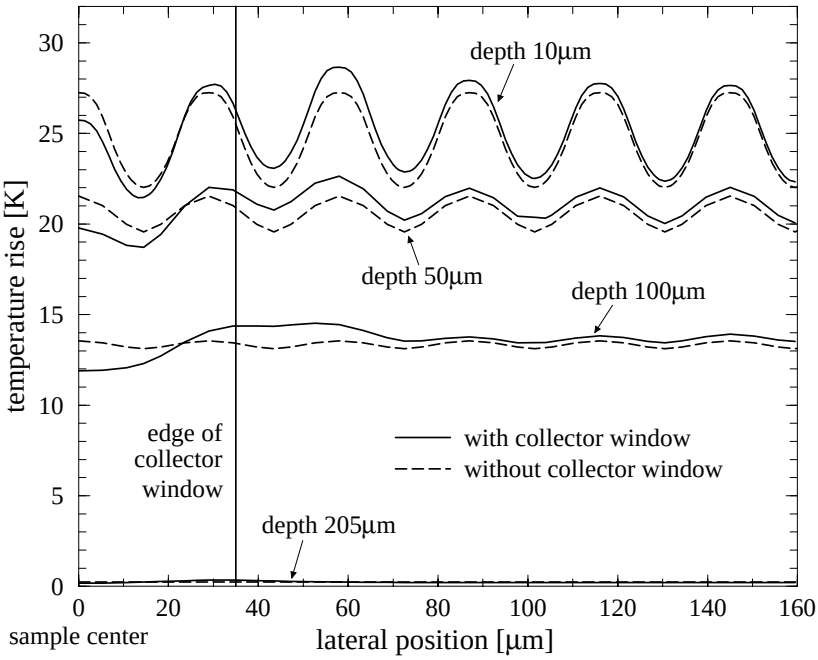
Figure 5.3: Temperature distribution in the interior of an IGBT sample with a window in the collector contact metallization layer (short circuit operation with 500 V collector–emitter voltage, ($t = 3\ \mu\mathrm{s}$).

becomes smaller as the distance to the rear surface increases, and it vanishes at the top surface. The integral of the carrier distribution along a vertical cutline — this is the quantity affecting the phase shift signal — is therefore raised by about 35 % due to the etching of the window.

As a consequence of the enhanced current density at the window edges, a local temperature increase is observed (cf. fig. 5.3). However, the profiles along lateral cutlines in various depths (cf. fig. 5.4) clearly reveal that this effect amounts to only some percent and is therefore negligible.

Figure 5.4: Temperature distribution in various depths within an IGBT sample prepared for Backside Laser Probing (short circuit operation with 500 V collector–emitter voltage, $t = 3\,\mu\mathrm{s}$). Due to symmetry, only the right hand half of the sample is shown.





In summary, etching a window in the rear metallization layer gives rise to current crowding at the contact edges. The resulting local increase of the carrier concentration enhances the carrier contribution to the phase shift signal by about 35 % whereas the modulations of the temperature distribution can be neglected. Since the latter is by far the most significant effect in samples with a high power dissipation as, e. g., IGBTs operating under short circuit conditions, the preparation effects need not to be taken into account for the evaluation of the measurement results in this case.

5.1.2 The measurement signal

The Backside Laser Probing technique employs a laser beam with an angular aperture of about 0.1 to 0.4. The resulting optical field distribution (cf. fig. 5.5) displays a minimum spot diameter of about $8\,\mu\mathrm{m}$ at the top surface. Due to the small interaction length of $2 \cdot 200\,\mu\mathrm{m}$, beam spreading increases the spot diameter to only $15\,\mu\mathrm{m}$ at the bottom surface, which promises an excellent spatial resolution.

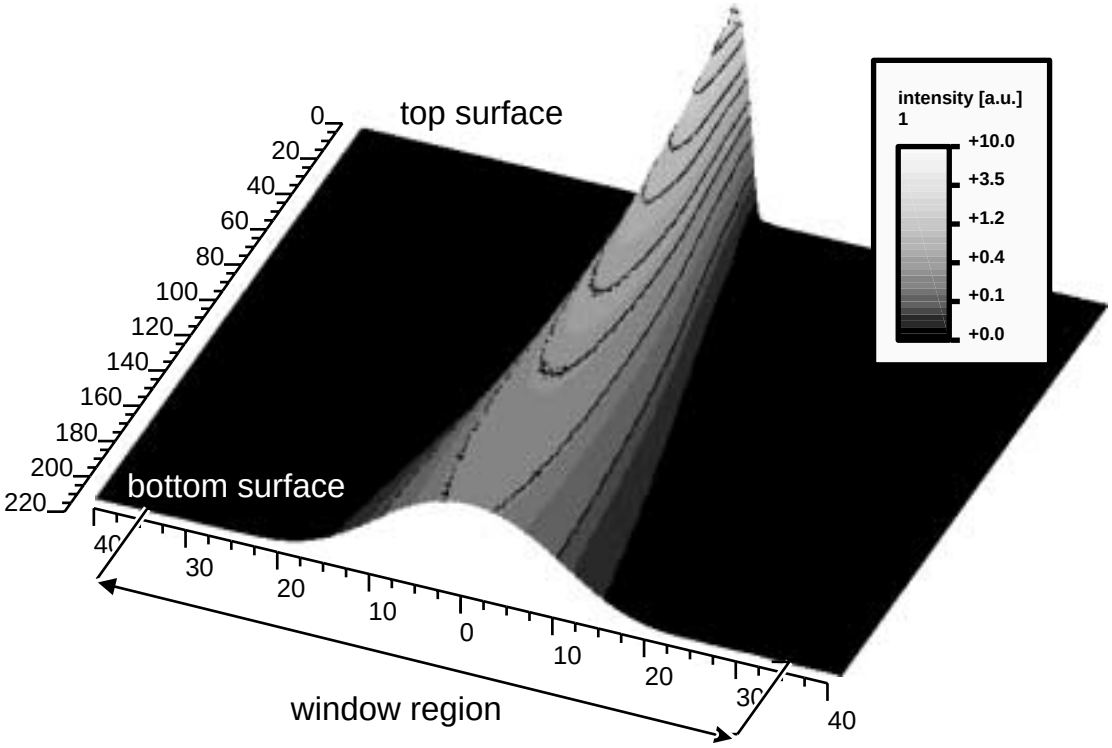


Figure 5.5: Optical field distribution in the interior of the investigated sample ($\eta_A = 0.25$).

Extracting the phase shift signal

As described in section (2.2.3), the positions of the probing beam and the reference beam are controlled by the driving frequencies of the acousto–optic modulator, ω_1 and ω_2, respectively. Since the diffraction within the modulator shifts the frequencies of the laser beams by ω_1 and ω_2, respectively, their electric fields on the detector are represented by [58]

$$\vec{E}_{pr}(t) = \vec{E}_{pr,0}\exp[i(\omega+2\omega_1)t + i\Delta\varphi(t)]; \quad \vec{E}_{ref}(t) = \vec{E}_{ref,0}\exp[i(\omega+2\omega_2)t] \tag{5.1}$$

where ω denotes the frequency of the original laser beam. Consequently, the following intensity signal is observed:

$$I_D(t) = |\vec{E}_{pr,0}|^2 + |\vec{E}_{ref,0}|^2 + 2\vec{E}_{pr,0}\vec{E}_{ref,0}\cos(2\Delta\omega t + \Delta\varphi(t)] \quad \text{with} \quad \Delta\omega = \omega_1 - \omega_2 \tag{5.2}$$

The desired phase shift $\Delta\varphi(t)$ can be extracted from the detected intensity signal $I_D(t)$ by the following ways (cf. appendix D): The unknown coefficients $|\vec{E}_{pr,0}|^2 + |\vec{E}_{ref,0}|^2$ and $2\vec{E}_{pr,0}\vec{E}_{ref,0}$ are determined from the extrema of the oscillations, thus facilitating a direct solution of equation (5.2). An alternative strategy is based on Fourier transformations and a suitable filtering in frequency space to eliminate the coefficients of equation (5.2) (cf. app. D.2).

As an example, the intensity signals calculated from virtual experiments with different angular apertures and the corresponding phase shift signals are shown in fig. (5.6). Without loss of generality, the difference $\Delta\omega$ of the driving frequencies can be set to zero for the theoretical investigations.

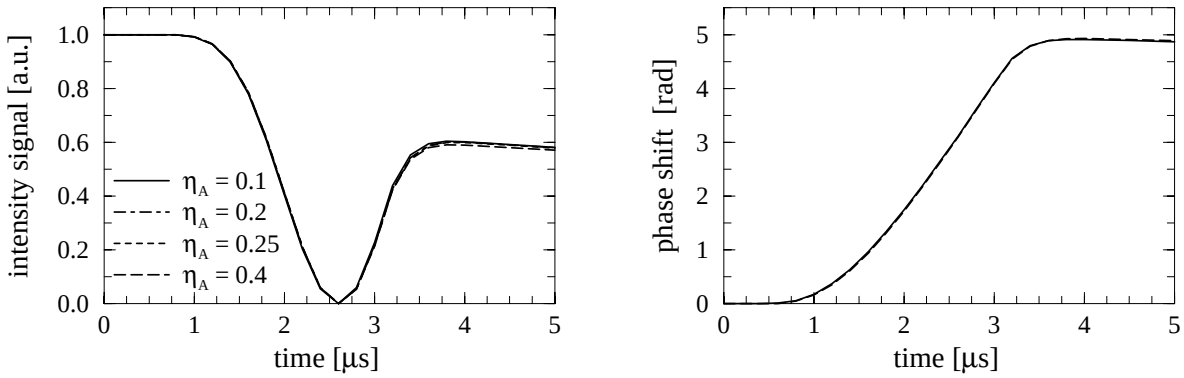


Figure 5.6: Short circuit operation of an IGBT sample (500 V collector–emitter voltage, pulse duration $3\ \mu\mathrm{s}$*): The left hand figure depicts the calculated detector signal with* $\Delta\omega = 0$ *while the right hand figure compares the phase shift signal extracted from the intensity modulation (solid line) with the integrated temperature profile (broken line).*

As it can be seen from the figure, employing different angular apertures results in nearly the same intensity signal. The major reason is that due to the rather small interaction length of $2 \cdot 200\ \mu\mathrm{m}$ the probing beam does not spread out significantly, even in case of a high angular aperture.

Interpretation of the phase shift signal

In case of a low power dissipation within the sample, the phase shift signal (cf. eq. 2.20) exhibits two contributions arising from transient modulations of the carrier concentration and the temperature distribution (cf. fig. 2.8). They are opposite in sign, but of the same order of magnitude [91]. It should be mentioned that on account of the window in the collector contact, the carrier contribution is about 35 % larger than that of an unprepared sample would be.

The capabilities of Backside Laser Probing appear when they are employed to investigate devices operating with large power dissipation. In this case, the carrier contribution to the phase shift signal is negligible in view of the thermal contribution. As it can be seen from fig. (5.6, right) the integral of the temperature distribution along the vertical path of the probing beam is in excellent agreement with the phase shift extracted from the intensity modulations, i. e. the lateral spreading of the laser beam does not introduce a detectable error. Consequently, the phase shift signal is a direct measure of the heat stored within the device [91]:

$$\Delta\varphi(t) = \frac{4\pi}{\lambda c_t A_D}\frac{\partial n_{Si}}{\partial T}\int_0^L c_t A_D \Delta T(x,t)\,dx = \frac{4\pi}{\lambda c_t A_D}\frac{\partial n_{Si}}{\partial T}\Delta Q(t) \tag{5.3}$$

Thus, although it provides integral information only, Backside Laser Probing constitutes a powerful characterization method for power devices since if offers a large measurement range as well.

Further information is gained if the temperature profile in an IGBT operating under short circuit conditions (cf. fig. 5.7, left) is approximated by a linear profile with vanishing temperature rise at the bottom. With this assumption, the integral in equation (5.3) can be expressed in terms of the

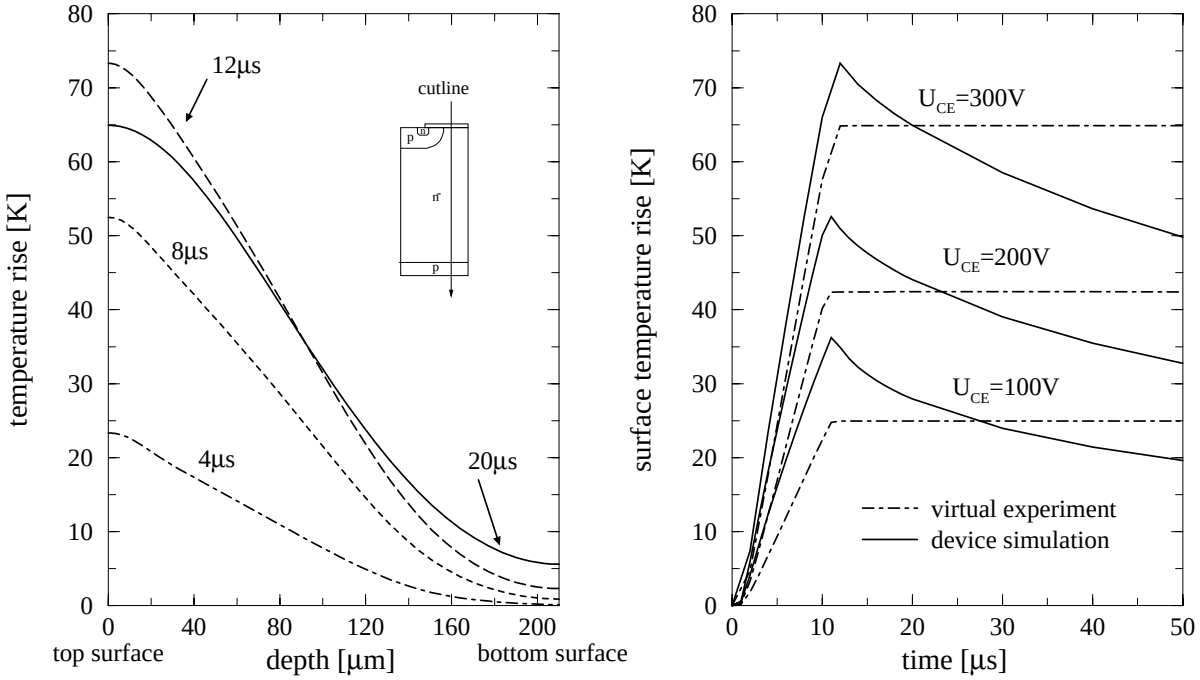


Figure 5.7: IGBT sample operating under short circuit conditions (pulse duration $10\ \mu$s). The left hand figure displays the temperature evolution at a collector–emitter voltage of 300 V. The right hand figure compares the surface temperature extracted from the phase shift signal according to eq. (5.4) with the result of an electrothermal device simulation.

surface temperature rise $\Delta T(z=0,t)$:

$$\Delta\varphi(t) = \frac{2\pi L}{\lambda}\,\frac{\partial n_{Si}}{\partial T}\Delta T(x=0,t) \tag{5.4}$$

The phase shift signal can thus be interpreted as a measure of the surface temperature. However, a comparison with the result of an electrothermal device simulation reveals that this simple consideration is only valid with an error of about 50 % (cf. fig. 5.4, right). Nevertheless, it clearly demonstrates that Backside Laser Probing also opens a way to get an approximate idea of the temperature rise in the hot spot of the investigated structure.

5.1.3 Differential Backside Laser Probing

A slight modification of the optical setup, namely Differential Backside Laser Probing [95], operates with both beams positioned in the window region. Thus, lateral temperature inhomogeneities can be directly extracted from the measurement signal.

To understand the shape of the differential phase shift signal on IGBTs, we have to examine which effects govern the temperature integrals along the paths of the two probing beams. On the one hand, the channel region and the space charge region at the reverse biased p–well of the IGBT constitute an area with a large heat generation rate. On the other hand, etching the window in the collector metallization results in current filaments at the edges of the contact (cf. fig. 5.8). Though the heat dissipation rate in this region is comparably small, it essentially contributes to the

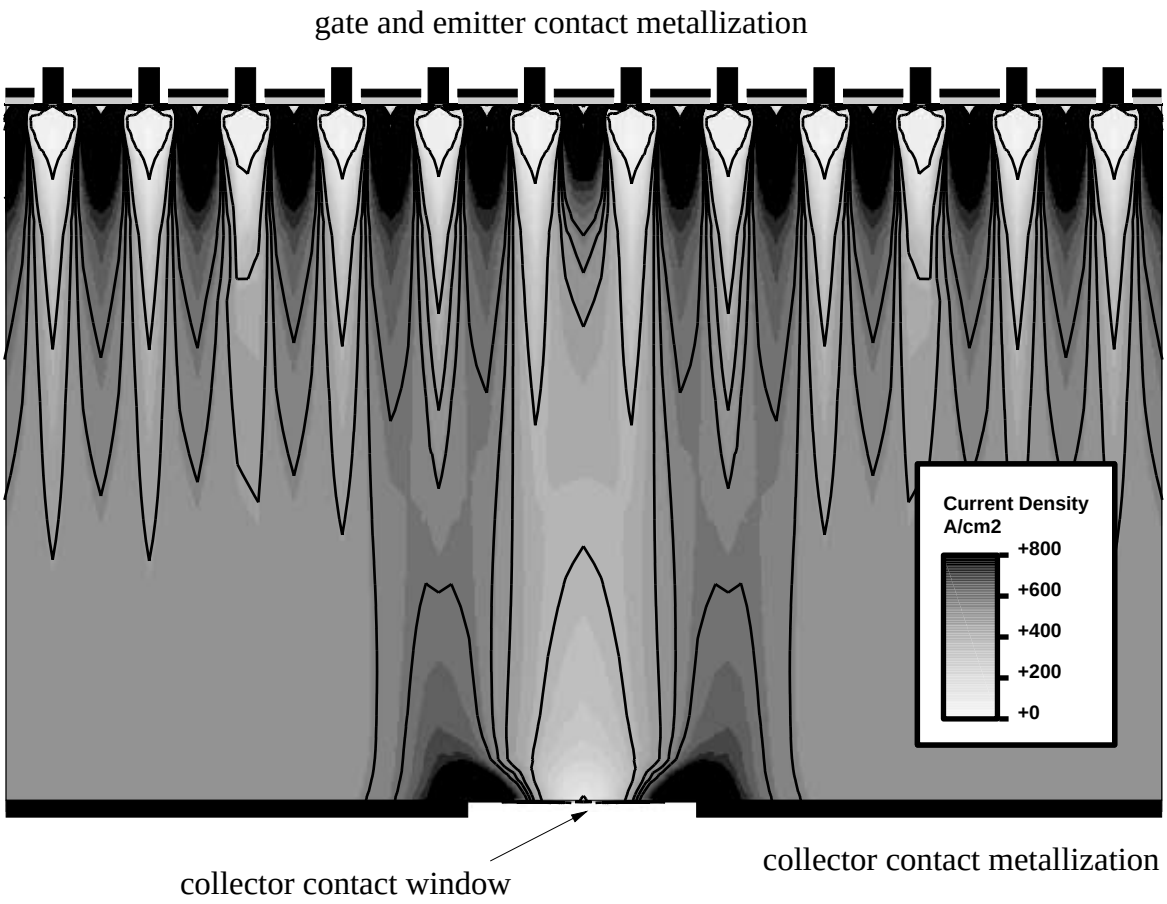


Figure 5.8: Current density distribution in an IGBT sample prepared for Backside Laser Probing (short circuit operation at 500 V collector–emitter voltage, $t = 3\ \mu\mathrm{s}$).

temperature integral due its rather large extension of more than $100\ \mu\mathrm{m}$. Consequently, the total absolute phase shift signal near the window edges is slightly larger than that in the center of the sample (cf. fig. 5.9).

Regarding the differential phase shift signal of two probing beams placed between two cells and at the edge of the p–well, respectively, (cf. fig. 5.10, left) we observe a positive peak during the current pulse. It is introduced by the probing beam penetrating the channel area and the space charge region at the reverse biased p–well, which exhibit a large power dissipation during the current pulse. After turn–off, however, the power dissipation in the current filaments at the window edges represents the most significant contribution. It is detected by the second probing beam and therefore results in a negative peak of the differential phase shift signal. Since the latter effect is introduced by the sample preparation the differential phase shift signal of an unprepared device would be completely different (cf. fig. 5.10, right): As electrothermal device simulation reveals, it comprises only one positive peak with two slight maxima due to the turn–on and turn–off losses, respectively. As thermal diffusion evens out the lateral temperature inhomogeneities arising during the current pulse, the differential signal drops to zero immediately after turn–off.

However, our major interest is focussed on the temperature distribution in the interior of unprepared samples. For that reason, the following analytical correction of the experimental signal is proposed to attain at least an approximate elimination of the preparation effect: As we know that the differential phase shift signal gained on unprepared samples would vanish after turn–off, a piecewise linear function is added to the signal measured on the prepared samples. This function is zero before turn–on, increases linearly during the current pulse up to the absolute value of the differential signal immediately before turn–off and decreases exponentially to zero afterwards. Hence, the corrected phase shift signal vanishes at the end of the current pulse and matches the fictitious phase shift signal of an unprepared sample with satisfying accuracy (cf. fig. 5.10, right). For further illustration, we have also included real measurement results in the figures. This demonstrates that despite the difficulties arising from the preparation effects, Differential Backside Laser Probing reveals useful information about the lateral temperature distribution. The shape of the experimental signals clearly confirms the predictions of the theoretical model, which emphasizes that simulating the measurement process constitutes a powerful strategy to support the interpretation of the measurement results.

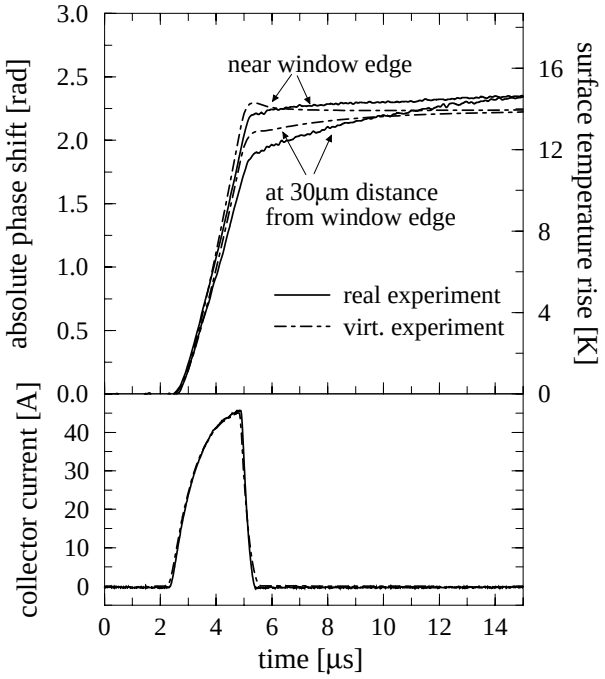


Figure 5.9: Absolute phase shift signal in the center and near the edge of the collector contact window.

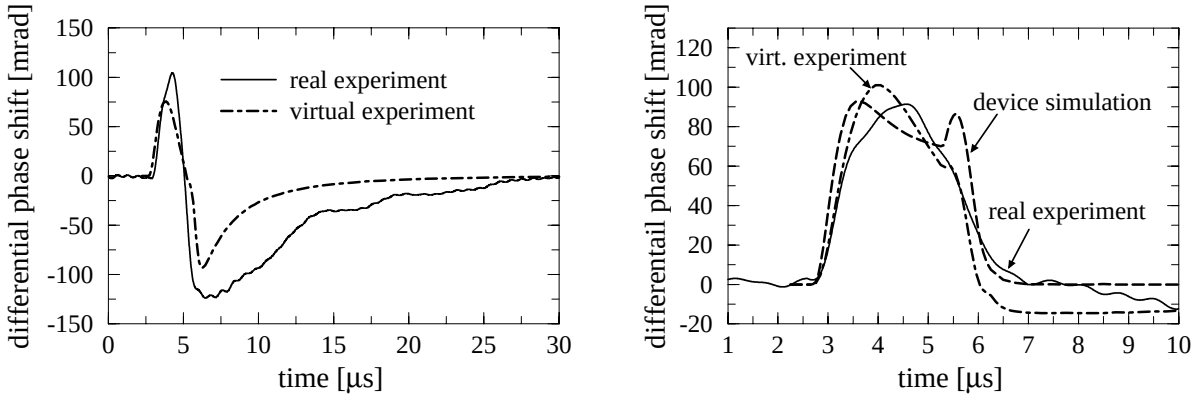


Figure 5.10: Differential phase shift signal of two probing beams positioned between two cells and at the edge of the p–well, respectively (Short circuit operation with 500 V collector–emitter voltage). Left: phase shift signals obtained on a prepared sample; right: The results of the real and the virtual experiment on prepared samples are corrected as mentioned in the text and compared to the phase shift obtained by an electrothermal device simulation of an unprepared sample.

5.1.4 Summary

Backside Laser Probing employs a vertically propagating laser beam and thus provides integral information about the carrier and temperature distribution. In particular, we can draw the following conclusions:

- Regardless of the image formation and the detector sensitivity, two adjacent interference extrema are observed if the phase of the probing beam is shifted by 2π. Consequently, the optical setup does not need to be calibrated.
- The necessary sample preparation (etching a window in the collector contact metallization) gives rise to current crowding at the contact edges. The enhanced carrier concentration affects the carrier contribution to the phase shift signal by about 35 %. The corresponding effect on the temperature distribution is negligible.
- Even in case of a high angular aperture the lateral convolution of the laser beam profile and the distribution of carrier concentration and temperature does not introduce a detectable error.
- In case of a large power dissipation, the phase shift signal is a measure for the heat stored in the device. Assuming a linear temperature profile, the top surface temperature can thus be extracted with an error of about 50 %.
- Positioning both laser beams in the window is the basic idea of Differential Backside Laser Probing. However, since the sample preparation alters the qualitative shape of the phase shift signal, an analytical correction of the measurement signal has to be applied.

As the gradients of the refractive index are parallel to the propagation direction of the probing beam, the latter is not deflected laterally. Consequently, Backside Laser Probing offers a very large measurement range and therefore constitutes a powerful technique to investigate the integrated temperature distribution within power devices.

5.2 Mach–Zehnder interferometry

While the Backside Laser Probing technique provides vertically integrated information only, we will now discuss a similar interferometric technique, namely Mach–Zehnder interferometry, which enables a vertical *scanning* as it employs a laterally impinging laser beam. Thermally or electrically induced modulations of the optical sample thickness affect the phase of the probing beam. This phase shift is detected by interference with a reference beam which is conducted around the sample (cf. section 2.2.4).

To study the capabilities of this technique the same situation as in section (4.3.7) is considered: A block of silicon ($200\,\mu\mathrm{m}$ wafer thickness, sample length $800\,\mu\mathrm{m}$) is heated by a constant heat flow of $2 \cdot 10^4\,\frac{\mathrm{W}}{\mathrm{cm}^2}$ through the top surface during the period $0 < t < 100\,\mu\mathrm{s}$. The resulting temperature distribution is plotted in fig. (4.30).

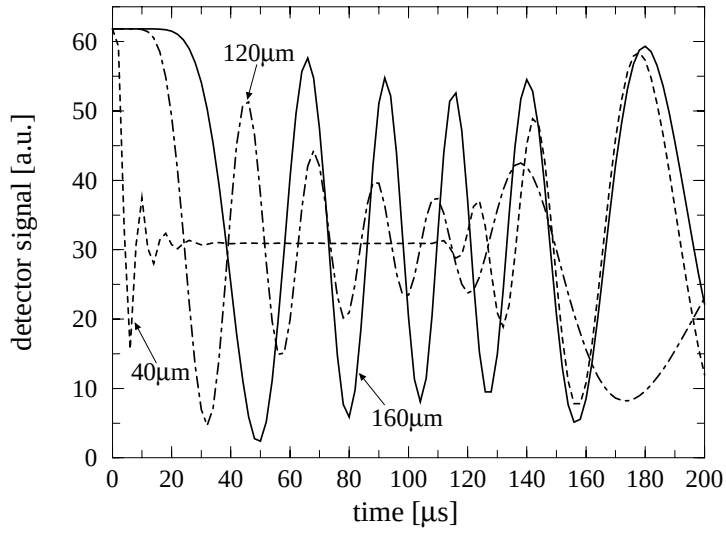


Figure 5.11: Detector signal of a Mach–Zehnder interferometer probing a sample which is heated by a constant heat flow of $2 \cdot 10^4 \, \frac{\mathrm{W}}{\mathrm{cm}^2}$*. The probing beam is incident at a depth of* $40 \, \mu\mathrm{m}$*,* $120 \, \mu\mathrm{m}$*, and* $160 \, \mu\mathrm{m}$*, respectively.*

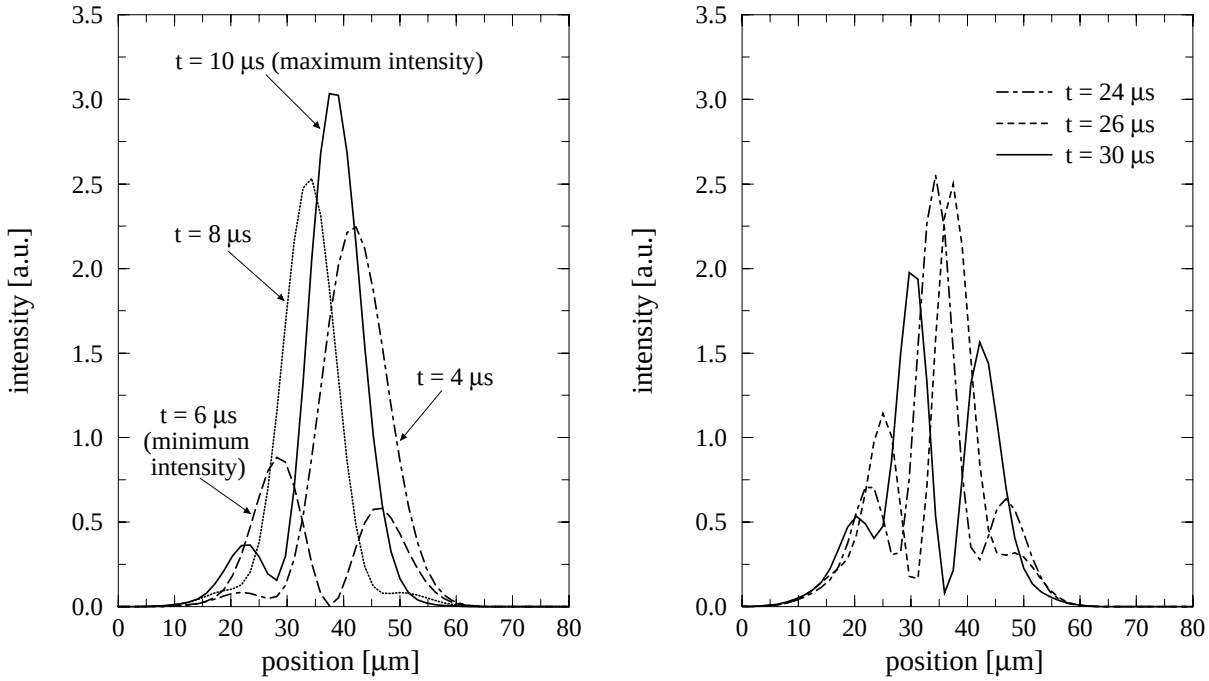


Figure 5.12: Intensity distribution on the detector of a Mach–Zehnder interferometer. The probing beam is incident at $40 \, \mu\mathrm{m}$*.*

The rising temperature increases the optical thickness of the sample. Due to the continuously increasing phase delay, the superposition of the probing beam and the reference beam alternates between constructive and destructive interference, resulting in oscillations of the detected intensity (cf. fig. 5.11). The temperature rise ΔT_{pp} between two peaks can be extracted by

$$\Delta T_{pp} = \frac{\lambda}{L \frac{\partial n_{Si}}{\partial T} + n_{Si} \frac{\partial L}{\partial T}} \; . \tag{5.5}$$

However, the thermally induced *gradients* of the refractive index lead to an increasing deflection of the probing beam which may amount to some tens of microns (cf. fig. 4.31). As a

consequence, the interference of the probing beam and the reference beam deteriorates due to their increasing distance on the detector. If the latter becomes too large, the field distribution merely displays different interference fringes with a constant integral intensity and the detector signal remains constant (cf. fig. 5.11 and 5.12). This undesired effect essentially limits the measurement range.

Broadening the field distribution of the reference beam reduces the sensitivity to a possible displacement of the probing beam and therefore helps to overcome this difficulty. In addition, a reduction of the interaction length and an appropriate design of test structures decrease the internal beam deflection and thus enhance the measurement range. However, since the Mach–Zehnder interferometer requires quite a large experimental effort, e. g. the facilities to split off and guide the reference beam as well as a careful sample preparation including the deposition of antireflective coatings, our investigations will focus on the much simpler Fabry–Perot transmission measurements, which will be discussed in the following section.

5.3 Fabry–Perot transmission measurements

Thermally induced Fabry–Perot oscillations emerge as an undesired effect during free carrier absorption measurements. However, since they reveal the optical thickness of the sample they can by exploited to extract the temperature evolution at the current beam position. In this section we will address the question of whether Fabry–Perot transmission measurements (cf. section 2.2.4) constitute a suitable method to ascertain the temperature distribution in the interior of power devices operating with a large power dissipation. Again, the theoretical study is based on the situation described in section (4.3.7), namely a constant heat flow through a thermal contact at the top surface. To achieve a large measurement range, we take into account specific test structures comprising an active area which is embedded in a silicon block of $800\ \mu\mathrm{m}$ in length (cf. fig. 4.32). We will investigate the probing of samples with different lengths of the active area and elaborate an evaluation rule for the measurement signal.

5.3.1 The detector signal

The intensity transmitted by a Fabry–Perot interferometer of thickness L writes [78] (cf. eq. 2.23)

$$I_t(t) = \frac{(1-r_f r_r)^2 \mathrm{e}^{-\alpha L}}{(1-r_f r_r \mathrm{e}^{-\alpha L})^2 + 4 r_f r_r \mathrm{e}^{-\alpha L} \sin^2[\varphi(t)]} I_0 \tag{5.6}$$

$$\varphi(t) = k_0 L n_{Si} + k_0 \cdot \left(n_{Si} \frac{\partial L}{\partial T} + L \frac{\partial n_{Si}}{\partial T} \right) \Delta T(t) \tag{5.7}$$

where α and $r_{f,r}$ denote the absorption coefficient within the sample and the reflectivity of the front and rear surface, respectively. Consequently, the temperature difference between two peaks is given by

$$\Delta T_{pp} = \frac{\lambda/2}{L \frac{\partial n_{Si}}{\partial T} + n_{Si} \frac{\partial L}{\partial T}} \ . \tag{5.8}$$

The intensity signal (cf. fig. 5.13) oscillates the more rapidly, the closer the beam position is to

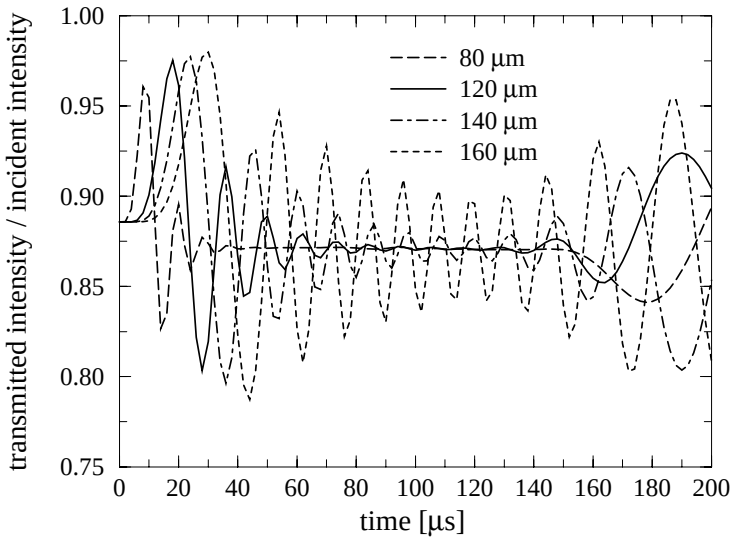


Figure 5.13: Calculated detector signal of Fabry–Perot transmission measurements (heat flow of $2 \cdot 10^4\,\mathrm{W/cm^2}$ during $0 < t < 100\,\mu\mathrm{s}$, length of active area $L_a = 800\,\mu\mathrm{m}$).

the upper surface which exhibits the greatest temperature rise (cf. fig. 4.30). Since the increasing internal deflection due to the evolving refractive index gradients impairs the superposition of the multiply reflected beams, the amplitude of the oscillations continuously decreases. Finally, the intensity signal converges to a constant value which no longer reveals any information about the temperature evolution, thus limiting the measurement range.

It is evident that a reduction of the interaction length helps to overcome this difficulty. Since shrinking the sample size is limited by practical constraints, we consider measurements on test structures with a smaller length of the active area (cf. fig. 5.14) and try to correct the intensity signal with respect to the lateral spreading of the temperature distribution (cf. fig. 4.32).

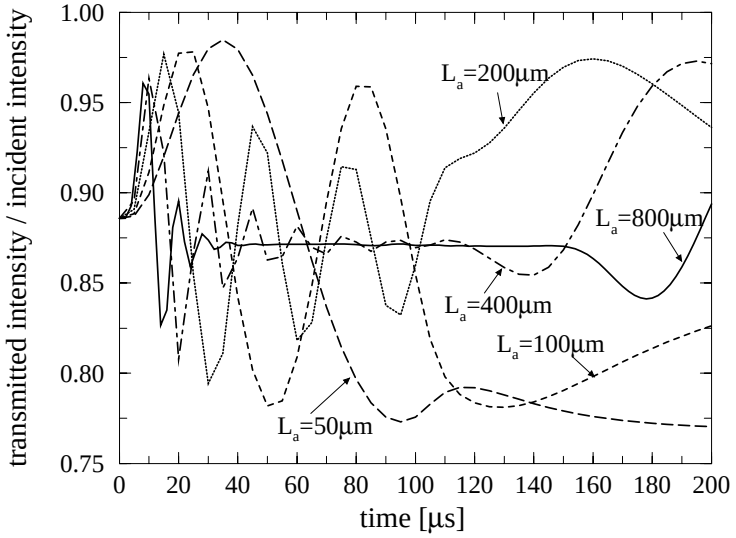


Figure 5.14: Calculated intensity signal of Fabry–Perot transmission measurements on test structures with various lengths L_a of the active region (heat flow of $2 \cdot 10^4\,\mathrm{W/cm^2}$ during $0 < t < 100\,\mu\mathrm{s}$, sample length $800\,\mu\mathrm{m}$, beam position $80\,\mu\mathrm{m}$).

5.3.2 Extracting the temperature evolution

A straightforward way for a quantitative evaluation is to count the interference extrema and calculate the temperature rise according to equation (5.8). A continuous extraction of the temperature evolution requires the assumption of an interpolation rule [98].

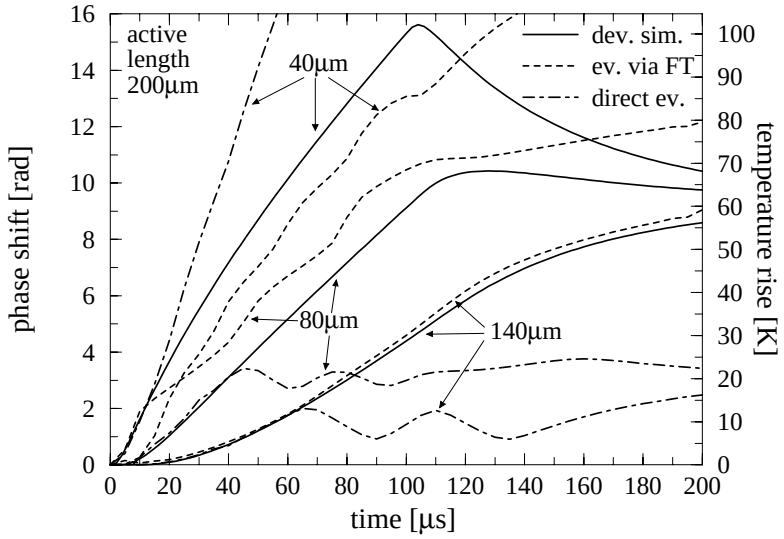


Figure 5.15: Phase shift signals (and temperature rises) extracted from the calculated intensity signals by a direct evaluation (dot–dashed lines) and by filtering in frequency space (dashed lines). The results of virtual experiments on test structures (active length $L_a = 200\,\mu\mathrm{m}$) are compared with the temperature profile obtained by a device simulation (solid lines).

As the expression (5.6) can be transformed to

$$\frac{1}{I_t(t)} = C_1 + 2C_2 \sin^2[\varphi(t)] = (C_1 + C_2) - C_2 \cos[2\varphi(t)] \tag{5.9}$$

with the two a priori unknown constants C_1 and C_2, we can also follow the strategy explained in appendix (D.2): The intensity signal is transformed to frequency space and subjected to band pass filtering. Thus, the phase can be extracted from the complex function obtained by transforming back to time space.

Since the formula (5.6) accounts for Fabry–Perot interference of non deflected beams the results presented in fig. (5.15) and (5.16) are intended to answer the two decisive questions: Which strategy is the better one for evaluating the intensity signals of deflected beams and which interaction length has to be inserted into equation (5.7) to extract the temperature evolution from the phase shift signal? Encouraged by the experiences with the deflection measurements (cf. section 4.3.7) which suggest to use the active length L_a in case of a linear dependence on the interaction length, we assume the latter to be equal to the length of the active area ($L = L_a$) in the following discussion. As mentioned above, the accuracy of the extracted temperature evolution is judged by a comparison with the reference profile. This is obtained by a one–dimensional solution of the heat flow equation since we are interested in the temperature distribution within large samples that lack of the lateral heat diffusion.

As it can be seen from fig. (5.15), the extraction strategy employing the Fourier transformations is superior if the probing beam traverses samples comprising a large active area. In this case, the intensity signal exhibits rapid oscillations with a decreasing amplitude, whose time–dependence is reflected in additional components of the frequency spectrum. Since they can be easily separated from the frequency of the basic oscillations, the Fourier transformation method accurately reveals the temperature evolution up to a temperature rise of about 70 K. However, the direct evaluation of equation (5.9) assuming an oscillation of constant amplitude naturally breaks down soon.

On the other hand, probing test structures comprising a smaller active area results in slower oscillations of the intensity signal, thus impairing the frequency distance between the oscillating

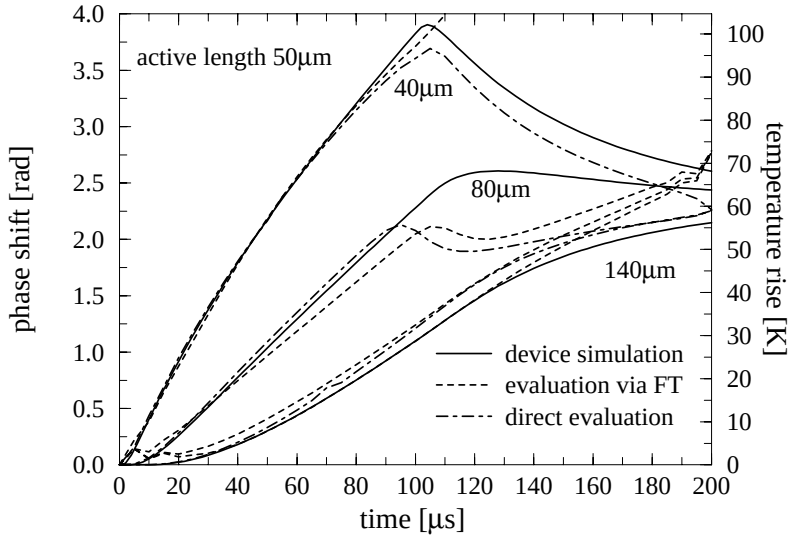


Figure 5.16: Phase shift signals (and temperature rises) extracted from the calculated intensity signals by a direct evaluation (dot–dashed lines) and by filtering in frequency space (dashed lines). The results of virtual experiments on test structures (active length $L_a = 50\,\mu\mathrm{m}$) are compared with the temperature profile obtained by a device simulation (solid lines).

component and the constant offset. Hence, the separation becomes more difficult and the extracted phase shift signal is not able to reproduce the smaller slope of the temperature evolution at the end of the heat current pulse. In this case, the direct extraction according to equation (5.9) is the better choice (cf. fig. 5.16) and works sufficiently accurate up to a temperature rise of about 90 K.

In any case, the satisfying agreement of the extracted temperature evolution with the reference temperature of a thermal simulation clearly demonstrates that the length L_a of the active area has to be inserted as the interaction length L. However, the reader should be reminded of the assumptions this study is based on (cf. the end of section 4.3.7). In particular, we have neglected a possible feedback of the temperature distribution on the electrical behavior of the investigated device. For that reason, the test structures with the active length of only $50\,\mu\mathrm{m}$ is not preferable. Nevertheless, even the absolute height of the detectable temperatures rises on test structure with $L_a = 200\,\mu\mathrm{m}$ is very promising. Therefore, Fabry–Perot transmission measurements obviously represent a valuable supplement to the existing laser probing techniques. This is especially true, as they can be performed together with laser absorption and deflection measurements in the same experimental setup which includes an optical switch to select either an incoherent or a coherent light source.

5.3.3 Summary

Fabry–Perot transmission measurements are expected to enable a vertical scanning of the temperature profiles in power devices operating with a large power dissipation. Due to their simpler experimental setup they are preferred to Mach–Zehnder interferometry. In summary, the most important results of this section are:

- The measurement range of Fabry–Perot transmission measurements is limited by the internal deflection of the probing beam due to refractive index gradients perpendicular to the optical axis.

- In case of low power dissipation or small interaction lengths the intensity signal exhibits only

a few oscillations. Fitting their amplitude and the constant offset, the temperature evolution can be extracted by solving the transmission formula for the phase shift.

- In case of a large power dissipation the increasing internal deflection impairs the superposition of the multiply reflected beams. The intensity signal exhibits rapid oscillations with decreasing amplitude. Fourier transformation and band pass filtering in the frequency space represent an accurate method to extract the phase shift signal.
- Designing test structures with a smaller active area embedded in a larger silicon die decreases the total deflection and thus enhances the measurement range up to a temperature rise of about 70 K. The temperature evolution extracted from the phase shift signal matches the reference temperature profile with excellent accuracy if the length of the active area is inserted as the interaction length into the evaluation formula.

Chapter 6

Electrothermal Device Simulation

The second part of this thesis addresses the validation and calibration of electrothermal device simulation models. Numerical simulations have emerged as a valuable method to investigate the physical processes in the interior of semiconductor devices. The fundamental mathematical structure of the numerical model consists of a set of partial differential equations representing an appropriate description with the desired range of validity. These equations are derived from basic physical theories, e. g. electrodynamics, thermodynamics, or quantum mechanics, and govern the evolution of the corresponding state variables.

From the physically rigorous point of view, the internal dynamics of semiconductors constitutes a many particle problem of quantum–mechanics. However, among the numerous degrees of freedom only a few macroscopic observables are of particular interest. A rigorous treatment of the many particle problem is therefore neither feasible nor desirable. Hence, the fundamental challenge in semiconductor device modeling is reducing complexity to a physically based set of equations with a limited number of variables.

Several semi–classical approaches to device simulation have been proposed which are based on a momentum expansion of Boltzmann's transport equation. Including the first three moments representing particle conservation, momentum conservation, and energy conservation, respectively, Bløtekjær [128] derived a set of equations which is known as the Hydrodynamic (HD) model today. Cook and Frey [129] simplified the HD model to the so–called Energy Balance (EB) model by neglecting the mean kinetic energy compared to the electron thermal energy. Assuming constant and equal temperatures of the electron gas, the hole gas and the host lattice, we obtain the Drift–Diffusion (DD) model as a special case of the EB model. From the historical point of view, the DD model constitutes the first approach to semiconductor device simulation and was reported by Van Roosbroeck [130] in 1950. It forms the original basis of many well known general purpose device simulators (e. g. PISCES, MEDICI [131], ATLAS [132]). Although an additional description of lattice heating was included into some of them afterwards, they usually lack of a self–consistent treatment of carrier and heat flow.

A different methodology was published by Wachutka who applied the principles of irreversible thermodynamics [133] and linear transport theory [134, 135] to derive a self–consistent electrothermal formulation, which is known as the Thermodynamic (TD) model [62, 136]. The following part of this thesis is dedicated to a comprehensive validation of the TD model, in particular to judge

its suitability for the accurate and predictive simulation of power devices. We will focus on the general purpose device simulator $\mathrm{DESSIS_{ISE}}$ [101] whose implemented set of partial differential equations is based on the TD model and will be discussed in the following chapter.

6.1 The electrothermal model – derivation from irreversible thermodynamics

Irreversible thermodynamics is founded on the concept of local thermal equilibrium: The whole structure is partitioned into small elements, each of which comprises a thermodynamic subsystem. These cells are supposed to be *locally* in thermal equilibrium, even if the structure *as a whole* is operated far away from its equilibrium state. At a fixed position $\vec{r}$ and time t, the structure is therefore characterized by the local thermodynamic equilibrium variables, such as e. g. energy density $u(\vec{r}, t)$ or particle density $n(\vec{r}, t)$.

The key idea for deriving a consistent treatment of carrier and heat flow in the semiconductor is considering each of the above mentioned elements a composite thermodynamic system which is formed by the subsystems of the electrons, the holes, and the host lattice. The dynamics of the whole structure is reflected by the space and time evolution of the state variables which is governed by the balance equations representing particle and energy conservation (cf. section 6.1.2). However, to obtain a closed system of dynamic equations, additional information is required: First, the particle and energy currents arising from a global inhomogeneity of the state variables are described by irreversible transport theory (cf. section 6.1.3). And second, so-called state equations relate the conjugate variables to the chosen state variables. For example, assuming a local equilibrium distribution of the carriers in k–space, the quasi–Fermi levels are expressed in terms of the carrier concentrations by the Fermi–Dirac distribution functions (cf. section 6.1.4).

6.1.1 Semiconductors as thermodynamic systems

For modeling the internal dynamics of semiconductors, three interacting thermodynamic systems are regarded, namely the subsystems of the electrons, the holes, and the host lattice. The two former are characterized, e. g., by the particle concentrations n and p and their temperatures T_n and T_p, respectively. Since the density of the lattice atoms is fixed, we only have to include one additional state variable, namely the lattice temperature T_L.

In a more general description which also accounts for the dynamic ionization of trap centers, in particular the donor and acceptor levels, additional subsystems with their specific particle densities and temperatures have to be regarded [137, 138]. However, since the ionization energies of the commonly used dopants in silicon are small compared with the thermal energy, we will restrict ourselves to the case of a complete ionization of all dopands.

In the thermodynamic model, these subsystems are assumed in local thermal equilibrium with each other, i. e. their specific temperatures are equal: $T_n = T_p = T_L =: T$. Hence, we obtain a set of three state variables which can be represented, e. g., by n, p, and T.

Even in the presence of an electric field, we do not have to introduce another state variable whose evolution is governed by a dynamic equation. Since the electromagnetic radiation field does

not affect the motion of the charge carriers, electromagnetic phenomena in non–optical devices are represented by the electrostatic field $\vec{E} = -\nabla\psi$. It can be calculated from Poisson's equation

$$\vec{\nabla}\left(\varepsilon\vec{\nabla}\psi\right) = q\left(n - p + N_A^- - N_D^+\right) \tag{6.1}$$

at any point of time and therefore does not represent an independent degree of freedom.

As will be shown below, the conditional equation for the temperature evolution is derived from the balance of the total energy. For that purpose, we have to determine the energy contributions of each subsystem. Applying the first law of thermodynamics, the fundamental Gibbs relations in the presence of an electrostatic field $\vec{E} = -\nabla\psi$ read:

$$\begin{aligned} du_n &= T\,ds_n + \mu_n^{(c)}\,dn && (6.2)\\ du_p &= T\,ds_p - \mu_p^{(c)}\,dp && (6.3)\\ du_L &= T\,ds_L && (6.4)\\ du_{el} &= \vec{E}\,d\vec{D} && (6.5) \end{aligned}$$

The electrostatic work $\vec{E}d\vec{D}$ can be split into two contributions which can be added to the works of the electrons and holes, respectively. For that purpose, the electrochemical potentials $\mu_{n,p}^{(ec)}$ and the quasi–Fermi levels $\phi_{n,p}$ are introduced by

$$\begin{aligned} -q\phi_n &:= \mu_n^{(ec)} := \mu_n^{(c)} - q\psi && (6.6)\\ -q\phi_p &:= \mu_p^{(ec)} := \mu_p^{(c)} - q\psi && (6.7) \end{aligned}$$

transforming the equations (6.2) to (6.5) to

$$\begin{aligned} du_n &= T\,ds_n - q\phi_n\,dn && (6.8)\\ du_p &= T\,ds_p + q\phi_p\,dp && (6.9)\\ du_L &= T\,ds_L && (6.10)\\ du_{el} &= 0 && (6.11) \end{aligned}$$

Expressing the entropy changes ds in terms of the chosen state variables and applying Maxwell's relations [133, ch. 7], we finally end up with

$$\begin{aligned} du_n &= c_n\,dT + q\left[T\left(\frac{\partial\phi_n}{\partial T}\right)_{n,p} - \phi_n\right]dn && (6.12)\\ du_p &= c_p\,dT - q\left[T\left(\frac{\partial\phi_p}{\partial T}\right)_{n,p} - \phi_p\right]dp && (6.13)\\ du_L &= c_L\,dT && (6.14) \end{aligned}$$

6.1.2 Balance equations

The dynamic evolution of each extensive parameter X is governed by the associated balance equation

$$\frac{\partial x(\vec{r},t)}{\partial t} + \operatorname{div}\vec{j}_x(\vec{r},t) = \Pi_x(\vec{r},t) \tag{6.15}$$

where x, $\vec{j}_x$, and Π_x denote the density of X, the corresponding current density, and the production rate, respectively. Integrating (6.15) over a small volume provides an intuitive interpretation: The average density of X increases by the net flow into the volume and by the production within this volume.

Since the generation G minus the recombination R constitutes the production rate of free carriers in semiconductors, the particle balances for electrons and holes become:

$$q\,\frac{\partial n(\vec{r},t)}{\partial t} - \operatorname{div}\vec{J}_n(\vec{r},t) \;=\; q\,(G-R) \tag{6.16}$$

$$q\,\frac{\partial p(\vec{r},t)}{\partial t} + \operatorname{div}\vec{J}_p(\vec{r},t) \;=\; q\,(G-R) \tag{6.17}$$

In these equations, the *particle* current densities $\vec{j}_{n,p}$ have been replaced by the *electrical* current densities $\vec{J}_n = -q\vec{j}_n$ and $\vec{J}_p = q\vec{j}_p$ which are commonly used for electrothermal device simulation.

Similarly, the conservation of the total energy is expressed by a balance equation with a vanishing energy production rate:

$$\frac{\partial u(\vec{r},t)}{\partial t} + \operatorname{div}\vec{J}_u(\vec{r},t) = 0 \tag{6.18}$$

6.1.3 Irreversible transport theory

As mentioned above, the description of global non–equilibrium states assumes that the structure can be partitioned into small cells in local thermal equilibrium. If the state variables of neighboring cells have different values, the tendency to evolve towards the global equilibrium originates particle and energy currents. These phenomena are the subject of Onsager's irreversible transport theory [134, 135]. According to his fundamental reciprocity theorem, the bilinear dependence $\Pi_s = \sum_{k,j} \vec{F}_k L_{kj} \vec{F}_j$ of the entropy production rate Π_s on the driving forces F_k is represented by a symmetric and positive definite matrix L_{kj}.

To apply Onsager's theorem, we first have to deduce the entropy production rate Π_s which is defined according to the generic balance (6.15) by $\Pi_s = \operatorname{div}\vec{J}_s + \partial s/\partial t$. As a first step, the total entropy density is derived from the equations (6.8) to (6.10). Hence, we obtain:

$$ds = ds_n + ds_p + ds_L = \frac{1}{T}\,du + \frac{q\phi_n}{T}\,dn - \frac{q\phi_p}{T}\,dp \tag{6.19}$$

Consequently, the entropy current density $\vec{J}_s$ is related to the energy current density $\vec{J}_u$ and the electrical current densities $\vec{J}_{n,p}$ by

$$\vec{J}_s = \frac{1}{T}\,\vec{J}_u - \frac{\phi_n}{T}\,\vec{J}_n - \frac{\phi_p}{T}\,\vec{J}_p\ . \tag{6.20}$$

Inserting these equations together with (6.16), (6.17), and (6.18) results in the following expression:

$$\Pi_s = \operatorname{div}\vec{J}_s + \frac{\partial s}{\partial t} = \vec{J}_u\vec{\nabla}\left(\frac{1}{T}\right) - \vec{J}_n\vec{\nabla}\left(\frac{\phi_n}{T}\right) - \vec{J}_p\vec{\nabla}\left(\frac{\phi_p}{T}\right) + \frac{q}{T}(G-R)(\phi_n-\phi_p)\ . \tag{6.21}$$

In the fundamental formalism [133, ch. 14], the contributions to the entropy production rate $\Pi_s = \sum_k \vec{F}_k\vec{J}_k$ are products of current densities J_k and the corresponding driving forces F_k, called

affinities. Equation (6.21) therefore identifies $\vec{\nabla}(1/T)$, $-\vec{\nabla}(\phi_n/T)$, $-\vec{\nabla}(\phi_p/T)$, and $(\phi_n - \phi_p)/T$ as the driving forces to the currents $\vec{J_u}$, $\vec{J_n}$, $\vec{J_p}$, and $q(G-R)$, respectively.

With a first order Taylor expansion in the vicinity of the thermodynamic equilibrium, the charge and energy current densities depend linearly on the driving affinities:

$$\begin{pmatrix} \vec{J_n} \\ \vec{J_p} \\ \vec{J_u} \end{pmatrix} = \begin{pmatrix} L_{nn} & L_{np} & L_{nu} \\ L_{pn} & L_{pp} & L_{pu} \\ L_{un} & L_{up} & L_{uu} \end{pmatrix} \begin{pmatrix} -\vec{\nabla}\left(\frac{\phi_n}{T}\right) \\ -\vec{\nabla}\left(\frac{\phi_p}{T}\right) \\ \vec{\nabla}\left(\frac{1}{T}\right) \end{pmatrix} \tag{6.22}$$

In general, the coefficients L_{XX} of the matrix are 3×3 submatrices, since the currents need not to be parallel with the driving forces. This effect is observed, for instance, in anisotropic material or in the presence of a magnetic field. In this thesis, however, we will restrict ourselves to the case where each of these submatrices reduces to a scalar transport coefficient multiplied by the unity matrix.

According to Onsager's reciprocity theorem [134, 135], the matrix of equation (6.22) is symmetric and positive definite, thus reducing the number of independent coefficients from 9 to 6.

It is convenient to express the current densities in terms of the driving forces $\vec{\nabla}\phi_{n,p}$ and $\vec{\nabla}T$. Equation (6.22) therefore becomes:

$$\vec{J_n} = -qn\mu_n(\vec{\nabla}\phi_n + P_n\vec{\nabla}T) - \sigma_{np}\vec{\nabla}\phi_p \tag{6.23}$$

$$\vec{J_p} = -qp\mu_p(\vec{\nabla}\phi_p + P_p\vec{\nabla}T) - \sigma_{pn}\vec{\nabla}\phi_n \tag{6.24}$$

$$\vec{J_u} = -\kappa_t\vec{\nabla}T + (TP_n + \phi_n)\vec{J_n} + (TP_p + \phi_p)\vec{J_p} + \sigma_{np}TP_p\vec{\nabla}\phi_n + \sigma_{pn}TP_n\vec{\nabla}\phi_p \tag{6.25}$$

The six independent coefficients are represented by the mobilities μ_n and μ_p, the coefficient $\sigma_{np} = \sigma_{pn}$, the thermoelectric powers P_n and P_p, and the thermal conductivity κ_t.

In the general theory, the electron current density may also depend on the gradient of the quasi–Fermi level of the holes, which has been subject of many discussions in literature (e. g. [139, 140]). Recently, internal laser probing techniques (cf. section 2.2.2) also have been employed for the experimental validation [141, 76]. However, the corresponding coefficient σ_{np} is small and will therefore be neglected in the following discussions, which is in agreement with the commonly accepted hypothesis. Hence, we obtain:

$$\vec{J_n} = -qn\mu_n(\vec{\nabla}\phi_n + P_n\vec{\nabla}T) \tag{6.26}$$

$$\vec{J_p} = -qp\mu_p(\vec{\nabla}\phi_p + P_p\vec{\nabla}T) \tag{6.27}$$

6.1.4 Carrier concentrations and Fermi levels

In order to complete the description of the electrothermal model, we finally have to specify a relation between the particle densities n and p and the associated quasi–Fermi levels ϕ_n and ϕ_p. Since the system of the electrons is assumed to be in local equilibrium, the electron distribution in k–space in the absence of an electric field is governed by the Fermi distribution function f_n with the quasi–Fermi energy $\mu_n^{(c)}$. The total concentration of free electrons is therefore obtained by the integral

$$n = \int_{E_{c0}}^{\infty} \rho_c(E) f_n(E)\, dE \tag{6.28}$$

where $\rho_c(E)$ denotes the density of states in the conduction band. Introducing the effective density of states N_c [142] and the Fermi integral $F_{1/2}$, the above equation is usually written as

$$n = N_c\, F_{1/2}\left(\frac{\mu_n^{(c)} - E_{c0}}{kT}\right) \approx N_c\, \exp\left(\frac{\mu_n^{(c)} - E_{c0}}{kT}\right) = N_c\, \exp\left(-\frac{E_{c0}}{kT}\right)\, \exp\left(\frac{q(\psi - \phi_n)}{kT}\right) \tag{6.29}$$

where Boltzmann statistics can be used for carrier concentrations up to $10^{18}\,\mathrm{cm}^{-3}$. This relation is also valid in the presence of an electric field if it is sufficiently weak, that all energies are shifted by $q\psi$, but the electron distribution in k–space remains nearly unaffected.

Solving Poisson's equation in an intrinsic semiconductor which is not exposed to an external electric field ($\psi = \phi_n = \phi_p = 0$), the term $N_c\, \exp(-E_{c0}/kT)$ is shown to be equal to the so–called intrinsic concentration [143]

$$n_i := \sqrt{N_c N_v}\, \exp\left(\frac{-E_g}{2kT}\right)\ . \tag{6.30}$$

Hence, we get $n = n_i \exp[\, q(\psi - \phi_n)/kT]$ and similarly $p = n_i \exp[\, q(\phi_p - \psi)/kT]$.

A convenient way to include Fermi statistics consists of introducing the effective intrinsic concentrations

$$n_{ie,n} := \frac{N_c\, F_{1/2}\left(\frac{-E_{c0}+q(\psi-\phi_n)}{kT}\right)}{\exp\left(\frac{q(\psi-\phi_n)}{kT}\right)} \quad \text{and} \quad n_{ie,p} := \frac{N_v\, F_{1/2}\left(\frac{E_{v0}+q(\phi_p-\psi)}{kT}\right)}{\exp\left(\frac{q(\phi_p-\psi)}{kT}\right)}\ . \tag{6.31}$$

While they approach n_i in the Boltzmann limit, approximate calculations for $n_{ie,\substack{p\\n}}$ are employed in case of large carrier concentrations. Thus, we finally obtain:

$$n = n_{ie,n}\, \exp\left(\frac{q(\psi - \phi_n)}{kT}\right) \tag{6.32}$$

$$p = n_{ie,p}\, \exp\left(\frac{q(\phi_p - \psi)}{kT}\right) \tag{6.33}$$

Solving for the quasi–Fermi levels results in the following state equations:

$$\phi_n(n,p,T) = \psi(n,p,T) - \frac{kT}{q} \ln\left(\frac{n}{n_{ie,n}(T)}\right) \tag{6.34}$$

$$\phi_p(n,p,T) = \psi(n,p,T) + \frac{kT}{q} \ln\left(\frac{p}{n_{ie,p}(T)}\right) \tag{6.35}$$

6.1.5 Governing equations

From the macroscopic point of view, the dynamics of the system is described by the time evolution of the state variables. The governing equations are therefore derived by inserting the current relations (6.26), (6.27), and (6.25) into the corresponding balance equations (6.16), (6.17), and (6.18), respectively. In addition, the state equations (6.34), (6.35), and (6.1) relate the conjugate variables ϕ_n and ϕ_p to the thermodynamic variables n, p, and T.

The time evolution of the carrier densities is governed by the particle balances (6.16) and (6.17):

$$q\,\frac{\partial n(\vec{r},t)}{\partial t} = \mathrm{div}\left[-qn\mu_n(\vec{\nabla}\phi_n + P_n\vec{\nabla}T)\right] - q\,(G-R) \tag{6.36}$$

$$q\,\frac{\partial p(\vec{r},t)}{\partial t} = -\mathrm{div}\left[-qp\mu_p(\vec{\nabla}\phi_p + P_p\vec{\nabla}T)\right] - q\,(G-R) \tag{6.37}$$

The key idea to obtain a conditional equation of the temperature evolution which is consistent with the drift–diffusion model consists in exploiting the conservation of the total energy. In other words, (6.12), (6.13), (6.14) and (6.25) are inserted into the balance equation (6.18), resulting in

$$\begin{aligned} c_t\,\frac{\partial T}{\partial t} - \vec{\nabla}\left(\kappa_t\vec{\nabla}T\right) &= -\vec{\nabla}\left[(\phi_n + P_nT)\vec{J}_n + (\phi_p + P_pT)\vec{J}_p\right] \\ &\quad + q\left(\phi_n - T\frac{\partial\phi_n}{\partial T}\right)\frac{\partial n}{\partial t} - q\left(\phi_p - T\frac{\partial\phi_p}{\partial T}\right)\frac{\partial p}{\partial t}\ . \end{aligned} \tag{6.38}$$

Although this equation does not reflect the heat balance from the thermodynamic point of view, it is usually referred as the "heat flow equation", whose right hand side represents the heat generation rate H.

In addition to these dynamic relations, the state equations (6.1), (6.34), and (6.35) have to be satisfied. On the one hand, the quasi–Fermi levels $\phi_{n,p}$ are explicitly given by (6.34) and (6.35) and can be directly inserted into (6.36) and (6.37). On the other hand, Poisson's equation

$$\vec{\nabla}\left(\varepsilon\vec{\nabla}\psi\right) = q\left(n - p + N_A^- - N_D^+\right) \tag{6.39}$$

cannot be solved for the potential ψ analytically. Therefore, ψ is introduced as an additional variable and the system of the four partial differential equations comprising (6.36), (6.37), (6.38), and (6.39) has to be solved numerically.

For the electrothermal simulations in this thesis the general purpose device simulator $\mathrm{DESSIS_{ISE}}$ [101] was employed, into which the electrothermal model as sketched above is implemented [144]. However, the heat flow equation (6.38) was simplified by neglecting the last two terms on the right hand side (cf. [101, eq. 14.13]). The resulting effects will be discussed in the following chapter.

6.1.6 Interpretation of the heat generation mechanisms

A familiar interpretation of the heat generation rate H is gained by evaluating the divergence operator on the right hand side of equation (6.38) and inserting the equations (6.16), (6.17), (6.26), and (6.27). Thus, we obtain

$$H = (H_{eJoule} + H_{hJoule}) + H_{rec} + (H_{Peltier} + H_{Thomson}) + H_{trans}\ . \tag{6.40}$$

First, we follow Wachutka [62] and identify

$$H_{eJoule} = \frac{\vec{J}_n^{\,2}}{qn\mu_n} \quad \text{and} \quad H_{hJoule} = \frac{\vec{J}_p^{\,2}}{qp\mu_p} \tag{6.41}$$

as the Joule heat of electrons and holes, respectively. Since these expressions always refer to a positive heat generation rate, this definition of Joule heat is preferable against the widely used formula $\vec{J}\vec{E}$. The latter becomes negative, for example, if the total current across a forward biased pn–junction is dominated by the diffusion current flowing in the direction opposite to the electric field.
Second, the recombination heat is represented by

$$H_{rec} = q(R-G)\left(\phi_p + TP_p - \phi_n - TP_n\right) \ . \tag{6.42}$$

As a third term, we obtain the Peltier/Thomson heat

$$H_{Peltier} + H_{Thomson} = -\vec{J}_n T\vec{\nabla}P_n - \vec{J}_p T\vec{\nabla}P_p \ . \tag{6.43}$$

It can be broken into the Thomson heat which applies to a current flow across a temperature gradient and the Peltier heat that arises from a current flowing across the gradients of the thermoelectric powers under isothermal conditions:

$$H_{Thomson} = -T\vec{J}_n\left(\frac{\partial P_n}{\partial T}\right)_{n,p}\vec{\nabla}T - T\vec{J}_p\left(\frac{\partial P_p}{\partial T}\right)_{n,p}\vec{\nabla}T \tag{6.44}$$

$$H_{Peltier} = -T\vec{J}_n\left(\frac{\partial P_n}{\partial n}\right)_{T,p}\vec{\nabla}n - T\vec{J}_p\left(\frac{\partial P_p}{\partial p}\right)_{T,n}\vec{\nabla}p \tag{6.45}$$

Finally, an additional heat production rate originates from transient modulations of the carrier concentrations:

$$H_{trans} = -q\left[P_nT + \left(\frac{\partial\phi_n}{\partial T}\right)_{n,p}T\right]\frac{\partial n}{\partial t} + q\left[P_pT + \left(\frac{\partial\phi_p}{\partial T}\right)_{n,p}T\right]\frac{\partial p}{\partial t} \tag{6.46}$$

As a consequence of simplifying the right hand side of the heat flow equation (cf. section 6.1.5), this contribution to the heat production rate in the simulator $\mathrm{DESSIS_{ISE}}$ becomes

$$H_{trans} = -q\left[P_nT + \phi_n\right]\frac{\partial n}{\partial t} + q\left[P_pT + \phi_p\right]\frac{\partial p}{\partial t} \ . \tag{6.47}$$

Note that this relation is physically unreasonable since the coefficients depend on the absolute value of the quasi–Fermi levels and are therefore affected by a shift of the reference level for the circuit potentials. Fortunately, this contribution is small in view of the other heat generation mechanisms and can therefore be neglected for practical purposes.

6.1.7 Boundary conditions

For specifying the boundary conditions of the electrical equations (6.1), (6.16), and (6.17), we have to distinguish between contacts and interfaces to insulating materials. The most common model for electric contacts is the so–called Ohmic contact [143] which assumes the quasi–Fermi levels to be equal to the externally applied bias V:

$$\phi_n = \phi_p = V \tag{6.48}$$

In addition, the carrier concentrations are supposed to be in thermal equilibrium, i. e. $np = n_{i,\mathit{eff}}^2$, resulting in the following boundary condition for the electrostatic potential:

$$\psi = V + \frac{kT}{q}\mathrm{arsinh}\left(\frac{N_D^+ - N_A^-}{2n_{i,\mathit{eff}}}\right) = V + V_{bi} \tag{6.49}$$

On the other hand, the boundary conditions at the interfaces to electrically insulating materials are derived by integrating (6.16), (6.17), and (6.39), respectively. Hence, we obtain

$$\begin{aligned} \varepsilon_{sc}\frac{\partial}{\partial\vec{n}}\psi\Big|_{sc} - \varepsilon_{ins}\frac{\partial}{\partial\vec{n}}\psi\Big|_{ins} &= \sigma_{Surf} && (6.50)\\ \vec{n}\vec{J}_n\Big|_{sc} &= -q\,R_{Surf} && (6.51)\\ \vec{n}\vec{J}_p\Big|_{sc} &= q\,R_{Surf} && (6.52) \end{aligned}$$

where $\vec{n}$, σ_{Surf}, R_{Surf} denote the normal vector to the interface, the interface charge density, and the surface recombination rate, respectively. If the interface charge and the surface recombination can be neglected, these conditions simplify to

$$\frac{\partial}{\partial\vec{n}}\psi\Big|_{sc} = 0 \qquad \frac{\partial}{\partial\vec{n}}\phi_n\Big|_{sc} = 0 \qquad \frac{\partial}{\partial\vec{n}}\phi_p\Big|_{sc} = 0\ . \tag{6.53}$$

As a consequence of energy conservation (6.18) the energy current density $\vec{n}\vec{J}_u$ across an interface has to be continuous. Thus, the following boundary condition holds [62]:

$$\kappa_{sc}\frac{\partial}{\partial\vec{n}}T\Big|_{sc} = h(T_{ext} - T) + T(P_n^{sc} - P_n^{ext})\vec{J}_n\vec{n} + T(P_p^{sc} - P_p^{ext})\vec{J}_p\vec{n} \tag{6.54}$$

Here, h denotes the heat exchange coefficient of the interface and T_{ext}, P_n^{ext} , and P_p^{ext} represent the temperature and the thermoelectric powers of the material beyond the boundary. In case of electrically insulating boundaries ($\vec{J}_n\vec{n} = \vec{J}_p\vec{n} = 0$) this equation includes two interesting special cases: On the one hand, perfect thermal coupling to a heat sink ($h \to \infty$) results in the Dirichlet condition $T = T_{ext}$. On the other hand, thermally insulating surfaces ($h = 0$) are reflected by the homogeneous Neumann boundary condition $\frac{\partial}{\partial\vec{n}}T\Big|_{sc} = 0$.

Finally, it should be noted that the implementation in the simulator $\mathrm{DESSIS_{ISE}}$ neglects the convective terms proportional to $\vec{J}_n\vec{n}$ and $\vec{J}_p\vec{n}$ on the right hand side of equation (6.54). The consequences of this simplification will be addressed in section (7.3).

6.1.8 Summary of the underlying assumptions

In summary, the thermodynamic model is based on the following preconditions:

- The whole structure is partitioned into small cells which are assumed in local thermal equilibrium. Each of them has to be small enough that spatial variations of all variables within the cell are negligible. However, a lower limit of the cell size originates from the following facts:

  - Each cell must contain a sufficient number of particles such that statistical averages are well defined. For example, demanding a minimum of 10^6 particles requires a cell volume of $1\,\mu\mathrm{m}^3$ at a carrier concentration of $10^{18}\,\mu\mathrm{m}^{-3}$.
  - The mean free path must be small with respect to the cell size. Otherwise, the Fermi–Dirac distribution function cannot be applied to account for the k–space distribution of the carriers.

- The distributions of the carriers in the k–space are governed by the local Fermi–Dirac distribution functions. In particular, since the local equilibrium is described by two quasi–Fermi levels, intraband relaxation and intervalence band relaxation must be much faster than interband relaxation.

- External forces must vary much slowlier than the largest relaxation time.

- The electron gas, the hole gas, and the host lattice are in thermal equilibrium with each other, i. e. their temperatures are equal.

- The currents are driven by the gradients of the state variables.

6.2 Modeling material properties

The model sketched in the previous section includes several material coefficients. These are the dielectric constant ε, the heat capacity c_t, the effective intrinsic density $n_{i,\mathit{eff}}$, the generation/recombination rate R, as well as the transport coefficients μ_n, μ_p, P_n, P_p, and κ_t. The models for the recombination rate, the carrier mobilities, and the intrinsic density have to cover a variety of physical phenomena and therefore usually exhibit a very complex dependence on the state variables. The following sections are dedicated to a summary of the most important effects. A more detailed overview can be found in [145].

The minimum demand on physical models is *consistency* and *transparency* [146]. Physical models have to be consistent with basic physical theories, such as, e. g., electrodynamics or quantum mechanics, and have to be consistent with each other. In addition, their parameters are claimed to be transparent, i. e. they have to enable a physical interpretation and have to be accessible to an experimental determination. However, some technological phenomena are too complex to be described by physical models of general validity until now. In this case, empirical relations are commonly used.

6.2.1 Effective intrinsic density

The definition (6.30) relates the intrinsic carrier density

$$n_i(T) = \sqrt{N_c(T)N_v(T)}\exp\left(\frac{-E_g(T)}{2kT}\right) \tag{6.55}$$

to the band gap E_g and the effective densities of states (DOS). The latter are calculated from the DOS effective masses m_c and m_v by assuming a parabolic band structure [142]:

$$N_c = 2\sqrt{\frac{m_c kT}{2\pi\hbar^2}}^3 \quad ; \quad N_v = 2\sqrt{\frac{m_v kT}{2\pi\hbar^2}}^3 \tag{6.56}$$

In heavily doped semiconductors, many–particle interactions and potential fluctuations in an increasingly disordered system result in a shrinkage of the effective energy gap. Empirical formulas of band gap narrowing have been reported by several authors: Based on electrical measurements of the np product in npn–transistors, two different models were proposed by Slotboom and de Graaff [147, 148, 149] and del Alamo [150, 151]. Bennett and Wilson [152] carried out a statistical comparison of different absorption and luminescence data for n–type silicon, which is represented by an analytical fit [145]. Details about the analytical structure of the models available in $\mathrm{DESSIS_{ISE}}$ can be found in [101].

Since the active area of power devices exhibits a weak doping level, band gap narrowing has to be taken into account in the highly doped emitter regions only. Together with the minority carrier lifetime, it constitutes a significant effect on the emitter efficiency.

6.2.2 Recombination and carrier lifetime

On of the crucial parameters of the electrothermal model is the total recombination rate R which exhibits a very complex dependence on the distributions of the charge carriers in phase space. With indirect semiconductors, transitions from the conduction band minimum to the valence band maximum (or vice versa) proceed by the assistance of a third particle for momentum transfer. Radiative recombination can be neglected in view of the recombination via traps in the energy gap (Shockley–Read–Hall recombination), the Auger recombination, and the impact ionization. Their physical background and the commonly employed models are summarized in the following paragraphs.

Shockley–Read–Hall recombination

An empirical model for the indirect recombination via a trap level in the energy gap was first derived by Shockley and Read [153] and Hall [154]. They balanced the following processes (cf. fig. 6.1):

- An electron from the conduction band is captured by an empty trap (electron capture)
- A trapped electron is released into the conduction band (electron emission)
- A trapped electron is emitted by neutralizing a hole in the valence band (hole capture)
- An electron in the valence band is captured by an empty trap, thus generating a hole in the valence band (hole emission)

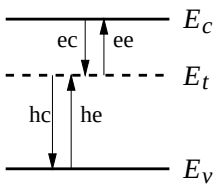


Figure 6.1: Capture and emission processes for recombination via a trap level.

The transition probabilities are expressed in terms of the trap concentration N_t and the capture cross sections $\sigma_{n,p}$. Their absolute values are measured by deep level transient spectrocopy (DLTS) which also provides information about the energy level of the recombination centers. An overview of experimental techniques can be found in [155].

Introducing the definitions

$$\tau_n := \frac{1}{\sigma_n \, v_{th} \, N_t} \quad \text{and} \quad \tau_p := \frac{1}{\sigma_p \, v_{th} \, N_t} \; , \tag{6.57}$$

the recombination rate finally becomes

$$R_{SRH} = \frac{np - n_i^2}{\tau_p(n + n_1) + \tau_n(p + p_1)} \; . \tag{6.58}$$

For the sake of simplicity, the transition rates have been expressed in terms of the total free carrier concentrations n and p. A more comprehensive analysis which also accounts for the distribution of the carriers in the conduction band and the valence band, respectively, was carried out by Shockley and Read [153]. In case of nondegenerate statistics, however, they obtain the same formula for the total recombination rate after integrating in the energy domain.

It should be mentioned that the equation (6.58) is based on the assumption that the above described capture and emission processes take place immediately. Consequently, the relaxation time t_r has to be small in view of both the variations $n/\dot{n}$ of the carrier concentrations as well as the "average occupation time" N_t/R_{SRH} of the traps. The former constraint arises from the assumption of steady–state conditions. It is violated if the transition rates lag behind the modulations of the carrier concentrations. In this case, the number of occupied traps has to be introduced as an independent variable which is governed by an additional balance equation [137, 138]. The second precondition, namely $t_r \ll N_t/R_{SRH}$, fails for very high injection levels. As the carrier concentration increases, the SRH formula predicts an infinitely increasing recombination rate. However, a saturation is expected in reality since the transition rates will be limited by the finite relaxation time for the capture and emission processes. This effect is included by an improved model proposed in [156].

It is worth mentioning some special cases of equation (6.58): Let $\Delta n := n - n_0$ and $\Delta p := p - p_0$ denote the excess carrier concentrations, i. e. the deviations from the equilibrium values n_0 and p_0, respectively. Within regions of charge neutrality, Δn and Δp have to be equal. We can therefore define the carrier lifetime by

$$\tau_{SRH} := \frac{\Delta n}{R_{SRH}} \; . \tag{6.59}$$

In highly n–doped regions ($n_0 \gg p_0, n_1, p_1, \Delta n$), the recombination rate (6.58) reduces to $R_{SRH} = \Delta n/\tau_p$. Hence, we can identify τ_p as the minority carrier lifetime in n–doped regions. Similarly, τ_n governs minority carrier recombination/generation in p–doped regions. However, under high injection conditions ($\Delta n \gg n_0, p_0, n_1, p_1$) the carrier lifetime is given by $\tau_{SRH} = \tau_n + \tau_p =: \tau_{amb}$ which is called the ambipolar carrier lifetime.

Doping dependence: As the doping concentration is increased, a decrease of the SRH lifetime is observed. This experimental result agrees with the theoretical expectation as the doping impurities introduce defects into the crystal lattice, thus enhancing the density of the recombination centers. Since a physically based model is not available yet, the doping dependence is usually represented by the so–called Scharfetter relation

$$\tau_{n,p} = \tau_{n,p}^{min} + \frac{\tau_{n,p}^{max} - \tau_{n,p}^{min}}{1 + \left(N_i / N_{n,p}^{ref}\right)^{\beta}} \ . \tag{6.60}$$

The variety of different experimental data (cf. [145, 101, 157, and references therein]) reflects the difficulty in finding a general model, which is in particular due to the fact that the carrier lifetimes and their doping dependence strongly depend on the specific processing conditions. The calibration of the carrier lifetime model with respect to the employed technology is therefore an important preparation step for reliable and predictive simulation results.

Temperature dependence: Similar to the doping dependence of the carrier lifetimes, there is no general agreement on their temperature dependence. For that reason, an empirical power law is commonly employed [157]

$$\tau(T) = \tau(300\,\mathrm{K}) \left(\frac{T}{300\,\mathrm{K}}\right)^{\alpha} . \tag{6.61}$$

Its parameters have to be adjusted by a separate calibration step.

Field dependence: The increased tunneling probabilities in a strong electric field can be modeled by a field–enhanced SRH recombination rate. Therefore, the carrier lifetime becomes field dependent. Since this effect is beyond the scope of this work, the reader is referred to the literature (e. g. [145]) for further details.

Surface recombination

At the interfaces between different materials and at the edges of the device, the presence of interface states increases the recombination/generation rate. Since the concentration of the recombination centers on an interface located at $z = 0$ can be written as $N_t(x, y, z) = N_S(x, y)\delta(z)$, the surface recombination rate is given by

$$R_{Surf} = \frac{np - n_i^2}{(n + n_1)/v_{sr,n} + (p + p_1)/v_{sr,p}} \tag{6.62}$$

which is of the same structure as the SRH formula. The surface recombination velocities $v_{sr,n}$ and $v_{sr,p}$ are in the order of $10\ \frac{\mathrm{cm}}{\mathrm{s}}$ to $10^7\ \frac{\mathrm{cm}}{\mathrm{s}}$.

Coupled defect level recombination

In order to optimize device performance, the carrier lifetime in power devices is adjusted by electron or ion irradiation or the diffusion of heavy metals, such as, e. g., platinum [158, 159]. This

processing step introduces a variety of different trap levels in the energy gap [160, 161]. In this case, the single–level Shockley–Read–Hall model constitutes a rough approximation only. A generalization for two coupled defect levels was derived by Schenk and Krumbein [162].

The total recombination rate is written as

$$\begin{aligned} R_{CDL} &= \left(\sqrt{R_{12}^2 + S_{12}} - R_{12}\right) \frac{\tau_{n1}\tau_{p2}(n+n_2)(p+p_1) - \tau_{n2}\tau_{p1}(n+n_1)(p+p_2)}{r_1 r_2} \\ &\quad + R_{SRH}(E_{t1}) + R_{SRH}(E_{t2}) \ . \end{aligned} \tag{6.63}$$

The first term accounts for the cascade transition from the conduction band to the valence band via the upper and the lower trap level. The additional single–level recombination rates via level 1 or level 2 are represented by the last two terms (cf. fig. 6.2).

It should be pointed out that — like the Shockley–Read–Hall formula — this model assumes instantaneous transitions between the different energy levels and is therefore limited to the same restrictions as mentioned above.

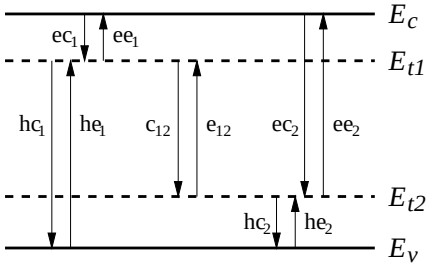


Figure 6.2: Capture and emission processes for two coupled defect levels (adopted from [162]).

Auger recombination

Auger recombination and generation is a three particle process: The energy of a direct transition between the conduction band and the valence band is transferred from or to another electron or hole in the conduction band or valence band, respectively. The net Auger recombination rate becomes:

$$R_{Au} = (C_n n + C_p p)\,(np - n_i^2) \tag{6.64}$$

Due to its cubic dependence on the carrier concentrations, Auger recombination gains increasing significance as the carrier concentrations rise above $10^{18}\,\mathrm{cm}^{-3}$.

Impact ionization

Additional electron–hole pairs are generated by free carriers that are sufficiently accelerated by a high electrical field, i. e. within a space charge region. If its width is larger than the mean free path between two ionizing collisions, impact ionization occurs. This effect leads to charge multiplication (avalanche generation) and is described by

$$G_{av} = \alpha_n n v_n + \alpha_p p v_p \tag{6.65}$$

where v_n and v_p denote the drift velocities. Various models of the ionization coefficients α_n and α_p have been proposed by different authors, e. g. [163, 164, 165, 166, 167].

Avalanche multiplication is one of the electrical breakdown mechanisms under reverse bias conditions and therefore limits the blocking capabilities of power devices.

6.2.3 Carrier mobility

The motion of carriers as a response to the driving affinities is limited by scattering events. The most important mechanisms in silicon are phonon scattering, impurity scattering, electron–hole scattering, and scattering by interface states. In case that the scattering events are independent and the $\vec{k}$–dependence of their relaxation times are equal, the corresponding scattering probabilities have to be summed up. The total low field mobility μ is therefore given by

$$\frac{1}{\mu_{low}} = \sum_i \frac{1}{\mu_i} \tag{6.66}$$

which is known as the Matthiessen rule. Although there are still some doubts about its accuracy, it is commonly used for mobility modeling. The most significant contributions arise from the bulk mobility μ_b accounting for phonon and impurity scattering, the surface contributions μ_{ac} and μ_{sr}, and the contribution μ_{eh} due to carrier–carrier scattering:

$$\frac{1}{\mu_{low}} = \frac{1}{\mu_b} + \left(\frac{1}{\mu_{ac}} + \frac{1}{\mu_{sr}} \right) + \frac{1}{\mu_{eh}} \tag{6.67}$$

For each contribution empirical formulas are proposed. The most common of them are briefly summarized below.

In strong electric fields, however, the linear dependence of the carrier drift velocities on the driving forces breaks down. Consequently, a saturation of the mobility is observed. This effect originates from the changes of the effective masses due to the non–parabolic band structure and deviations from the equilibrium distribution functions in high electric fields. An empirical model for the high–field saturation was proposed by Caughey and Thomas [168] and generalized by Canali [169] who introduced temperature dependent parameters. Thus, the field dependent mobility is written as

$$\mu(E_\parallel, T) = \frac{\mu_{low}(T)}{\left[1 + \left(\frac{\mu_{low}(T) E_\parallel}{v_{sat}(T)} \right)^{\beta(T)} \right]^{1/\beta(T)}} \tag{6.68}$$

where the low field mobility μ_{low} is given by equation (6.67).

Bulk mobility

Two mechanisms limit the bulk mobility, namely phonon scattering and ionized impurity scattering. As temperature rises, the lattice vibrations increase and cause a decrease of the mobility. Theoretical analysis predicts a temperature dependence of $T^{-3/2}$. On the other hand, the carriers moving faster at higher temperatures are less effectively scattered by the impurity potentials. Therefore, impurity scattering is expected to be proportional to $T^{3/2}/N_i$ [170]. An empirical model of both effects has been proposed by Caughey and Thomas [168] and extended by Masetti [171]:

$$\mu_b = \mu_{min1}\, \mathrm{e}^{-\frac{P_c}{N_i}} + \frac{\mu_{const} - \mu_{min2}}{1 + (N_i/C_r)^\alpha} - \frac{\mu_1}{1 + (C_S/N_i)^\beta} \tag{6.69}$$

The Masetti formula includes the "lattice mobility" $\mu_L = \mu_b(N_i = 0)$ and the "impurity scattering mobility" μ_I which is obtained in the limit $N_i \to \infty$. Although the interpolation is mostly empirical, it is one of the most commonly employed bulk mobility models.

Scattering mechanisms at surfaces and interfaces

Especially in the channel region of MOS structures, additional scattering mechanisms significantly decrease the mobility. The Lombardi model [172] accounts the scattering by surface acoustic phonons and surface roughness. To describe the former, the deformation potential theory of surface phonon scattering [173] was applied to express the electron mobility μ_{ac} in terms of the normal electric field:

$$\mu_{ac}(E_\perp, T) = T^{-1}\left(B\,\frac{T}{E_\perp} + C\,\frac{1}{E_\perp^{1/3}}\right) \tag{6.70}$$

In this relation, the effective thickness of the inversion layer is represented by the two parameters B and C.

Surface roughness scattering is modeled by [174]

$$\mu_{sr}(E_\perp) = \frac{\delta}{E_\perp^2}\ . \tag{6.71}$$

Since surface roughness scattering strongly depends on the crystal orientation and the specific processing conditions, the parameter δ has to be adjusted for each device technology.

Carrier carrier scattering

In case of high carrier concentrations a significant mobility decrease arises from electron–hole scattering. Applying the screening theory of Conwell and Weisskopf [175], Fletcher [176] and Choo [177] proposed the following model:

$$\mu_{eh} = \frac{D\cdot(T/T_0)^{3/2}}{\sqrt{np}\ \ln\left[1 + F\cdot(T/T_0)^2(pn)^{-1/3}\right]} \tag{6.72}$$

6.3 Numerical methods

Solving the coupled system of the partial differential equations (6.36), (6.37), (6.38), and (6.39) involves the following steps: Discretization in space and time transforms the partial differential equations to a nonlinear system of algebraic equations. The whole system is either simultaneously solved by the direct Newton algorithm or decoupled and subjected to the iterative Gummel scheme. In both cases, the update of the current vector of unknowns requires the solution of a linear system of equations. To tackle this problem, a variety of iterative and direct algorithms have been proposed in the literature (cf. e. g. [178, 143] for an overview).

6.3.1 Discretization

The space discretization is based on the box discretization method [179] (cf. fig. 6.3). Its greatest advantage is the ability to handle arbitrary geometries and nonuniform meshes. The perpendicular bisectors of the edges to the surrounding vertices form the boundary of a box around each vertex.

Then, the partial differential equations are integrated over the volume of the box, applying the theorem of Gauss to transform the integrated divergence operators. For example, Poisson's equation (6.39) becomes

$$\sum_{j \neq i} d_{ij} \varepsilon (\psi_i - \psi_j)/l_{ij} = V_i \, q \, (n_i - p_i + N^-_{A,i} - N^+_{D,i}) \quad (6.73)$$

where V_i represents the volume of the box (in two dimensions, the area of the box), d_{ij} and l_{ij} denote the area (in two dimensions, the length) of the perpendicular bisector and the length between vertex i and j, respectively.

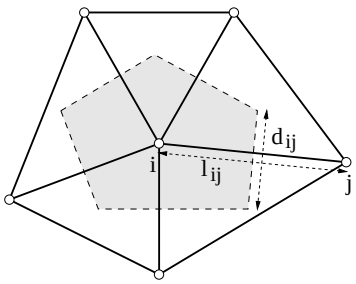


Figure 6.3: Box discretization around vertex i.

The time discretization scheme is based on the TR/BDF2 method [180]. A trapezoidal differentiation formula is applied to determine the step towards an artificial point of time $t_n + \gamma \Delta t_n$. Based on this interpolation, the real time step from $t = t_n$ to $t = t_{n+1} = t_n + \Delta t_n$ is calculated. Details can be found in [101, 180].

6.3.2 Solution of nonlinear equations

The solution of a nonlinear system of equations is based on the Newton algorithm. Summarizing the unknowns in a vector x, the update Δx_k of the iteration no. k is calculated from

$$f'(x_k) \Delta x_k = -f(x_k) \quad (6.74)$$

where the vector function f represents the system of equations.

Two iterative methods are commonly employed for solving the coupled system of the discretized equations (6.39), (6.36), (6.37), and (6.38): During the decoupled Gummel scheme [181], the update of the electrostatic potential is calculated from Poisson's equation, while the quasi–Fermi levels and the temperature are fixed. Then, the particle balances are solved for the new quasi–Fermi levels, keeping the electrostatic potential and the temperature fixed. Finally, the temperature distribution is updated and the cycle is repeated with the just computed values as the new initial solution. During one iteration, the above described Newton algorithm is therefore invoked four times where each dimension of x is equal to the number of grid nodes. This method works well if the coupling between the equations is weak.

A direct coupled solution is gained if all the unknowns are summarized in x and simultaneously updated by applying the Newton algorithm to the total system of equations. In this case, the coupling between the equations is included in the Jacobian f'. Therefore, this method is more stable and faster than the Gummel scheme, but requires the solution of a fourfold larger linear system of equations.

Obviously, various modifications of these methods can be imagined [101], for example, a Gummel scheme of two components, namely the heat flow equation and the coupled system of the electrical equations. In addition, the equations may be solved in different orders or in terms of another set of variables. However, choosing the variables n, p, T, and ψ and the direct and coupled solution of all four equations constitutes the fastest and most stable strategy.

Chapter 7

Analysis of Electrothermal Effects in Power Devices

The first part of this thesis has been dedicated to a comprehensive discussion of the capabilities of different internal laser probing techniques. The analysis supported the development of these methods and significantly improved the accuracy and the reliability of the experimental results.

Thus, the availability of internal probing techniques has opened up promising chances for the development of power devices and the validation of electrothermal simulation models. While measuring terminal behavior during static operation or transient switching has been state of the art in device characterization for many years, these novel techniques reveal the internal carrier concentration and temperature distribution. Thus, they provide valuable experimental information about the physical processes in the interior of semiconductor devices and therefore facilitate a much more comprehensive validation of the theoretical models which are employed for electrothermal device simulation. This chapter is focussed on a thorough discussion of major electrothermal models, in particular those governing the electrical behavior in the forward conducting state and the corresponding heat generation. Based on various experimental evidence, we will judge the suitability of these models for the predictive simulation of power devices.

It should be mentioned that the analysis of the measurement processes incorporates an electrothermal device simulation (cf. section 3.1) whose calibration, in turn, is based on the results of these probing techniques. However, it is evident that a first guess calibration (cf. e. g. [126]) is enabled by the well–established electrical characterization methods supplemented by the currently available optical measurement techniques [57, 49, 96, 58, 92]. A recalibration with enhanced accuracy can be carried out by employing the optimized setups of the existing methods and the additional probing techniques which are currently under development.

7.1 Electrical behavior

Regarding the physical mechanisms which affect the electrical behavior of power devices, we will now address the following questions: First, which kind of information can be extracted from the experimental results? And second, does the electrothermal model accurately predict the distribution and the motion of the carriers? In the discussion, the experimental reference comprises

the measured carrier concentration profile as well as the terminal characteristics during static and transient operating conditions. As an example of bipolar devices, the investigations apply to a pin diode that has been subjected to platinum diffusion for lifetime control.

7.1.1 Carrier distribution under high injection conditions

As the first experimental result gained from the optical probing techniques, the carrier distribution in the weakly doped drift region of the device have been ascertained from laser absorption measurements. It will be demonstrated in this section that this measured carrier concentration profile provides valuable information about carrier lifetimes and emitter efficiencies.

In the forward conducting state, power devices are operating under high injection conditions. In this case, the carrier distribution is approximately governed by the ambipolar diffusion equation

$$\Delta n - \frac{n}{L_D^2} = 0 \ . \tag{7.1}$$

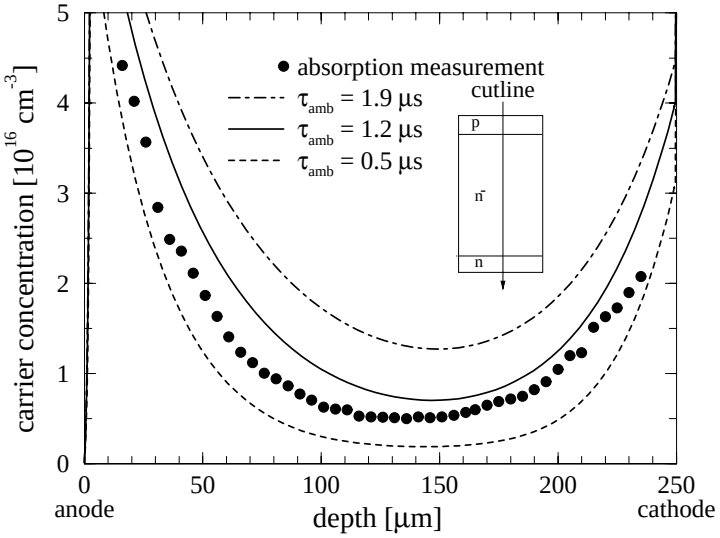


Figure 7.1: Carrier distribution in a pin diode (current density $150\ \mathrm{A/cm^2}$). The simulation results are calculated with different ambipolar carrier lifetimes [78].

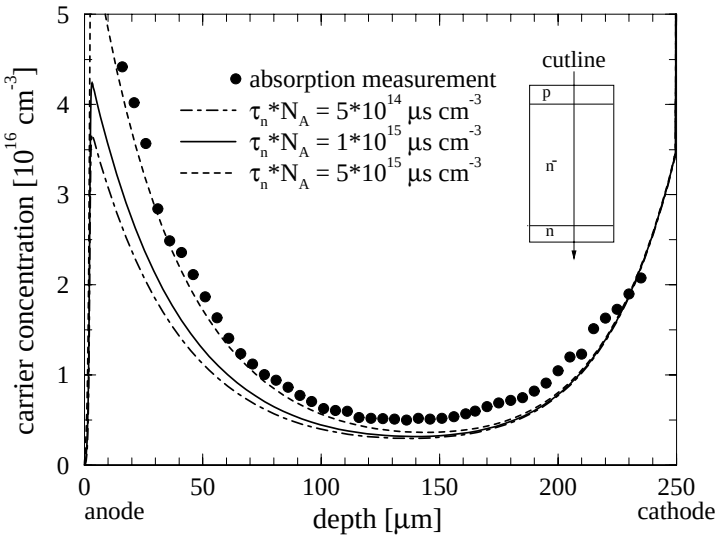


Figure 7.2: Carrier distribution in a pin diode (current density $150\ \mathrm{A/cm^2}$). The simulation results are calculated with different minority carrier lifetimes τ_n in the p–emitter [78].

In one dimension, it has the general solution

$$n(x) = n_{min} \cosh\left(\frac{x - x_{min}}{L_D}\right) \quad . \tag{7.2}$$

Hence, the 1D carrier distribution represents a cosh–function whose curvature reveals the ambipolar diffusion length L_D. Exploiting the relation

$$L_D = \sqrt{\tau_{amb}\, 2\, \frac{kT}{q}\, \frac{\mu_n \mu_p}{\mu_n + \mu_p}} \; , \tag{7.3}$$

we can directly determine the ambipolar carrier lifetime τ_{amb}. As it can be seen from figure (7.1), the uncertainty of the extracted carrier lifetime τ_{amb} is less than a factor of 2, which constitutes excellent error bounds concerning the measurement of carrier lifetimes.

As mentioned above, a clever adjustment of the charge carrier distribution in a pin diode helps to improve the trade–off between the forward conductivity and the turn–off behavior. Hence, we are in particular interested in the symmetry of the carrier concentration profile which is affected by the anode and cathode emitter efficiencies. For example, as the minority carrier lifetime of the electrons in the anode increases, the anode emitter efficiency is enhanced, thus raising the carrier concentration in the upper region of the device (cf. fig. 7.2). However, it can not be discriminated from the carrier distribution whether a higher emitter efficiency results from an increased doping dose or an enhanced minority carrier lifetime. The distinction becomes possible only if the additional information from the transient temperature distribution is included in the investigation. Further details are discussed in section (7.2.2). Nevertheless, in view of the fact that the figures are plotted on a *linear* scale, it is evident that the calibrated simulation is able to reproduce the carrier concentration profile with very good accuracy.

7.1.2 Forward terminal characteristics

In the forward conducting state, the total voltage drop at the typical rating currents of $50\,\mathrm{A/cm^2}$ to $150\,\mathrm{A/cm^2}$ is in the order of about 2 V. A significant fraction of it — typically 50 % to 80 %

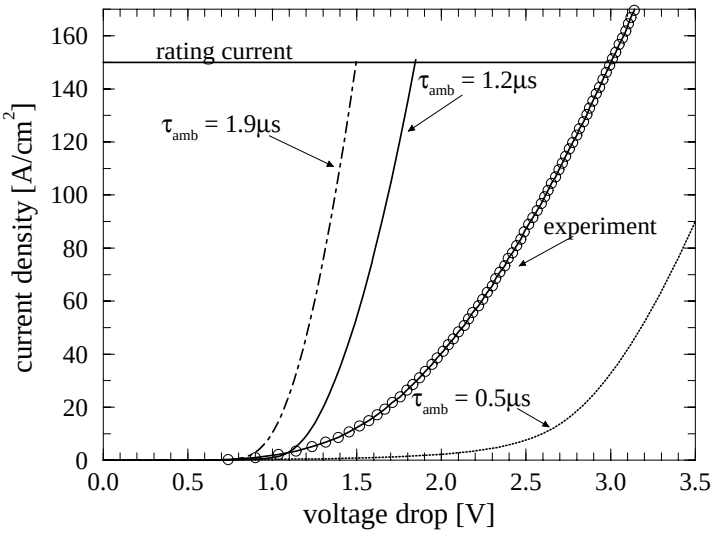


Figure 7.3: Forward characteristics of a pin diode. The simulation results are calculated with different ambipolar carrier lifetimes [78].

— drops across the substrate region operating under high injection conditions. Consequently, the forward characteristics is shifted towards lower voltage drops (cf. fig. 7.3) as the ambipolar carrier lifetime increases, since the smaller recombination rate results in a higher overall carrier concentration, thus enhancing the substrate conductivity.

The carrier concentration profile of the investigated device reveals an ambipolar carrier lifetime of about $1.2\,\mu\mathrm{s}$ (cf. fig. 7.1). However, the corresponding simulated forward characteristics exhibits a smaller voltage drop than the measured characteristics. Since the mobilities in weakly doped substrate material are well known, i. e. the experimental uncertainty is too small to account for the difference, the real device may include additional resistive mechanisms in the heavily doped emitters or at the contacts which are not covered by the simulation model. Regarding the different values of the voltage drop, the additional resistance should be about several $\mathrm{m\Omega\,cm^2}$, which is a reasonable order of magnitude. For that reason, the simulation results can be considered satisfying as they *simultaneously* reflect the measured carrier concentration profile and the forward characteristics with acceptable accuracy.

7.1.3 Transient switching conditions

As the last, but not least important, issue of the electrical properties, the transient switching characteristics will be discussed in this section. In practical applications, one of the most important features of switching diodes is the reverse recovery behavior, i. e. the turn–off behavior when a negative voltage is suddenly applied to the forward biased device. In this case, the initial slope of the current transient is governed by the wire inductivities [6]. Unfortunately, the calibrated simulator is not able to reproduce the measured switching behavior (cf. fig. 7.4). Consequently, we also observe a large discrepancy of the extracted charges, i. e. the integrals over the period of negative current: The calculated and the measured results differ by nearly a factor of 3. This means that within the real device a larger fraction of the stored charge recombines during the switching process.

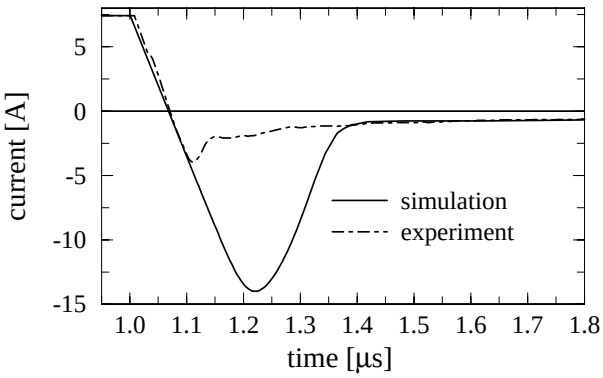


Figure 7.4: Reverse recovery behavior of a pin diode (reverse current slope $1.2\,\mathrm{kA/cm^2}\mu\mathrm{s}$) [182].

A possible reason could be an unproper modeling of the recombination via the trap centers. As platinum introduces a variety of impurity levels in the band gap [160, 158, 161] the single–level Shockley–Read–Hall formula (cf. eq. 6.58) constitutes a very rough approximation only. Employing the coupled–defect–level (CDL) model (cf. eq. 6.63 in section 6.2.2), it would be possible to take into account at least the two most significant levels of platinum. This model would therefore be preferable although it comprises a lot of additional parameters most of which are not known with sufficient accuracy. However, it has been demonstrated that including the CDL–model does not result in a better description of the reverse recovery behavior [182]. The major

reason is that both the SRH–model and the CDL–model are valid in case of stationary operating conditions only, since the capture and emission processes of the trap centers are assumed to take place immediately. To overcome this restriction, several approaches to allow for the dynamic trapping have been reported (e. g. [137, 138, 183, 184, 185]). The corresponding simulation results exhibit a delayed turn–off process and a significantly reduced reverse recovery current [183, 184, 185], thus constituting a much better description of the switching behavior.

7.2 Heat generation mechanisms

The interaction of the charge carriers with the host lattice partly converts the electrical work to heat, thus raising the temperature within the structure. After the simulation results for the electrical properties have been discussed above, it will now be investigated whether the electrothermal model accurately reflects the thermal behavior of the device.

It has been the subject of many discussions in literature which formula represents the heat generation rate of a self–consistent electrothermal extension of the drift–diffusion model. A variety of different models have been reported (e. g. [186, 187, 188, 189, 62]). In addition, the electrothermal models which are implemented into commercially available device simulators [101, 131, 132] display considerable differences. For that reason, this section addresses the question of which heat generation mechanisms are most significant in semiconductor power devices and therefore have to be included in the heat flow equation. The investigations are carried out by employing the general purpose device simulator $\mathrm{DESSIS_{ISE}}$ into which the model proposed in [62] is implemented.

7.2.1 Devices without lifetime control

First, we will focus on devices that have not been subjected to special processing steps for lifetime control. As a representative example, a Non–Punch–Through IGBT (NPT–IGBT, cf. fig. 1.2) is investigated, in particular its operation with ohmic load and under short circuit conditions.

At the top surface, adiabatic boundary conditions are assumed since the heat flow across the bond wires can be neglected. The mounting onto the copper substrate is reflected by a boundary condition of the third kind at the rear surface of the device (cf. eq. 6.54). However, the model implemented into the simulator $\mathrm{DESSIS_{ISE}}$ neglects the convective terms and assumes the following simplified boundary condition:

$$\kappa_{sc} \frac{\partial}{\partial \vec{n}} T \Big|_{sc} = h(T_{ext} - T) \tag{7.4}$$

At first, the temperature evolution during transient switching with ohmic load is investigated. During the on–state, the device conducts a current of an average density of $150\,\mathrm{A/cm^2}$. On top of the structure, the simulated temperature distribution (cf. fig. 7.5) exhibits the greatest temperature rise which is accurately confirmed by Internal Laser Deflection measurements. However, a considerable qualitative difference is observed at the bottom of the device: Whereas the simulated temperature profile continuously drops to minimum temperature rise at the rear surface, the experimental result exhibits a second temperature maximum there. It obviously originates from an additional heat source which is not included in the simulation and will be investigated thoroughly in section (7.3).

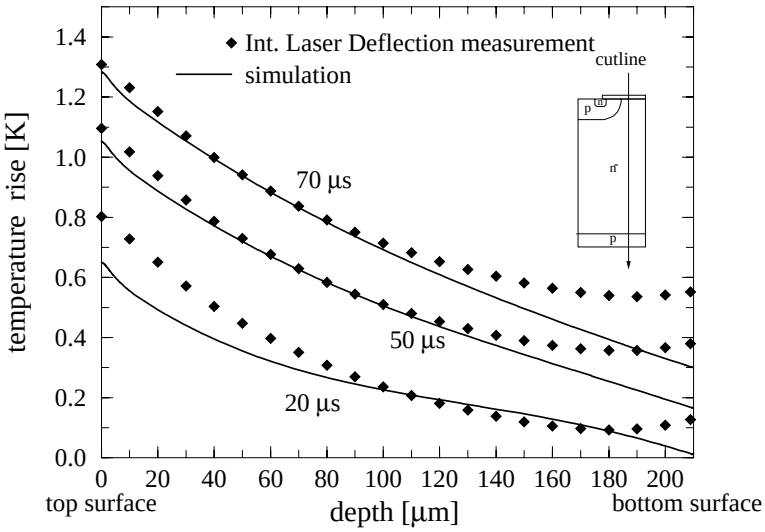


Figure 7.5: Temperature rise in an NPT–IGBT during a current pulse of $150\,\mathrm{A/cm^2}$ *(pulse duration* $70\,\mu\mathrm{s}$*).*

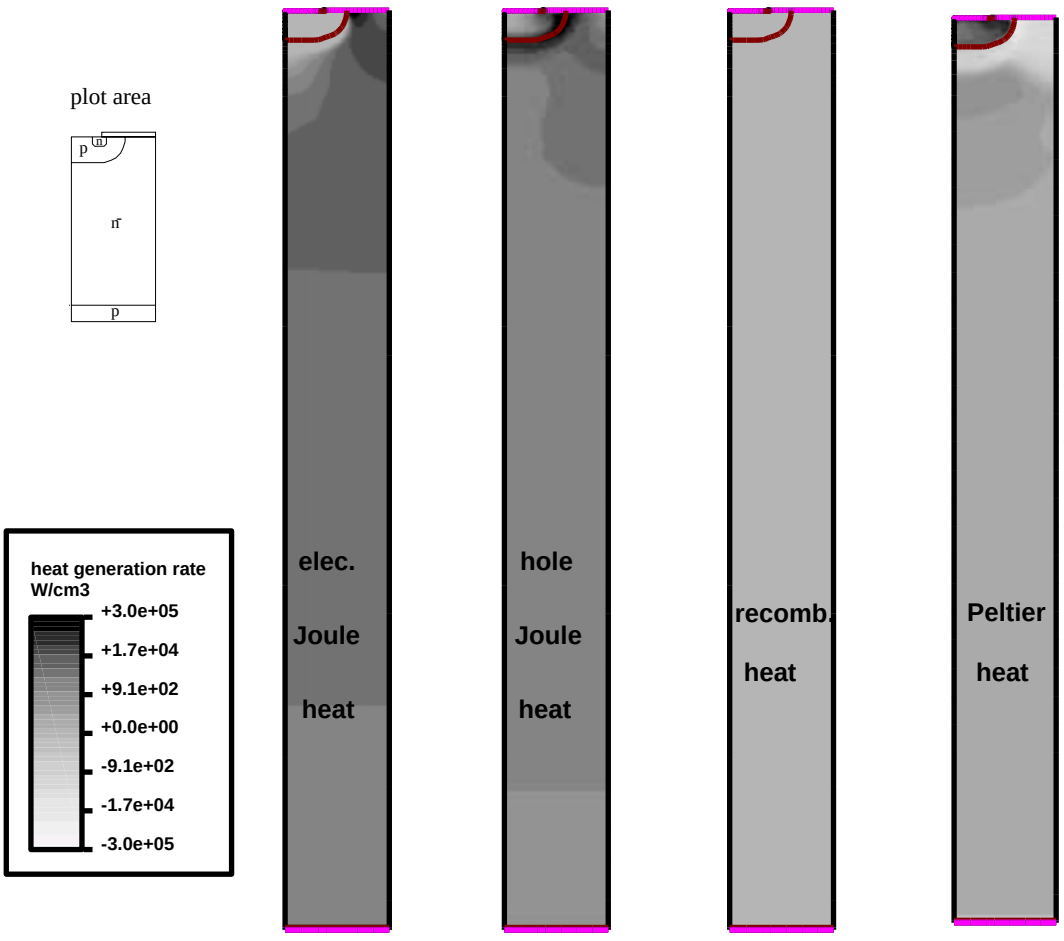


Figure 7.6: Heat generation within an NPT–IGBT operating with ohmic load at a current density of $95\,\mathrm{A/cm^2}$*. Due to symmetry only the right hand half of the IGBT cell is shown. Note, that a vanishing heat generation rate is plotted in gray. Hence, areas in dark gray refer to a positive heat generation rate whereas areas in light gray correspond to a negative heat generation rate.*

The major effect on the temperature profile arises from Joule heating of the electrons in the MOS channel and Joule heating of the holes in the space charge region around the reverse biased junction of the p–well and the n^-–substrate (cf. fig. 7.6). Due to the large carrier lifetime, the recombination heat in this kind of devices is negligible. Another contribution originates from Peltier heating. Since the Peltier coefficients depend on the carrier density, a positive or negative heat generation arises from a current flow across a gradient of the carrier concentration. Thus, contributions of both signs are observed at the p–well/n^-–substrate junction. However, their absolute values are about 1 to 2 orders of magnitude smaller than the Joule heat, which therefore constitutes the governing effect on the temperature evolution within the device.

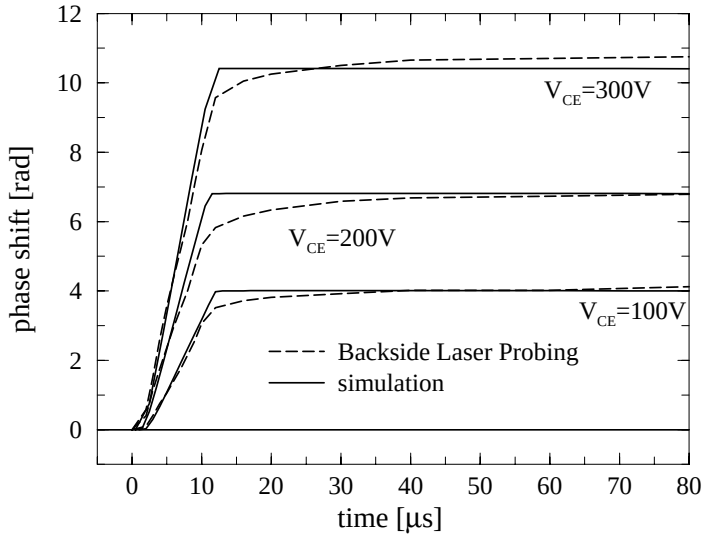


Figure 7.7: Phase shift signal (stored heat) for Backside Laser Probing of an NPT–IGBT operating under short circuit conditions with various collector–emitter voltages.

As the temperature rise during short circuit operation is too large to be detected by Internal Laser Deflection measurements, Backside Laser Probing has been employed to ascertain the heat which is stored within the device (cf. eq. 5.3). Since the Joule heats exhibit a quadratic dependence on the current density they are expected to be the dominant heat generation mechanism during short circuit operation. As it can be seen from fig. (7.7), the calculated phase shift signal and with it the dissipated heat is in excellent agreement with the experimental data, thus giving evidence that the implemented electrothermal model accurately reproduces the temperature profile in case of operation with a large power dissipation.

7.2.2 Devices subjected to heavy metal diffusion

In order to improve the trade–off between the turn–off losses and the voltage drop during forward conduction, power devices may be subjected to specific processing steps for lifetime control, e. g. proton or helium irradiation or the diffusion of heavy metals such as gold or platinum [158, 159, 190, 191].

As a representative of this kind of devices, a pin diode with platinum doping is investigated. The temperature profile displays two peaks (cf. fig. 7.8) which are most clearly visible immediately after turn–on. A narrow peak is located at the top surface while the broader maximum can be found in the interior of the structure. Analyzing the various contributions to the total heat generation rate (cf. fig. 7.9) reveals that the temperature rise in the bulk of the device originates from Joule heating

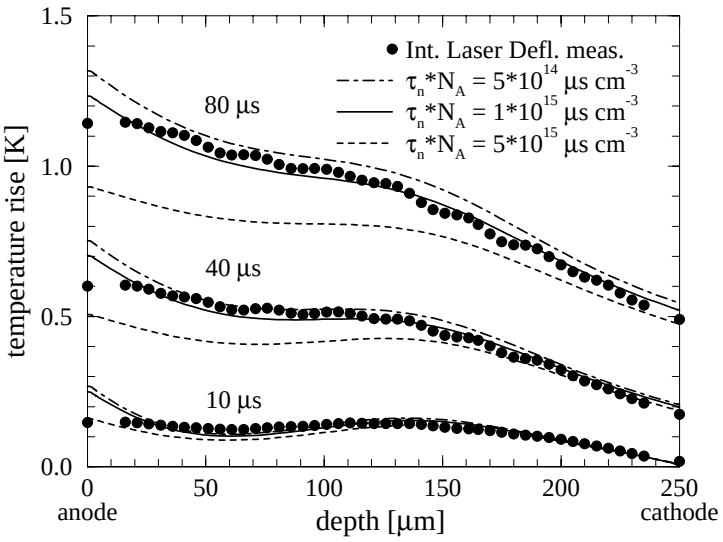


Figure 7.8: Temperature distribution within a pin diode subjected to a current pulse of $150\,\mathrm{A/cm^2}$. *The simulation results are calculated with different minority carrier lifetimes* τ_n *in the anode.*

of electrons and holes. Since the carrier distribution is at its minimum in the interior of the device (cf. fig. 7.1) the corresponding Joule heat attains a maximum there. On the other hand, the most significant heat generation at the cathode and anode junctions is due to recombination heat since the carrier lifetime in the highly doped emitters is much smaller than the ambipolar carrier lifetime in the bulk. Consequently, the resulting temperature peak reveals valuable information about the minority carrier lifetime in the emitter region (cf. fig. 7.8). At the cathode side, recombination heat can not effect a local temperature rise due to the large heat capacity of the copper substrate the sample is mounted onto.

Similar results are obtained for a Punch–Through IGBT (PT–IGBT) (cf. fig. 8.9) incorporating

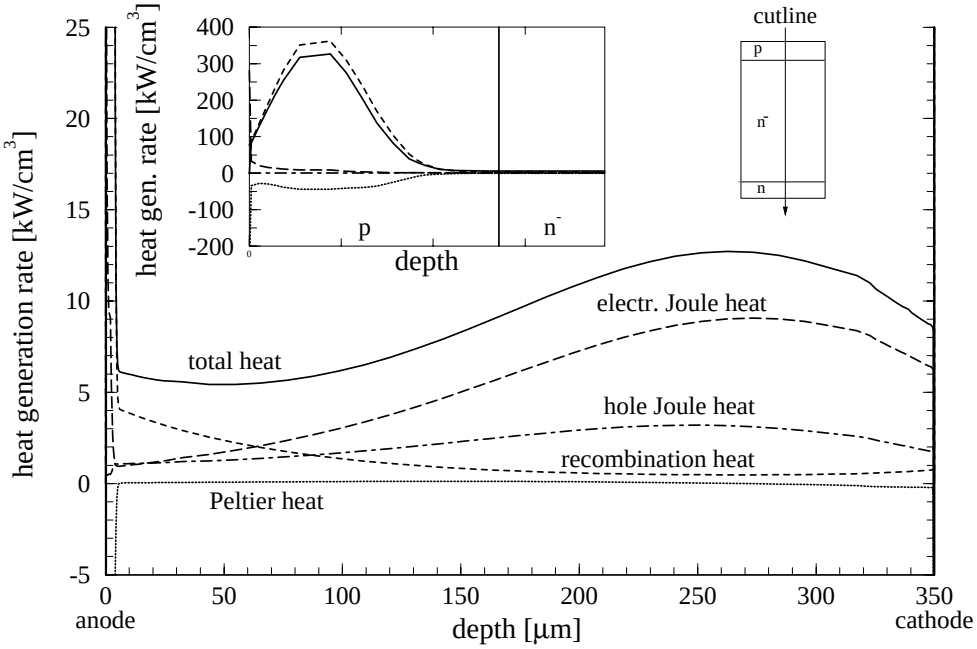


Figure 7.9: Heat generation within a pin diode subjected to a current pulse of $150\,\mathrm{A/cm^2}$. *The insert depicts the heat generation rate at the* $\mathrm{p/n^-}$ *junction.*

a platinum doping for lifetime control. The temperature profile (cf. fig. 7.10) is governed by Joule heating of the electrons in the channel and the holes in the space charge region of the p–well. In addition, a shallow temperature peak is observed at the n–buffer which originates from recombination heat. Unfortunately, this inhomogeneity is below the resolution of Internal Laser Deflection measurements. Nevertheless, the calculated temperature profile is in excellent agreement with the experimental result, thus validating the electrothermal model.

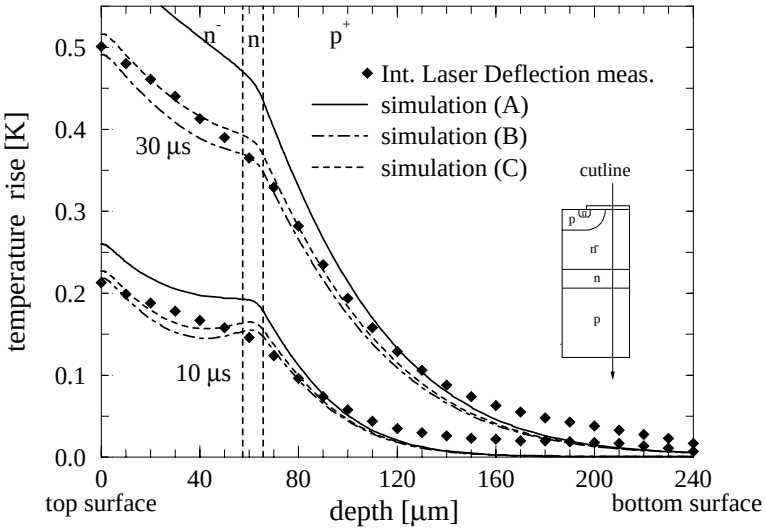


Figure 7.10: Temperature evolution within a PT–IGBT with lifetime control by platinum doping. The device is subjected to a current pulse of $110\ \mathrm{A/cm^2}$ *(pulse duration* $70\ \mu\mathrm{s}$*). The simulations are carried out with different parameters of the carrier lifetime model (cf. table 8.1 on page 129). Set C represents the calibrated model.*

7.3 Peltier heating

As mentioned above, an additional heat source at the bottom of the previously investigated NPT–IGBT must be present which effects the second temperature peak that is detected by the Internal Laser Deflection measurements (cf. fig. 7.5). As it can be seen from a closer look at the heat generation in the vicinity of the rear p–emitter (cf. fig. 7.11), the total heat generation rate is negative and the most significant contribution arises from Peltier heating due to the current flow across a forward biased pn–junction. However, an additional heat generation takes place at the metal/semiconductor interface. Figure (7.12) illustrates the various contributions of the carriers flowing across the pn–junction and the metal/semiconductor interface. An electron in the conduction band moving across the $\mathrm{n^-/p}$ junction gains energy which is extracted from the crystal heat by the absorption of phonons. At the transition into the metal contact, the energy difference to the Fermi level is released, thus giving off heat. The net effect is a positive heat generation rate. Similarly, the holes which are emitted from the Fermi level at the semiconductor/metal interface gain energy during the transition to the valence band and also while flowing across the $\mathrm{n^-/p}$ junction. Since the p–emitter is much thinner than the thermal diffusion length, the temperature profile is affected by the sum of all heat generation rates, which is either positive or negative if the electron contribution is greater or smaller than the hole contribution, respectively.

The numerical model allows for the Peltier heating in the interior of the device by an appropriate contribution in the heat generation rate of the heat flow equation (cf. section 6.1.6). However, Peltier heating at the metal contact needs to be reflected in the thermal boundary conditions. Due

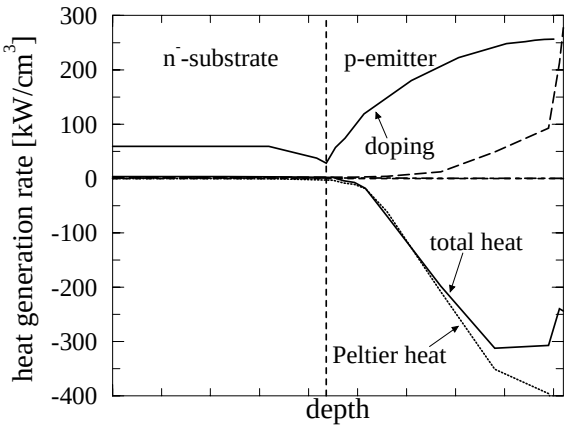


Figure 7.11: Heat generation at the rear emitter of an NPT–IGBT operating with ohmic load at a current density of $150\ \mathrm{A/cm^2}$.

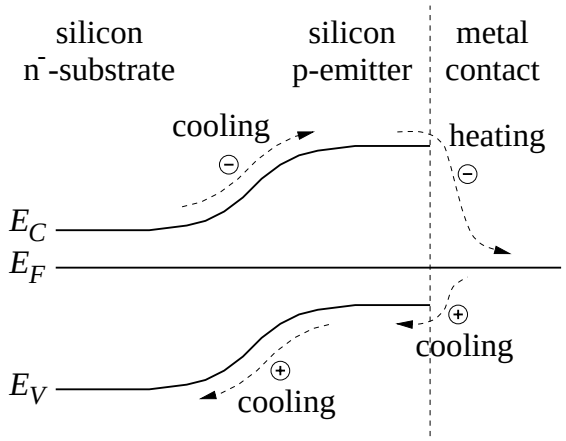


Figure 7.12: Particles gaining or giving off energy at the metal/semiconductor interface and the p–emitter/n^-*–substrate junction.*

to the simplification of the latter [101] (cf. eqs. 6.54 and 7.4) the corresponding heat generation effect is not covered by the present implementation of the electrothermal model. If the Peltier coefficients in the metal are neglected since they are about three orders of magnitude smaller than those of silicon, the missing contribution can be calculated analytically. As the electron current density is much greater than the hole current density it is positive in sign and amounts to

$$H_{Peltier}^{Contact} = TP_n^{sc}\vec{J}_n\vec{n} + TP_p^{sc}\vec{J}_p\vec{n} = +15\,\frac{\mathrm{W}}{\mathrm{cm}^2} \tag{7.5}$$

at the current density under consideration [123] (Note that P_n is negative). Although it is interesting that this produces exactly the necessary amount of heat to explain the observed difference between the calculated and the measured temperature profile, there are still some doubts since, for example, any additional resistances might also give rise to an additional heat generation.

Another experimental evidence is gained from Backside Laser Probing. The examination is based on the fact that the integral contribution of Peltier heating vanishes under stationary operating conditions, if the recombination is negligible and the temperature distribution is approximately homogeneous:

$$\begin{aligned}
\int_V H_{Peltier}\,dV &= -\int_V \vec{J}_n T\vec{\nabla}P_n\,dV - \int_V \vec{J}_p T\vec{\nabla}P_p\,dV \\
&= -\int_{\partial V} P_n T\vec{J}_n\,d\vec{\sigma} + \int_V P_n T\nabla\vec{J}_n\,dV \quad - \int_{\partial V} P_p T\vec{J}_p\,d\vec{\sigma} + \int_V P_p T\nabla\vec{J}_p\,dV \\
&= -TP_n^{Contact}\int_{\partial V}\vec{J}_n\,d\vec{\sigma} = 0
\end{aligned} \tag{7.6}$$

In other words, as long as we are merely interested in the integral of the temperature distribution we can add a possibly missing contribution of the Peltier heat by omitting the Peltier effect at all.

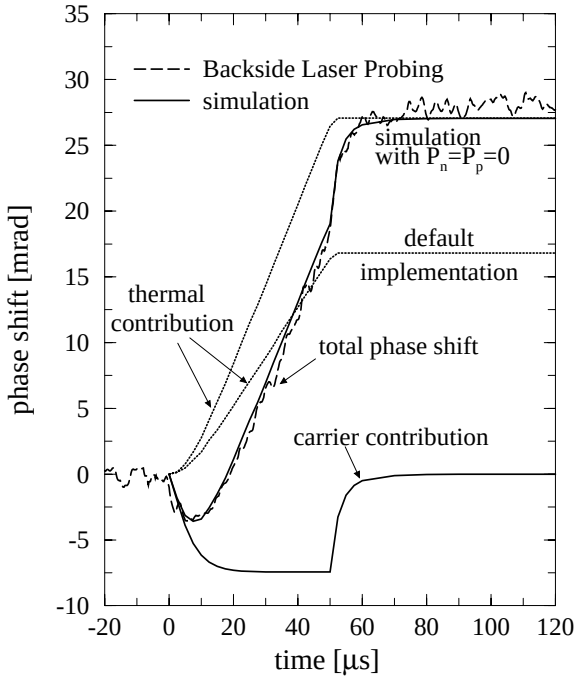


Figure 7.13: Phase shift signal (stored heat) from Backside Laser Probing of an NPT–IGBT operating at low power dissipation (collector–emitter voltage 2 V, pulse duration $50\ \mu s$*).*

Operating the NPT–IGBT at a small collector–emitter voltage of 2 V, the heat within the device is extracted from the thermal contribution to the phase shift signal of Backside Laser Probing (cf. fig. 7.13). On the one hand, the default model of the calibrated simulator (cf. fig. 7.5 and 7.7) yields a thermal contribution which is about 40 % too small (cf. fig. 7.13). On the other hand, merely setting the Peltier coefficients to zero results in an excellent agreement with the measured phase shift signal, thus providing another experimental evidence that the missing heat contribution originates from the neglected Peltier heat at the semiconductor/metal interface.

7.4 Contact resistances and thermal boundary conditions

Finally, a specific issue of the boundary conditions is discussed, namely a possible electrical or thermal series resistance. While the former is due to a contact resistance, the latter reflects the thermal coupling to the copper substrate which can be regarded as a heat reservoir.

It has been demonstrated above that the temperature peaks near the emitters of bipolar devices originate from recombination heat and therefore reveal information about the corresponding minority carrier lifetimes in the emitters. However, the temperature maximum may also arise from the Joule heating due to a contact resistance. This assumption would be supported as the voltage drop of the simulated forward characteristics is smaller than the measured one (cf. fig. 7.3). However, if the corresponding series resistance is included at a contact with adiabatic thermal boundary conditions, the simulation produces a large temperature rise which is not in agreement with the measured temperature profile [182] (cf. fig. 7.14). Hence, we conclude that a significant contact resistance can only be present at contacts that are thermally coupled to a heat reservoir. In this case, most of the dissipated power is absorbed by the heat reservoir and the simulated temperature profile agrees with the experimental result (cf. fig. 7.14). The ambiguity of whether a contact resistance has to

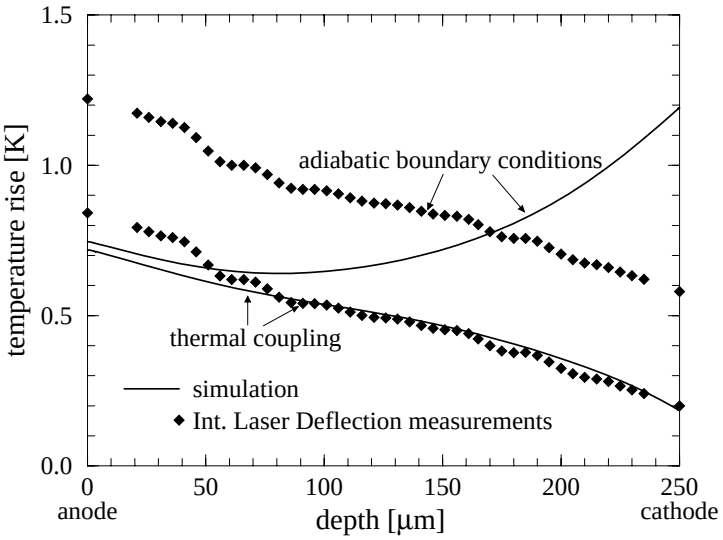


Figure 7.14: Temperature evolution within a pin diode (current density $150\,\mathrm{A/cm^2}$). At the cathode, a contact resistance is assumed which is determined from the difference of the simulated and the measured voltage drop [182]. The integration constant for the measured temperature profile is extracted from different boundary conditions.

be assumed in the simulation cannot be clarified by Internal Laser Deflection measurements only. The reason is that they are sensitive to the refractive index *gradients* and therefore require the a priori knowledge of the boundary conditions to determine the integration constant for temperature profile. However, as the absolute temperature rise significantly depends on the assumed boundary condition (cf. fig. 7.14), the distinction should become possible by employing a technique which directly detects the temperature rise. For example, Fabry–Perot transmission measurements (cf. section 2.2.4) are a suitable method.

The effect of a contact resistance on the forward characteristics is plotted in figure (7.15). While the simulated voltage drop at a specific current density can be adjusted by assuming an additional series resistance, the curvature of the total characteristics is not accurately reproduced. We therefore have to conclude that the observed difference cannot be due to a constant contact resistance only. For an improved description of the forward characteristics, we rather have to include additional effects, e. g. a current–dependent contact resistance, an injection–dependent carrier lifetime, or a current–dependent emitter efficiency.

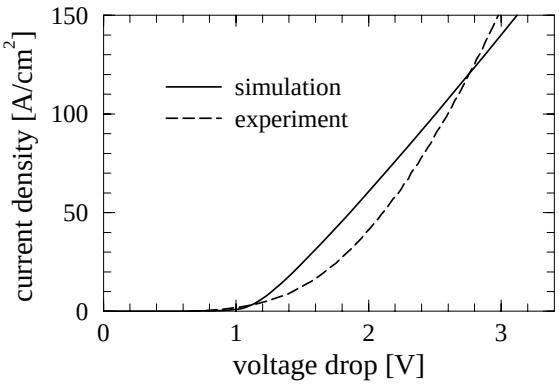


Figure 7.15: Forward characteristics of a pin diode with a contact resistance at the cathode.

7.5 Summary

The experimental results of internal laser probing techniques enable a more comprehensive validation of the electrothermal device simulation models. The major results are:

- The carrier concentration profile in the intrinsic region of power devices reveals the ambipolar carrier lifetime under high injection conditions and provides a first estimation of the

adjacent emitters' efficiencies.

- Calibrating the carrier lifetime model with respect to the carrier concentration profile, the simulated forward characteristics of bipolar devices with a heavy metal doping exhibits a smaller voltage drop than the measured characteristics. Although a contact resistance can explain the difference at a specific current density, the curvature of the whole characteristics originates from additional effects which are not covered by the present model, e. g. an injection–dependent carrier lifetime or a current–dependent contact resistance.

- Employing stationary recombination models, e. g. the single–level SRH–model or the coupled–defect–level (CDL) model, the simulated reverse recovery current of pin diodes is much larger than the measured one. Dynamic models which take into account the transient capture and emission kinetics of the impurities promise a better description of the switching behavior.

- Joule heating of electrons and holes is the most significant heat generation mechanism in power devices. It constitutes the major contribution to the total heat generation rate in the channel region of MOS structures, in the space charge region of reverse biased pn–junctions, and in the intrinsic drift region which is under high injection conditions. In addition, recombination heat originates a significant contribution in devices with lifetime control, in particular in the vicinity of heavily doped emitters. Thus, the local temperature rise reveals information about the minority carrier lifetime in the adjacent emitter.

- Although the total heat generation is predominantly affected by the Joule heats and the recombination heat, Peltier heating constitutes an additional contribution, especially at the n^-–substrate/p–emitter junction of NPT–IGBTs and at the interface between the semiconductor and the metal contact. However, the latter contribution is not covered by the model implemented into the simulator $\mathrm{DESSIS_{ISE}}$ due to a simplification of the thermal boundary conditions. Various results from sensitive laser probing techniques provide experimental evidence that this simplification is the reason for the observed difference between the measured and the simulated temperature profile, thus validating the rigorous model reported in [62].

- Despite this simplification the incorporated model is able to predict the electrothermal behavior of power devices with satisfying accuracy. Since Peltier heating is of minor importance during operation with large power dissipation, no differences are observed between the simulated and the measured results for short circuit operating conditions.

Chapter 8

Model Calibration for Predictive Simulation

Since it has been demonstrated that the electrothermal model is suitable for the accurate and predictive simulation of power devices, this chapter is dedicated to the question of model calibration. As the previously described internal laser probing techniques provide valuable information about the internal carrier concentration and temperature profiles, the calibration procedure can be founded on a variety of experimental results, thus promising a more general validity. As representative examples, we will discuss the model calibration of two different types of commercially available IGBT devices, namely a 1200 V NPT–IGBT (cf. fig. 1.2) and a test structure of a 600 V PT–IGBT (cf. fig. 8.9) which has been subjected to platinum diffusion for lifetime control.

The following sections address the most important parameters and their effects on the various characteristics. Special attention is paid to the forward terminal characteristics and the internal carrier concentration profile, which reflect most of the required information for the model calibration. As the final check of the calibrated model, the simulated temperature profiles during transient switching are compared with the corresponding measurement results. The entire calibration sequence is summarized in the last section of this chapter.

8.1 Defining the structure

As a first step, the composition of the investigated structure has to be known, i. e. the layout, the material distribution, and the doping profiles. Since 3D simulations still require a high computational effort, device engineers are interested in meaningful 2D simulations for product development and optimization. However, the complexity of today's semiconductor devices requires a careful choice of a representative cross section and a suitable coordinate system.

Second, the doping profiles have to be known. Although specific measurement techniques, e. g. secondary ion mass spectrometry (SIMS) or spreading resistance measurements, provide detailed quantitative information, the remaining uncertainty is still quite large, admitting significant variations of many characteristic features of the device performance. Hence, the following paragraphs will also address the question of which experimental results facilitate a fine adjustment of the doping distribution the simulations are based on.

8.1.1 Choosing the simulation domain

The structure of commercially available IGBT chips is illustrated in figure (8.1). While the p–emitter at the bottom surface is manufactured by a non–structured diffusion process, the top surface comprises an array of MOS devices. Thus, the whole IGBT chip can be thought to be composed of many thousands of identical cells. Each of them is formed by a cuboid with a square base. On the other hand, the MOS structures are approximately circular in shape since the source of the MOSFET and the p–well defining the channel length are formed by underdiffusion.

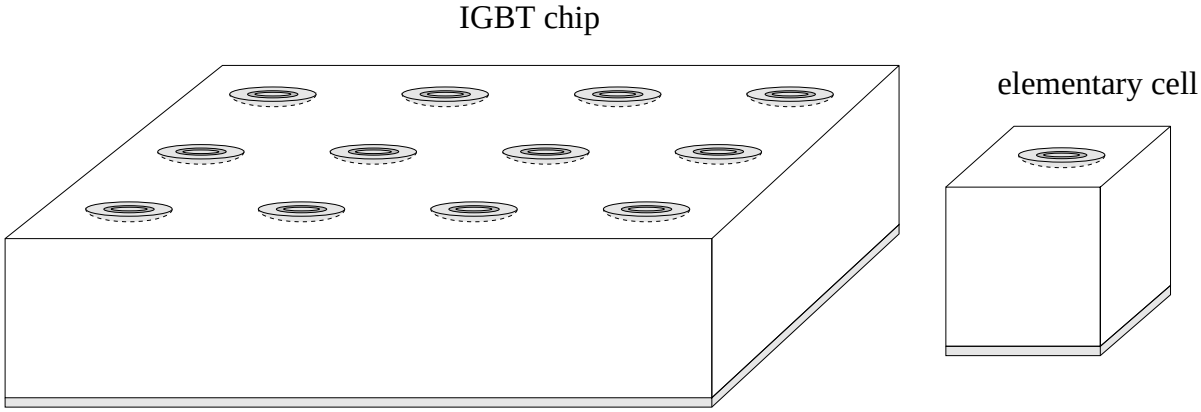


Figure 8.1: Multi–cell structure of an IGBT chip.

Obviously, performing the simulations in terms of cylindrical coordinates constitutes the most appropriate description of the MOS structures. However, the square base of the elementary cell is thus approximated by a circle. The choice of its radius is a crucial issue since it relates the internal current density distribution to the total terminal current. A correct mapping represents one of the

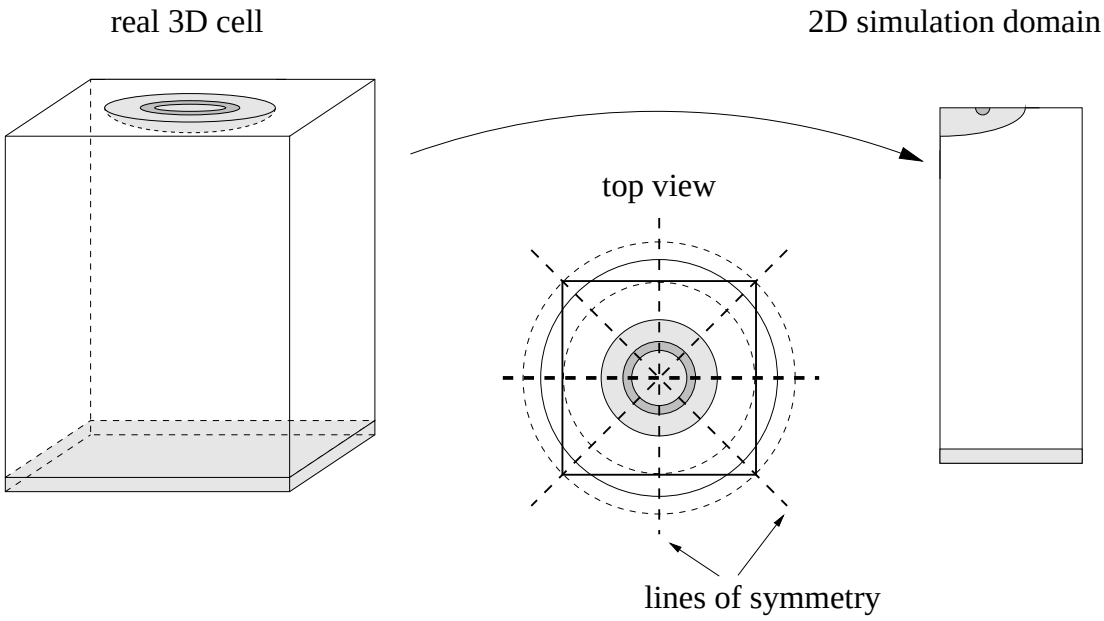


Figure 8.2: Transforming the real 3D device to the 2D simulation domain.

basic problems in simulating 3D devices on 2D cross sections and has been the subject of many discussions. On the one hand, the emitter current is composed of the electron current through the MOS channel and the hole current collected by the electric field at the p–well/n–substrate junction. It therefore depends on the geometry of the MOS structure, but is rather *independent* of the total cell size. On the other hand, the collector current is equal to the product of the internal current density and the total base area since the current distribution at the bottom surface is homogeneous.

The radius of the simulation domain can be chosen half of an edge length or half of the diagonal length, for example. In the top view, the resulting base area conforms with the inscribed circle or the circumscribed circle, respectively. Although each approach accurately reflects the distance to the adjacent cell along the corresponding axis of symmetry (cf. fig. 8.2) they do not conserve the total cell area. As a consequence, at any given current density either the terminal current or the internal current density does not agree with that of the real IGBT device. As a third possibility, the radius of the simulation domain can be chosen as $1/\sqrt{\pi}$ of the edge length, thus forming a circle whose area matches the square base area of the 3D elementary cell. In the next sections, we will compare the simulation results obtained by these approaches with different simulation domains and coordinate systems.

The coordinate system

At first, designing a 2D simulation model includes choosing an appropriate cross section and a coordinate system. They have to reflect both the geometry of the MOS region as well as the total cell area as properly as possible. As the *second* step, the parameters of the carrier lifetime and the mobility models are adjusted and therefore depend on the former choice of the coordinate system. For that reason, a meaningful comparison of the different approaches must not be carried out with the same set of parameters. We rather have to perform a separate calibration procedure as described below for each simulation domain and coordinate system. The suitability of the approaches is then judged from the results of the separately calibrated simulations.

Figure (8.3) compares the simulation results for the carrier concentration within the NPT–IGBT at a current density of $145\,\mathrm{A/cm^2}$. Due to the very large carrier lifetime, we observe a *linear*

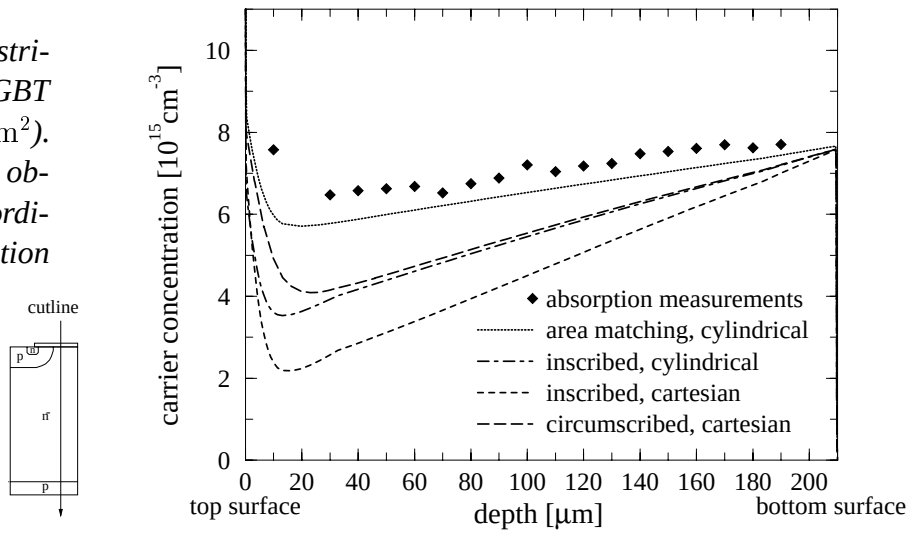


Figure 8.3: Carrier distribution within an NPT–IGBT (current density $145\,\mathrm{A/cm^2}$). The simulation results are obtained with different coordinate systems and simulation domains.

profile in the bulk of the device. Hence, it can be characterized by the two carrier densities at the rear p–emitter and in the top region ($x \approx 20\,\mu\mathrm{m}$), respectively. Two tendencies are clearly visible from the figure: First, increasing the cell size with a fixed geometry of the MOS structure raises the upper carrier concentration, thus decreasing the slope of the profile. This is a consequence of the two–dimensional current distribution [192]. At a given average current density, the larger cell has to conduct the larger current. In order to drive the enhanced hole current towards the p–well junction, a larger *lateral* concentration gradient at the top surface must be present. Consequently, at a given depth a larger average concentration is observed within the larger cells.

As the second trend, a lower carrier density at the top surface is obtained from a calculation in terms of cartesian coordinates. The latter corresponds to a stripe geometry in 3D whose effective channel width per cell area is greater than those of the array structure. Since enlarging the channel width per unit area is equivalent to shrinking the cell size with a fixed MOSFET geometry, the smaller carrier concentration is due to the same reason as mentioned above.

As it can be seen from figure (8.3), the best agreement with the measured carrier distribution is obtained from the calculation in terms of cylindrical coordinates if the radius of the simulation domain is adjusted to reflect the area of the real cell. This finding is comprehensible as the cylindrical coordinates best reflect the approximately rotational symmetry of the MOS structure. Additionally, choosing an adapted simulation radius allows not only the total chip area but also the number of cells connected in parallel and the channel width per unit area to be equal to those of the real device.

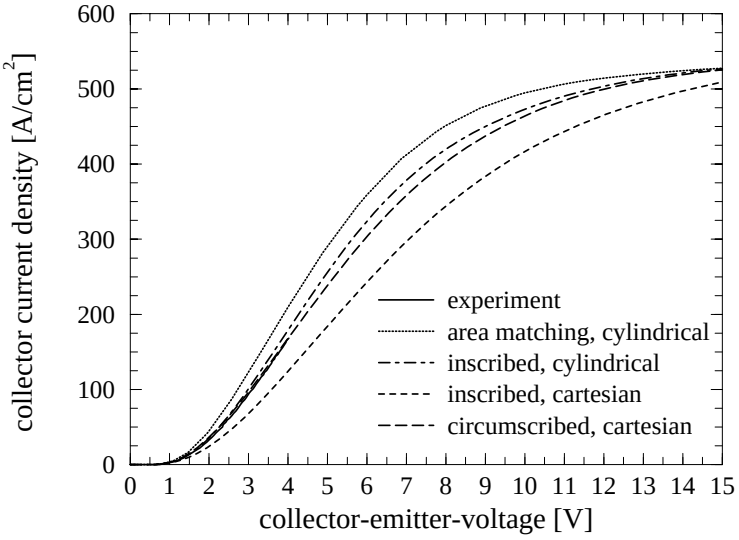


Figure 8.4: Forward characteristics of the NPT–IGBT at 15 V gate bias. The simulation results are obtained with different coordinate systems and simulation domains.

The forward characteristics of the different simulation models are compared with the experimental data in figure (8.4). At a given current density, the lower voltage drop is observed with those simulation results which correspond to the larger simulation domain and/or the use of cylindrical coordinates. This tendency originates from the above discussed effects on the carrier distribution since an increased carrier concentration enhances the substrate conductivity. As contact resistances are not included in the simulation, the calculated voltage drop must not be higher than the measured one. For that reason, the simulation in terms of the cylindrical coordinates with the effective radius of the simulation domain reproduces the experimental characteristics with very good accuracy.

Channel discretization

Although we aspire that the simulation results are independent of the chosen discretization, a remarkable effect of the underlying mesh is sometimes unavoidable. For instance, the current through the channel of a MOS device does not only depend on the mobility model but is also affected by the mesh in the channel region, in particular by the shape of the elements and the chosen discretization depth, i. e. the distance between the gate oxide and the first grid line parallel to the interface (cf. fig. 8.5). Though this dependence saturates for a grid spacing below 0.1 nm, the corresponding mesh refinement is not desirable for the sake of computational economy. In addition, we should be aware that the current distribution in the channel is governed by quantum effects which are not covered by the classical transport model. However, although the *current density distribution* in the channel is not properly reflected, an accurate simulation of the overall IGBT behavior can be expected as long as the *total channel current* is calculated accurately. But we should keep in mind that the parameters of the calibrated mobility model depend on the chosen discretization and therefore their validity cannot be generalized.

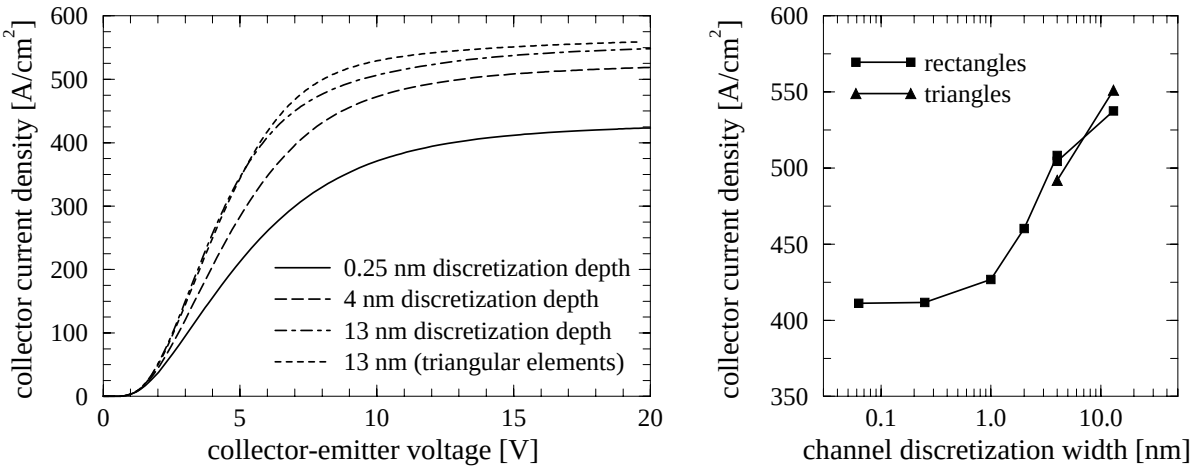


Figure 8.5: Forward characteristics of the NPT–IGBT at 15 V gate bias. The simulations are carried out with fixed parameters of the mobility models, but different discretization depths in the channel region using rectangular or triangular meshes, respectively. The right hand figure depicts the dependence of the collector currents at $V_{GE} = V_{CE} = 15\,\mathrm{V}$ on the discretization depth.

8.1.2 Structure and doping profiles

The exact knowledge of the doping distribution is an indispensable precondition of accurate device simulations. However, the measured doping profiles suffer from quite a large uncertainty and a limited spatial resolution. In many cases, this constitutes a significant problem as some of the device characteristics sensitively depend on specific properties of the doping distribution. This section therefore deals with the fine adjustment of the doping profiles which are assumed to be preadjusted within the experimental accuracy of the available probing techniques.

Doping distribution of the MOSFET

The electrical behavior of the MOS structures depends on the doping profile in the channel region which is formed by the diffusion of the p–well. In particular, its peak doping concentration significantly affects the threshold voltage, as it can be seen from the transfer characteristics with shorted gate and collector, i. e. $V_{GE} = V_{CE}$ (cf. fig. 8.6). Even a variation of 10 % of the peak concentration, which is quite below the experimental uncertainty of the doping profile, results in a parallel shift of the transfer characteristics of about 0.4 V, which constitutes a remarkable effect on the terminal behavior. Hence, the first step of the calibration procedure is the precise adjustment of the p–well's peak concentration with respect to the threshold voltage of the MOS structure.

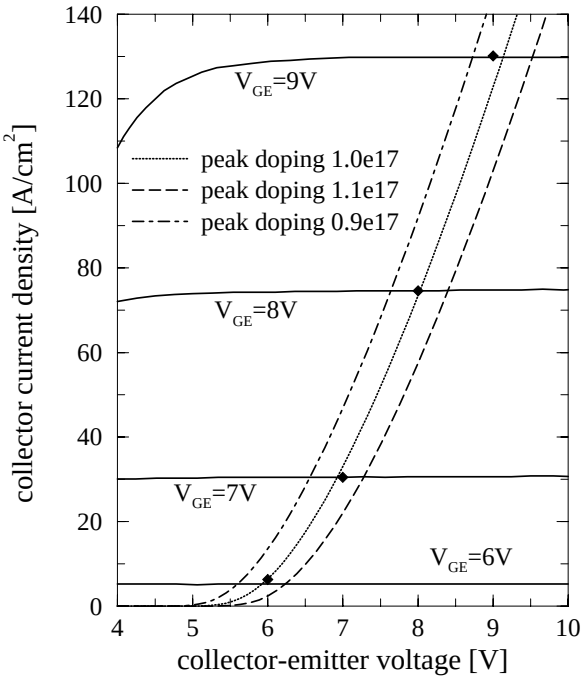


Figure 8.6: Adjusting the peak concentration of the p–well forming the channel of the MOS structure. The broken lines represent the simulation results with $V_{GE} = V_{CE}$. The solid lines refer to the experimental characteristics with a fixed gate bias. In this case, the dots indicate those operating conditions where the collector–emitter voltage is equal to the gate–emitter voltage [193].

It should be mentioned that the MOS characteristics depend on two additional parameters, namely the channel mobility and the channel length, which is determined by the lateral diffusion of the p–well. As a longer channel leads to a larger channel resistance, a larger voltage drop is observed from the terminal characteristics. Although a similar tendency originates from a decreased channel mobility, these effects can be separated by the combined measurement of the transfer characteristics and the carrier concentration profile. Increasing the peak doping concentration of the p–well results in a *parallel shift*, while a lower channel mobility decreases the *slope* of the transfer characteristics. On the other hand, a shorter channel length raises the carrier concentration in the upper region of the device (cf. section 8.1.1). For further details the reader is referred to [193].

Doping of emitters

Similar to the doping concentration of the p–well, also the doping profile of the rear p–emitter is usually not known with the desired accuracy. Since a variation of 25 % effects a detectable change of the carrier distribution (cf. fig. 8.7), the doping profile can be adjusted with respect to the carrier

concentration profile. Note that in case of a shallow and weakly doped emitter the minority carrier recombination is negligible. Thus, the emitter becomes transparent to the minority carrier current and its efficiency is hardly affected by the corresponding carrier lifetime.

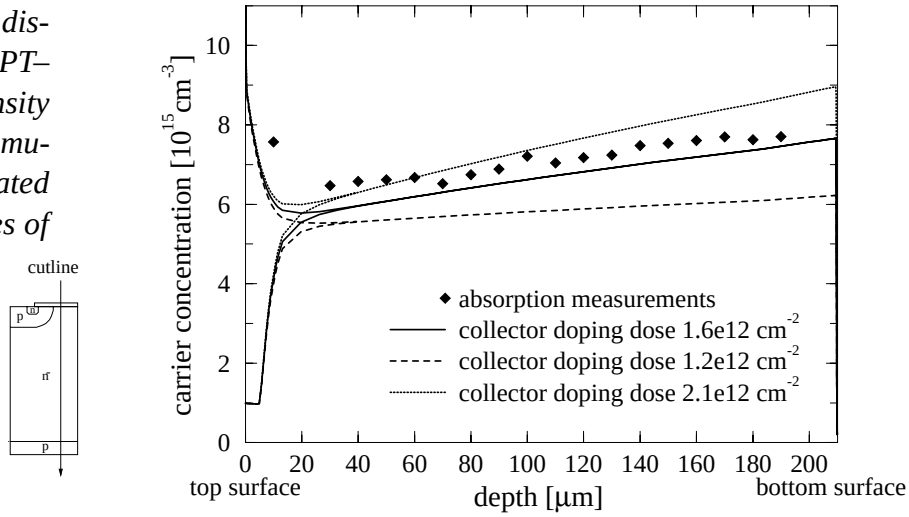


Figure 8.7: Carrier distribution within the NPT–IGBT at a current density of $145\,\mathrm{A/cm^2}$. *The simulation results are calculated with different doping doses of the rear p–emitter [193].*

8.2 Carrier mobility

Among the most important parameters are those of the carrier mobility models. While the bulk mobilities are well known, there is a large uncertainty concerning the channel mobility in MOS structures [194, 195, 196, 197]. This is especially true since they sensitively depend on the processing conditions. For that reason, the technology–dependent parameter δ of the Lombardi model (cf. eq. 6.71 in section 6.2.3) is calibrated with respect to the saturation currents of the forward characteristics (cf. fig. 8.8).

In practice, this can be done very efficiently by performing a simulation which ramps up the gate bias with shorted gate and collector. Then, the calculated results have to reproduce the dots

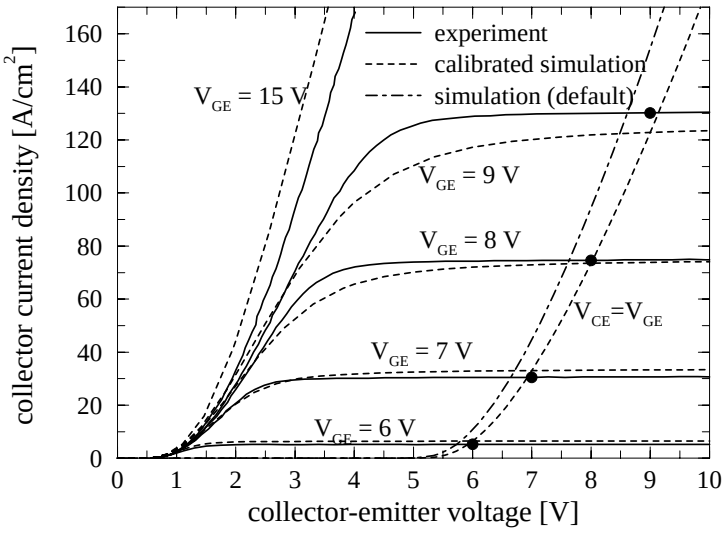


Figure 8.8: Forward characteristics of an NPT–IGBT for various gate biases. The dots refer to the operating points with equal gate and collector bias [126].

in fig. (8.8) which mark the operating conditions with equal gate and collector bias. Note that modifying the channel conductance changes the slope of the transfer characteristics whereas a variation of the p–well's peak concentration gives rise to a parallel shift (cf. fig. 8.8 and 8.6).

8.3 Calibration of the carrier lifetimes

Since the flow of carriers is significantly affected by the distribution of the recombination centers, the calibration of the carrier lifetime models is a crucial preparation step for predictive device simulations. However, major difficulties arise from a possible heavy metal diffusion for lifetime control since these impurities introduce a variety of trap levels with different capture cross sections. For that reason, complicated models of the local lifetime reduction would have to be employed for a physically rigorous treatment.

However, the goal of this section is to follow another strategy: For the sake of computational economy, the simple SRH–model is preferable although it merely takes into account a single impurity level. Nevertheless, since the particle balances (cf. eq. 6.36 and 6.37) do not directly depend on the carrier lifetimes but on the total recombination rate, an accurate simulation of the device behavior is attained if the total recombination rate R is reflected properly. Thus, the basic challenge consists in finding a suitable description of the spatially dependent carrier lifetime $\tau := \Delta n/R$.

Based on experimental facts and the theoretical conclusion that the solubility of a fundamental defect is strongly correlated to the doping density, the local lifetime reduction is commonly described by the empirical Scharfetter formula

$$\tau_{n,p} = \frac{\tau_{n,p}^0}{1 + \left(N_i / N_{n,p}^{ref}\right)} \ . \tag{8.1}$$

Thus, the spatial dependence of the carrier lifetime is governed by the doping profile.

In the following section, however, we will apply another methodology: The fundamental idea is to replace the four parameters $\tau_{n,p}^0$ and $N_{n,p}^{ref}$ by a different set which much better reflects the physics of the investigated device. In case of a PT–IGBT (cf. fig. 8.9), for example, the recombination in the intrinsic region is governed by the ambipolar carrier lifetime τ_{amb}, the recombination in the buffer is governed by the minority carrier lifetime $\tau_p(N_{D,Buffer})$ of the holes, and the recombination in the substrate is governed by the minority carrier lifetime $\tau_n(N_{A,Substrate})$ of the electrons (cf. section 6.2.2). Choosing the ratio τ_n^0/τ_p^0 as the fourth parameter, the following transformation is obviously appropriate:

$$\begin{array}{rcl} \tau_n^0 & & \tau_{amb} := \tau_n^0 + \tau_p^0 \\ \tau_p^0 & \Longleftrightarrow & \tau_n^0/\tau_p^0 \\ N_n^{ref} & & \tau_S := \tau_n(N_{A,Substrate}) \\ N_p^{ref} & & \tau_B := \tau_p(N_{D,Buffer}) \end{array} \tag{8.2}$$

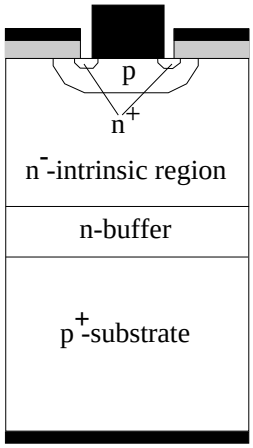


Figure 8.9: Schematic sketch of a PT–IGBT.

The following investigations refer to the ambipolar carrier lifetime in the intrinsic region, the minority carrier lifetimes in heavily doped emitters, and the ratio τ_n^0/τ_p^0 as the independent parameters. From these, the values of $\tau_{n,p}^0$ and $N_{n,p}^{ref}$ are calculated which are required for evaluating equation (8.1) during the simulation.

8.3.1 NPT–IGBTs

As a first example, the NPT–IGBT is investigated which has not been subjected to specific processing steps for lifetime control. Consequently, we expect a very large carrier lifetime. This is confirmed by the carrier distribution (cf. fig. 8.10) which exhibits a negligible curvature, thus revealing that τ_{amb} is large. Since the device operation is not very sensitive to the exact values of the carrier lifetimes, they can be easily adjusted with the required accuracy. Assuming a ratio τ_n^0/τ_p^0 of 5 which is the common experience for this device technology [198], we find $\tau_n^0 = 70\,\mu\mathrm{s}$ and $\tau_p^0 = 14\,\mu\mathrm{s}$ as suitable set. Furthermore, as the rear emitter is very shallow and weakly doped, the minority carrier lifetimes have little or no effect. For that reason, the parameters $N_{n,p}^{ref}$ are kept

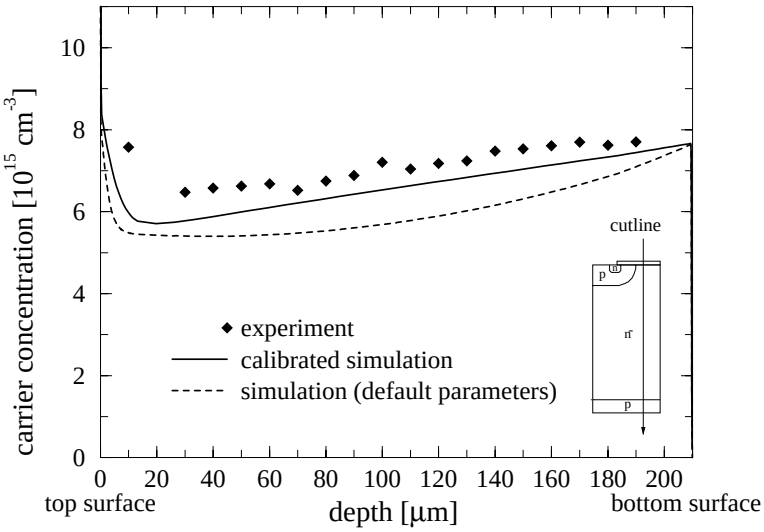


Figure 8.10: Carrier distribution in an NPT–IGBT at a current density of $145\,\mathrm{A/cm^2}$ *for different ambipolar carrier lifetimes [126].*

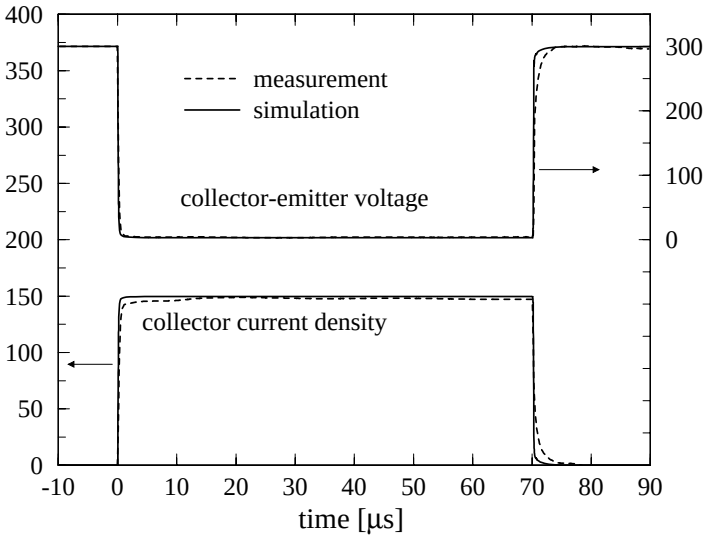


Figure 8.11: Transient evolution of the terminal current and the voltage drop during operation with ohmic load (current density $145\,\mathrm{A/cm^2}$ *during the on–state, pulse duration* $70\ \mu\mathrm{s}$*).*

at their default values. The thus calibrated model reproduces the internal carrier distribution (cf. fig. 8.10), the forward characteristics at different gate biases (cf. fig. 8.8), as well as the terminal behavior (cf. fig. 8.11) and the temperature profiles during transient switching (cf. fig. 7.5 and 7.7).

8.3.2 Devices subjected to heavy metal diffusion

As an example of devices with lifetime control by platinum diffusion, we will discuss a PT–IGBT test structure. It exhibits a very interesting feature, namely a region of negative–differential–resistance (NDR) in the forward characteristics, which occurs at the transition from low injection to high injection conditions [199, 200]. At a low current density, the forward voltage drops primarily across the highly resistive intrinsic region while the voltage drop across the combined n–buffer/p^+–substrate emitter structure is low. As the current density increases, the voltage drop across the emitter structure increases and the voltage drop across the intrinsic region decreases due to the enhanced injection of carriers. If the latter effect prevails against the former, the total voltage drop decreases.

However, this phenomenon can only be observed on small test structures. Within large devices, even small lateral inhomogeneities originate that most of the cells operate at a current density below or above the NDR region, since the NDR region itself corresponds to unstable operating conditions. As a consequence, ramping up the total current will increase the fraction of cells operating in high injection, thus causing a monotonically increasing voltage drop in the forward characteristics of the *whole* structure.

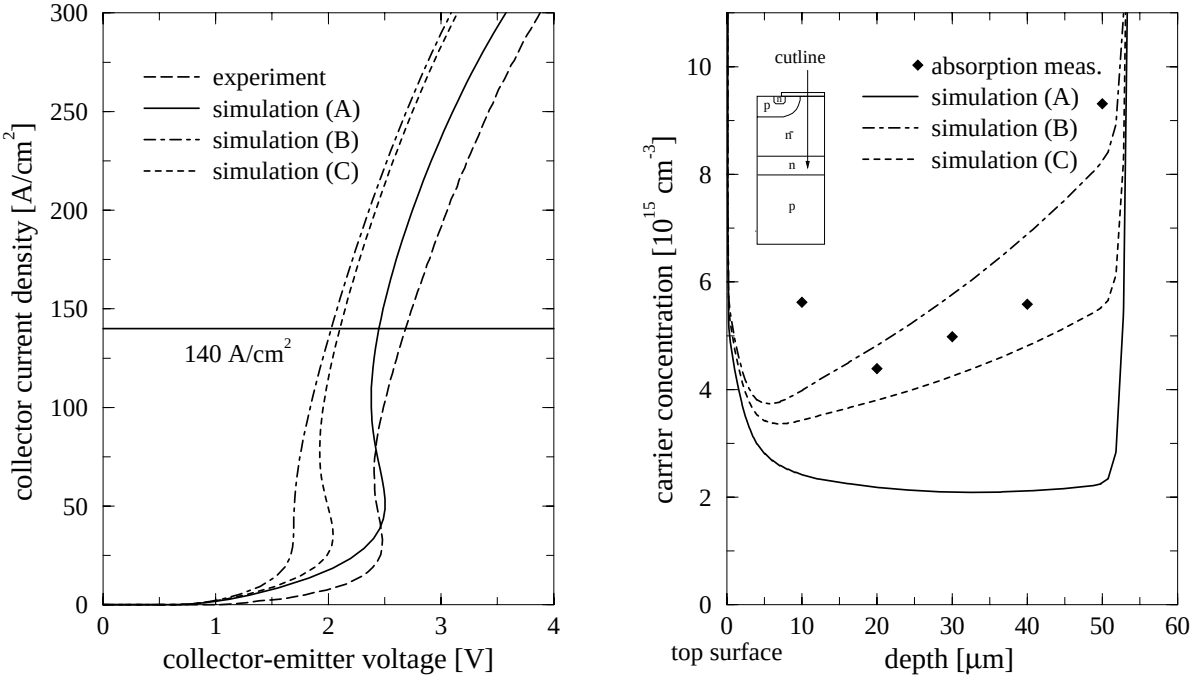


Figure 8.12: Forward characteristics at 15 V gate bias (left) and carrier concentration at a current density of $140\ \mathrm{A/cm^2}$ *of a PT–IGBT test structure subjected to platinum diffusion [126].*

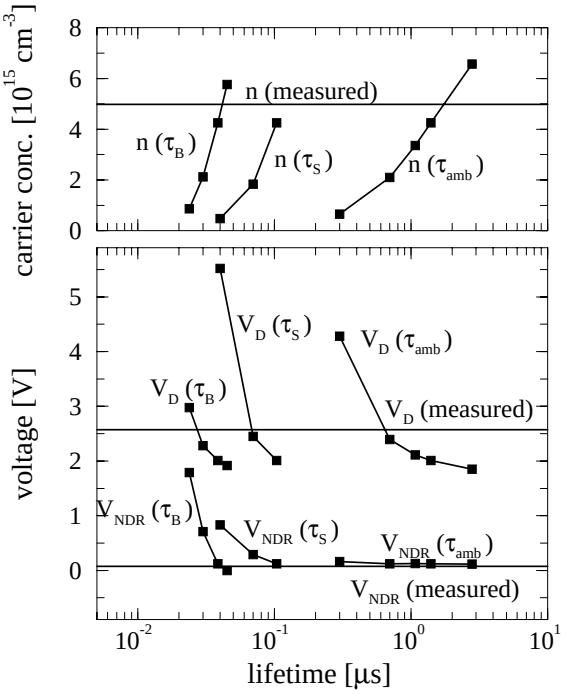


Figure 8.13: Dependence of characteristic quantities of the discussed PT–IGBT test structure on the parameters of the carrier lifetime model [126].

As mentioned above, the physics of the device is reflected more properly, if the parameters of Scharfetter's relation (8.1) are replaced by the ambipolar carrier lifetime, the minority carrier lifetimes in the heavily doped regions, and the ratio τ_n^0/τ_p^0. The transformation is given by equation (8.2).

At first, we realize that the simulation results do not depend on the ratio τ_n^0/τ_p^0, thus reducing the dimension of the parameter space to three.

Model calibration with pure reference to the terminal characteristics leads to parameters (set A) that are not at all confirmed by the measured carrier distribution (cf. fig. 8.12). Hence, we follow another strategy for the calibration which is based on combined measurement results. For that purpose, some characteristic features of the device behavior are chosen, namely the voltage drop V_D at the rating current density of $140\,\mathrm{A/cm^2}$, the width V_{NDR} of the NDR region (i. e. the difference between the snapback and the holding voltage), and the carrier concentration in the intrinsic region at the rating current density. Figure (8.13) illustrates the dependence of the chosen quantities on the parameters of the carrier lifetime model, thus indicating the strongest correlations. Proceeding this way yields a consistent set of parameters (set C) which reproduces all experimental results during stationary operation (cf. fig. 8.12) and transient switching (cf. fig. 7.10 and 8.14) with satisfying accuracy.

version	$\tau_{amb}\,[\mu\mathrm{s}]$	τ_n^0/τ_p^0	$\tau_B\,[\mu\mathrm{s}]$	$\tau_S\,[\mu\mathrm{s}]$
A	0.7	13:1	0.10	0.038
B	1.4	13:1	0.04	0.038
C	1.4	13:1	0.10	0.038

Table 8.1: Simulation parameters for the PT–IGBT.

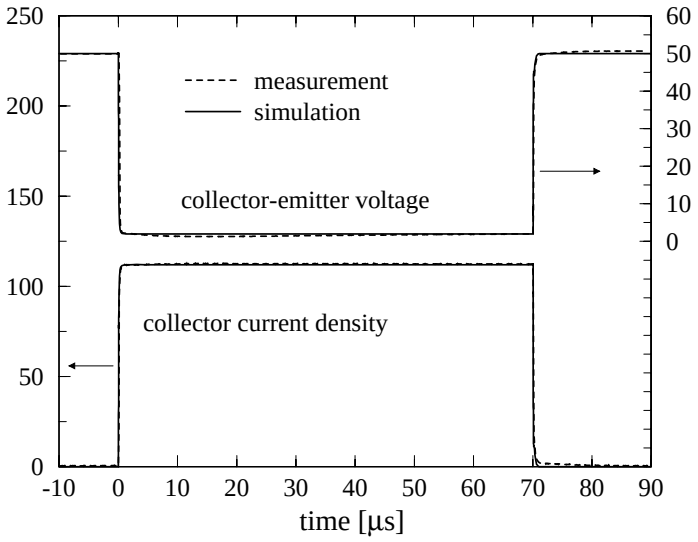


Figure 8.14: Transient evolution of the terminal current and the voltage drop during operation with ohmic load (current density $110\,\mathrm{A/cm^2}$ *during the on–state, pulse duration* $70\ \mu\mathrm{s}$*).*

8.4 Summary

The most important results of this chapter can be summarized by the following statements:

- Since the MOS structures of IGBTs are formed by lateral diffusion, they are approximately circular in shape. Therefore, the use of cylindrical coordinates is suggested. The best choice for the width of the simulation domain is $1/\sqrt{\pi}$ of the edge length. Thus, the base area of the simulated cell is equal to that of the real device.

- NPT–IGBT devices which have not been subjected to specific processing steps for lifetime control exhibit a linear carrier concentration profile. Its bottom concentration is determined by the efficiency of the rear p–emitter. With a fixed geometry of the MOS structure, the upper concentration value decreases with shrinking cell size. Simulating in terms of cartesian coordinates produces a carrier concentration profile with a lower density in the top region of the device.

- The calibration of the channel mobility model is not generally valid since the simulated channel conductivity sensitively depends on the discretization width. In addition, due to neglecting quantum effects the classical model cannot claim to produce the correct distribution of physical quantities in the channel, e. g. the carrier densities. However, the performance of the whole IGBT only depends on the integral conductivity of the MOS structure which acts as a voltage controlled current source. Therefore, accurate simulation results for the bipolar regions of the device are attainable.

The local reduction of the carrier lifetime can be modeled by the simple Scharfetter relation if its parameters are chosen carefully. For that purpose, they are replaced by a different set of parameters which comprise the ambipolar lifetime and the minority carrier lifetimes in the heavily doped regions of the device. Employing combined measurement results, namely the forward terminal characteristics, the carrier distribution, and the temperature profiles during transient switching facilitates a comprehensive calibration according to the following strategy:

- First, the peak concentration of the p–well is adjusted with reference to the threshold voltage of the MOS structure.
- Second, the ambipolar carrier lifetimes are extracted from the curvature of the carrier concentration profiles.
- The channel mobility is calibrated from the saturation currents of the forward characteristics.
- The p–emitter doping of NPT devices can be extracted from the bottom concentration of the carrier distribution profile.
- The minority carrier lifetimes in PT devices with platinum diffusion is extracted from specific features of the forward characteristics and the carrier concentration profile.

As a final check, the calculated terminal behavior and the temperature distribution during transient switching is compared with the experimental data. The latter are accurately reproduced without any further calibration steps, thus proving the proposed procedure to be suitable.

Chapter 9

Conclusion and Outlook

As their common physical principle, the laser probing techniques discussed in this thesis exploit the dependence of the complex refractive index on carrier concentration and lattice temperature. Thus, they provide space–resolved and time–resolved information about the carrier and temperature distribution in the interior of power devices. The first part of this thesis is dedicated to a comprehensive analysis of these probing techniques. A physically rigorous numerical model has been developed which facilitates the simulation of the entire measurement process. It includes an electrothermal device simulation, the calculation of the beam propagation through the sample, the lenses, and aperture holes, and finally the simulation of the detector response. Together with analytical investigations, the numerical model has been employed to analyze the capabilities of these methods and to support their development with respect to a high resolution, a minimum experimental error, a large measurement range, and a minimum sensitivity to parasitic effects.

The second part of this thesis deals with the validation and calibration of electrothermal device simulation models, in particular the self–consistent electrothermal extension of the drift–diffusion model which is derived from the principles of irreversible thermodynamics. It has been demonstrated that the optimized internal laser probing techniques constitute a valuable supplement to the well–established electrical characterization methods, thus enabling an advanced and more profound investigation of the governing model equations and their boundary conditions. In addition, we have elaborated a thorough calibration strategy for the models of the most important material properties, namely the carrier mobility and the carrier lifetime.

In summary, the following conclusions can be drawn:

9.1 Optical probing techniques

First of all, some general statements apply to all of the investigated probing techniques: One of the most important results of the theoretical study is that the lateral spreading of the probing beam does not introduce a significant error, i. e. the interaction of the beam profile with the distribution of the complex refractive index is accurately reflected by regarding the one–dimensional carrier concentration and temperature integrals along the beam path. On the other hand, the major source of error originates from an oblique propagation of the probing beam which may result from the internal beam deflection or a possible misalignment of the sample. In addition, except for the

Fabry–Perot transmission measurements, the detector signals are corrupted by multiple reflections at the polished sample surfaces unless suitable precautions are taken, e. g. depositing an antireflective coating or employing an incoherent light source.

Thus, the currently available optical setup facilitates the space–resolved and time–resolved detection of the carrier and temperature distribution with excellent accuracy. Designing the samples according to the theoretically predicted optimum geometry promises absorption measurements which reveal the carrier concentration profile with an error of only a few percent. Additionally, Internal Laser Deflection measurements constitute a sensitive and accurate probing technique for scanning vertical temperature profiles. However, due to the saturation of the detector response function, the measurement range is limited to temperature rises of only a few Kelvins for typical sample lengths of some millimeters. For that reason, various demanding problems in physical research and in the development of semiconductor devices cannot be addressed as they require a large measurement range rather than a high sensitivity. The most effective way to overcome this restriction is measuring specific test structures comprising a small active area within a silicon die of larger size. It has been demonstrated that the resulting heat spreading effects can be approximately corrected by introducing an effective interaction length which is related to the sample geometry by simple analytical formulas.

As a useful extension of the existing setup, interferometric techniques with a laterally propagating probing beam can be included. However, their fundamental problem is the internal beam deflection. As it impairs the superposition of the probing beam and the reference beam, it represents the limiting effect of the measurement range. Nevertheless, probing the above mentioned

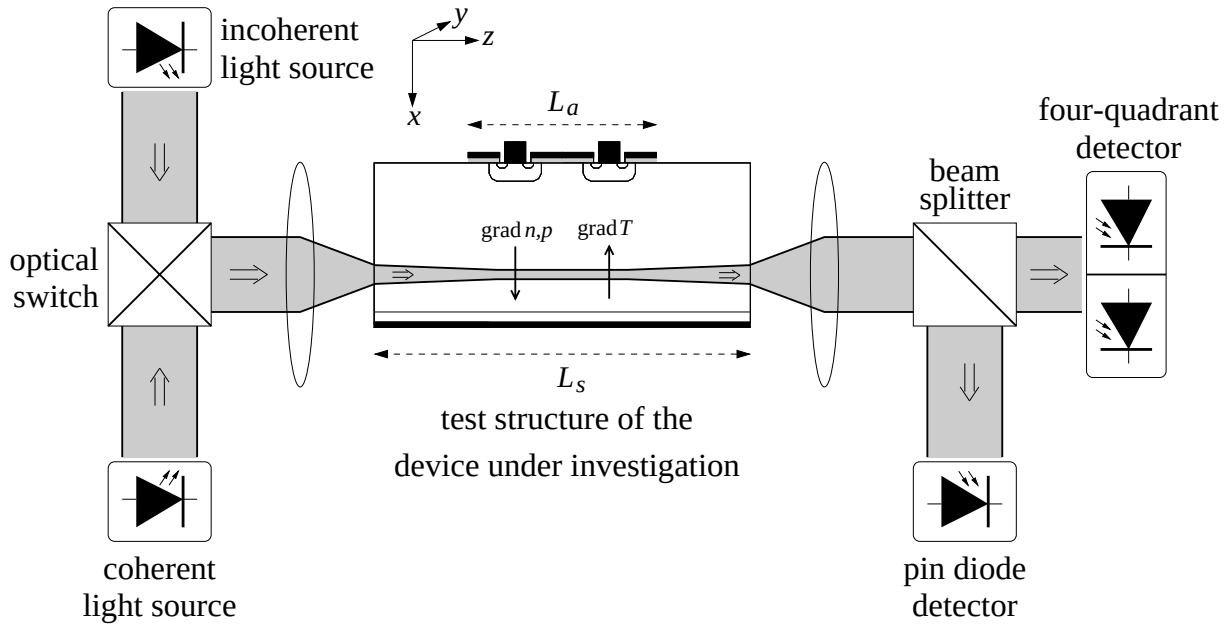


Figure 9.1: Extension of the experimental setup to facilitate free carrier absorption measurements (incoherent light source and pin diode detector), laser deflection measurements (incoherent light source and four–quadrant diode detector), and Fabry–Perot transmission measurements (coherent light source and pin diode detector).

test structures enables the desired detection of large temperature rises which typically occur during short circuit operation of power devices. Since interferometric techniques do not require a calibration of the optical setup, as, e. g., the detector response function, they constitute a valuable supplement to the available methods. Fabry–Perot transmission measurements emerge as the most appropriate technique which can be easily included in the current setup (cf. fig. 9.1): Employing the incoherent light source enables simultaneous absorption and deflection measurements. To avoid distortions of the absorption signal due to the beam displacement on the detector, the transmitted intensity is detected by a separate pin diode while the deflection signal is gained from a four–quadrant detector. Thus, the gradients of carrier concentration and lattice temperature are determined together with the absolute value of the carrier concentration. Switching to the coherent light source results in Fabry–Perot oscillations of the pin diode's photo current from which the temperature evolution can be extracted.

In addition to the lateral measurement techniques, Backside Laser Probing has been discussed. It operates by means of a vertically propagating probing beam and therefore does not permit a space–resolved scanning of vertical profiles. However, the refractive index within the sample exhibits merely small gradients perpendicular to the optical axis, resulting in a negligible internal deflection of the probing beam. Since the oscillating interference signal therefore lacks of saturation effects, Backside Laser Probing is capable of detecting temperature rises of some hundreds of Kelvins. Hence, although this technique provides merely integral information on carrier concentration and temperature, it constitutes one of the most powerful methods for investigating devices subjected to operating conditions with a very large power dissipation.

The presented model for the numerical simulation of the entire measurement process opens up interesting perspectives for further investigations: For example, it can be extended to enable inverse modeling: The internal carrier and temperature distribution is represented by a sufficiently large number of unknowns, whose initial values are determined by an electrothermal device simulation. Next, the optical module is employed for calculating the detector signal. Comparing it with the real measurement signal yields the necessary update of the carrier concentration and temperature distribution. This procedure is iterated until the extracted updates reproduce the current values. Although this strategy requires a lot of computational effort, it can be helpful for the evaluation of measurements on extremely small structures or the determination of 3D carrier and temperature distributions by two 2D scans in x–y–direction and in x–z–direction (cf. fig. 9.1).

Furthermore, the optical model does not only provide a powerful strategy for the simulation of probing techniques but is also capable of calculating the operating characteristics of optical sensors exploiting the thermo–optical or the electro–optical effect [201]. Such kind of devices have been proposed for monitoring the temperature evolution within microstructures [202, 203, 204]. The Fabry–Perot effect can also be exploited to detect any physical quantity affecting the refractive index, such as, for example, mechanical stress, provided the corresponding dependence is sufficiently pronounced. A complete simulation of such a technique is enabled by the presented model (cf. fig. 3.1) if the electrothermal device simulation is extended to a consistent electro–thermo–mechanical simulation.

9.2 Electrothermal modeling

In the second part of this thesis, a thorough validation and calibration of device simulation models are presented. They are based on the combined evaluation of electrical and optical characterization methods. The self–consistent electrothermal extension of the drift–diffusion model [62] has proven to be a suitable description for the accurate and predictive simulation of power devices. Except for the reverse recovery behavior of pin diodes, the electric properties of the investigated devices are accurately reproduced by the calibrated simulation, both in stationary operation as well as under transient switching conditions. The cited models of the carrier mobility turn out to enable an appropriate simulation of carrier transport. However, major difficulties emerge in modeling carrier recombination, especially within devices which have been subjected to specific processing steps for lifetime control. Although the well–known Shockley–Read–Hall formula leads to rather satisfying results for IGBTs if it is calibrated carefully, the simulated turn–off behavior of bipolar devices differs from the measured characteristics. To keep pace with today's technological progress the available numerical methods therefore have to be supplied by a physical description of the capture and emission kinetics of the individual trap centers.

As measured temperature profiles have become available, also the thermal simulation results can be validated directly. The effects of the most significant heat generation mechanisms on the temperature distribution are confirmed by the experimental results with excellent accuracy. Minor simplifications of the implementation in the general purpose simulator $\mathrm{DESSIS_{ISE}}$ [101], e. g. the neglecting of the convective terms in the boundary conditions of the heat flow equation, originate specific deviations of the simulated and the measured temperature distribution. These differences are clearly detectable by the sensitive probing techniques. However, they amount to only some Kelvin and are therefore negligible for practical applications. This is especially true in case of operating conditions with a large power dissipation where the missing contribution is superimposed by the properly modeled Joule heating. Consequently, no differences between the simulated and the measured results are observed at all and the cited model accurately reflects the coupling between the electrical and the thermal domain.

*

With shrinking device dimensions, additional physical phenomena will become more and more important, thus demanding novel physical models. Their validation will gain essential benefit from the availability of probing techniques which provide space–resolved and time–resolved information about the internal distribution of carriers and heat. The demonstrated interaction of experimental and theoretical investigations has shown to be indispensable for a continuous improvement of both the characterization methods as well as the theoretical models. Thus, more and more powerful tools can be developed to tackle the exciting challenges emerging from the increasing physical and technological complexity of future devices and systems.

Appendix A

Derivation of the Propagator Matrix

This appendix shall describe the discretization of the equations (3.9) and (3.10) and the derivation of the propagator matrix P_n.

First, the differential equations (3.9) and (3.10) introduced in chapter (3)

$$\begin{aligned}\frac{\partial}{\partial z}E^F(x,z) &= \frac{i\omega^2}{2c^2\kappa}\varepsilon_R(x,z)E^F(x,z)+\frac{i}{2\kappa}\frac{\partial^2}{\partial x^2}E^F(x,z)-\frac{i\kappa}{2}E^F(x,z) \qquad (A.1)\\ &\quad +\frac{i\omega^2}{2c^2\kappa}\varepsilon_R(x,z)\,e^{-2i\kappa z}E^B(x,z)+\frac{i}{2\kappa}e^{-2i\kappa z}\frac{\partial^2}{\partial x^2}E^B(x,z)-\frac{i\kappa}{2}e^{-2i\kappa z}E^B(x,z)\\ \frac{\partial}{\partial z}E^B(x,z) &= -\frac{i\omega^2}{2c^2\kappa}\varepsilon_R(x,z)E^B(x,z)-\frac{i}{2\kappa}\frac{\partial^2}{\partial x^2}E^B(x,z)+\frac{i\kappa}{2}E^B(x,z) \qquad (A.2)\\ &\quad -\frac{i\omega^2}{2c^2\kappa}\varepsilon_R(x,z)\,e^{2i\kappa z}E^F(x,z)-\frac{i}{2\kappa}e^{2i\kappa z}\frac{\partial^2}{\partial x^2}E^F(x,z)+\frac{i\kappa}{2}e^{2i\kappa z}E^F(x,z)\end{aligned}$$

are integrated over the interval $[z_n, z_{n+1}]$, assuming a linear interpolation of $E^{F,B}$ and ε_R. For example, the first term of the first equation becomes

$$\begin{aligned}\frac{i\omega^2}{2c^2\kappa}\int_{z_n}^{z_{n+1}}\varepsilon_R(x,z)E^F(x,z)\,dz &= E^F(x,z_{n+1})\,\frac{i\omega^2\Delta z_n}{6c^2\kappa}\left[\varepsilon_R(x,z_{n+1})+\frac{1}{2}\,\varepsilon_R(x,z_n)\right]\\ &\quad +E^F(x,z_n)\,\frac{i\omega^2\Delta z_n}{6c^2\kappa}\left[\frac{1}{2}\,\varepsilon_R(x,z_{n+1})+\varepsilon_R(x,z_n)\right] \qquad (A.3)\end{aligned}$$

Similarly, the other summands comprise products of $E^{F,B}(x,z_{n,n+1})$, $\varepsilon_R(x,z_{n,n+1})$, and $\mathrm{e}^{i\kappa z_{n,n+1}}$. Sorting by $E^{F,B}(x,z_n)$ and $E^{F,B}(x,z_{n+1})$ transforms (A.1) and (A.2) to a matrix equation of the following form:

$$\begin{aligned}&\begin{bmatrix} A_n(x)-\dfrac{i\Delta z_n}{4\kappa}\,\partial_x^2 & B_n(x)+K_n\,\partial_x^2 \\ C_n(x)+K_n^*\,\partial_x^2 & D_n(x)+\frac{i\Delta z_n}{4\kappa}\,\partial_x^2\end{bmatrix}\begin{bmatrix}E^F(x,z_{n+1})\\ E^B(x,z_{n+1})\end{bmatrix}\\ &=\begin{bmatrix} \tilde{A}_n(x)+\dfrac{i\Delta z_n}{4\kappa}\,\partial_x^2 & \tilde{B}_n(x)+\tilde{K}_n\,\partial_x^2 \\ \tilde{C}_n(x)+K_n^*\,\partial_x^2 & \tilde{D}_n(x)-\frac{i\Delta z_n}{4\kappa}\,\partial_x^2\end{bmatrix}\begin{bmatrix}E^F(x,z_n)\\ E^B(x,z_n)\end{bmatrix} \qquad (A.4)\end{aligned}$$

In this equation, the constants K_n and $\tilde{K}_n$ are functions of ω, κ, z_n, and z_{n+1}, while $A_n(x)$, $B_n(x)$, $C_n(x)$, $D_n(x)$, $\tilde{A}_n(x)$, $\tilde{B}_n(x)$, $\tilde{C}_n(x)$, and $\tilde{D}_n(x)$ additionally depend on $\varepsilon_R(x,z_n)$ and $\varepsilon_R(x,z_{n+1})$.

The second step is the lateral discretization which will be illustrated for the finite difference discretization scheme:

$$\left[\frac{\partial^2}{\partial x^2} f(x)\right]_{x=x_j} = \frac{2}{(x_{j+1}-x_{j-1})(x_{j+1}-x_j)} f(x_{j+1}) - \frac{2}{(x_{j+1}-x_j)(x_j-x_{j-1})} f(x_j) + \frac{2}{(x_{j+1}-x_{j-1})(x_j-x_{j-1})} f(x_{j-1}) \tag{A.5}$$

Discretizing the differential operator ∂_x^2 in the matrices of equation (A.4) at $x = x_j$ therefore yields three terms at each side of the equation which refer to the positions x_{j-1}, x_j, and x_{j+1}:

$$\begin{aligned} &M^1_{n,j}\, u(x_{j+1}, z_{n+1}) \;+\; M^2_{n,j}\, u(x_j, z_{n+1}) \;+\; M^3_{n,j}\, u(x_{j-1}, z_{n+1}) \\ &= M^4_{n,j}\, u(x_{j+1}, z_n) \;+\; M^5_{n,j}\, u(x_j, z_n) \;+\; M^6_{n,j}\, u(x_{j-1}, z_n) \end{aligned} \tag{A.6}$$

with

$$u(x_j, z_n) = \begin{bmatrix} E^F(x_j, z_n) \\ E^B(x_j, z_n) \end{bmatrix} . \tag{A.7}$$

The complex 2×2 matrices $M^k_{n,j}$ can be easily derived from (A.4) by inserting the discretization rule (A.5). Hence, we finally end up with the desired relation

$$M_{1,n}\vec{u}(z_{n+1}) = M_{2,n}\vec{u}(z_n) \tag{A.8}$$

in which the fields have been summarized in a vector

$$\begin{aligned} \vec{u}(z_n) &= \left[E^F(x_1, z_n), E^B(x_1, z_n), \ldots, E^F(x_{N_x}, z_n), E^B(x_{N_x}, z_n)\right]^T \\ &= \left[u^T(x_1, z_n), \ldots, u^T(x_{N_x}, z_n)\right]^T \end{aligned} \tag{A.9}$$

and $M_{2,n}$ and $M_{1,n}$ represent band structured matrices with only one occupied superdiagonal and subdiagonal line:

$$M_{1,n} = \begin{pmatrix} \ddots & \ddots & \ddots & & & \\ & M^1_{n,j-1} & M^2_{n,j-1} & M^3_{n,j-1} & & \\ & & M^1_{n,j} & M^2_{n,j} & M^3_{n,j} & \\ & & & \ddots & \ddots & \ddots \end{pmatrix} \tag{A.10}$$

$$M_{2,n} = \begin{pmatrix} \ddots & \ddots & \ddots & & & \\ & M^4_{n,j-1} & M^5_{n,j-1} & M^6_{n,j-1} & & \\ & & M^4_{n,j} & M^5_{n,j} & M^6_{n,j} & \\ & & & \ddots & \ddots & \ddots \end{pmatrix} \tag{A.11}$$

Thus, the propagator matrix $P_n = M^{-1}_{2,n} M_{1,n}$ can be calculated very efficiently.

Appendix B

Formulation of Radiation Boundary Conditions

The boundary conditions of the differential equations (A.1) and (A.2) are extracted from the refraction law at the interfaces which relates the computation variables E^F and E^B to the incident waves (cf. e. g. eq. 3.24). Three approaches are proposed in section (3.3.3) to find a representation of $k_z^X = \sqrt{\varepsilon_X k_0^2 - k_x^2}$ ($X = a, b, l, r$) in real space: First, k_z^X can be approximated by $\sqrt{\varepsilon_X} k_0$ which may be regarded as the leading term of a Taylor expansion in terms of k_x. Additionally including the linear term, we obtain the approximation (3.26). And finally, an exact representation is gained by the transformation to k_x–z–space (cf. eq. 3.29).

In order to compare the accuracy of these approaches, the refraction at a silicon/air interface (cf. fig. B.1) is considered. The reflected and transmitted waves are calculated by employing different formulations of the radiation boundary conditions. As it can be seen from figure (B.2) the implementation including the Fourier transformations leads to results which are in excellent agreement with the analytical Fresnel formula [119]. On the other hand, merely taking into account the constant term in the Taylor expansion of k_z^X gives rise to significant deviations for angles of incidence beyond 5°, whereas additionally including the linear term of the expansion constitutes a sufficiently accurate approximation for angles of incidence up to about 15°. In this case, the latter approach is preferable against the Fourier transformation method since the matrix representations of the Taylor expansions comprise sparse matrices.

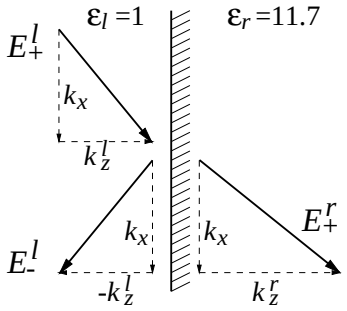


Figure B.1: Incident and outgoing waves at an interface.

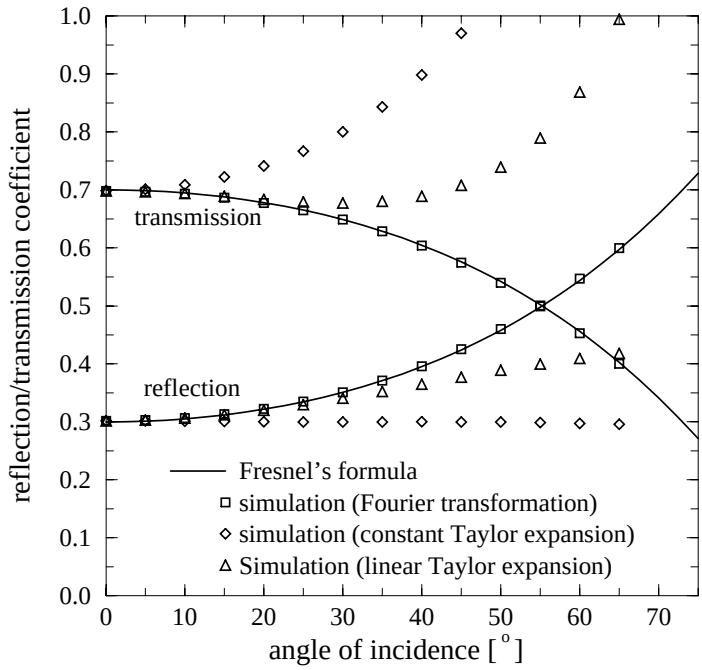


Figure B.2: Reflection and transmission coefficient at a silicon/air interface. Simulation results obtained with different implementations of the boundary conditions are compared with Fresnel's formula.

Appendix C

Analytical Calculation of the Deflection Signal

C.1 Calculation of the internal deflection

Applying the principles of geometrical ray tracing, the internal deflection within a given distribution of the refractive index $n_{Si}(x,z)$ is calculated according to Fermat's principle, i. e. we are looking for the beam path $x(z)$ minimizing the Hamilton integral

$$S = \int_0^L \mathcal{L}(x, x', z)\, dz \tag{C.1}$$

with the Lagragian [205]

$$\mathcal{L}(x, x', z) = n_{Si}(x,z)\sqrt{1 + x'^2} \qquad \text{where } x' = \frac{dx}{dz}\ . \tag{C.2}$$

In the special case that $n_{Si}(x,z)$ is independent of z, the Euler integral constitutes a conserved quantity:

$$const = \frac{\partial \mathcal{L}}{\partial x'} x' - \mathcal{L} = -\frac{n_{Si}(x)}{\sqrt{1 + x'^2}} =: C \tag{C.3}$$

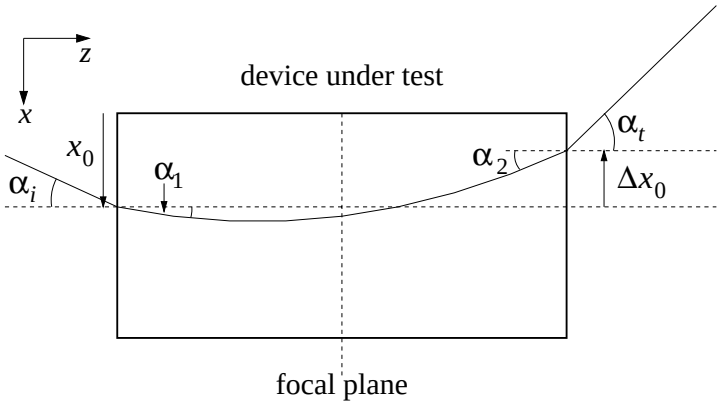


Figure C.1: Internal deflection of a probing beam.

Its value is determined from the boundary condition

$$x'(z=0) = \tan(\alpha_1) = \frac{\sin(\alpha_i)}{\sqrt{n_{Si}^2(x_0) - sin^2(\alpha_i)}} \tag{C.4}$$

where α_i denotes the angle of incidence. Hence, we get $C = -\sqrt{n_{Si}^2(x_0) - \sin^2(\alpha_i)}$. And finally, solving the first order partial differential equation (C.3) yields

$$\boxed{\pm z = \int_{x_0}^{x(z)} \frac{dx}{\sqrt{\frac{n_{Si}^2(x)}{n_{Si}^2(x_0) - \sin^2(\alpha_i)} - 1}}} \; . \tag{C.5}$$

Either the positive or the negative sign is valid, if the refractive index increases or decreases as x increases, respectively.

In case of a constant refractive index gradient

$$n_{Si}(x) = n_{Si}(x_0) + \frac{dn_{Si}}{dx}(x - x_0) \; , \tag{C.6}$$

a complete analytical solution can be derived by carrying out the above integration

$$\boxed{x - x_0 = l \cdot \cosh\left(\frac{z - z_0}{l}\right) - l \cdot \frac{n_{Si}(x_0)}{\sqrt{n_{Si}^2(x_0) - \sin^2(\alpha_i)}}} \tag{C.7}$$

where

$$l := \frac{\sqrt{n_{Si}^2(x_0) - \sin^2(\alpha_i)}}{dn_{Si}/dx}; \quad z_0 := l \cdot \operatorname{arcosh}\left(\frac{n_{Si}(x_0)}{\sqrt{n_{Si}^2(x_0) - \sin^2(\alpha_i)}}\right) \; . \tag{C.8}$$

For example, a temperature gradient of $1\,\mathrm{K}/\mu\mathrm{m}$ leads to $l = 21.4\,\mathrm{mm}$, $z_0 = 0$ for $\alpha_i = 0$ and leads to $l = 21.1\,\mathrm{mm}$, $z_0 = 3.1\,\mathrm{mm}$ for $\alpha_i = 30^\mathrm{o}$.

In case of normal incidence ($\alpha_i = 0$), the total beam shift Δx_0 and the angular deflection α_2 (cf. fig. C.2) become:

$$\Delta x_0 = l \cdot [\cosh(L/l) - 1] = \frac{L^2}{2l} + O\left(\frac{L}{l}\right)^4 = \frac{dn_{Si}}{dx} \cdot \frac{L^2}{2n_{Si}(x_0)} + O\left(\frac{L}{l}\right)^4 \tag{C.9}$$

$$\tan(\alpha_2) = \frac{dx}{dz}|_{z=L} = \sinh(L/l) = \frac{L}{l} + O\left(\frac{L}{l}\right)^3 = \frac{dn_{Si}}{dx} \cdot \frac{L}{n_{Si}(x_0)} + O\left(\frac{L}{l}\right)^3 \tag{C.10}$$

If we are only interested in the field distribution at the rear surface, an equivalent point of view is to consider a laser beam emerging from the focal plane at an angle α_2 and propagating without being deflected. This beam leads to the same transmitted field distribution if its origin is shifted by $\Delta\tilde{x}$ from the optical axis (cf. fig. C.2):

$$\begin{aligned} \Delta\tilde{x} &= \Delta x_0 - \frac{L}{2}\tan\alpha_2 = l \cdot \cosh\left(\frac{L}{l}\right) - l - \frac{L}{2} \cdot \sinh\left(\frac{L}{l}\right) \\ &= -\frac{L^4}{24\,l^3} + l \cdot O\left(\frac{L}{l}\right)^6 \; . \end{aligned} \tag{C.11}$$

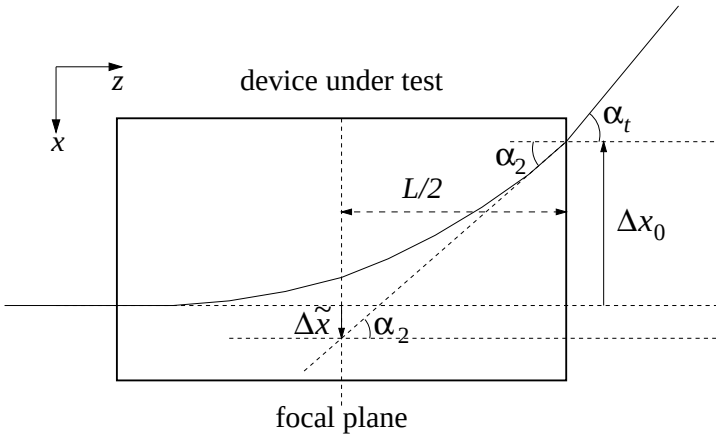


Figure C.2: Normal incidence of the probing beam.

C.2 Image formation by a thin lens

In this section, we consider the image formation by a thin lens (cf. fig. C.3). Assuming the paraxial approximation, Fourier optics (cf. section 3.4) [121] is employed to derive an analytical relation of the field distribution at plane 0 and the field distribution at plane 3. For simplicity reasons, we will restrict ourselves to a 2D calculation.

C.2.1 Image formation by a thin lens

The propagation in free space over the distance d_1 is described by a phase modulation in k_x–space:

$$\begin{aligned}\vec{E}_1(k_x) &= \exp(-i\sqrt{k_0^2 - k_x^2}\, d_1)\vec{E}_0(k_x) \\ &\approx \mathrm{e}^{-ik_0 d_1} \exp\left(i\frac{k_x^2}{2k_0}d_1\right)\vec{E}_0(k_x) \qquad \text{(C.12)}\end{aligned}$$

Inserting the Fourier transformations

$$\vec{E}_0(k_x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{-ik_x x}\vec{E}_0(x)\,dx\ , \qquad \text{(C.13)}$$

$$\vec{E}_1(x) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} \mathrm{e}^{ik_x x}\vec{E}_1(k_x)\,dk_x\ , \qquad \text{(C.14)}$$

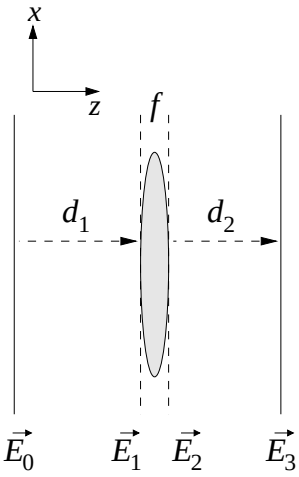


Figure C.3: Image formation by a thin lens of focal length f.

we obtain by lengthy, but straightforward, algebraic manipulations

$$\begin{aligned}\vec{E}_1(x) &= \frac{1}{2\pi}\iint_{-\infty}^{\infty} \mathrm{e}^{ik_x x}\mathrm{e}^{-ik_0 d_1}\mathrm{e}^{i\frac{k_x^2}{2k_0}d_1}\mathrm{e}^{-ik_x x'}\vec{E}_0(x')\,dx'\,dk_x \\ &= \textit{const}\cdot\mathrm{e}^{-ik_0 d_1}\int_{-\infty}^{\infty}\exp\left(-i\frac{k_0(x-x')^2}{2d_1}\right)\vec{E}_0(x')\,dx'\ . \qquad \text{(C.15)}\end{aligned}$$

Since the effect of the thin lens is modeled by a phase shift in real space

$$\vec{E}_2(x) = \exp\left(i\frac{k_0}{2f}x^2\right)\,\vec{E}_1(x) \tag{C.16}$$

the following relation holds:

$$\begin{aligned}\vec{E}_3(x) \;\;=\;\; \textit{const}\cdot \mathrm{e}^{-ik_0(d_1+d_2)} \iint_{-\infty}^{\infty} \exp\left(-i\frac{k_0(x-x')^2}{2d_2}\right)\,\exp\left(i\frac{k_0}{2f}x'^2\right) \\ \cdot\exp\left(-i\frac{k_0(x'-x'')^2}{2d_1}\right)\vec{E}_0(x'')\,dx''\,dx'\;.\end{aligned} \tag{C.17}$$

It is finally transformed to

$$\boxed{\vec{E}_3(x) = \textit{const}\cdot \mathrm{e}^{-ik_0(d_1+d_2)}\exp\left(-i\frac{k_0V_t}{2f}x^2\right)\int_{-\infty}^{\infty}\exp\left[-i\frac{k_0}{a}(x'+V_tx)^2\right]\vec{E}_0(x')\,dx'} \tag{C.18}$$

where

$$V_t := \frac{1}{d_2/f-1} \quad\text{and}\quad a := \frac{2d_1d_2f}{d_2-f}\left(\frac{1}{f}-\frac{1}{d_1}-\frac{1}{d_2}\right) = 2[d_1-f(1+V_t)] = 2[d_1-V_td_2]\;. \tag{C.19}$$

The parameter V_t performs a scaling of the x–axis and therefore constitutes the inverse of the transverse magnification. On the other hand, the parameter a allows for the focussing conditions and vanishes if plane 3 is the image of plane 0. A physical interpretation is gained if the lens is considered to form an image of plane 3. Since the latter is located at $f(1+V_t)$, the parameter a represents twice the distance between plane 0 and the image of plane 3.

C.2.2 Image of a Gaussian field distribution

To calculate the field distribution on the detector during laser deflection measurements, we consider the following model: A Gaussian field distribution rotated by an angle α_2 and shifted by $\Delta\tilde{x}$ (cf. fig. C.4) is emerging from plane 0 and is projected onto plane 3 by a thin lens (cf. fig. C.3). This situation reflects the optical setup of laser deflection measurements, if plane 0 is chosen as the focal plane of the focussing lens, plane 3 represents the detector plane, and the distance d_1 is equal to $d_1 = F + L/2n_{Si}$ (cf. fig. 4.22).

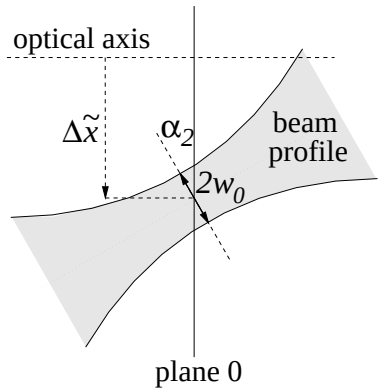


Figure C.4: Gaussian beam propagating at angle α_2.

The field distribution at plane 0 becomes approximately

$$\vec{E}_0(x) = \vec{E}\,\exp\left[-\frac{(x-\Delta\tilde{x})^2}{w_0^2/\cos^2(\alpha_2)}\right]\exp\left[ik_0\sin(\alpha_2)\cdot x\right]\;. \tag{C.20}$$

The resulting image distribution at plane 3

$$|\vec{E}_3(x)| = \textit{const} \cdot \exp\left[-\frac{\left(x - \frac{a\sin(\alpha_2)}{2V_t} + \frac{\Delta\tilde{x}}{V_t}\right)^2}{\frac{w_0^2}{V_t^2\cos^2(\alpha_2)}\left(1 + \frac{a^2\cos^4(\alpha_2)}{k_0^2 w_0^4}\right)} \right] \tag{C.21}$$

represents again a Gaussian profile $\exp[-(x-\Delta x)^2/w^2]$. Its spot radius w and the shift Δx are given by

$$\Delta x = \frac{a}{2V_t}\sin(\alpha_2) - \frac{\Delta\tilde{x}}{V_t} \qquad w = \frac{w_0}{V_t\cos(\alpha_2)}\sqrt{1 + \frac{a^2\cos^4(\alpha_2)}{k_0^2 w_0^4}}\ . \tag{C.22}$$

While the total beam displacement in general is affected by both the internal shift $\Delta\tilde{x}$ and the angular deflection α_2, two interesting special cases can be adjusted: First, if the image of the focal plane in the sample (plane 0) is located exactly on the detector ($a = 0$), the setup is only sensitive to the internal shift $\Delta\tilde{x}$. This condition, however, is hard to realize since a is in the order of some mm and $a \cdot \sin(\alpha_2)$ has to be much smaller than $\Delta\tilde{x}$, which is only some microns. Second, if plane 3 is located at the focal plane of the lens ($d_2 = f$), we obtain $V_t \to \infty$, $a \to \infty$, $a/2V_t \to -f$ and the setup is sensitive to the angular deflection only.

Finally, it should be mentioned that the standardized displacement

$$\frac{\Delta x}{w} = \frac{-\Delta\tilde{x} + \frac{a}{2}\sin(\alpha_2)}{\frac{w_0}{\cos(\alpha_2)}\sqrt{1 + \frac{a^2\cos^4(\alpha_2)}{k_0^2 w_0^4}}} \longrightarrow \pm\frac{k_0 w_0}{2}\tan(\alpha_2) \tag{C.23}$$

is independent of the magnification V_t. In case of a large parameter a, it converges to a limit which is completely independent of all parameters d_1, d_2 and f of the optical setup.

Appendix D

Extracting the Phase Shift Signal

Interferometric techniques are sensitive to phase modulations of the probing beam. As temperature rises within the sample, oscillations of the detected light intensity are observed (cf. fig. D.1). Their evolution is typically described by the relation

$$I(t) = C_1 + C_2 \cos[\varphi(t)] \qquad \text{(D.1)}$$

in which the coefficients C_1 and C_2 may additionally display a weak time dependence. It originates from the deflection of the probing beam on account of the refractive index gradients perpendicular to the optical axis (cf. sections 5.2 and 5.3). The desired information about the evolution of the measurands, i. e. the temperature and the carrier concentration, is included in the phase shift $\varphi(t)$. It is obvious, that $\varphi(t)$ is shifted by π if two adjacent extrema of the detector signal are observed. However, we are interested in a continuous extraction of $\varphi(t)$ which can be achieved by the following strategies:

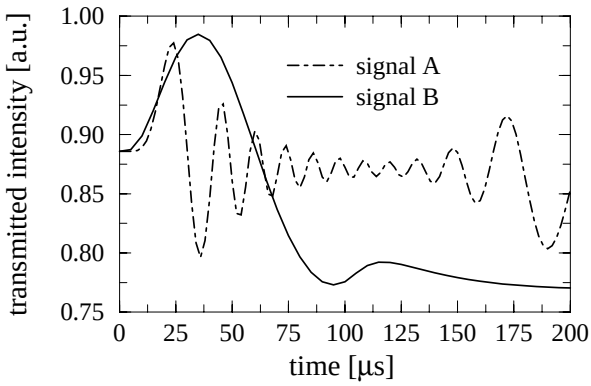


Figure D.1: Measurement signals of interferometric techniques. The sample is subjected to a current pulse during $0 < t < 100\ \mu$s.

D.1 Direct extraction

Equation (D.1) can be solved for the desired phase shift signal

$$\varphi(t) = \arccos\left(\frac{I(t) - C_1}{C_2}\right) \qquad \text{(D.2)}$$

provided that the coefficients C_1 and C_2 are known. They can be approximately extracted from the minimum and maximum detected intensity. However, any kind of noise will thus introduce an error to the phase shift signal. In addition, in case of a decreasing amplitude (cf. signal A in fig. D.1) the formula must not be applied to the whole period of time. In fact, within each interval

between two adjacent extrema a separate cos–function has to be fitted and inverted, thus raising additional effort.

A similar strategy to evaluate Fabry–Perot oscillations is reported in [98]: In order to interpolate between two extrema, a qualitative function (e. g. polynomial or exponential) describing the temperature evolution is assumed. Its parameters are extracted by exploiting the symmetry of the oscillations around their extrema.

D.2 Transformation to frequency space

A very comfortable method to evaluate oscillating signals with several periods employs Fourier transformations and an appropriate filtering in frequency space [206]. To illustrate the principle, let's consider an intensity signal $I(t) = C_1 + C_2 \cos(\varphi)$ with an approximately linearly increasing phase $\varphi(t) = \omega_\varphi \cdot t$. Its Fourier spectrum contains three components, namely the constant offset at the frequency 0 and two peaks at the frequencies $\pm\omega_\varphi$. The latter is due to the fact that our original signal is *real*. Multiplying with a band pass filter around ω_φ cancels the offset and the component with the negative frequency. Thus, the transformation back to time space yields a *complex* signal $\tilde{I}(t) = C_2 \exp[i\varphi(t)]$, the phase of which can be extracted by

$$\varphi(t) = \arctan\left(\frac{\mathrm{Im}[\tilde{I}(t)]}{\mathrm{Re}[\tilde{I}(t)]}\right) \quad . \tag{D.3}$$

Note that an appropriate choice of the filter function is the key for a successful and accurate determination of the phase evolution.

This strategy implies two remarkable advantages: First, it also works if C_1 and C_2 are slightly time–dependent as these Fourier components are observed at different frequencies. And second, it can be transformed to an automatically applicable algorithm which is preferable against manual counting of the interference extrema. However, it fails in case of a slow phase modulation (cf. signal B in fig. D.1) since it is difficult to separate the constant offset from the component ω_φ in frequency space.

List of Symbols

a	cm	$a := 2(L/2n_S i + F - V_t d_2)$, (cf. eq. 4.26, fig. 4.22, eq. C.19, fig. C.3)
A	–	standardized absorption signal, i. e. AC component of the transmitted intensity scaled to its DC component
A_D	cm^2	active area of the device under test
A_j	cm^{-3}	expansion coefficient (cf. eq. 4.16)
$A_n, \tilde{A}_n$	–	for the derivation of the propagator matrix: functions depending on ω, κ, z_n, z_{n+1}, $\varepsilon(x, z_n)$, and $\varepsilon(x, z_{n+1})$ (cf. eq. A.4)
b	cm	$b := F - V_t d_2 = F - f d_2/(d_2 - f)$, (cf. fig. 4.22)
$B_n, \tilde{B}_n$	–	for the derivation of the propagator matrix: functions depending on ω, κ, z_n, z_{n+1}, $\varepsilon(x, z_n)$, and $\varepsilon(x, z_{n+1})$ (cf. eq. A.4)
B_x	$\mathrm{Vs/cm}^2$	x–component of the magnetic field
B_y	$\mathrm{Vs/cm}^2$	y–component of the magnetic field
B_z	$\mathrm{Vs/cm}^2$	z–component of the magnetic field
$B^a_\pm$	–	$N_x \times 2N_x$ matrices extracting the forward and backward propagating wave at the entrance plane, respectively, from the vector $\vec{u}(z_1)$ of the computation variables in real space
$\tilde{B}^a_\pm$	–	matrix representation of $B^a_\pm$ in k_x–z–space (cf. eq. 3.23)
$B^b_\pm$	–	$N_x \times 2N_x$ matrices extracting the forward and backward propagating wave at the emergence plane, respectively, from the vector $\vec{u}(z_{N_z})$ of the computation variables in real space
$\tilde{B}^b_\pm$	–	matrix representation of $B^b_\pm$ in k_x–z–space (cf. eq. 3.23)
c	$\mathrm{cm/s}$	velocity of light
c_n	$\mathrm{J/K\,cm}^3$	heat capacity of the electron gas

c_p	$\mathrm{J/K\,cm^3}$	heat capacity of the hole gas
c_L	$\mathrm{J/K\,cm^3}$	heat capacity of the host lattice
c_t	$\mathrm{J/K\,cm^3}$	total heat capacity per volume
C	F	capacitance
C_n	$\mathrm{cm^6/s}$	Auger coefficient of the electrons
$C_n, \tilde{C}_n$	$-$	for the derivation of the propagator matrix: functions depending on ω, κ, z_n, z_{n+1}, $\varepsilon(x, z_n)$, and $\varepsilon(x, z_{n+1})$ (cf. eq. A.4)
C_p	$\mathrm{cm^6/s}$	Auger coefficient of the holes
D	cm	cell width of an IGBT (cf. fig. 4.7)
$\vec{D}$	$\mathrm{As/cm^2}$	electrical displacement
D_{amb}	$\mathrm{cm^2/s}$	ambipolar diffusion coefficient
$D_n, \tilde{D}_n$	$-$	for the derivation of the propagator matrix: functions depending on ω, κ, z_n, z_{n+1}, $\varepsilon(x, z_n)$, and $\varepsilon(x, z_{n+1})$ (cf. eq. A.4)
d_1	cm	distance between the preimage plane and the imaging lens (cf. fig. C.3)
d_2	cm	distance between the imaging lense and the detector plane (cf. fig. 4.22 and C.3)
$\vec{E}$	$\mathrm{V/cm}$	electric field
E^B	$\mathrm{V/cm}$	computational variable representing the envelope of the backward propagating wave
E_c	eV	conduction band minimum
E_{c0}	eV	conduction band minimum in the absence of an external electric field
E^F	$\mathrm{V/cm}$	computational variable representing the envelope of the forward propagating wave
E_g	eV	energy gap
E_i	$\mathrm{V/cm}$	electric field (y–component) of the incident laser beam at the entrance plane
$\vec{E}_i$	$\mathrm{V/cm}$	in the discretized equations: vector comprising the incident beam's electric fields (y–components) at the grid nodes of the entrance plane

$\vec{E}_{pr}$	$\mathrm{V/cm}$	electric field of the probing beam
$\vec{E}_{pr,0}$	$\mathrm{V/cm}$	amplitude of the electric field of the probing beam
$\vec{E}_r$	$\mathrm{V/cm}$	in the discretized equations: vector comprising the reflected beams's electric fields (y–components) at the grid nodes of the entrance plane
$\vec{E}_r(x,y,t)$	$\mathrm{V/cm}$	field distribution of the reflected wave on the front surface of the sample at time t
$\vec{E}_{ref}$	$\mathrm{V/cm}$	electric field of the reference beam
$\vec{E}_{ref,0}$	$\mathrm{V/cm}$	amplitude of the electric field of the reference beam
E_{rel}	$-$	for virtual experiments: average relative error (cf. eq. 4.8)
$E_{rel,opt}$	$-$	for virtual experiments: minimum of E_{rel}
E_t	eV	in the band diagram of a semiconductor: energy level of a trap
$\vec{E}_t$	$\mathrm{V/cm}$	in the discretized equations: vector comprising the transmitted beams's electric fields (y–components) at the grid nodes of the sample's rear surface
$\vec{E}_t(x,y,t)$	$\mathrm{V/cm}$	field distribution of the transmitted wave on the rear surface of the sample at time t
E_v	eV	valence band maximum
E_{v0}	eV	valence band maximum in the absence of an external electric field
$\vec{e}_x$	$-$	unity vector in x–direction
E_x	$\mathrm{V/cm}$	x–component of the electric field
$\vec{e}_y$	$-$	unity vector in y–direction
E_y	$\mathrm{V/cm}$	y–component of the electric field
$\vec{e}_z$	$-$	unity vector in z–direction
E_z	$\mathrm{V/cm}$	z–component of the electric field
$E_{\parallel}$	$\mathrm{V/cm}$	electric field component parallel to the current density
$E_{\perp}$	$\mathrm{V/cm}$	electric field component perpendicular to the current density
f	cm	focal length of a lens, in particular the focal length of the imaging lens for the laser deflection measurements

F	cm	distance between the rear surface of the sample and the imaging lens
F_k	different	driving affinity
$f_n(E)$	eV^{-1}	Fermi distribution function
$F_{1/2}(x)$	$-$	Fermi integral
g	cm	width of the gap between two segments of a four–quadrant detector (cf. fig. 4.23)
G	$\mathrm{cm}^{-3}\mathrm{s}^{-1}$	total carrier generation rate
G_{av}	$\mathrm{cm}^{-3}\mathrm{s}^{-1}$	Avalanche generation rate
h	$\mathrm{W/K\,cm}^2$	heat exchange coefficient of an interface
$\hbar$	eVs	Planck's constant
H	$\mathrm{W/cm}^3$	total heat generation rate
H_{eJoule}	$\mathrm{W/cm}^3$	Joule heat of the electrons (cf. eq. 6.41)
H_{hJoule}	$\mathrm{W/cm}^3$	Joule heat of the holes (cf. eq. 6.41)
$H_{Peltier}$	$\mathrm{W/cm}^3$	Peltier heat (cf. eq. 6.43)
H_{rec}	$\mathrm{W/cm}^3$	recombination heat (cf. eq. 6.42)
$H_{Thomson}$	$\mathrm{W/cm}^3$	Thomson heat (cf. eq. 6.43)
H_{trans}	$\mathrm{W/cm}^3$	heat generation due to transient carrier injection (cf. eq. 6.46)
i	$-$	imaginary unit
I	A	terminal current
$I_D(x,y)$	$\mathrm{W/cm}^2$	intensity distribution of the laser light on the detector
$I_k(t)$	A	photo current of the kth segment of a four–quadrant photo detector
I_k^{off}	A	photo current of the kth segment of a four–quadrant photo detector during the off–state of the device under test
I_{off}	$\mathrm{W/cm}^2$	light intensity transmitted by the sample during its off–state
I_{on}	$\mathrm{W/cm}^2$	light intensity transmitted by the sample during its on–state
I_t	$\mathrm{W/cm}^2$	transmitted light intensity

I_0	$\mathrm{W/cm^2}$	incident light intensity
j_n	$\mathrm{cm^{-2}s^{-1}}$	particle current density of the electrons
J_n	$\mathrm{A/cm^2}$	electrical current density of the electrons
j_p	$\mathrm{cm^{-2}s^{-1}}$	particle current density of the holes
J_p	$\mathrm{A/cm^2}$	electrical current density of the holes
j_Q	$\mathrm{W/cm^2}$	heat current density
J_s	$\mathrm{W/K\,cm^2}$	entropy current density
J_u	$\mathrm{W/cm^2}$	energy current density
k	$\mathrm{eV/K}$	Boltzmann's constant
$K_n, \tilde{K}_n$	$\mathrm{cm^2}$	for the derivation of the propagator matrix: constants depending on ω, κ, z_n, and z_{n+1} (cf. eq. A.4)
k_x	$1/\mathrm{cm}$	for the simulation of beam propagation: component of the wave vector perpendicular to the optical axis
k_z	$1/\mathrm{cm}$	for the simulation of beam propagation: component of the wave vector along the optical axis
k_z^a	$1/\mathrm{cm}$	$k_z^a = \sqrt{\varepsilon_a k_0^2 - k_x^2}$, i. e. z–component of the wave vector in front of the entrance plane ($z < z_1$)
k_z^b	$1/\mathrm{cm}$	$k_z^b = \sqrt{\varepsilon_b k_0^2 - k_x^2}$, i. e. z–component of the wave vector behind the sample's rear surface ($z > z_{N_z}$)
k_0	$1/\mathrm{cm}$	absolut value of the wave vector of the probing beam
l	cm	for calculation of the beam deflection within a constant gradient dn_{Si}/dx of the refractive index: $l := \frac{n_{Si}(x_0)}{dn_{Si}/dx}$
L	cm	interaction length
L_a	cm	length of the active area
L_D	cm	ambipolar diffusion length
$L_{eff,lin}$	cm	effective interaction length in case of a linear dependence (cf. eq. 4.31)
$L_{eff,qu}$	cm	effective interaction length in case of a quadratic dependence (cf. eq. 4.32)

L_{kj}	different	3×3 submatrices forming the coeffcients of the Onsager matrix (cf. eq. 6.22)
L_j	cm	$\frac{1}{L_j} := \sqrt{\frac{1}{L_D^2} + \gamma_j^2 \frac{4}{L^2}}$, cf. eq. (4.16)
L_s	cm	sample length
M	$-$	standardized deflection signal (difference between the photo currents of two opposite segments scaled to the sum of all photo currents of a four–quadrant detector)
m_n^*	kg	effective mass of an electron at the conduction band minimum
m_p^*	kg	effective mass of a hole at the valence band maximum
$M_{n,j}^k$	$-$	for the derivation of the propagator matrix: 2×2 matrices whose coefficients depend on ω, κ, z_n, z_{n+1}, $x_{j+1} - x_j$, $x_j - x_{j-1}$, $\varepsilon(x_j, z_n)$, and $\varepsilon(x_j, z_{n+1})$; range of the indices: $1 \leq k \leq 6$, $1 \leq n \leq N_z - 1$, $1 \leq j \leq N_x$ (cf. eq. A.6)
$M_{1,n}, M_{2,n}$	$-$	for the derivation of the propagator matrix: band structured matrices whose coefficients are complex 2×2 submatrices (cf. eq. A.8)
n	cm^{-3}	concentration of free electrons
N	$-$	for virtual experiments: number of calculated detector signals
$N(x)$	cm^{-3}	with absorption measurements: $N(x) := \frac{1}{L} \int_{-L/2}^{L/2} n(x,z)\, dz$
$\vec{n}$	$-$	unity vector normal to an interface
Δn	cm^{-3}	excess electron concentration $\Delta n := n - n_0$
N_A	cm^{-3}	concentration of acceptors
$N_{A,Substrate}$	cm^{-3}	acceptor concentration of the p^+–substrate of a PT–IGBT
N_A^-	cm^{-3}	concentration of ionized acceptors
N_c	cm^{-3}	effective density of states in the conduction band
N_D	cm^{-3}	concentration of donors
$N_{D,Buffer}$	cm^{-3}	donor concentration of the n^+–buffer of a PT–IGBT
N_D^+	cm^{-3}	concentration of ionized donors

n_{extr}	cm^{-3}	electron concentration (=carrier concentration) extracted from the detector signal of an internal laser probing technique
n_i	cm^{-3}	intrinsic carrier concentration
$n_{i,\mathit{eff}}$	cm^{-3}	effective intrinsic carrier concentration
$n_{ie,n}$	cm^{-3}	effective intrinsic electron concentration (cf. eq. 6.31)
$n_{ie,p}$	cm^{-3}	effective intrinsic hole concentration (cf. eq. 6.31)
N_i	cm^{-3}	impurity concentration
n_L	–	refractive index of a lens
n_{min}	cm^{-3}	minimum concentration of the carrier distribution
N_n^{ref}	cm^{-3}	reference concentration in the Scharfetter relation for the electrons (cf. eq. 6.60)
N_p^{ref}	cm^{-3}	reference concentration in the Scharfetter relation for the holes (cf. eq. 6.60)
n_r	–	static refractive index
n_{ref}	cm^{-3}	electron concentration (=carrier concentration) obtained by an electrothermal device simulation (reference concentration for the virtual experiment)
n_{Si}	–	real part of the optical refractive index at the laser frequency of the probing beam
$\left(\frac{\partial n_{Si}}{\partial C}\right)_T$	cm^3	$\left(\frac{\partial n_{Si}}{\partial C}\right)_T := \left(\frac{\partial n_{Si}}{\partial n}\right)_{p,T} + \left(\frac{\partial n_{Si}}{\partial p}\right)_{n,T}$
$n_{Si,0}$	–	real part of the optical refractive index at the laser frequency of the probing beam during the off–state of the investigated device
N_t	cm^{-3}	density of trap centers
N_v	cm^{-3}	effective density of states in the valence band
N_x	–	for the simulation of beam propagation: number of grid nodes in x–direction, i. e. perpendicular to the optical axis
N_z	–	for the simulation of beam propagation: number of grid nodes in z–direction, i. e. along the optical axis
n_0	cm^{-3}	equilibrium concentration of the electrons
n_1	cm^{-3}	$n_1 = n_{i,\mathit{eff}} \exp[(E_t - E_F)/kT]$

p	cm^{-3}	concentration of holes
P	–	for the simulation of beam propagation: propagator matrix between the entrance plane and the emergence plane (cf. eq. 3.14)
Δp	cm^{-3}	excess hole concentration $\Delta p := p - p_0$
P_n	–	for the simulation of beam propagation: propagator matrix between the planes $z = z_n$ and $z = z_{n+1}$ (cf. eq. 3.14 and appendix A)
P_n	$\mathrm{V/K}$	thermoelectric power of the electrons
$\vec{P}_n$	$\mathrm{C/cm}^2$	electric polarization due to a displacement of the electron gas
P_n^{ext}	$\mathrm{V/K}$	at the boundaries of the simulation domain: electron thermoelectric power of the external region
P_n^{sc}	$\mathrm{V/K}$	at the boundaries of the simulation domain: electron thermoelectric power of the semiconductor
P_p	$\mathrm{V/K}$	thermoelectric power of the holes
$\vec{P}_p$	$\mathrm{C/cm}^2$	electric polarization due to a displacement of the hole gas
P_p^{ext}	$\mathrm{V/K}$	at the boundaries of the simulation domain: hole thermoelectric power of the external region
P_p^{sc}	$\mathrm{V/K}$	at the boundaries of the simulation domain: hole thermoelectric power of the semiconductor
p_0	cm^{-3}	equilibrium concentration of the holes
p_1	cm^{-3}	$p_1 = n_{i,eff} \exp[(E_F - E_t)/kT]$
q	C	elementary charge
Q	–	slope of the detector response as function of the standardized beam displacement, i. e. $Q := \frac{\partial M}{\partial(\Delta x/w)}\|_{\Delta x=0}$
$\tilde{Q}$	cm^{-1}	slope of the detector response as function of the absolute beam displacement, i. e. $\tilde{Q} := \frac{\partial M}{\partial \Delta x}\|_{\Delta x=0} = Q/w$
ΔQ	J	heat stored within the device during the duty cycle
Q_S	C	excess charge stored within a power devices during its forward conducting state
$\vec{r}$	cm	position of a particle

R	$\mathrm{cm}^{-3}\mathrm{s}^{-1}$	total carrier recombination rate
R_{Au}	$\mathrm{cm}^{-3}\mathrm{s}^{-1}$	Auger recombination rate
R_{CDL}	$\mathrm{cm}^{-3}\mathrm{s}^{-1}$	coupled defect level recombination rate
r_f	$-$	reflection coefficient at the front surface of a Fabry–Perot resonator
R_{FP}	$-$	reflectivity of a Fabry–Perot resonator
r_r	$-$	reflection coefficient at the rear surface of a Fabry–Perot resonator
R_{SRH}	$\mathrm{cm}^{-3}\mathrm{s}^{-1}$	Shockley–Read–Hall recombination rate
R_{Surf}	$\mathrm{cm}^{-2}\mathrm{s}^{-1}$	surface recombination rate
R_1	cm	radius of the front curvature of a spherical lens
R_2	cm	radius of the rear curvature of a spherical lens
s	$\mathrm{J/K\,cm}^3$	entropy density of a thermodynamic system
s	cm	for optical probing techniques: segment width of a four–quadrant detector (cf. fig. 4.23)
s_L	$\mathrm{J/K\,cm}^3$	entropy density of the host lattice
s_n	$\mathrm{J/K\,cm}^3$	entropy density of the electron gas
s_p	$\mathrm{J/K\,cm}^3$	entropy density of the hole gas
t	s	time
T	K	absolute temperature
ΔT	K	temperature rise
T_{ext}	K	at the boundaries of the simulation domain: temperature of the adjacent region
T_{extr}	K	absolute temperature extracted from the detector signal of an internal laser probing technique
T_{FP}	$-$	transmittance of a Fabry–Perot resonator
T_L	K	lattice temperature
T_n	K	temperature of the electron gas
T_p	K	temperature of the hole gas

ΔT_{pp}	K	temperature rise causing two adjacent maxima of thermally induced Fabry–Perot oscillations
T_{ref}	K	absolute temperature obtained by an electrothermal device simulation (reference temperature for the virtual experiment)
u	$\mathrm{J/cm^3}$	energy density of a thermodynamic system
$\vec{u}(z_n)$	$\mathrm{V/cm}$	for the simulation of beam propagation: vector comprising the unknown field components at the plane $z = z_n$ (cf. eq. 3.11)
u_{el}	$\mathrm{J/cm^3}$	energy density of the electrostatic field
u_L	$\mathrm{J/cm^3}$	energy density of the host lattice
u_n	$\mathrm{J/cm^3}$	energy density of the electron gas
u_p	$\mathrm{J/cm^3}$	energy density of the hole gas
V	V	voltage
V	$\mathrm{cm^3}$	volume of the device under test
∂V	$\mathrm{cm^2}$	boundary of the volume V
V_{CE}	V	collector–emitter voltage
V_{GE}	V	gate–emitter voltage
v_n	$\mathrm{cm/s}$	drift velocity of the electrons
v_p	$\mathrm{cm/s}$	drift velocity of the holes
v_{sat}	$\mathrm{cm/s}$	saturation drift velocity of a carrier
v_{sr}	$\mathrm{cm/s}$	surface recombination velocity
$v_{sr,n}$	$\mathrm{cm/s}$	surface recombination velocity of the electrons
$v_{sr,p}$	$\mathrm{cm/s}$	surface recombination velocity of the holes
V_t	$-$	inverse of the transverse magnification of an optical image formation (cf. eq. C.19)
v_{th}	$\mathrm{cm/s}$	thermal velocity
w	cm	spot radius of the probing beam on the detector
W	cm	average spot radius of a Gaussian beam

$w(z)$	cm	spot radius of a Gaussian beam
W_{opt}	cm	optimum average spot radius of a Gaussian beam which minimizes the experimental error (cf. eq. 4.10)
w_0	cm	spot radius of a Gaussian beam in the focal plane
$w_{0,opt}$	cm	optimum spot radius in the focal plane which minimizes the experimental error
Δx	cm	displacement of the probing beam on the detector (cf. fig. 4.22)
$\Delta\tilde{x}$	cm	displacement of the probing beam if the internal deflection is extrapolated onto the focal plane of the focussing lens (cf. fig. 4.22)
x_i	cm	for the simulation of beam propagation: x–coordinate of the $i.$ vertex perpendicular to the optical axis
x_{min}	cm	position where the carrier concentration is at its minimum
Δx_0	cm	displacement of the probing beam on the rear surface of the sample (cf. fig. 4.22)
x_{N_x}	cm	for the simulation of beam propagation: x–coordinate of the right hand boundary perpendicular to the optical axis
x_0	cm	x–coordinate of the incident probing beam (cf. fig. 4.22)
x_1	cm	for the simulation of beam propagation: x–coordinate of the left hand boundary perpendicular to the optical axis
y_0	cm	y–coordinate of the incident probing beam
z_i	cm	for the simulation of beam propagation: z–coordinate of the $i.$ vertex on the optical axis
z_{N_z}	cm	for the simulation of beam propagation: z–coordinate of the emergence plane
z_1	cm	for the simulation of beam propagation: z–coordinate of the entrance plane
α	$1/\mathrm{cm}$	absorption coefficient at the laser frequency of the probing beam
$\left(\frac{\partial\alpha}{\partial C}\right)_T$	cm^2	$\left(\frac{\partial\alpha}{\partial C}\right)_T := \left(\frac{\partial\alpha}{\partial n}\right)_{p,T} + \left(\frac{\partial\alpha}{\partial p}\right)_{n,T}$
α_n	$1/\mathrm{cm}$	ionization coefficient of the electrons in the avalanche generation

α_p	$1/\mathrm{cm}$	ionization coefficient of the holes in the avalanche generation
α_t	–	angle between the emerging probing beam and the rear surface of the sample (cf. fig. 4.22)
α_0	$1/\mathrm{cm}$	absorption coefficient at the laser frequency of the probing beam for the off–state of the investigated device
α_2	–	angle between the internal beam path and the rear surface of the sample (cf. fig. 4.22)
γ	–	spatial frequency of the expansion (4.12)
γ_j	–	discrete solutions for γ (cf. eq. 4.15)
$\delta(x)$	$1/\mathrm{cm}$	Dirac's delta distribution
Δ	cm	spatial resolution of free carrier absorption measurements (cf. section 4.2.5)
$\Delta(x,y)$	cm	thickness of a lens at position (x,y) (cf. fig. 3.4)
Δ_0	cm	maximum thickness of a lens (cf. fig. 3.4)
ε	$\mathrm{C/V\,cm}$	static dielectric constant
ε_a	–	relative dielectric constant in front of the entrance plane, i. e. for $z < z_1$
ε_b	–	relative dielectric constant behind the emergence plane, i. e. for $z > z_{N_z}$
ε_{ins}	$\mathrm{C/V\,cm}$	at a semiconductor/insulator interface: static dielectric constant of the insulator
ε_R	–	relative dielectric constant at the laser frequency of the probing beam
ε_{sc}	$\mathrm{C/V\,cm}$	at a semiconductor/insulator interface: static dielectric constant of the semiconductor
ε_0	$\mathrm{C/V\,cm}$	absolute dielectric constant
η_A	–	angular aperture of the probing beam
$\eta_{A,opt}$	–	optimum angular aperture of the probing beam which minimizes the experimental error
$\Theta(x)$	–	unit step function: $\Theta(x) = 0$ for $x < 0$ and $\Theta(x) = 1$ for $x \geq 0$
κ	$1/\mathrm{cm}$	for the simulation of beam propagation: global parameter for the transformation to the computational variables (cf. eq. 3.7)

κ_{sc}	$\mathrm{W/K\,cm}$	at the boundaries of the simulation domain: thermal conductivity of the semiconductor
κ_t	$\mathrm{W/K\,cm}$	total thermal conductivity
λ	cm	wave length of the probing beam
μ_{ac}	$\mathrm{cm^2/Vs}$	contribution to the total mobilbity due to surface acoustic phonon scattering
μ_b	$\mathrm{cm^2/Vs}$	bulk mobility of a carrier
μ_{eh}	$\mathrm{cm^2/Vs}$	contribution to the total mobilbity due to carrier–carrier scattering
μ_i	$\mathrm{cm^2/Vs}$	contribution to the total mobilbity due to scattering mechanism i
μ_{low}	$\mathrm{cm^2/Vs}$	mobility of a carrier in a weak electric field
μ_n	$\mathrm{cm^2/Vs}$	electron mobility
$\mu_n^{(c)}$	eV	chemical potential of the electron gas
$\mu_n^{(ec)}$	eV	electrochemical potential of the electron gas
μ_p	$\mathrm{cm^2/Vs}$	hole mobility
$\mu_p^{(c)}$	eV	chemical potential of the hole gas
$\mu_p^{(ec)}$	eV	electrochemical potential of the hole gas
μ_{sr}	$\mathrm{cm^2/Vs}$	contribution to the total mobilbity due to surface roughness scattering
Π_s	$\mathrm{W/K\,cm^3}$	entropy production rate $\Pi_s := \mathrm{div}\,\vec{J}_s + \partial s/\partial t$
σ_n	$\mathrm{cm^2}$	capture cross section of an empty trap capturing an electron from the conduction band
σ_p	$\mathrm{cm^2}$	capture cross section of an occupied trap capturing a hole from the valence band
σ_{Surf}	$\mathrm{C/cm^2}$	interface charge density
τ_{amb}	s	ambipolar carrier lifetime
τ_B	s	minority carrier lifetime of the holes in the $\mathrm{n^+}$–buffer (cf. eq. 8.2)
τ_{eff}	s	effective carrier lifetime $\tau_{\mathit{eff}} = Q_S/I$
τ_n	s	recombination lifetime of the electrons

τ_n^c	s	average collision time of the electrons
τ_n^0	s	electron lifetime in an intrinsic semiconductor (cf. eq. 8.1)
τ_p	s	recombination lifetime of the holes
τ_p^c	s	average collision time of the holes
τ_p^0	s	hole lifetime in an intrinsic semiconductor (cf. eq. 8.1)
τ_S	s	minority carrier lifetime of the electrons in the p^+–substrate (cf. eq. 8.2)
τ_{SRH}	s	Shockley–Read–Hall carrier lifetime (cf. eq. 6.59)
$\varphi(t)$	–	phase shift between the light rays of the orders n and $n+1$ of an interferometer
$\varphi(x,y)$	–	phase shift of the wavefront within a thin lens
$\Delta\varphi$	–	phase shift of the probing beam with reference to the reference beam
ϕ_n	V	quasi–Fermi level of the electron gas
ϕ_p	V	quasi–Fermi level of the hole gas
Φ	$^\mathrm{o}$	with oblique transition of the probing beam: angle of incidence
Φ_{Si}	$^\mathrm{o}$	with oblique transition of the probing beam: angle of the propagation direction in the silicon
ψ	V	electrostatic potential
ω	$1/\mathrm{s}$	laser frequency of the probing beam
$\Delta\omega$	$1/\mathrm{s}$	$\Delta\omega := \omega_1 - \omega_2$
ω_1	$1/\mathrm{s}$	driving frequency of the acousto–optic modulator which determines the diffraction of the probing beam
ω_2	$1/\mathrm{s}$	driving frequency of the acousto–optic modulator which determines the diffraction of the reference beam

Bibliography

[1] J. Bardeen and W. H. Brattain, “The transistor, a semiconductor triode,” Phys. Rev., vol. 74, pp. 230, 1948.

[2] W. Shockley, “The theory of pn junction in semiconductors and pn junction transistor,” Bell Syst. Tech. J., vol. 28, pp. 435, 1949.

[3] B. J. Baliga, “Modern power devices,” John Wiley & Sons Inc., New York, USA, 1987.

[4] B. J. Baliga, “Trends in power semiconductor devices,” IEEE Trans. on Electron Devices, vol. 43, pp. 1717, 1996.

[5] A. Porst, F. Auerbach, H. Brunner, G. Deboy, and F. Hille, “Improvement of the diode characteristics using emitter–controlled principles (emcon–diode),” Proc. of ISPSD, Weimar, Germany, pp. 213, 1997.

[6] P. C. Mourick, “Das Abschaltverhalten von Leistungsdioden,” Dissertation, Technische Universität Berlin, Germany, 1988.

[7] G. Sölkner, “Optical beam testing and its potential for electronic device characterization,” Microelectronic Engineering, vol. 24, pp. 341, 1994.

[8] J. M. Soden and R. E. Anderson, “IC failure analysis: techniques and tools for quality and reliability measurements,” Microelectronic Reliability, vol. 35, pp. 429, 1995.

[9] S. Lindsay, “Laser light spot mapping of depletion in power semiconductor devices,” Phys. Stat. Solidi (A), vol. 53, pp. 311, 1979.

[10] R. Stengl, “Physikalische Untersuchungen an hochsperrenden planaren pn-Übergängen,” Dissertation, Friedrich–Alexander–Universität Erlangen–Nürnberg, Germany, 1986.

[11] H. Schink, “Elektronenoptische Darstellung von Feldverteilungen in Halbleiterbauelementen,” Dissertation, Technische Universität München, Germany, 1984.

[12] G. Wachutka, warning to the students during the lectures about numerical modeling and simulation.

[13] A. Hoffmann and K. Schuster, “An experimental determination of the carrier lifetime in pin diodes from the stored carrier charge,” Solid State Electronics, vol. 7, pp. 717, 1964.

[14] H. Schlangenotto and W. Gerlach, “On the effective carrier lifetime in psn rectifiers at high injection level,” Solid State Electronics, vol. 12, pp. 267, 1969.

[15] H. J. Kuno, “Analysis and characterization of pn junction diode switching,” IEEE Trans. on Electron Devices, pp. 8, Jan. 1964.

[16] H. Benda and E. Spenke, “Reverse recovery processes in silicon power rectifiers,” Proc. IEEE, vol. 55, pp. 1331, 1967.

[17] Y. C. Kao and J. R. Davis, “Correlations between reverse recovery time and lifetime of pn junction driven by a current ramp,” IEEE Trans. on Electron Devices, vol. 17, pp. 652, 1970.

[18] F. Berz, “Step recovery of pin diodes,” Solid State Electronics, vol. 22, pp. 927, 1979.

[19] F. Berz, “Ramp recovery in pin diodes,” Solid State Electronics, vol. 23, pp. 783, 1980.

[20] K. P. Brieger, “Untersuchungen an pin–Dioden im Hochstrombereich zur Bestimmung der Trägerlebensdauer, der h–Parameter, der Summe der Beweglichkeiten sowie der mittleren Ladungsträgerkonzentration,” Dissertation, Technische Universität Berlin, Germany, 1985.

[21] B. R. Gossick, “Post–injection barrier electromotive force of pn junctions,” Phys. Rev., vol. 91, pp. 1012, 1953.

[22] S. R. Lederhandler and L. J. Giacoletto, “Measurement of minority carrier lifetime and surface effects in junction devices,” Proc. IRE, vol. 43, pp. 477, 1955.

[23] L. W. Davies, “The use of pln structures in investigations of transient recombination from high injection levels in semiconductors,” Proc. IEEE, vol. 51, pp. 1637, 1963.

[24] S. C. Choo and R. G. Mazur, “Open circuit voltage decay behavior of junction devices,” Solid State Electronics, vol. 13, pp. 553, 1970.

[25] W. Gerlach and H. Schlangenotto, “On the post–injection voltage decay of psn rectifiers at high injection levels,” Solid State Electronics, vol. 15, pp. 393, 1972.

[26] C. Boit, “Über die quantitative Erfassung der Hochinjektionseffekte in flußbelasteten pin–Dioden mit Hilfe der Rekombinationsstrahlungs–Meßmethode,” Dissertation, Technische Universität Berlin, Germany, 1987.

[27] J. R. Haynes and H. B. Briggs, “Radiation produced in germanium and silicon by electron–hole–recombination,” Phys. Rev., vol. 86, pp. 647, 1952.

[28] J. R. Haynes and W. C. Westphal, “Radiation resulting from recombination of holes and electrons in silicon,” Phys. Rev., vol. 101, pp. 1676, 1956.

[29] W. Gerlach, “Untersuchungen über den Einschaltvorgang des Leistungsthyristors,” Telefunken–Zeitung 39, pp. 301, 1966.

[30] W. Gerlach, H. Schlangenotto, and H. Maeder, "On the radiative recombination rate in silicon," Phys. Stat. Solidi (A), vol. 13, pp. 277, 1972.

[31] F. Dannhäuser and J. Krausse, "Die räumliche Verteilung der Rekombination in legierten Silizium–psn–Gleichrichtern bei Belastung in Durchlaßrichtung," Solid State Electronics, vol. 16, pp. 861, 1973.

[32] H. Schlangenotto, H. Maeder, and W. Gerlach, "Temperature dependence of the radiative recombination coefficient in silicon," Phys. Stat. Solidi (A), vol. 21, pp. 357, 1974.

[33] G. Schierwater, "Untersuchung der optischen Absorption an freien Ladungsträgern und der Rekombinationsstrahlung am Elektron–Loch–Plasma von pin–Dioden," Dissertation, Technische Universität Berlin, Germany, 1975.

[34] W. Seifert, "Untersuchungen von psn–Dioden bei starker Injektion," Dissertation, Technische Universität Berlin, Germany, 1975.

[35] R. Jörgens, "Möglichkeiten und Grenzen der Bestimmung von Trägerverteilungen in durchlaßbelasteten psn–Dioden durch Messung der Rekombinationsstrahlung, Teil I: Beschreibung der Meßmethode," Technisches Messen, 49. Jahrgang, Heft 9, pp. 313, 1982.

[36] R. Jörgens, "Möglichkeiten und Grenzen der Bestimmung von Trägerverteilungen in durchlaßbelasteten psn–Dioden durch Messung der Rekombinationsstrahlung, Teil II: Grenzen des Meßverfahrens," Technisches Messen, 49. Jahrgang, Heft 11, pp. 409, 1982.

[37] R. Jörgens, "Untersuchungen des stationären und dynamischen Verhaltens von Ladungsträgern in durchlaßbelasteten psn–Dioden durch Auswerten der Rekombinationsstrahlung," Dissertation, Rheinisch–Westfälische Technische Hochschule Aachen, Germany, 1984.

[38] H. Bleichner, M. Rosling, J. Vobecky, M. Lundqvist, and E. Nordlander, "A comparative study of the carrier distributions in dynamically operating GTO's by means of two optically probed measurement methods," Proc. of ISPSD, Tokyo, Japan, pp. 1095, 1990.

[39] H. B. Briggs and R. C. Fletcher, "Absorption of infrared light by free carriers in germanium," Phys. Rev., vol. 91, pp. 1342, 1953.

[40] N. J. Harrick, "Use of infrared absorption to determine carrier distribution in germanium and surface recombination velocity," Phys. Rev., vol. 101, pp. 491, 1956.

[41] D. E. Houston, S. Krishna, and E. D. Wolley, "Study of charge dynamics in high speed power devices using free carrier absorption measurement," IEDM Tech. Digest, pp. 504, 1976.

[42] R. W. Cooper and D. H. Paxman, "Measurement of charge carrier behavior in pin diodes using a laser technique," Solid State Electronics, vol. 21, pp. 865, 1978.

[43] W. Görtz, “Ein Beitrag zur Bestimmung des Ladungsträgerverhaltens in psn–Dioden im Fall starker Injektion unter Verwendung der Absorptions–Meßmethode,” Dissertation, Rheinisch–Westfälische Technische Hochschule Aachen, Germany, 1984.

[44] H. Bleichner, E. Nordlander, M. Rosling, and S. Berg, “A time–resolved optical system for spatial characterization of the carrier distribution in a gate–turn–off thyristor (GTO),” IEEE Trans. on Instrumentation and Measurement, vol. 39, pp. 473, 1990.

[45] Y. Martin and E. A. Ash, “Photodisplacement techniques for defect detection,” Phil. Trans. Roy. Soc. London, vol. 320, pp. 257, 1986.

[46] M. B. Suddendorf, M. Liu, and M. G. Somekh, “A new high–resolution dual–probe system for detecting and imaging thermal and plasma waves,” Scanning, vol. 14, pp. 247, 1992.

[47] W. Claeys, S. Dilhaire, and V. Quintard, “Laser probing of thermal behavior of electronic components and its application in quality and reliability testing,” Microelectronic Engineering, vol. 24, pp. 411, 1994.

[48] P.-W. Epperlein, “Micro–temperature measurements on semiconductor laser mirrors by reflectance modulations: A newly developed technique for laser characterization,” Jpn. J. Appl. Phys., vol. 32, pp. 5514, 1993.

[49] G. Deboy, G. Sölkner, E. Wolfgang, and W. Claeys, “Absolute measurement of transient carrier concentration and temperature gradients in power semiconductor devices by internal IR–laser deflection,” Microelectronic Engineering, vol. 31, pp. 299, 1996.

[50] A. Rosencwaig, “Thermal wave characterization and inspection of semiconductor materials and devices,” in ”Photoacoustic and Thermal Wave Phenomena in Semiconductors” (ed. by A. Mandelis), North Holland, Amsterdam, 1987.

[51] V. Quintard, S. Dilhaire, D. Lewis, T. Phan, and W. Claeys, “Laser beam thermography of circuits in the particular case of passivated semiconductors,” Microelectronic Engineering, vol. 31, pp. 291, 1996.

[52] A. C. Boccara, D. Fournier, and J. Badoz, “Thermo-optical spectroscopy: Detection by the ”mirage effect”,” Appl. Phys. Lett., vol. 36, pp. 130, 1980.

[53] W. B. Jackson, N. M. Amer, A. C. Boccara, and D. Fournier, “Photothermal deflection spectroscopy and detection,” Appl. Optics, vol. 20, pp. 1333, 1981.

[54] J. C. Murphy and L. C. Aamodt, “Photothermal spectroscopy using optical beam probing: Mirage effect,” J. Appl. Phys., vol. 51, pp. 4580, 1980.

[55] A. Salazar and A. Sanchez-Lavega, “Low temperature thermal diffusivity measurements of gases by the mirage technique,” Rev. Scientific Instruments, vol. 70, pp. 98, 1999.

[56] D. Fournier, C. Boccara, A. Skumanich, and N. M. Amer, “Photothermal investigation of transport in semiconductors: Theory and experiment,” J. Appl. Phys., vol. 59, pp. 787, 1986.

[57] G. Deboy, "Charakterisierung von Leistungshalbleitern durch Interne Laserdeflektion," Dissertation, Technische Universität München, Germany, 1996.

[58] N. Seliger, "Characterization of semiconductor devices by laser interferometry," Dissertation, Vienna University of Technology, Austria, 1998.

[59] H. K. Heinrich, D. M. Bloom, and B. R. Hemmenway, "Noninvasive sheet charge density probe for integrated silicon devices," Appl. Phys. Lett., vol. 48, pp. 1066, 1986.

[60] M. Goldstein, G. Sölkner, and E. Gornik, "Heterodyn interferometer for the detection of electric and thermal signals in integrated circuits through the substrate," Rev. Scientific Instr., vol. 64, pp. 3009, 1993.

[61] M. Goldstein, G. Sölkner, and E. Gornik, "Heterodyn interferometer for the detection of electric and thermal signals in integrated circuits through the substrate," Microelectronic Engineering, vol. 24, pp. 431, 1994.

[62] G. Wachutka, "Rigorous thermodynamic treatment of heat generation and conduction in semiconductor device modeling," IEEE Trans. on Computer Aided Design, vol. 9, pp. 1141, 1990.

[63] J. C. Sturm and C. M. Reaves, "Silicon temperature measurement by infrared absorption: fundamental processes and doping effects," IEEE Trans. on Electron Devices, vol. 39, pp. 81, 1992.

[64] R. A. Soref and B. R. Bennett, "Electrooptical effects in silicon," IEEE J. Quantum Electronics, vol. 23, pp. 123, 1987.

[65] N. W. Ashcroft and N. D. Mermin, "Solid state physics," Holt, Rinehart and Winston, Philadelphia, USA, 1976.

[66] H. C. Huang, S. Yee, and M. Soma, "Quantum calculations of the change of refractive index due to free carriers in silicon with nonparabolic band structure," J. Appl. Phys., vol. 67, pp. 2033, 1990.

[67] B. K. Ridley, "Quantum processes in semiconductors," Oxford University Press, New York, USA, 1982.

[68] W. Kowalsky, "Dielektrische Werkstoffe der Elektrotechnik und Photonik," B. G. Teubner, Stuttgart, Germany, 1993.

[69] H. Y. Fan and M. Becker, "Infrared absorption of silicon," Phys. Rev., vol. 78, pp. 178, 1950.

[70] W. Spitzer and H. Y. Fan, "Infrared absorption in n–type silicon," Phys. Rev., vol. 108, pp. 268, 1957.

[71] D. K. Schroeder, R. N. Thomas, and J. C. Schwartz, "Free carrier absorption in silicon," IEEE Trans. on Electron Devices, vol. 25, pp. 254, 1978.

[72] P. E. Schmid, “Optical absorption in heavily doped silicon,” Phys. Rev. B, vol. 23, pp. 5531, 1981.

[73] C. M. Horwitz and R. M. Swanson, “The optical (free–carrier) absorption of a hole–electron plasma in silicon,” Solid State Electronics, vol. 23, pp. 1191, 1980.

[74] F. Hille, Dissertation, Munich University of Technology, in preparation.

[75] C. Mehnert, “Untersuchungen zur Infrarot–Absorption an freien Ladungsträgern in Leistungshalbleitern,” Diplomarbeit, Lehrstuhl für Technische Elektrophysik, Technische Universität München, Germany, 2000.

[76] A. E. Schlögl, “Theorie and Validierung zur Modellbildung bipolarer Leistungshalbleiter im Temperaturbereich von 100 K bis 400 K,” Dissertation, Technische Universität München, Germany, 2000.

[77] L. S. Yu, Q. Z. Liu, Z. F. Guan, and S. S. Lau, “Direct measurements of the refractive index change in silicon with optically injected carriers,” Appl. Phys. Lett., vol. 68, pp. 1546, 1996.

[78] F. Hille and R. Thalhammer, “Die Eignung der Verfahrens der Internen Laserdeflexion als Charakterisierungsmethode für Leistungsbauelemente,” Techn. Report 1997/5, Inst. für Technische Elektrophysik, Technische Universität München, Germany, 1997.

[79] G. G. Macfarlane, T. P. McLean, J. E. Quarrington, and V. Roberts, “Fine structure in the absorption–edge spectrum of Si,” Phys. Rev., vol. 111, pp. 1245, 1958.

[80] J. C. Sturm and C. M. Reaves, “Fundamental mechanisms and doping effects in silicon infrared absorption for temperature measurement by infrared transmission,” SPIE, vol. 1393, pp. 309, 1990.

[81] H. Rogne, P. J. Timans, and H. Ahmed, “Infrared absorption in silicon at elevated temperatures,” Appl. Phys. Lett., vol. 69, pp. 2190, 1996.

[82] G. E. Jellison and H. H. Burke, “The temperature dependence of the refractive index of silicon at elevated temperatures at several wavelengths,” J. Appl. Phys., vol. 60, pp. 841, 1986.

[83] H. W. Icenogle, B. C. Platt, and W. L. Wolfe, “Refractive indexes and temperature coefficients of germanium and silicon,” Applied Optics, vol. 15, pp. 2348, 1976.

[84] A. N. Magunov, “Temperature dependence of the refractive index of silicon single–crystal in the 300 – 700 K range,” J Opt. Spectroscopy, vol. 73, pp. 205, 1992.

[85] J. A. M. Caulley, V. M. Donnelly, M. Vernon, and I. Taha, “Temperature dependence of the near–infrared refractive index of silicon, gallium arsenide, and indium phosphide,” Phys. Rev. B, vol. 49, pp. 7408, 1994.

[86] N. Seliger, P. Habaš, and E. Gornik, “Time–domain characterization of lattice heating in power VDMOSFETs by means of an interferometric laserprobe technique,” Proc. of ESSDERC, Bologna, Italy, pp. 847, 1996.

[87] M. Bertolotti, V. Bogdanov, A. Ferrari, A. Jascow, N. Nazorova, A. Pikhtin, and L. Schirone, "Temperature dependence of the refractive index in semiconductors," J. Opt. Soc. Am. B, vol. 7, pp. 918, 1990.

[88] J. Linnros, P. Norlin, and A. Hallén, "A new technique for depth resolved carrier recombination measurements applied to proton irradiated thyristors," IEEE Trans. on Electron Devices, vol. 40, pp. 2065, 1993.

[89] O. Tornblad, B. Breitholtz, and M. Östling, "Measurements and simulations of excess carrier distributions in 3.3 kV IGBTs during static conditions and turn–on," Proc. of ISPSD, Weimar, Germany, pp. 81, 1997.

[90] F. Hille, L. Hoffmann, H.-J. Schulze, and G. Wachutka, "Carrier lifetime characterization using an optimized free carrier absorption technique," Proc. of ISPSD, Toulouse, France, pp. 299, 2000.

[91] R. Thalhammer, C. Fürböck, N. Seliger, G. Deboy, E. Gornik, and G. Wachutka, "Internal characterization of IGBTs using the backside laser probing technique – interpretation of measurement by numerical simulation," Proc. of ISPSD, Kyoto, Japan, pp. 199, 1998.

[92] N. Seliger, P. Habaš, and E. Gornik, "A study of backside laser–probe signals in MOSFETs," Microelectronic Engineering, vol. 31, pp. 87, 1996.

[93] N. Seliger, E. Gornik, C. Fürböck, D. Pogany, P. Habaš, R. Thalhammer, and M. Stoisiek, "Characterization of semiconductor devices by infrared laser interferometry," E&I, Sonderheft "Trends in der Mikroelektronik", vol. 115, no. 7/8, pp. 403, 1998.

[94] C. Fürböck, N. Seliger, D. Pogany, M. Litzenberger, E. Gornik, M. Stecher, H. Goßner, and W. Werner, "Backside laserprober characterization of thermal effects during high current stress in smart power ESD protection devices," IEDM Tech. Digest, pp. 691, 1998.

[95] C. Fürböck, R. Thalhammer, M. Litzenberger, N. Seliger, D. Pogany, E. Gornik, and G. Wachutka, "A differential backside laserprobing technique for the investigation of the lateral temperature distribution in power devices," Proc. of ISPSD, Toronto, Canada, pp. 193, 1999.

[96] M. Goldstein, "Heterodyn–Interferometer zur Detektion elektrischer und thermischer Signale in integrierten Schaltungen durch die Substratrückseite," Dissertation, Technische Universität Wien, Austria, 1993.

[97] N. Seliger, D. Pogany, C. Fürböck, P. P. Habaš, E. Gornik, and M. Stoisiek, "A laser beam method for evaluation of thermal time constant in smart power devices," Microelectron. Reliab., vol. 37, pp. 1727, 1997.

[98] D. Pogany, N. Seliger, E. Gornik, M. Stoisiek, and T. Lalinský, "Analysis of the temperature evolution from the time–resolved thermo–optical interferometric measurements with few Fabry–Perot peaks," J. Appl. Phys., vol. 84, pp. 4495, 1998.

[99] N. Seliger, D. Pogany, C. Fürböck, P. Habaš, E. Gornik, and M. Stoisiek, “A study of temperature distribution in SOI–smart power devices in transient conditions by optical interferometry,” Proc. of ESSDERC, Stuttgart, Germany, pp. 512, 1997.

[100] R. Thalhammer, F. Hille, and G. Wachutka, “Numerical simulation of infrared laser probing techniques,” Proc. of SISPAD, Leuven, Belgium, pp. 276, 1998.

[101] R. Escoffier, U. Krumbein, E. Lyumkis, B. Polsky, A. Schenk, B. Schmithüsen, C. Steiner, and W. Fichtner, “DESSIS 4.0 manual,” ISE Integrated Systems Engineering AG, Zurich, Switzerland, 1996.

[102] T. B. Koch, “Computation of wave propagation in integrated optical devices,” Dissertation, University College, London, GB, 1989.

[103] C. Hafner, “Beiträge zur Berechnung der Ausbreitung elektromagnetischer Wellen in zylindrischen Strukturen mit Hilfe des ’Point Matching Verfahrens’,” Dissertation 6683, Eidgenössische Technische Hochschule Zürich, Switzerland, 1980.

[104] C. Hafner, “The generalized multipole technique for computational electromagnetics,” Artech House Books, 1990.

[105] T. O. Körner, “Rigorous simulation of light propagation in semiconductor devices,” Dissertation, Eidgenössische Technische Hochschule Zürich; in: Series in Microelectronics, vol. 81, Hartung–Gorre, Konstanz, Germany, 1999.

[106] K. S. Yee, “Numerical simulation of initial boundary value problems involving Maxwell’s equations in isotropic media,” IEEE Trans. on Antennas and Propagation, vol. 14, pp. 302, 1966.

[107] J. W. Thomas, “Numerical partial differential equations: Finite difference methods,” vol. 22 of ”Texts in Applied Mathematics”, Springer, Germany, 1995.

[108] P. E. Lagasse and R. Baets, “Application of propagating beam methods to electromagnetic and acoustic wave propagation problems: a review,” Radio Science, vol. 22, pp. 1255, 1987.

[109] M. Niederhoff, “Feldberechnung in Hochleistungslaserdioden,” Dissertation, Technische Universität München, Germany, 1996.

[110] M. D. Feit and J. A. Fleck, “Light propagation in graded–index optical fibers,” Appl. Optics, vol. 17, pp. 3990, 1978.

[111] J. V. Roey, J. van der Donk, and P. E. Lagasse, “Beam–propagation method: analysis and assessment,” J. Opt. Soc. Am., vol. 71, pp. 803, 1981.

[112] R. Z. Yahel and I. Last, “Numerical simulation of laser beam propagation in three–dimensional random media: beam splitting and patch formation,” Waves in Random Media, vol. 2, pp. 81, 1992.

[113] C. B. Burckhardt, “Diffraction of a plane wave at a sinusoidally stratified dielectric grating,” J. Opt. Soc. Am., vol. 56, pp. 1502, 1966.

[114] C.-M. Yuan and A. J. Strojwas, “Modeling optical microscope images of integrated–circuit structures,” J. Opt. Soc. Am. A, vol. 8, pp. 778, 1991.

[115] R. Petit, “Electromagnetic theory of gratings,” Springer, Berlin, Germany, 1980.

[116] M. S. Yeung, “Modeling high numerical aperture optical litography,” Proc. of SPIE Optical/Laser Microlitography, vol. 922, pp. 149, 1988.

[117] H. Kirchauer and S. Selberherr, “Rigorous three–dimensional photoresist exposure and development simulation over nonplanar topography,” IEEE Trans. on Computer Aided Design, vol. 16, pp. 1431, 1997.

[118] T. G. Moore, J. G. Blaschak, A. Taflove, and G. A. Kriegsmann, “Theory and application of radiation boundary operators,” IEEE Trans. on Antennas and Propagation, vol. 36, pp. 1797, 1988.

[119] E. Hecht and A. Zajac, “Optics,” Addison–Wesley, USA, 1974.

[120] G. Mur, “Absorbing boundary conditions for the finite–difference approximation of the time–domain electromagnetic–field equations,” IEEE Trans. on Electromagnetic Compatibility, vol. 23, pp. 377, 1981.

[121] J. W. Goodman, “Introduction to fourier optics,” McGraw–Hill, New York, USA, 1968.

[122] R. Thalhammer, F. Hille, and G. Wachutka, “Design and interpretation of laser absorption measurements for power devices,” Proc. of ISDRS, Charlottesville, VA, USA, pp. 531, 1999.

[123] R. Thalhammer, G. Deboy, W. Keilitz, U. Müller, and G. Wachutka, “Electrothermal effects in semiconductor power devices analyzed by numerical simulation and internal laser deflection measurement,” Proc. of ISDRS, Charlottesville, VA, USA, pp. 51–54, 1995.

[124] U. Müller, “Charakterisierung von Leistungsdioden durch interne IR–Lasermeßverfahren,” Diplomarbeit, Technische Universität München, Germany, 1995.

[125] E. Knauf, “Vergleich verschiedener IGBT–Konzepte mit Hilfe der Internen Laserdeflexion,” Diplomarbeit, Technische Universität München, Germany, 1996.

[126] R. Thalhammer, G. Deboy, E. Knauf, E. Kühbandner, and G. Wachutka, “Calibration of electrothermal power device models using combined characterization techniques,” Proc. of ISPSD, Weimar, Germany, pp. 181, 1997.

[127] D. Pogany, N. Seliger, T. Lalinsky, J. Kuzmik, P. Habaš, P. Hrkut, and E. Gornik, “Study of thermal effects in GaAs micromachined power sensor microsystems by an optical interferometer technique,” Microelectronic Journal, vol. 29, pp. 191, 1998.

[128] K. Bløtekjær, "Transport equations for electrons in two–valley semiconductors," IEEE Trans. on Electron Devices, vol. 17, pp. 38, 1970.

[129] R. K. Cook and J. Frey, "An efficient technique for two dimensional simulation of velocity overshot effects in Si and GaAs devices," Compel, vol. 1, pp. 65, 1982.

[130] W. vanRoosbroeck, "Theory of flow of electrons and holes in germanium and other semiconductors," Bell System Tech. J., vol. 29, pp. 560, 1950.

[131] Technology Modeling Associates Inc., "MEDICI User's Manual," Palo Alto, CA, USA, 1992.

[132] Silvaco International, "ATLAS User's Manual," Santa Clara, CA, USA, 1997.

[133] H. B. Callen, "Thermodynamics and an introduction to thermostatistics," 2nd ed., John Wiley and Sons, New York, USA, 1985.

[134] L. Onsager, "Reciprocal relations in irreversible processes I.," Phys. Rev., vol. 37, pp. 405, 1931.

[135] L. Onsager, "Reciprocal relations in irreversible processes II.," Phys. Rev., vol. 38, pp. 2265, 1931.

[136] G. Wachutka, "Unified framework for thermal, electrical, magnetic, and optical semiconductor device modeling," Compel, vol. 10, pp. 311, 1991.

[137] G. Wachutka, "Consistent treatment of carrier emission and capture kinetics in electrothermal and energy transport models," Microelectronics J., vol. 26, pp. 307, 1995.

[138] M. Lades, "Modeling and simulation of wide bandgap semiconductor devices: 4H/6H-SiC," Dissertation, Munich University of Technology, Germany, 2000.

[139] T. T. Mnatsakanov, I. L. Rostovtsev, and N. I. Philatov, "Investigation of the effect of nonlinear physical phenomena on charge carrier transport in semiconductor devices," Solid State Electronics, vol. 30, pp. 579, 1987.

[140] T. T. Mnatsakanov, D. Schröder, and A. E. Schlögl, "Effect of high injection level phenomena on the feasibility of diffusive approximation in semiconductor device modeling," Solid State Electronics, vol. 42, pp. 153, 1998.

[141] A. E. Schlögl, T. T. Mnatsakanov, and D. Schröder, "Temperature dependent behavior of silicon power semiconductors — a new physical model validated by device–internal probing between 400 K and 100 K," Proc. of ISPSD, Kyoto, Japan, pp. 383, 1998.

[142] S. M. Sze, "Semiconductor devices, physics and technology," John Wiley and Sons, New York, USA, 1985.

[143] S. Selberherr, "Analysis and simulation of semiconductor devices," Springer Verlag Wien, Austria, 1984.

[144] K. Kells, "General electrothermal semiconductor device simulation," Dissertation, Eidgenössische Technische Hochschule Zürich, Switzerland; in: Series in Microelectronics, vol. 37, Hartung–Gorre, Konstanz, Germany, 1994.

[145] A. Schenk, "Advanced physical models for silicon device simulation," Springer, Wien, Austria, 1998.

[146] G. Wachutka, "Tailored modeling: a way to the 'virtual microtransducer fab'?," Sensors and Actuators A, vol. 46–47, pp. 603, 1995.

[147] J. W. Slotboom and H. C. deGraff, "Measurements of bandgap narrowing in silicon bipolar transistors," Solid State Electronics, vol. 19, pp. 857, 1976.

[148] J. W. Slotboom, "The pn–product in silicon," Solid State Electronics, vol. 20, pp. 279, 1977.

[149] J. W. Slotboom and H. C. deGraff, "Bandgap narrowing in silicon bipolar transistors," IEEE Trans. on Electron Devices, vol. 24, pp. 1123, 1977.

[150] J. del Alamo, S. Swirhun, and R. M. Swanson, "Measuring and modeling minority carrier transport in heavily doped silicon," Solid State Electronics, vol. 28, pp. 47, 1985.

[151] J. del Alamo, S. Swirhun, and R. M. Swanson, "Simultaneous measuring of hole lifetime, hole mobility and bandgap narrowing in heavily doped n–type silicon," IEDM Tech. Digest, pp. 290, 1985.

[152] H. S. Bennett and C. L. Wilson, "Statistical comparisons of data on band–gap narrowing in heavily doped silicon: Electrical and optical measurements," J. Appl. Phys., vol. 55, pp. 3582, 1984.

[153] W. Shockley and W. T. Read, "Statistics of the recombinations of holes and electrons," Phys. Rev., vol. 87, pp. 835, 1952.

[154] R. N. Hall, "Electron–hole recombination in germanium," Phys. Rev., vol. 87, pp. 387, 1952.

[155] D. K. Schroder, "Semiconductor material and device characterization," John Wiley and Sons, New York, USA, 1990.

[156] S. R. Dhariwal, L. S. Kothari, and S. C. Jain, "On the recombination of electrons and holes at traps with finite relaxation time," Solid State Electronics, vol. 24, pp. 749, 1981.

[157] M. S. Tyagi and R. van Overstraeten, "Minority carrier recombination in heavily–doped silicon," Solid State Electronics, vol. 36, pp. 577, 1983.

[158] K. P. Lisiak and A. G. Milnes, "Platinum as a lifetime–control deep impurity in silicon," J. Appl. Phys., vol. 46, pp. 5229, 1975.

[159] B. J. Baliga, "Recombination level selection criteria for lifetime reduction in integrated circuits," Solid State Electronics, vol. 21, pp. 1033, 1978.

[160] S. A. Azimov, L. I. Islamov, and N. A. Sultanov, "Investigation of the influence of heat treatment on the electrical properties of platinum–doped silicon," Sov. Phys. Semicond., vol. 8, pp. 758, 1974.

[161] S. D. Brotherton, P. Bradley, and J. Bicknell, "Electrical properties of platinum in silicon," J. Appl. Phys., vol. 50, pp. 3396, 1979.

[162] A. Schenk and U. Krumbein, "Coupled defect–level recombination: Theory and application to anomalous diode characteristics," J. Appl. Phys., vol. 78, pp. 3185, 1995.

[163] A. G. Chynoweth, "Ionization rates for electrons and holes in silicon," Phys. Rev., vol. 109, pp.1537, 1958.

[164] J. J. Moll and R. van Overstraten, "Charge multiplication in silicon pn junctions," Solid State Electronics, vol. 6, pp. 147, 1963.

[165] R. van Overstraten and H. D. Man, "Measurement of the ionization rates in diffused silicon pn junctions," Solid State Electronics, vol. 13, pp. 583, 1970.

[166] Y. Okuto and C. R. Crowell, "Threshold energy effects on avalanche breakdown voltage in semiconductor junctions," Solid State Electronics, vol. 18, pp. 161, 1975.

[167] T. Lackner, "Avalanche multiplication in semiconductors: A modification of Chynoweth's law," Solid State Electronics, vol. 34, pp. 33, 1991.

[168] D. M. Caughey and R. E. Thomas, "Carrier mobilities in silicon empirically related to doping and field," Proc. IEEE, pp. 2192, Dec. 1967.

[169] C. Canali, G. Majni, R. Minder, and G. Ottaviani, "Electron and hole drift velocity measurements in silicon and their empirical relation to electric field and temperature," IEEE Trans. on Electron Devices, vol. 22, pp. 1045, 1975.

[170] H. Brooks, "Scattering by ionized impurities in semiconductors," Phys. Rev., vol. 83, pp. 879, 1951.

[171] G. Masetti, M. Severi, and S. Solmi, "Modeling of carrier mobility against carrier concentration in arsenic–, phosphorus– and boron–doped silicon," IEEE Trans. on Electron Devices, vol. 30, pp. 764, 1983.

[172] C. Lombardi, S. Manzini, A. Saporito, and M. Vanzi, "A physically based mobility model for numerical simulation of nonplanar devices," IEEE Trans. on Computer Aided Design, vol. 7, pp. 1164, 1988.

[173] C. T. Sah, T. H. Ning, and L. K. Tschopp, "The scattering of electrons by surface oxide charges and by lattice vibrations at the silicon–silicon dioxide interface," Surface Sci., vol. 32, pp. 561, 1972.

[174] A. Hartstein, T. H. Ning, and A. B. Fowler, "Electron scattering in silicon inversion layers by oxide and surface roughness," Surface Sci., vol. 58, pp. 178, 1976.

[175] E. Conwell and V. F. Weisskopf, "Theory of impurity scattering in semiconductors," Phys. Rev., vol. 77, pp. 388, 1950.

[176] N. H. Fletcher, "The high current limit for semiconductor junction devices," Proc. of Institution of Radio Engineers, vol. 45, pp. 862, 1957.

[177] S. C. Choo, "Theory of forward–biased diffused–junction pln rectifier. part I: Exact numerical solutions," IEEE Trans. on Electron Devices, vol. 19, pp. 954, 1972.

[178] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, "Numerical recipes in C," Cambridge University Press, USA, 1988.

[179] R. E. Bank, D. J. Rose, and W. Fichtner, "Numerical methods for semiconductor device simulation," IEEE Trans. on Electron Devices, vol. 30, pp. 1031, 1983.

[180] R. E. Bank, W. M. Coughran, W. Fichtner, E. H. Grosse, D. J. Rose, and R. K. Smith, "Transient simulation of silicon devices and circuits," IEEE Trans. on Computer Aided Design, vol. 4, pp. 436, 1985.

[181] H. K. Gummel, "A self–consistent iterative scheme for one–dimensional steady state transistor calculations," IEEE Trans. on Electron Devices, vol. 11, pp. 455, 1964.

[182] A. Schneider, "Untersuchung von Leistungsdioden durch Messung der Rekombinationsstrahlung und numerische Simulation," Techn. Report 1998/1, Inst. für Technische Elektrophysik, Technische Universität München, Germany, 1998.

[183] L. Colalongo, M. Valdinoci, and M. Rudan, "An efficient numerical method to solve the time–dependent semiconductor equations including trapped charge," Proc. of SISPAD, Erlangen, Germany, pp. 82, 1995.

[184] R. Siemieniec, D. Schipanski, W. Südkamp, and J. Lutz, "Simulation und experimental results of irradiated power diodes," Proc. of EPE, Lausanne, Switzerland, 1999.

[185] W. Südkamp, R. Siemieniec, and J. Lutz, "Untersuchung der Temperaturabhängigkeit der Trägerlebensdauer in elektronenbestrahlten Dioden mit der DLTS- und Speicherzeitmessung," 28. Kolloquium 'Halbleiter-Leistungsbauelemente und Materialgüte von Silizium', Freiburg, Germany, pp. 20–1, 1999.

[186] S. P. Gaur and D. H. Navon, "Two–dimensional carrier flow in a transistor structure under nonisothermal conditions," IEEE Trans. on Electron Devices, vol. 23, pp. 50, 1976.

[187] D. K. Sharma and K. V. Ramanathan, "Modeling thermal effects on MOS I–V–characteristics," IEEE Electron Device Letters, vol. 4, pp. 362, 1983.

[188] M. S. Adler, “Accurate calculations of the forward drop and power dissipation in thyristors,” IEEE Trans. on Electron Devices, vol. 25, pp. 16, 1978.

[189] A. Chryssafis and W. Love, “A computer–aided analysis of one–dimensional thermal transients in npn power transistors,” Solid State Electronics, vol. 22, pp. 249, 1979.

[190] J. Vobecky and P. Hazdra, “Future trends in local lifetime control,” Proc. of ISPSD, Hawaii, USA, pp. 181, 1996.

[191] J. Vobecky, P. Hazdra, and J. Homola, “Optimization of power diode characteristics by means of ion irradiation,” IEEE Trans. on Electron Devices, vol. 43, pp. 2283, 1996.

[192] W. Feiler, W. Gerlach, and U. Wiese, “Two–dimensional analytic models of the carrier distribution in the on–state of the IGBT,” Solid State Electronics, vol. 38, pp. 1781, 1995.

[193] E. Kühbandner, “Elektrothermische Simulation von Insulated Gate Bipolar Transistoren,” Diplomarbeit, Technische Universität München, Germany, 1996.

[194] T. Yamanaka, S. J. Fang, H.-C. Lin, J. P. Snyder, and C. R. Helms, “Correlation between inversion layer mobility and surface roughness measured by AFM,” IEEE Electron Device Letters, vol. 17, pp. 178, 1996.

[195] G. Reichert and T. Ouisse, “Relationship between empirical and theoretical mobility models in silicon inversion layers,” IEEE Trans. on Electron Devices, vol. 43, pp. 1394, 1996.

[196] D. Vasileska and D. K. Ferry, “Scaled silicon MOSFETs: Universal mobility behavior,” IEEE Trans. on Electron Devices, vol. 44, pp. 577, 1997.

[197] A. Pirovano, A. L. Lacaita, G. Zandler, and R. Oberhuber, “Explaining the dependence of electron and hole mobilities in Si MOSFET’s inversion layer,” IEDM Tech. Digest, pp. 527, 1999.

[198] G. Deboy, private communication.

[199] I. Son, T.-W. Tang, and D. H. Navon, “Modeling of bistable device I–V characteristic resulting from conductivity modulation in semiconductors,” IEEE Trans. on Electron Devices, vol. 35, pp. 450, 1988.

[200] M. R. Simpson, “Analysis of negative differential resistance in the I–V characteristic of shorted–anode LIGBTs,” IEEE Trans. on Electron Devices, vol. 38, pp. 1633, 1991.

[201] R. Thalhammer, F. Hille, P. Scheubert, and G. Wachutka, “Physically rigorous modeling of sensing techniques exploiting the thermo–optical and electro–optical effect,” Proc. of MSM, San Juan, Puerto Rico, USA, pp. 683, 1999.

[202] G. Cocorullo, F. G. D. Corte, M. Iodice, I. Rendina, and P. M. Sarro, “A temperature all–silicon micro–sensor based on the thermo–optic effect,” IEEE Trans. on Electron Devices, vol. 44, pp. 766, 1997.

[203] G. Cocorullo, F. G. D. Corte, I. Rendina, and P. M. Sarro, "Thermo–optic effect exploitation in silicon microstructures," Sensors and Actuators A, vol. 71, pp. 19, 1998.

[204] T. Liu, G. F. Fernando, Z. Y. Zhang, and K. T. V. Grattan, "Simultaneous strain and temperature measurements in composites using extrinsic Fabry–Perot interferometric and intrinsic rare–earth doped fiber sensors," Sensors and Actuators A, vol. 80, pp. 208, 2000.

[205] D. Marcuse, "Light transmission optics," Van Nostrand, New York, USA, 1982.

[206] M. Litzenberger, "Thermal characterization of smart power electrostatic discharge protection devices by backside laserprobing," Diploma Thesis, Institute for Solid State Electronics, Vienna University of Technology, Austria, Feb. 1999.

Acknowledgement

First of all, I wish to express my gratitude to Prof. Dr. Gerhard Wachutka for enabling me to carry out this challenging research work. I remember many educational discussions and appreciate his continuous interest and support. Prof. Wachutka has created an infrastructure at the Institute for Physics of Electrotechnology that facilitates research of international reputation and close collaboration with industrial partners.

I am also very grateful to my co–examiner Prof. Dr. Frederick Koch, whose interest and constructive comments contributed encouraging ideas to my studies and the associated diploma theses.

I like to thank my colleague Dipl.–Phys. Frank Hille for his commitment in improving the experimental facilities and performing most of the laser absorption and deflection measurements. I will miss our intense discussions and wish him all the best for his own Ph.D. thesis. I also want to acknowledge the contributions of my former diploma students, in particular Dipl.–Phys. Frank Hille, Dipl.–Phys. Elmar Knauf, Dipl.–Phys. Erich Kühbandner, and Dipl.–Phys. Andreas Schneider.

I am indebted to our industrial partners at the Siemens labs, in particular Prof. Dr. Eckehard Wolfgang, Dr. Gerald Sölkner, Dr. Ludwig Hoffmann, Dr. Gerald Deboy, Dr. Andreas Schlögl, and Prof. Dr. Michael Stoisiek. Dr. Deboy introduced me to the field of power electronics and optical measurement techniques during the beginning of my Ph.D. studies. Dr. Sölkner's continuous interest and many fruitful discussions about optical problems and laser probing techniques have made a valuable contribution to the success of this thesis.

I would also like to express my appreciation to the "laserprober group" at the Vienna University of Technology, namely Prof. Dr. Erich Gornik, Dr. Dionyz Pogany, Dr. Norbert Seliger, and Dipl.–Ing. Christoph Fürböck. Their dedicated participation in our common projects, the elaboration of reliable and precise measurement results from Backside Laser Probing, and their immediate response to any question have made our collaboration a great pleasure and the source of many interesting results and publications.

Special thanks to Prof. Dr. Roland Schiek for many stimulating discussions about optics and modeling of beam propagation and to Dipl.–Ing. Peter Scheubert for numerous useful information about numerical algorithms.

I would like to thank Dr. Gerald Sölkner, Dr. Markus König, Dr. Norbert Seliger, Dipl.–Phys. Gabriele Schrag, Dipl.–Math. Eva–Renate König, Dipl.–Phys. Frank Hille, Dipl.–Phys. Siegfried Ramminger, Dipl.–Phys. Winfried Kaindl, Dipl.–Phys. Robert Sattler, and M. A. Yvonne Müller for graciously proofreading parts of the manuscript. I am especially grateful to Stefan Schaub for his continuous helpfulness and his technical assistance, in particular when I was unable to circumvent the agonizing struggles with Microsoft's software products. Last, but not least, I particularly wish to thank my colleagues at the Institute for Physics of Electrotechnology for the excellent and friendly atmosphere which has significantly contributed to the pleasant memories of my years at the institute.