

Investigation of Carrier Dynamics in Self-Organized Quantum Dots for Memory Devices

vorgelegt von

Diplom-Physiker

Martin Paul Geller

aus Wilhelmshaven

von der Fakultät II – Mathematik und Naturwissenschaften
der Technischen Universität Berlin
zur Erlangung des akademischen Grades

Doktor der Naturwissenschaften

– Dr. rer. nat. –

genehmigte Dissertation

Promotionsausschuss:

Vorsitzender: Prof. Dr. T. Möller

Berichter: Prof. Dr. D. Bimberg

Berichter: Prof. Dr. M. Grundmann

Tag der wissenschaftlichen Aussprache: 12. April 2007

Berlin 2007

D83

Abstract

In this work, the charge carrier dynamics of self-organized semiconductor quantum dots is investigated by using time-resolved capacitance spectroscopy (deep level transient spectroscopy – DLTS). Two main issues are studied in detail, that are of importance for using quantum dots in future memory devices: Carrier storage and capture. A storage time of milliseconds at room temperature is a basic prerequisite in order to compete with today’s DRAM (dynamic random access memory). The carrier capture time determines significantly the possible write time. The capacitance spectroscopy is refined within this thesis in order to study emission and capture dynamics of charge carriers for different quantum dot states, as well as storage time and localization energy of the ground state.

Time-resolved capacitance measurements of tunneling emission is used for the first time to determine the localization energy of the electron and hole ground state in InGaAs/GaAs quantum dots to 290 meV and 210 meV, respectively. The average storage time for electrons and holes in this quantum dot system is estimated to be about 200 ns and 0.5 ns at room temperature. A longer storage time of about 1 μ s is obtained for holes in the ground state of GaSb/GaAs quantum dots – for a localization energy of 450 meV. The use of InAs/GaAs quantum dots in combination with a Al_{0.6}Ga_{0.4}As barrier yields a further increase in storage time for an enhanced localization energy of 560 meV. The important milestone of a storage time of 5 ms at room temperature is demonstrated, in the order of the typical DRAM refresh time. The connection between localization energy and storage time enables to estimate retention times for different material combinations of quantum dots and surrounding matrix. A retention time of more than 10 years is predicted for (In)(Ga)Sb quantum dots in an AlAs matrix, hence, the feasibility of a quantum-dot-based non-volatile Flash memory is shown.

Another newly developed method, the charge-selective DLTS, is used to investigate the carrier capture into InGaAs/GaAs quantum dots. In capacitance transients the hole capture is observed and activation energies and capture cross-sections for emission *and* capture are determined. The experimental findings lead to the assumption, that carrier capture into quantum dots in an applied electric field is also controlled by phonon-assisted tunneling, as observed previously for the emission. A large capture cross-section of about 7×10^{-12} cm² is obtained and a capture time of ~ 0.3 ps at room temperature is estimated, which is more than four orders of magnitude below the write time of a DRAM cell.

From the experimental results, the realization of a future quantum-dot-based memory, that combines the advantages of today’s semiconductor memories (DRAM and Flash), seems to be feasible: A long storage time of more than 10 years (Flash), high endurance of more than 10^{15} write/erase cycles (DRAM) in combination with a fast read/write and erase time below 20 ns (DRAM). A concept of a quantum-dot-based memory is developed and proposed within this thesis.

Zusammenfassung

Im Rahmen dieser Arbeit wird die Ladungsträgerdynamik in selbstorganisierten Quantenpunkten mittels zeitaufgelöster Kapazitätsspektroskopie (deep level transient spectroscopy – DLTS) untersucht. Zwei Kernpunkte werden genauer betrachtet, die von entscheidender Bedeutung für die Verwendung von Quantenpunkten in zukünftigen Halbleiter-Speichern sind: Die Speicherung und der Einfang von Ladungsträgern. Eine Speicherzeit von Millisekunden bei Raumtemperatur ist die Grundvoraussetzung, um mit der heutigen DRAM-Zelle (Dynamic Random Access Memory) konkurrieren zu können. Die Geschwindigkeit des Ladungsträgereinfangs bestimmt maßgeblich die maximal mögliche Schreibgeschwindigkeit. Die Kapazitätsspektroskopie wird innerhalb dieser Arbeit dahingehend weiterentwickelt, dass Untersuchungen hinsichtlich Einfangs- und Emissionsdynamik in unterschiedliche Quantenpunktzustände sowie die Bestimmung der Speicherzeit und Lokalisierungsenergie des Grundzustandes ermöglicht werden.

Bei der Untersuchung von InGaAs/GaAs-Quantenpunkten wird eine neu entwickelte Methode – die zeitaufgelöste Kapazitätsmessung der Tunnelemission – verwendet, um erstmals eine Lokalisierungsenergie von 290 meV/210 meV für Elektronen und Löcher experimentell bestimmen zu können. Daraus lässt sich eine mittlere Speicherzeit von 200 ns für Elektronen bzw. 0.5 ns für Löcher bei Raumtemperatur abschätzen. Eine längere Speicherzeit von ungefähr 1 μ s wird für Löcher im Grundzustand von GaSb/GaAs-Quantenpunkten beobachtet – für eine Lokalisierungsenergie von 450 meV. Eine weitere Verlängerung der Speicherzeit ergibt sich durch die Verwendung einer Al_{0,6}Ga_{0,4}As-Barriere in Kombination mit InAs/GaAs-Quantenpunkten, bei erneuter Zunahme der Lokalisierungsenergie auf 560 meV. Die Arbeit demonstriert erstmals eine Speicherzeit von 5 ms bei Raumtemperatur, welche in der Größenordnung der DRAM-Refresh-Zeit liegt. Die Abhängigkeit der Speicherzeit von der Lokalisierungsenergie erlaubt eine Abschätzung der mittleren Speicherzeit für verschiedene Materialkombinationen von Quantenpunkten und umgebender Matrix. Eine Speicherzeit von über 10 Jahren wird für (In)(Ga)Sb-Quantenpunkte in einer AlAs Matrix vorhergesagt, wodurch ein quantenpunktbasierter nicht-flüchtiger Flash-Speicher realisierbar erscheint.

Mittels der ebenfalls neu entwickelten Methode der ladungsselektiven DLTS wird der Einfangprozess von Löchern in InAs/GaAs-Quantenpunkte untersucht. In Kapazitätstransienten wird der Locheinfang beobachtet und es können Aktivierungsenergien und Einfangquerschnitte auch aus den Einfangtransienten bestimmt werden. Die Ergebnisse lassen vermuten, dass der Einfang von Ladungsträgern in Quantenpunkte in einem elektrischen Feld ebenfalls durch phononen-assistiertes Tunneln geschieht, wie bereits in früheren Arbeiten für die Emission gezeigt. Es ergibt sich ein Einfangquerschnitt von $\sim 7 \times 10^{-12}$ cm². Daraus wird eine mittlere Einfangzeit der Löcher von ~ 0.3 ps bei Raumtemperatur abgeschätzt, die mehr als vier Größenordnungen unterhalb der Schreibzeit heutiger DRAM-Speicher liegt.

Auf Basis der Ergebnisse dieser Arbeit erscheint die Realisierung eines zukünftigen quantenpunktbasierten Flash-Speichers möglich, der die Vorteile der beiden heute gängigen Halbleiterspeicher (DRAM und Flash) vereinigt: Lange Speicherzeit von mehr als 10 Jahren (Flash), hohe Haltbarkeit von mehr als 10^{15} Schreib-/Löschzyklen (DRAM) in Kombination mit schnellem Schreib-/Lese- und Löschzugriff unterhalb von 20 ns (DRAM). Ein Konzept eines solchen quantenpunktbasierten Speichers wurde entwickelt und wird innerhalb dieser Arbeit vorgestellt.

Contents

1	Introduction	1
2	Semiconductor structures and devices	5
2.1	Heterostructures	5
2.2	Quantum Dots	7
2.2.1	Fabrication Techniques	7
2.2.2	Electronic Properties	9
2.2.3	Charge Carrier Emission	12
2.3	Memories	16
2.3.1	Dynamic Random Access Memory	16
2.3.2	Non-volatile Flash Memory	18
2.3.3	A QD-based Memory Cell	20
3	Carrier Dynamics	25
3.1	Deep levels in Semiconductors	25
3.1.1	Carrier Capture and Emission	26
3.1.2	Detailed Balance Condition in Thermal Equilibrium	27
3.1.3	Temperature Dependence of the Emission Rate	29
3.1.4	Temperature Dependence of the Capture Cross-Section	30
3.1.5	Majority and Minority Carrier Traps	31
3.1.6	General Solution of the Rate Equation	31
3.2	Quantum Dots in Semiconductors	34
3.2.1	Thermal emission rate	34
3.2.2	Electric Field Enhanced Emission	37
3.2.3	Influence of a Barrier Below the Quantum Dots	40
3.2.4	Charge Carrier Capture	42
4	Capacitance Spectroscopy	43
4.1	Depletion Region	43
4.1.1	Schottky Contact	43
4.1.2	p-n Junction	45
4.1.3	Width of the Depletion Region	46
4.1.4	Depletion Layer Capacitance	48

4.2	Static Capacitance-Voltage Profiling	48
4.3	Capacitance Transient Spectroscopy	49
4.3.1	Measurement Principle	49
4.3.2	Charge-Selective Deep Level Transient Spectroscopy	52
4.3.3	Time-Resolved Tunneling Capacitance Measurement	54
5	Storage and Emission of Charge Carriers	57
5.1	Thermal Emission from InGaAs/GaAs QDs	57
5.1.1	Sample Structures	58
5.1.2	Capacitance-Voltage Measurements	59
5.1.3	DLTS Measurements	61
5.1.4	Carrier Storage Times in an Electric Field	65
5.2	Tunneling Emission from InGaAs Quantum Dots	66
5.2.1	Check for Temperature Dependence	66
5.2.2	Hole Tunneling	68
5.2.3	Electron Tunneling	70
5.2.4	Localization Energies	70
5.2.5	Discussion: Electronic Properties and Storage Times in InAs QDs	72
5.3	Hole Storage in GaSb/GaAs Quantum Dots	74
5.3.1	Sample Structures	74
5.3.2	Capacitance-Voltage Measurements	77
5.3.3	DLTS Measurements	78
5.3.4	Summary and Discussion	81
5.4	InAs/GaAs Quantum Dots with AlGaAs Barrier	83
5.4.1	Sample Structures	83
5.4.2	Capacitance-Voltage Measurements	84
5.4.3	DLTS Measurements: Reference Sample A	85
5.4.4	DLTS Measurements: AlGaAs Barrier Sample B	90
5.4.5	Summary	96
5.5	Storage Time in Quantum Dots	96
6	Carrier Capture into Quantum Dots	99
6.1	DLTS Measurements on Hole Capture	99
6.1.1	Emission and Capture Transients	101
6.2	Activation Energies and Capture Cross-Sections	102
6.2.1	Phonon-Assisted Tunneling During Hole Capture	104
6.3	Hole Capture Without a Barrier	105
6.4	Discussion and Summary	107
6.4.1	Discussion: Carrier Capture and Relaxation into QD States . . .	107
6.4.2	Summary	109
7	Conclusion	111
8	Outlook	115

A Experimental Details	117
A.1 DLTS Setup	117
A.2 Sample Processing	119
A.2.1 p-n and n-p Diodes	119
A.2.2 Ohmic Contacts	120
A.2.3 Sulfur Passivation	121
Publications	135
Danksagung	139

“COMPUTERS IN THE FUTURE MAY WEIGH NO MORE THAN 1.5 TONS.”

POPULAR MECHANICS, AMERICAN MAGAZINE (1949)

Chapter 1

Introduction

Storage and handling of information is one of the fundamental issues of our modern information society. The need for faster processing of an increasing amount of data is the driving force in information technology over the last decades. The main strategy of the semiconductor industry to improve the performance of microelectronic devices is scaling down the feature size, that means integration of more and more components per unit area. This strategy allowed to increase the computation speed and memory capacity in logic and memory devices, while simultaneously decreasing the power consumption and costs. The ongoing success is also a result of the enormous progress in semiconductor physics.

The exponentially progressing reduction in feature size was predicted by Gordan Moore [1], known nowadays as “Moore’s Law”: “The number of components per unit area on a chip doubles about every 12 months.” In 1965, the early days of integrated circuits, he made this prediction for the next 10 years. At this time, the Moore group was able to integrate about 60 components on a chip. Using semilogarithmic paper and starting from the first transistors, he extrapolated to 65,000 components on a chip in 1975. After 10 years had passed, he was amazed how precisely his prediction had been met. With some modifications, the development continued since then, going on beyond every physical limit earlier predicted [2, 3].

The ongoing demand for further miniaturization has lead to a typical feature size in 2005 of about 90 nm. The “International Roadmap for Semiconductors” [4] (which tries to identify the technological challenges and needs the semiconductor industry might encounter) predicts a further shrinkage down to 14 nm in 2020. This rapid reduction in feature size requires a huge amount of research and the development of new technologies, as the industry is touching a regime where quantum mechanics dominates the physical properties. For instance, for a DRAM cell (dynamic random access memory), it has become increasingly difficult to maintain such a decrease in the feature size, because the signal becomes less immune to leakage current and internal noise, due to tunnelling through ultra-thin barriers in transistors and capacitors.

The fabrication of smaller and smaller structures with optical lithography techniques will be technologically extremely challenging. This guides us to a second relation, some-

times referred to as “Moore’s second law”: “The cost of building fabrication facilities to manufacture chips has also been increased exponentially, by about a factor of two every chip generation.” In 1970, to build a single fabrication facility (or “fab”) it took about \$5 million. In the year 2005, a fab already costed approximately \$2 billion. By the year 2010, the costs may reach more than \$15 billion, if Moore’s second law continues to hold.

The origin of this exponential growth lies in the nature of the technological production process, the so-called “top-down approach”. It starts at the wafer level and patterning proceeds by lithography and etching to obtain the desired features on a micrometer or nanometer scale. The extension to smaller feature sizes is predominantly facilitated by reducing the wavelength of photo-lithography, using extreme UV lithography or x-ray lithography. As the required precision in machinery and resolution of the fabrication tools is increasing, the effort and the installation costs grow dramatically.

Alternatively, the fabrication of nanosized structures can be started from individual atoms or molecules which are ordered physically or react chemically to obtain the desired features without lithography. This is called the “bottom-up approach”. Bottom-up techniques are typically based on self-organization processes, which are coordinated processes of individual atoms or molecules to produce larger and ordered structures. In a paper by Stranski and Krastanow [5], already published in 1938, the possibility of self-organized island formation was described. Starting with a flat surface and depositing a layer of material with a different lattice constant leads to a strained wetting layer on the lattice surface. Under certain circumstances the strain energy can be reduced by forming small islands from the material of the wetting layer. These tiny structures with base length in the order of tens of nanometers can be very regular in size and shape and can confine charge carriers in all three dimensions. Due to this zero-dimensional confinement for charge carriers in a semiconductor material, they act as “artificial atoms” and are referred to as self-organized or self-assembled quantum dots (QDs) [6]. The first realizations by modern epitaxial growth techniques were reported just about 15 years ago [7, 8].

Concerning devices based on QDs: Lasers and optical amplifier are the most obvious and widely investigated applications [9–14]. They will certainly be the first ones to enter the consumer market, but they may not be the only ones. In the long term it can be expected that this new class of self-organized semiconductor structures will be exploited in a wide range of sectors in electronics and optoelectronics, like already first single-photon devices are based on QDs [15, 16]. One of the exciting applications could be the use of self-organized materials in nanoelectronic logic and memory devices.

Today’s semiconductor memory industry focuses mostly on two concepts: The DRAM and the Flash [17]. Both memory concepts have their advantages and disadvantages in speed, endurance, cost and storage time. In a DRAM [18], the information is stored by means of a transistor and a capacitor. The advantages of a DRAM are fast access times below 20 nanoseconds and a high endurance in the order of 10^{15} write/erase cycles. A disadvantage is the number of involved electrons per cell in the order of 10^5 , resulting in a high power consumption. Furthermore, DRAMs are volatile memories,

i. e. the stored information has to be refreshed every tens of milliseconds, otherwise it is lost.

A Flash memory is a non-volatile storage device, this means, the average storage time is in the order of >10 years, and the electron number involved to store one bit of information is up to two orders of magnitude smaller (in the range of 10^3). DRAMs draw power continuously, even when data is not being read or written, and lose all information when the power is switched off. In contrast, Flash memories consume power only when a read or write process occurs. Hence, they are the ideal memory concept for mobile electronic devices, like digital cameras, mobile phones, personal digital assistants (PDAs), memory sticks, and MP3-players. However, Flash memories are not the ultimate memory concept yet. Flash needs at least a microsecond for the write process [17] and has a very poor endurance in the order of 10^6 write/erase cycles.

The combination of the advantages of a DRAM and a Flash memory cell is one of the greatest challenge for the semiconductor memory industry. Such an ultimate (or universal) memory should be non-volatile with a storage time in the order of years and exhibit a fast read/write and erase time of a few nanoseconds (or below) with an endurance better than 10^{15} read/erase cycles. In addition, it should be a single-electron memory cell, where one charge carrier represents one bit of information.

QDs have ideal properties to act as building-blocks in future memories. Billions of QDs are formed simultaneously in a single technological step, allowing a massive parallel production from a bottom-up approach and offering an elegant method to create huge ensembles of electronic traps. They can store just a few or even single charge carriers with a retention time depending on the material combination – possibly up to years at room temperature. Furthermore, the carrier relaxation process into QDs is in the order of sub-picoseconds [19, 20], an important prerequisite for a very fast write access time in such memories. The implementation of QDs could lead to a non-volatile memory with high storage densities combined with a fast read/write access time.

Content of this Work

In this work, the carrier dynamics of QDs are studied by time-resolved capacitance spectroscopy (deep level transient spectroscopy – DLTS). Two substantial issues, which are of importance for using QDs in future memory devices, are highlighted: The storage and capture of charge carriers. DLTS is used and refined within this thesis in order to probe emission and capture of charge carriers in different QD states, as well as storage time and localization energy of the ground state.

In Chapter 2, the basics concerning semiconductor structures and memory devices are briefly recapitulated. After an introduction into the electronic properties and carrier dynamics of QDs, the chapter focuses on a description of the two main semiconductor memories. DRAM and Flash are compared in order to discover the needs for an ultimate memory cell. At the end of this chapter, a concept of a QD-based memory is proposed, which was developed within this thesis. In the following Chapter 3, the carrier dynamics of deep levels and QDs are presented in more detail with deriva-

tions of rate equations for emission and capture. These equations are of importance for the evaluation of the capacitance measurements concerning activation energies and capture cross-sections. The depletion-layer capacitance spectroscopy is summarized in Chapter 4. After briefly revising the fundamentals of static capacitance-voltage measurements, the time-resolved capacitance transients experiments are described. In addition, two newly developed capacitance measurement techniques are presented, the charge-selective DLTS and the time-resolved capacitance tunneling measurement. In the following two chapters, the experimental results are presented. Chapter 5 deals with the storage and emission of electrons and holes in different QD systems. In Chapter 6, the charge-selective DLTS method is used to study the carrier capture into QDs. Finally, the main results are recapitulated in the conclusion (Chapter 7) and the next steps and perspectives in the direction of a possible QD memory device are discussed in the outlook (Chapter 8).

“THE IMPORTANT THING IN SCIENCE IS NOT SO MUCH
TO OBTAIN NEW FACTS AS TO DISCOVER NEW WAYS OF
THINKING ABOUT THEM.”
SIR WILLIAM HENRY BRAGG

Chapter 2

Semiconductor structures and devices

In this chapter, the basics of semiconductor heterostructures, QDs and memories are briefly summarized. To begin with, a general overview of semiconductor heterostructures is given in Section 2.1. This section focuses on the III-V compounds, like GaAs, AlAs and GaSb, their bandgap and band alignment. Among others, bandgap and alignment determine the localization energy and storage time of charge carriers in the investigated III-V self-organized QDs. Fabrication techniques, electronic properties and emission processes of these QDs are described in more detail in Section 2.2. In Section 2.3, the most common semiconductor memories, DRAM and Flash, are described and compared in order to discover the needs for a future QD-based memory. Finally, a concept of such memory is proposed in Section 2.3.3.

2.1 Heterostructures

Metal Organic Chemical Vapor Deposition (MOCVD) or Molecular Beam Epitaxy (MBE) are epitaxial growth techniques to deposit defect-free different semiconductor materials with different electronic properties (e. g. band offset) on each other. Such a semiconductor crystal (made of more than one material) is referred to as “heterostructure”, the interface between the two materials as “heterojunction”. Amongst others, the electronic properties of the entire heterostructure are determined by this heterojunction [21]. Therefore, heterostructures offer more possibilities for designing devices with certain desired optical or electronic properties.

The use of heterojunctions in the design and fabrication of semiconductor devices has led to the development of new structures with improved performance. For instance, one of the most exploited system is the double heterostructure (DHS) laser, invented in 1963 by Alferov and Kroemer [22, 23] and honored by the Nobel price in physics in the year 2000. Alferov built the first DHS semiconductor laser using gallium arsenide (GaAs) and aluminum arsenide (AlAs) in 1969. Especially these III-V semiconductor alloys play until now an important role in all kinds of optoelectronic devices, like

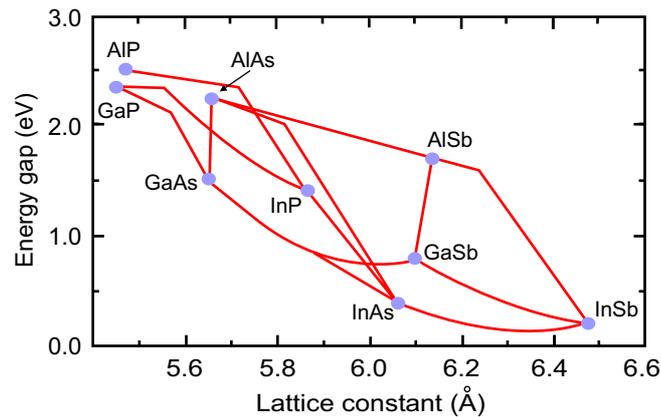


Figure 2.1: Energy bandgap as a function of lattice constant for different III-V compound semiconductors (blue points) and their ternary alloys (red lines) at zero temperature. After [28].

lasers [24, 25] and amplifiers [26, 27].

By using epitaxial growth techniques, it is also easily possible to create ternary alloys, for instance $\text{Al}_x\text{Ga}_{1-x}\text{As}$, which exhibit a bandgap accordingly between the both binary alloys. Figure 2.1 shows the energy bandgap as a function of lattice constant for different III-V compound semiconductors (blue points) and their ternary alloys (red lines) at zero temperature. The III-V compound semiconductors and their ternary alloys offer a large variety of material combinations in heterostructure devices with the possibility of a controlled band-structure engineering.

For a semiconductor heterostructure, besides the energy bandgap, the relative energetic position of the valence and conduction band at the interface is of central importance: the so-called band alignment. The band alignment and the bandgap control the band offsets in the conduction and valence band. The band offset is simply defined as the discontinuity in the band edges at the interface between the two semiconductors. Figure 2.2 shows the various types of band alignments that can arise in semiconductor heterojunctions [29]: Type I, type II staggered, and type II broken-gap. In a type I alignment, the bandgap of the second semiconductor E_g^2 lies completely within the bandgap of the first one E_g^1 , shown in Fig. 2.2(a). This band alignment occurs in a large number of heterojunctions, e. g. InAs/GaAs, GaAs/AlGaAs and GaSb/AlSb. A type II staggered band alignment is given for the situation, that the bandgap of the second semiconductor does not completely lies inside the first one, however, the bandgaps of the two material still overlap, Fig. 2.2(b). A type II staggered band alignment is characteristic for the GaSb/GaAs system, which plays an important role in this work and which will be discussed in Section 5.3 in more detail. A type II broken-gap alignment occurs, when the bandgaps of the two materials do not overlap at all in energy.

Of particular interest for this work is the already mentioned $\text{Al}_x\text{Ga}_{1-x}\text{As}$ ternary semiconductor alloy [28–30]. It is used as a energetic barrier to increase the carrier storage time at room temperature in InAs QDs in a GaAs matrix (see Section 5.4). The GaAs/ $\text{Al}_x\text{Ga}_{1-x}\text{As}$ interface is a type I heterojunction. AlGaAs in GaAs leads

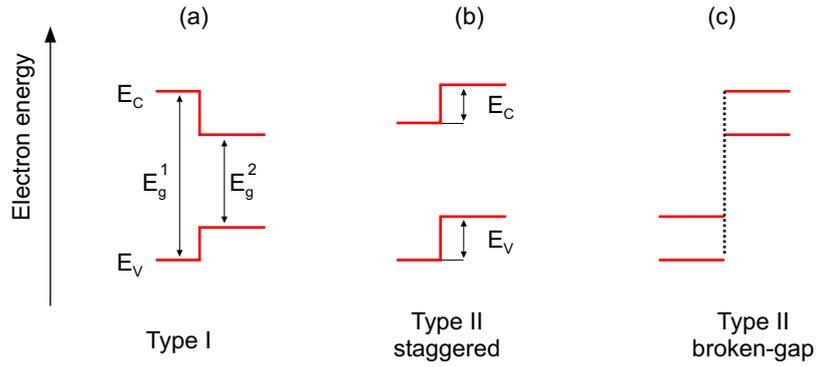


Figure 2.2: Types of band alignment at a heterojunction consisting of two semiconductors with different bandgap E_g^1 and E_g^2 . The conduction and valence band edges have been labeled with E_C and E_V , while the conduction and valence band offsets have the label ΔE_C and ΔE_V , respectively.

to a potential barrier for electrons and holes in the conduction and valence band, respectively. A variation of the composition x in the $\text{Al}_x\text{Ga}_{1-x}\text{As}$ material adjusts the electronic band structure and the conduction and valence band offsets between the two binary components GaAs and AlAs. Figure 2.3 displays the conduction ΔE_C and valence band offset ΔE_V of the $\text{Al}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ interface as function of the aluminum composition. The AlGaAs alloy has a direct bandgap up to $x = 0.45$ (like GaAs), then becomes an indirect bandgap semiconductor (like AlAs). The dependence of the conduction band offset is shown for the X -point (ΔE_C^X) and the Γ -point (ΔE_C^Γ). The maximum band offset at the important Γ -point is 1.06 eV/550 meV in the conduction/valence band.

2.2 Quantum Dots

Self-organized quantum dots (QDs) are low-dimensional heterostructures [6], that confine electrons and/or holes in all three spatial directions. Such a zero-dimensional system behaves in the electronic and optical properties almost like an “artificial atom” and offers lots of possibilities for new applications, e. g. in electronic devices such as field effect transistors [31–33] or memory cells [32, 34–36] as well as in optoelectronic devices such as lasers [9, 10, 13], detectors [37, 38], amplifiers [39], high frequency optical devices [14, 40, 41], and single-photon sources [15, 16, 42].

2.2.1 Fabrication Techniques

First efforts to produce a zero-dimensional confinement were devoted to lithography techniques in a “top-down” approach. Optical lithography and holography, X-ray lithography, electron beam lithography (EBL), and scanning tunneling microscopy (STM) were used (see Reference [6] and references therein) for patterning and etching of a quantum well structure. The two-dimensional quantum well has a smaller bandgap than the surrounding matrix (e. g. InAs quantum well in a GaAs matrix) and

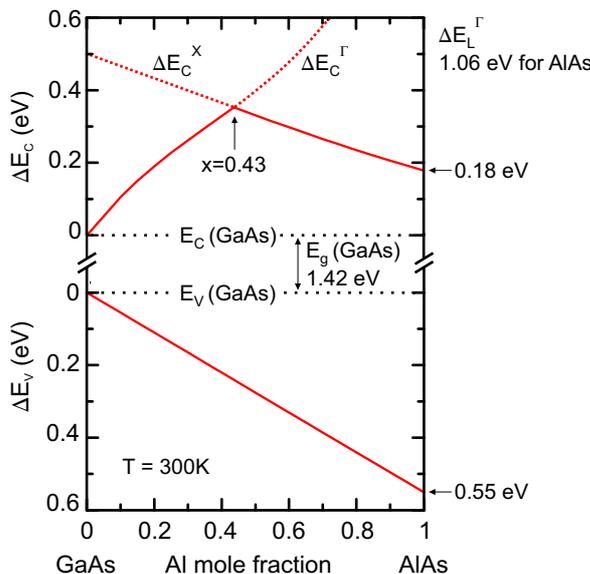


Figure 2.3: Conduction (ΔE_C) and valence band offset (ΔE_V) in dependence of the Al mole fraction x for a GaAs/ $\text{Al}_x\text{Ga}_{1-x}\text{As}$ heterostructure. The dependence of the conduction band offset is shown for the X -point (ΔE_C^X) and the Γ -point (ΔE_C^Γ). After [28–30].

confines the charge carriers in one dimension. The lithography technique is used in a second step for etching away the surrounding matrix within a nanometer scale: A QD structure with a zero-dimensional confinement is obtained. However, this top-down approach suffers from insufficient lateral resolution and interface damage caused by the patterning and etching procedure.

Alternatively, the fabrication of QDs can be done by self-organization in a bottom-up approach. Already in 1938 Stranski and Krastanow published [5] such a method, which offers the possibility of self-organized island formation. On a flat substrate surface a layer of material with a different lattice constant is deposited, leading to a strained wetting layer. The strain energy can be reduced by forming small islands from the material of the wetting layer. These small structures with dimensions in the nanometer range can be very regular in size and shape and can confine carriers in all three dimensions. In the following years the term Stranski-Krastanow growth mode (SK-growth) was used in heteroepitaxy for the formation of islands on an initially two-dimensional layer. The formation of coherent, i. e. defect free, islands as a result of SK-growth is today exploited to fabricate QDs [7, 8].

Figure 2.4(a) shows a three-dimensional STM image (STM: scanning tunneling microscopy) of an uncovered InAs QD, grown on an a GaAs[001] surface. After the QD formation the QDs are normally covered with GaAs for optical or electrical purposes. Figure 2.4(b) shows a cross-sectional TEM image (TEM: Transmission Electron Microscopy) of a typical InGaAs QD covered with GaAs and AlGaAs, forming the surrounding matrix. The wetting layer is also visible in this TEM image. Self-organized

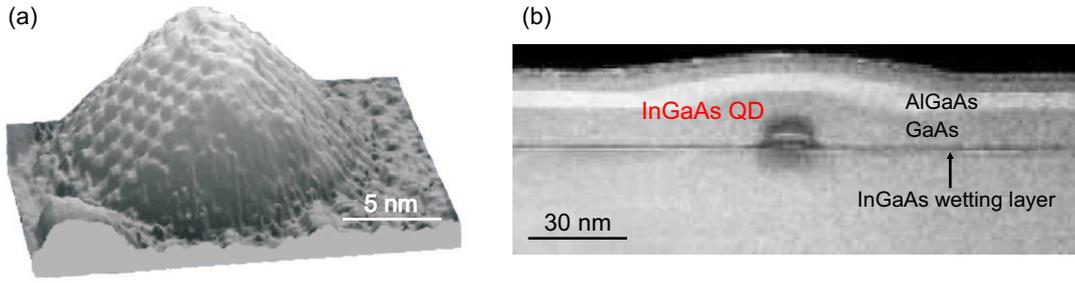


Figure 2.4: (a) Three dimensional top-view scanning tunneling microscopy (STM) image of an uncovered InAs QD, grown on a GaAs[001] surface [44]. (b) Cross-sectional transmission electron microscopy (TEM) image of a GaAs-covered InGaAs QD [45]. The InGaAs wetting layer is also visible.

QDs are supposed to have a pyramidal or truncated pyramidal shape [43] with typical base width and height in the order of 10-30 nm and 2-5 nm, respectively, depending on the growth conditions and material combination.

In contrast to the SK-growth mode, where the QDs are formed on a wetting layer which remains after the QD formation, the Volmer-Weber growth mode leads also to QD formation *without* a wetting layer. A third well-known growth mode is the Frank-van der Merwe mode, which represents a layer-by-layer growth (two dimensional) without QD formation. The particular growth mode for a given system depends on the interface energies and the lattice mismatch. For detailed information, see for instance, Reference [6] and references therein.

2.2.2 Electronic Properties

Charge carriers in a semiconductor bulk material can, in principle, move freely in all three directions. A classical band structure with a continuous dispersion of energy as function of momentum is observed and the density of states $D(E)$ increases proportional to the square root of the energy: $D(E) \sim \sqrt{E}$; schematically depicted in Fig. 2.5. If the carrier motion is limited in one direction in the order of the carrier deBroglie wavelength, by using heteroepitaxial techniques, one observes quantization effects in the direction of size reduction. The deBroglie wavelength is given by:

$$\lambda = \frac{h}{p} = \frac{h}{\sqrt{3m^*k_B T}}, \quad (2.1)$$

where m^* is the effective mass of the charge carrier. For instance, the deBroglie wavelength for electrons and heavy holes in GaAs is 25 nm/9 nm at room temperature, respectively.

The reduction of free carrier motion in one direction leads to a so-called quantum well with a discrete energy spectrum in the direction of reduction. The two-dimensional density of states $D(E)$ of such a system is constant in the two directions of free motion. In the third (reduced) direction, the density of states is a step-like function, see Fig. 2.5.

That means, the density of states remains constant to a certain energy and increases step-like to the next localized sub-band.

The complete reduction of the infinite bulk extension leads to charge carrier localization in all three directions and complete breakdown of the band structure within a QD. As in a real atom, the resulting energy level structure is discrete and so is the density of states:

$$D_{QD}(E) = \sum_i \delta(E - E_i). \quad (2.2)$$

The QD provides a confining potential with finite depth for electrons and/or holes. Figure 2.6 shows schematically the two types of QDs: Type I QDs (like InAs/GaAs) confine electrons and holes, Fig. 2.6(a), while a type II QD confines either electrons or holes. The depicted type II GaSb/GaAs QD system has a localization potential for holes only and provides a barrier for the electrons in the conduction band. The localization energy represents the energy difference of the confined electron or hole state with respect to the conduction or valence band edge, equivalent to the binding energy.

The single particle states in self-organized QDs are influenced by three main structural properties: size, shape and chemical composition. To be precise, the gradient of chemical composition inside the dot [46] influences the electronic properties, too. The influence of the QD size on the electron and hole can be understood in principle from the very simple particle-in-a-box model. For an infinite potential well the quantum-mechanical solution predicts energy separation and a non-zero energy for the lowest level, depending on the size of the box. In contrast, the confining barrier in buried QDs is of finite height, given by the band offsets. In first approximation, only a finite number of bound states exist, depending on the QD size and shape. As the QD becomes smaller, the separations between the energy levels increase – quantum size effect

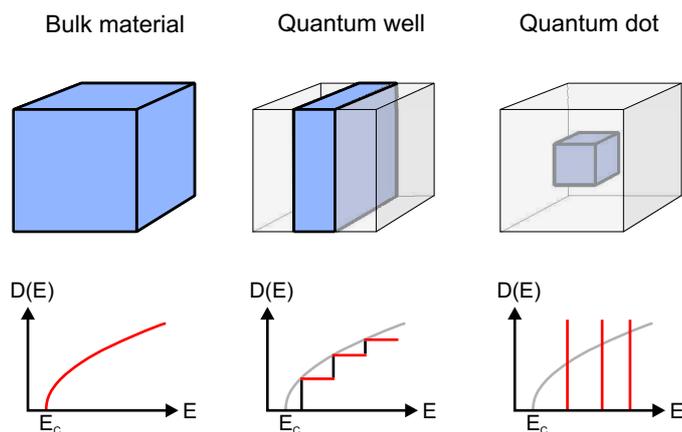


Figure 2.5: Top row: Schematic illustration of the reduction of the infinite extension of a bulk material to a quantum well in one dimension and a QD in three dimensions. Bottom row: Schematic figures of the density of electronic states.

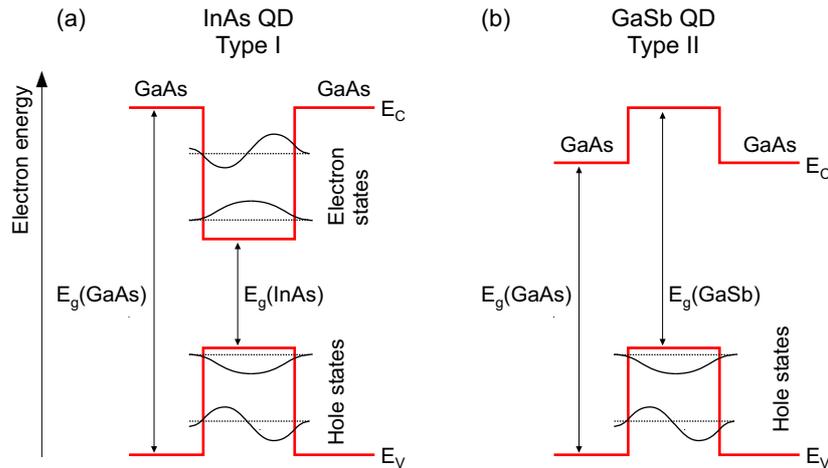


Figure 2.6: Schematic picture of the electron and hole states in a type-I InAs/GaAs (a) and in a type-II GaSb/GaAs (b) QD system. A type-I QD confines electrons as well as holes, while a type-II QD confines either electrons or holes. The depicted type-II GaSb/GaAs QD confines only holes and has a potential barrier for electrons in the conduction band.

– and all states are lifted in energy towards the valence/conduction band edge. Below a certain size no bound electron and/or hole states exist.

Besides the size, shape and composition the strain distribution inside the QD is of central importance. For QDs formed by SK-growth from two lattice-mismatched materials, the confinement potential is not constant inside the dot and an inhomogeneous strain distribution is present [47]. An additional contribution to the confinement potential often arises from piezoelectric charging of the QD heterostructure, scaling approximately linearly with size.

Modelling of QD States

Starting with the more or less known structural properties of a QD (as its size, shape and composition) and using a continuum mechanical (CM) or a valence force field model (VVF) to model the strain distribution inside the dot, the 8-band $\mathbf{k}\cdot\mathbf{p}$ theory can be used to calculate the single particle states with the corresponding wavefunctions [48,49]. Within this model, the conduction band and the three valence bands (heavy hole, light hole, and spin-orbit split off) with their coupling interactions are taken into account. By including the electron and hole spins, 8-bands are used in this model. A more detailed derivation of this theory and its application to QDs can be found in Reference [49].

Figure 2.7(a) shows the electron and hole energies for a pyramidal QD shape for different base lengths. The corresponding wave functions (probability density isosurfaces for $p=65\%$) are displayed in Fig. 2.7, too. Here, a continuum mechanical model was used to obtain the strain distribution. Clearly visible is the strong dependence of the electron and hole state energy on the base length of the QD. For increasing dot size an increasing localization energy for the electron/hole ground state up to ~ 315 meV

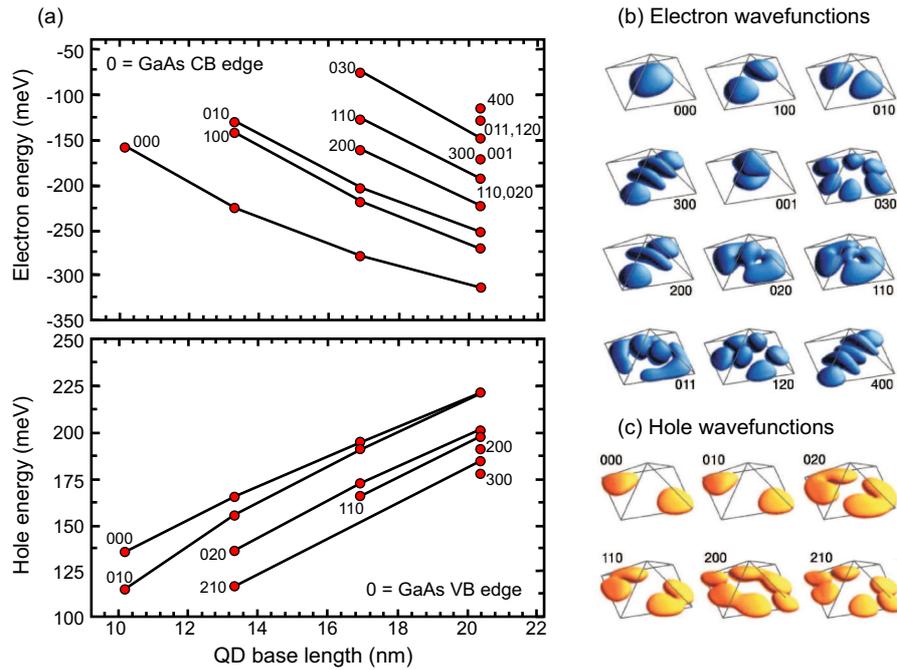


Figure 2.7: Calculated bound electron and hole state energy in a pyramidal InAs/GaAs QD, obtained by using 8-band $\mathbf{k}\cdot\mathbf{p}$ theory. Panel (a) shows the electron/hole energy in dependence of the QD base length. The lines connect states having the same wave function symmetry. The 000 state is the electron and hole ground state, respectively. Corresponding probability density isosurfaces ($p = 65\%$) are shown for electron (b) and hole (c) states for a QD base length of 20.4 nm [48, 50].

and ~ 225 meV is obtained from the calculation for a QD base length of 20.4 nm.

Besides the 8-band $\mathbf{k}\cdot\mathbf{p}$ -theory, the empirical pseudopotential theory has been established to calculate the electronic properties of self-organized QDs. This method starts with supercells containing hundred of thousands of atoms to obtain a QD potential. In contrast the 8-band $\mathbf{k}\cdot\mathbf{p}$ -theory starts with bulk properties (the 8 bands) to obtain the electronic structure of the QD. A description of the pseudo-potential theory can be found, for instance, in Reference [51].

2.2.3 Charge Carrier Emission

Self-organized QDs can trap electrons or holes, storing these charge carriers for a certain time depending on the emission processes involved. Therefore, QDs act as a information storage unit: a memory cell. In Section 2.3 the most common semiconductor memories are described and a QD-based memory is proposed. Previously, in this section the main carrier escape mechanisms from QD states into the conduction or valence band are discussed. These mechanisms limit the storage time in such a QD-based memory and are, hence, of importance.

Confined charge carriers can escape by the following four mechanisms: Thermal activation, tunneling, phonon-assisted tunneling, and optical activation. Phonon-assisted

tunneling is a combination process of tunneling and thermal activation. The charge carrier is thermally activated into a higher energy state and tunnels subsequently into the conduction or valence band through the triangular barrier. The mechanisms are schematically depicted in Fig. 2.8 for electron escape into the conduction band and briefly described in the following. A more detailed description with derivations of the rate equations is presented in Chapter 3.

Thermal Activation

Thermal activation is one of the major process for carrier emission from self-organized QDs as well as for deep levels in semiconductors. Thermal activation from deep levels and QDs have been widely studied, see e. g. [52–54]. Besides tunneling, it is the major emission process, which would limit the charge carrier storage time in a QD-based memory.

A QD, acting as a deep level in a semiconductor with a thermal activation energy E_A can emit electrons (n) or holes (p), with a net rate of emission $R_n^e(t)$ and $R_p^e(t)$, respectively. The net emission rate is the number of charge carrier per unit-time, emitted from a QD or trap per volume. This definition holds also for the electron and hole net capture rate $R_n^c(t)$ and $R_p^c(t)$. In thermal equilibrium, the net rate of emission and capture must be equal. This principle of detailed balance yields in a steady state situation an equation for the thermal emission rate of the electrons

$$e_n(T) = \gamma_n \frac{g_0}{g_1} T^2 \sigma^n \exp\left(-\frac{E_A^n}{k_B T}\right), \quad (2.3)$$

and holes

$$e_p(T) = \gamma_p \frac{g_1}{g_0} T^2 \sigma^p \exp\left(-\frac{E_A^p}{k_B T}\right), \quad (2.4)$$

where g_0 (g_1) is the degeneracy of the empty (occupied) state, σ^n (σ^p) is the capture cross-section at a certain temperature, E_A^n (E_A^p) the electron (hole) activation energy

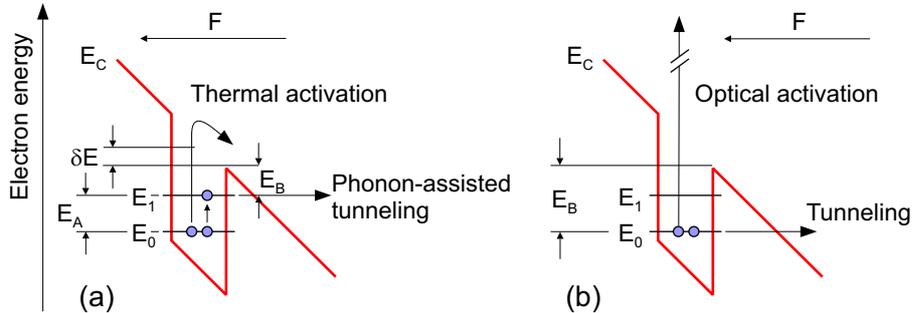


Figure 2.8: The possible carrier emission processes in an applied electric field: Thermal activation and phonon-assisted tunneling (a), tunneling and optical activation (b) are shown for electron escape from a QD ground state into the conduction band.

and γ_n (γ_p) a temperature-independent constant for the electron (hole) state, evaluated in Section 3.1.3.

Experimentally, it is found that the thermally activated electron or hole emission process fits very well an equation like Equation (2.3) or (2.4). However, considerable care is required in the physical interpretation of the thermal activation energy E_A . Especially in QDs the emission process can be more complicated and the thermal activation energy is not always equal to the localization (binding energy) of the QD state. More details can be found in Chapter 3.

Pure Tunneling

If the band structure is tilted due to the presence of an electric field, the emission rate of charge carriers can be enhanced due to three effects: (a) the Poole-Frenkel effect, (b) pure tunneling, and (c) phonon-assisted tunneling.

Poole-Frenkel is a classical effect, Frenkel [55] already described in 1938 in a one-dimensional model. The charge carrier is still emitted due to thermal activation over the top of the potential barrier, however, the barrier is lowered by the presence of the electric field F , cf. Fig. 2.8(a). Based on the expression for the emission rate, Equation (2.3), without an electric field, e_n one commonly writes

$$e_n(F) = e_n \cdot \exp \left[\frac{\delta E(F)}{k_B T} \right], \quad (2.5)$$

where $\delta E(F)$ is the effective change in the activation energy of the level, which is for a square well potential with radius r and the elementary charge e :

$$\delta E(F) = eFr. \quad (2.6)$$

The decrease in the barrier height δE has also been calculated for a Coulomb well to $\delta E = e\sqrt{(eF)/(\pi\epsilon)}$ [56], with the dielectric constant ϵ of the surrounding matrix. From more accurate three-dimensional calculation for a point defect, more complicated solutions have been found [57, 58]. However, Equation (2.6) is a good first approximation and is normally used to describe the Poole-Frenkel energy.

Besides the Poole-Frenkel effect, the charge carrier emission rate can be enhanced by pure tunneling to the conduction or valence band, Fig. 2.8(b). This well-known quantum mechanical effect has been calculated for a Dirac well with a height E_B into a continuous band of energy states in [56, 59], yielding an equation for the tunneling rate e^T (or emission probability) through a triangular barrier:

$$e^T = \frac{eF}{4\sqrt{2m^*E_B}} \exp \left[-\frac{4}{3} \sqrt{\frac{2m^*}{\hbar^2}} \cdot \frac{E_B^{3/2}}{eF} \right]. \quad (2.7)$$

The exponential factor is the common expression for the transparency of the triangular barrier, whereas the pre-exponential factor is a result of the wavefunction in the confining potential.

Using an one-dimensional (1D) Wentzel-Kramer-Brillouin (WKB) calculation [60] for an 1D potential with a confining potential in z-direction L , the tunneling rate e^T has a different pre-exponential factor:

$$e^T = \frac{\hbar\pi}{2m^*L^2} \exp \left[-\frac{4}{3} \sqrt{\frac{2m^*}{\hbar^2}} \frac{E_B^{3/2}}{eF} \right]. \quad (2.8)$$

Phonon-Assisted Tunneling

A combination process of thermal activation and tunneling is the phonon-assisted tunneling, schematically illustrated in Fig. 2.8(a). For a confined energy level in an applied electric field, the phonon-assisted tunneling represents a two step emission process: thermal activation to a higher confined energy state and subsequent tunneling through the remaining triangular barrier. The emission rate for this combined process depends on the temperature, the potential shape and depth, and the electric field.

The emission rate of phonon-assisted tunneling from a Dirac well to a continuum of states has been calculated in a semi-classical [56] and in a full quantum mechanical [59, 61] approach. The latter yields an expression of the form

$$e^P = \sum_{m=-\infty}^{\infty} W_m \cdot e^T(E_A + m\hbar\omega), \quad (2.9)$$

where the sum includes all involved phonon modes, and $e^T(E_A + \hbar\omega)$ is the expression from Equation (2.7). W_m is the statistical weight for the phonon mode m , that can be found in [54, 61].

The more simplified semi-classical approach [56] yields an expression, where the total emission rate e^P for phonon-assisted tunneling is a product of the transparency of the triangular barrier, Equation (2.8), and the thermal activation rate for electrons, Equation (2.3), see Ref. [62]:

$$e_n^P = \gamma_n \frac{g_0}{g_1} T^2 \sigma^n \exp \left(-\frac{E_A^n}{k_B T} \right) \cdot \left[-\frac{4}{3} \sqrt{\frac{2m^*}{\hbar^2}} \frac{E_B^{3/2}}{eF} \right]. \quad (2.10)$$

Vincent et al. [56] demonstrated, that for a spatially extended potential, like the Coulomb potential of deep levels or the confinement of QDs, both the Poole-Frenkel effect and the phonon-assisted tunneling are important for typical electric fields in the order of $10^6 \dots 10^8$ V/m. The influence of phonon-assisted tunneling has been studied experimentally for deep levels [63–65] and observed for QDs [62, 66–68] in various material systems.

Optical Activation

Carrier emission due to optical activation plays no role within this thesis, however, it is mentioned here to complete the list of possible emission processes. Carriers can

escape from confined states by optical activation (also called photoionization or intraband absorption), if the sample is irradiated with photons having a sufficient energy, Fig. 2.8(b).

For QDs, a few theoretical studies of intraband absorption have been reported so far, mainly on intraband absorption from ground to excited states inside the QD [48]. Zhang et al. [69] calculated intraband absorption for pyramidal InAs/GaAs QDs for transitions from the ground or excited state to wetting layer and conduction band states for photon energies up to 600 meV.

2.3 Memories

The semiconductor memory industry focuses essentially on two memories: The DRAM and the Flash. Both memory concepts have their advantages and disadvantages in speed, endurance, cost and storage time. A memory concept which adds the advantages of a DRAM to a Flash is one of the major challenge for the industry and would combine high storage densities, fast read/write access, good endurance and low production cost. In addition, for portable applications like mobile phones and mp3-players, low power consumption in combination with long storage time in the order of years (non-volatility) is desired. A crowd of new concepts has been developed in the past, like magnetic random access memory (MRAM) and ferro-electric random access memory (FRAM) [17]. However, the ultimate breakthrough is not in sight yet.

Especially QDs provide a number of interesting advantages for a new generation of memories. Billions of self-organized QDs are formed simultaneously in a single technological step, allowing massive parallel production in a bottom-up approach and offering an elegant method to create huge ensembles of electronic traps. They can store just a few or even single charge carriers with a retention time depending on the material combination – possibly up to years at room temperature. With an area density of up to 10^{11} per cm^{-2} an enormous storage density in the order of 1 TBit/inch² could be possible, if each QD would represent one information bit. Epitaxial investigations have already shown, that it is possible to produce ordered arrays of dots on an atomically flat surface [70], a crucial requirement for electrical addressing in memory structures. Furthermore, the carrier relaxation process into QDs is in the order of sub-picoseconds [19, 20], an important prerequisite for a very fast write access time in such memories. Hence, the implementation of QDs could lead to a non-volatile memory with high storage densities combined with a fast read/write access time.

This chapter concentrates on the basic principles for using QDs as building blocks in future memories. In Section 2.3.1 and 2.3.2, the most common semiconductor memories are described. In Section 2.3.3 a ultimate/universal memory cell based on self-organized QDs is proposed, which could combine the advantages of DRAM and Flash.

2.3.1 Dynamic Random Access Memory

A DRAM is the main memory in a personal computer and consists of a transistor and a capacitor, Fig. 2.9(a). The capacitor stores the information by means of electric charge

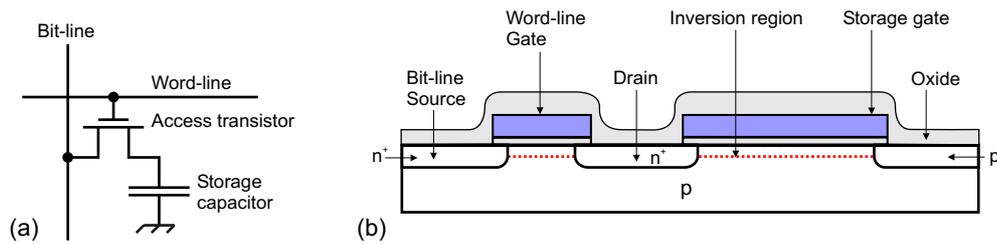


Figure 2.9: Schematic picture of a single-transistor DRAM cell with a storage capacitor (a). Cell layout for a planar DRAM cell (b).

and the charge state is given by the voltage level on the capacitor. The stored charge will be removed typically in a few milliseconds mainly due to leakage and recombination currents. Dynamic memories require periodic “refreshing” of the stored charge, they are volatile and even after a read process the information has to be written again into the DRAM cell. The refresh process is controlled automatically by chip on the integrated circuit and this continuously read and write operation has led to the name “Dynamic”.

Figure 2.9(b) shows schematically the layout of a planar – the simplest – DRAM cell [71]. The storage capacitor uses the inversion channel region as one plate, the storage gate (which is polysilicon) as the other plate, and the gate oxide as the dielectric. The access transistor is a metal oxide semiconductor field effect transistor (MOSFET) with a source, drain and a gate contact. The drain contact is linked to the storage capacitor. The source contact is connected to the bit-line and the gate contact to the word-line. Via word and bit-line a single cell can be addressed in an organized matrix of DRAM cells: this so-called “random access” provides short access times independently of location of the data (see Ref. [18] for more details). For a write/read operation, a voltage pulse is passed through the selected bit and word-line. Only at the crossing point of a certain DRAM cell, the access transistor switches to “open” and the capacitor is charged with electrons.

DRAM cells provide a fast read, write and erase access time in the order of ~ 20 ns in combination with a very good endurance of more than 10^{15} write/erase cycles. The endurance is defined as the minimum number of possible write/erase operations until the memory cell is destroyed. On the other hand, DRAM memory cells draw power continuously due to the refresh process and the information is lost after switching off the computer. One information bit is also stored with a relatively large number of electrons in the order of 10^5 , shown in Fig. 2.10 (where the number of electrons per cell is plotted versus the feature size for DRAM and Flash and different single and few-electron memories). A permanent high power consumption is the result. The volatility and the power consumption make DRAMs unsuitable for mobile applications.

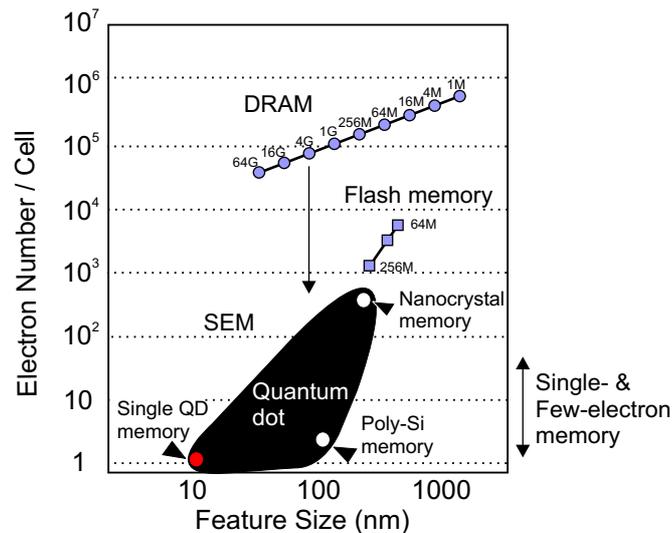


Figure 2.10: Electron number per one memory cell versus minimum feature size for conventional memories and single- and few-electron memories. After [72].

2.3.2 Non-volatile Flash Memory

Static RAMs (SRAMs), used for the first and second level cache in a personal computer, are like DRAMs volatile. SRAM have the advantage that they do not need periodic refreshing, however, they lose their stored data when switching off the power. Non-volatile memories (NVMs), on the other hand, can retain their data for typically more than 10 years without power consumption. The most important non-volatile memory is the Flash-EEPROM (“Electrically erasable and programmable read-only memory”), in short just Flash. Flash has the possibility of repeatable electrical read, write and programming. The Flash memory market is the fastest growing memory market today, as it is the ideal memory device especially for portable applications. In a mobile phone it holds the instructions and data needed to send and receive calls, and stores phone numbers. But not only portable applications are ideal for Flash. A Flash chip tells a computer how to boot up. Electronic products of all types, from microwaves ovens to industrial machinery, store their operating instructions in Flash memories.

The most common Flash memory is a silicon based floating gate structure. In this memory concept [73], the charge carriers are trapped inside the floating-gate, which is a poly-silicon structure embedded in two SiO_2 barriers. The band structure of such a memory concept is schematically depicted in Fig. 2.11(a). The SiO_2 barriers have a thickness in the order of 10 nm and an enormous energetic height of more than 3 eV to maintain a storage time of more than 10 years at room temperature. However, the two main disadvantages of a Flash cell have its origin in these energetically extreme high barriers: The slow write time in the order of milliseconds and the poor endurance in the order of 10^6 write/erase cycles. The SiO_2 barriers are fixed in their energy height and the charge carriers are injected into the floating gate over one of these barriers

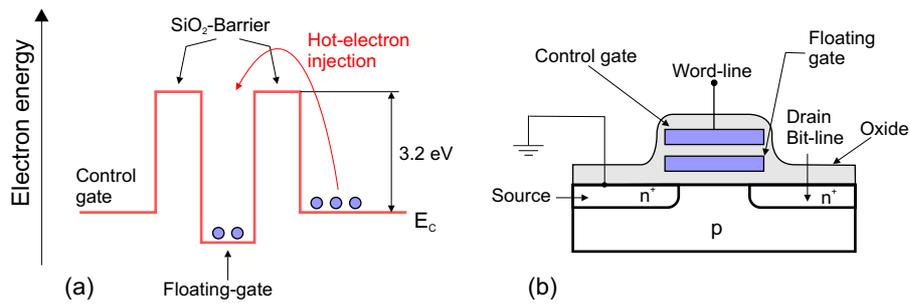


Figure 2.11: Schematic bandstructure of a floating-gate Flash memory based on Si as matrix material and SiO₂ for the barriers (a). Cell layout of a floating-gate Flash (b). After [73].

Memory	Write	Read	Storage	Electrons	Endurance
DRAM 	~20 ns good	~20 ns good	~ms poor	>10000 poor	>10 ¹⁵ good
Flash 	~10 ms poor	~20 ns good	>10 years good	~1000 good	~10 ⁶ poor

Table 2.1: Comparison of DRAM and Flash concerning some selected important properties. The DRAM cell provides a fast write and read access time (~ 20 ns) but has only a storage time in the order of milliseconds at room temperature (volatile). In comparison, the Flash cell is non-volatile with a storage time of more than 10 years. However, a Flash cell has a poor endurance (10^6) and a slow write time (milliseconds).

during the write process.

The write process of the information is realized by means of “hot-electron injection” or Fowler-Nordheim tunneling. During the hot-electron injection process, a small voltage is applied between the source contact and the bit-line (drain contact). The word-line (control gate) is set to a high positive bias of 10-20 V. Now, the MOSFET transistor is set to “open” and electrons can flow through the inversion region from source to drain and are, in addition, accelerated in the direction of the floating-gate due to a high electric field in the order of 10^7 V/cm. The electrons in the inversion region, which are accelerated to the floating gate are getting a high kinetic energy. These “hot-electrons” have enough kinetic energy to reach the floating gate over one of the SiO₂ barriers, but destroy step by step this barrier. This is the reason for the poor endurance of a Flash memory cell. In addition, the slow write time is due to a low probability for a relaxation process of these hot-electrons into the floating gate.

The read-out of the stored information is normally done by measuring the resistance of the two-dimensional electron gas (2DEG) of the MOSFET structure – the inversion channel. Again, a bias is applied at the gate contact, forming a thin channel (the 2DEG) between source and drain. Stored electrons in the floating gate reduce the mobility of the 2DEG and a higher resistance between source and drain is measured.

The non-volatility and the smaller number of charge carriers involved per bit, in the order of 1000 electrons for a floating gate Flash cell, yields a reduced power consumption than in a DRAM. Table 2.1 summarizes the main properties of DRAM and Flash for comparison.

2.3.3 A QD-based Memory Cell

An ultimate or universal memory cell should provide long storage times like Flash and long endurance ($> 10^{15}$) in combination with even better read/write- and erase times than DRAM (< 1 ns). In addition, the ultimate memory cell should store one information bit by means of a single charge carrier, either electron or hole. Such a memory offers the possibility to operate in all kinds of applications like the main memory (nowadays DRAM), mass storage devices (nowadays hard-disk drive), mp3-players (nowadays Flash), digital cameras (nowadays Flash), memory sticks (nowadays Flash), mobile phones (Flash), etc. A QD-based Nano-Flash could fulfill the requirements and have a lot of advantages in many kinds of applications, like:

- The realization of a non-volatile PC main memory is possible. Such a memory with fast read/write access time could provide the entire data of a main memory just direct after switch on. The time-consuming boot-up progress of a personal computer, the data transfer from the hard disk drive to the main memory, would not be necessary anymore.
- The DRAM main memory has an average access time (read/write) of approximately 20 ns, while the central processing unit (CPU) has nowadays a clock frequency of about 3 GHz, that means an average access time in the order of 0.33 ns, more than 2 orders of magnitude faster. This processor-memory bottleneck [74] widens every year by about 5 percent, becoming more and more a primary obstacle to improve computer performance. The computation speed of a modern PC is limited by the access time of the DRAM main memory. Intel and AMD already introduced double core processor for parallel computing [75] to circumvent this obstacle, as a further increase of the processor speed with constant DRAM access time would be useless. A fast semiconductor memory with read/write and erase times below a nanosecond could close the processor-memory bottleneck again.
- As already mentioned, a crucial point for future memory devices is the power consumption. DRAM and Flash need approximately 10^3 and 10^5 charge carriers per information bit. The power consumption could be reduced by a reduction of the number of charge carriers involved, whereas a single charge carrier memory is the ultimate goal. Self-organized QDs can trap only few or single electrons/holes. A single charge carrier memory is, hence, possible if one information bit is represented by one QD.

- Today's DRAM and Flash memories have a storage density of about 10 Gbit/inch² for a feature size of 90 nm [4]. By using QDs in a memory where every single QD represents an information bit, storage density of up to 1 Tbit/inch² can be reached for area densities of 10¹¹ QDs per cm². A prerequisite for such high storage density is the possibility to address every single QD individually.

QDs in a p-n Diode

The concept of a QD-based memory [76] consists of three parts: This first part is the use of QDs as storage units for the charge carriers. Both states of an information bit ("0" and "1") are represented by the charge state of the QDs. The second part is to place the QDs inside a p-n or Schottky diode structure, nearby or inside a depletion region. The depletion region can be used to realize three different memory operations: storage, writing and erasing of the data. The third part realizes the read operation. Therefore, a 2DEG is placed below the QD layer.

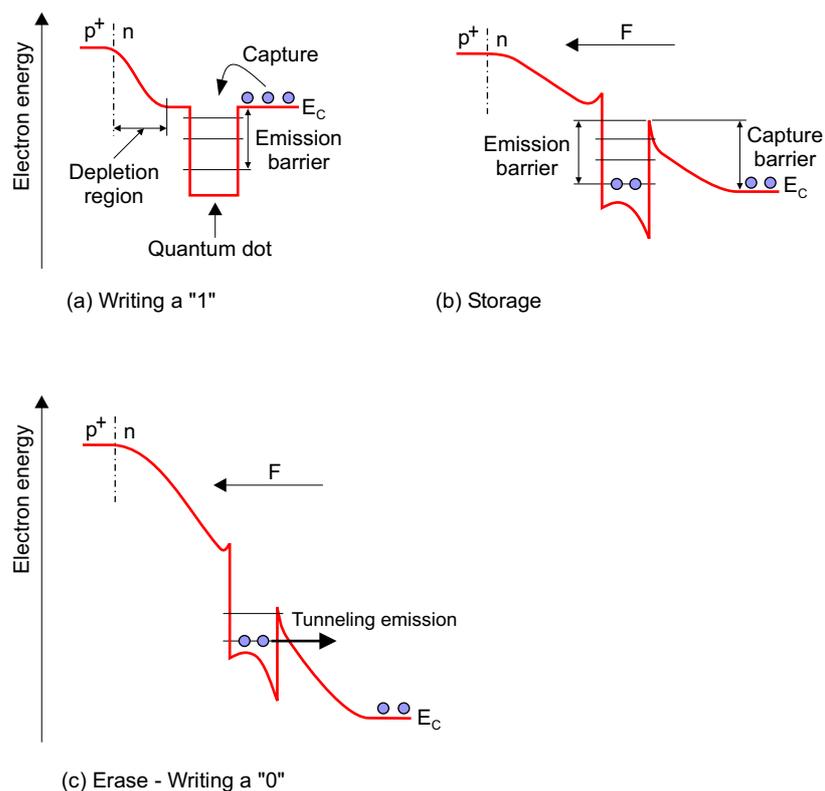


Figure 2.12: Schematic illustration of the write, erase, and storage process in a proposed QD-based Flash memory [76]. A very fast write time of a "1" state in the order of picoseconds is expected since the carrier capture process occurs directly from the conduction or valence band edge (a). The storage situation is achieved by placing the QDs in the depletion-region of a p-n diode (b). The information can be erased by applying a high electric field in reverse direction, which leads to tunneling of the charge carriers (c).

The charge carrier storage and the write/erase process works as follows and is schematically depicted in Fig. 2.12: For the storage operation, an emission and capture barrier is needed. The emission barrier is necessary to store a “1”, the capture barrier to store a “0”. The emission barrier prevents the stored charge carriers inside the QDs, representing a “1” state, from thermally activated emission and tunneling into the conduction or valence band. In this memory concept, the emission barrier is realized by the localization energy of the charge carriers and can be easily adjusted by the material combination of QDs and surrounding matrix. Without a capture barrier, charge carriers would relax directly into the QDs and flip a “0” to a “1” state. The capture barrier is maintained by the band bending across the QDs, if they are placed inside the depletion region of a p-n diode.

In contrast to the fixed and extremely high barriers in the Flash memory, in this concept, a doped matrix material (like Al(Ga)As) should be used. Then, the capture barrier is provided by the depletion region of the p-n diode and can be adjusted easily by an external bias. Free electrons or holes exist at the conduction or valence band and can relax directly into the QD states during the write process, Fig. 2.12(a), without a capture barrier like in a Flash memory cell. The disadvantages of a Flash memory due to the energetically extreme high SiO₂ barriers are eliminated. Fast write operations in the order of the charge carrier relaxation time of picoseconds [19, 20] in QDs are in principle possible, more than four orders of magnitude faster than the typical DRAM write access time (~ 20 ns). In addition, a very good endurance of 10^{15} write/erase operations (like in a DRAM cell) should be feasible.

The width of the depletion region can be adjusted by an applied external bias (see Chapter 4) to realize the write/erase and storage situation. By a forward bias, the width of the depletion can be reduced, so that the QDs are outside the depletion region, Fig. 2.12(a), and charge carrier capture into the QD states occurs. This represents the write operation of a “1”. For zero bias, the doping concentration and the location of the QDs inside the p-n diode is adjusted accordingly, so that the QDs are inside the depletion region. This represents the storage situation, Fig. 2.12(b). To erase the information (this operation is equal to writing a “0” state), a high reverse bias is applied to the p-n diode, Fig. 2.12(c). The QDs are inside the depletion region and a strong electric field (up to 10^7 V/cm) leads to tunneling of the charge carriers through the triangular barrier. The erase time depends on the emission barrier height and the electric field, which again depends on the doping concentration and applied reverse bias.

One possibility to read-out the stored information is the usage of a 2DEG below the QD layer. As already described for the Flash memory device (Section 2.3.2), by measuring the resistance of the 2DEG the stored information can be obtained. Stored electrons in the QDs reduce the mobility of the 2DEG, hence, a higher resistance is measured. This situation is attributed to a stored “1”. A low resistance of the 2DEG is attributed to a stored “0”. In contrast to the Flash memory, the 2DEG can be formed by a In(Ga)As quantum well and a δ -doping below the QD layer, shown in Fig. 2.13. The δ -doping provides charge carriers, which relax into the quantum well and form

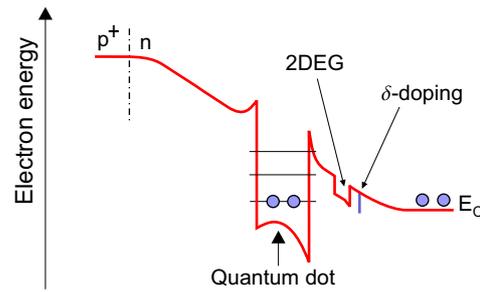


Figure 2.13: Schematic illustration of a possible realization of the read process in the QD-based Flash memory. Below the QD layer a two-dimensional electron gas (2DEG) is placed. The resistance of the electron gas depends on the number of charge carriers inside the QDs, hence, by measuring the resistance the charge state (the information) can be obtained.

a 2DEG at the interface. This read-out process has the advantage to maintain the information state within the memory cell, which is in contrast to the DRAM read-out mechanism, where the charge state of the capacitor is destroyed and has to be rewritten again.

Device structure

Figure 2.14 shows schematically the device structure of such a QD-Flash, where the QDs are charged with holes to represent an information bit. The doping sequence of the p- and n-regions would be vice versa for an electron storage device.

The QDs are embedded in a n^+ -p diode structure below the p-n junction. The distance to the junction is adjusted, such that the QDs are inside the depletion region for zero bias. The appropriate distance between p-n junction and QD layer depends on

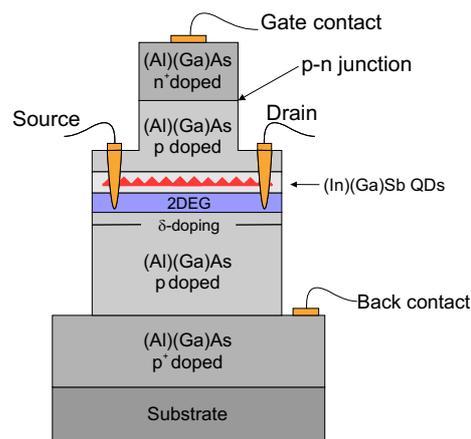


Figure 2.14: Schematic picture of the layer sequence for a possible QD-based Flash memory. The QDs are located below the p-n junction and below the QD layer a 2DEG is placed to detect the charge state.

the doping concentration of the matrix surrounding the QDs. It can easily be calculated by using Equation (4.19) for the width of the depletion region in Section 4.1.3. The matrix material should consist of $\text{Al}_x\text{Ga}_{1-x}\text{As}$ with a doping concentration in the order of $10^{16}\dots 10^{17}\text{ cm}^{-3}$. $(\text{In})(\text{Ga})\text{Sb}$ is the most promising material for the QDs to realize a strong hole localization, leading to a long storage time at room temperature in the order of years. The relation between storage time and material combination is the one of the main focuses in Chapter 5.

The write and erase operation can be achieved by an appropriate external bias, applied between gate and back contact. Both contacts are placed on highly p- and n-doped ($> 10^{18}\text{ cm}^{-3}$) layers. The 2DEG is situated 10-50 nm below the QD layer and should be outside the depletion region at zero bias, so that it is filled with charge carriers, provided by the additional δ -doping. The 2DEG is connected to the source and drain contact to measure the resistance which depends on the charge state of the QDs.

“TODAY’S SCIENTISTS HAVE SUBSTITUTED MATHEMATICS FOR EXPERIMENTS, AND THEY WANDER OFF THROUGH EQUATION AFTER EQUATION, AND EVENTUALLY BUILD A STRUCTURE WHICH HAS NO RELATION TO REALITY.”
NIKOLA TESLA

Chapter 3

Carrier Dynamics

In this chapter, the carrier dynamics (the capture and emission of majority and minority charge carriers) of deep levels and QDs is discussed. The storage time in QDs, investigated in Chapter 5, is limited by charge carrier emission due to thermal activation and/or tunneling. The capture process, studied in Chapter 6, limits the possible write time in a QD-based memory.

Section 3.1 concentrates on carrier capture and emission of deep levels in semiconductors and the common procedure for deriving the thermal emission rate from a detailed balance argument, Section (3.1.2). The carrier dynamics of deep levels has been intensively investigated experimentally and theoretically [53, 77]. Experimental investigations on QDs in semiconductors with capacitance spectroscopy [78–80] have demonstrated, that the theoretical assumptions are in general also valid for these low-dimensional systems. In first approximation, they can be treated like deep states in a semiconductor. Of course, some differences exist, hence, Section 3.2 describes the carrier dynamics for QDs in more detail. This chapter is also a prerequisite for understanding the time-resolved measurement method, the capacitance spectroscopy, presented in Chapter 4.

3.1 Deep levels in Semiconductors

Besides the effective doping due to shallow donor and acceptor states close to the band edges, the electrical properties of semiconductors are also determined by energetically deeper states. While the shallow doping states are mostly ionized for practical temperatures (>50 K), deep levels are usually referred to as “deep” if their binding energy relative to the conduction band or valence band, respectively, is larger than several $k_B T$. As a consequence, these deep levels can be charged with electrons or holes even up to room temperature and beyond. In this section, the basic description of the carrier dynamics of such a deep level system is developed in a conventional rate equation approach. The discussion of the carrier dynamics refers to deep levels or more briefly “traps”, without taking into account the particularities of a QD systems, which are discussed in Section 3.2.

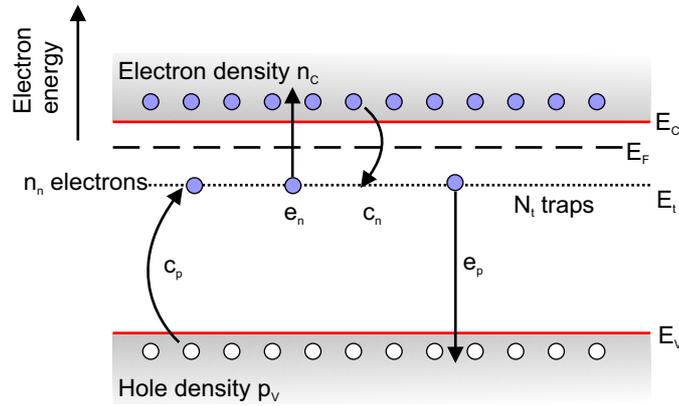


Figure 3.1: Electron and hole capture and emission processes for N_t trap states with an energy E_t . The number of trapped electrons is n_n . The semiconductor has a Fermi level E_F , an electron density n_C in the conduction band E_C and a hole density p in the valence band E_V .

3.1.1 Carrier Capture and Emission

A trap level in the bandgap of a semiconductor can be charged and release charge carriers by four different dynamic processes, which are capture and emission of electrons and holes, Fig. 3.1:

- Capture of an electron from the conduction band into the trap state (c_n).
- Emission of an electron from the trap state into the conduction band (e_n).
- Capture of a hole from the valence band into the trap state (c_p).
- Emission of a hole from the trap state into the valence band (e_p).

The subscripts “n” and “p” denote electron and hole-related properties, respectively.

The electron capture rate c_n is defined analogously to the concept employed in nuclear physics and characterized by a capture cross-section σ^n times the carrier flux, which is determined by the density of carriers n times the root-mean-square thermal carrier velocity $\langle v \rangle$:

$$c_n = \sigma^n \langle v_n \rangle n_C. \quad (3.1)$$

The hole capture rate is accordingly:

$$c_p = \sigma^p \langle v_p \rangle p_V, \quad (3.2)$$

with the hole density p_V and the hole capture cross section σ^p .¹

¹In general, the capture cross-section is temperature-dependent, cf. Section 3.1.4 and it is only a proportional constant that connects the number of charge carriers n_C times the thermal velocity with the capture rate. It describes empirically the capture and relaxation into a deep state without knowing the physical mechanisms involved. A more detailed discussion can be found in Section 6.4.

The population of the deep level is determined by the interplay of the emission and capture processes of electrons and holes, Fig. 3.1. N_t shall be the trap density per volume, n_n the density of occupied trap states and $N_t - n_n$ the density of unoccupied traps. With the emission and capture rates for electrons and holes e_n , c_n , e_p , and c_p , the net rates of change in the electron occupation n_n of the deep levels due to electron/hole emission and capture are:

- Net emission rate of electrons from occupied trap states $n_n(t)$ into the conduction band: $R_n^e(t) = e_n n_n(t)$.
- Net capture rate of electrons from the conduction band into unoccupied trap states $[N_t - n_n(t)]$: $R_n^c(t) = c_n [N_t - n_n(t)]$.
- Net emission rate of holes from unoccupied trap states $[N_t - n_n(t)]$ into the valence band: $R_p^e(t) = e_p [N_t - n_n(t)]$.
- Net capture rate of holes from the valence band into occupied trap states $n_n(t)$: $R_p^c(t) = c_p n_n(t)$.

The quantity called the “emission rate” e_n is in fact the “emission rate per trapped electron”, representing the probability per unit time that a particular electron is emitted, which is a characteristic of the trap; it is constant throughout the emission process, that means time-independent. In this simple situation, e_n is also the rate constant. Similarly, c_n is defined as the capture rate per unoccupied state and everything holds also for the hole emission e_p and capture rate c_p . In contrast, the *net rates* R_n^c , R_n^e , R_p^c , and R_p^e are *time-dependent*. That means, the net rates in the trap population represent the number of charge carriers per unit time that are emitted or captured.

The net rate of change for N_t traps in an n-doped material is the sum of the net emission and capture rates:

$$\begin{aligned} \frac{dn_n(t)}{dt} &= R_n^c(t) - R_n^e(t) + R_p^e(t) - R_p^c(t) \\ &= (c_n + e_p) [N_t - n_n(t)] - (e_n + c_p) n_n(t). \end{aligned} \quad (3.3)$$

This differential Equation (3.3) is a rate equation, which describes the population of the trap states with charge carriers. For a weak concentration of deep states N_t compared to the net doping, the influence of the emission and capture processes on the free carrier concentrations n_C and p_V in the semiconductor is negligible. Therefore, n_C and p_V can be considered as constant and solutions of the rate Equation (3.3) may be obtained subject to the boundary conditions.

3.1.2 Detailed Balance Condition in Thermal Equilibrium

For the following examination of the time-dependent behavior of the deep states population, it should be noted here: Whatever the initial conditions might be, the system will strive to reach thermodynamic equilibrium: This implies the principle of detailed balance. Therefore, the net electron capture and emission rates R_n^c and R_n^e must be

equal. The same must hold for the capture and emission of holes with the net rates R_p^c and R_p^e . These requirements have to be fulfilled in addition to the steady state condition $\frac{dn_n(t)}{dt} = 0$. Using Equation (3.3), this leads to the average trap population in thermal equilibrium, i. e. $n_n(t)$ for $t = \infty$:

$$\frac{n_n(\infty)}{N_t} = \left(\frac{c_n}{c_n + e_n} \right) = \left(\frac{e_p}{e_p + c_p} \right). \quad (3.4)$$

The population of the traps in thermal equilibrium on the other hand is defined by the Fermi-Dirac distribution [81]. For a deep level with energy E_t and degeneracy g_0 if empty, and g_1 when occupied, the population is [52, 77]:

$$\frac{n_n(\infty)}{N_t} = \frac{1}{1 + \frac{g_0}{g_1} \exp\left(\frac{E_t - E_F}{k_B T}\right)}. \quad (3.5)$$

For a common spin-degenerate level holds $\frac{g_0}{g_1} = \frac{1}{2}$. Combining Equation (3.4) and Equation (3.5) yields for electrons

$$\frac{e_n}{c_n} = \frac{g_0}{g_1} \exp\left(\frac{E_t - E_F}{k_B T}\right), \quad (3.6)$$

and analogously for holes

$$\frac{e_p}{c_p} = \frac{g_1}{g_0} \exp\left(\frac{E_F - E_t}{k_B T}\right). \quad (3.7)$$

Equation (3.6) and (3.7) imply for $E_t < E_F$ that $c_n > e_n$ and $e_p > c_p$. Hence, from Equation (3.4) results in thermal equilibrium, that the trap state will have an average population with electrons $\frac{n_n(\infty)}{N_t}$ that is greater than one half, or simplified speaking: The state will be occupied by electrons in thermal equilibrium. For example, let $c_n = 2e_n$. Then, Equation (3.4) yields an average population with electrons of $2/3$. Accordingly, from $E_t > E_F$ results that $c_n < e_n$ and $e_p < c_p$, and the state in equilibrium will have an average population with electrons smaller than one half: The state is occupied with holes in thermal equilibrium.

It is worth noticing here, that the capture rate c_n and c_p depend on the doping concentration, whereas the capture cross-sections σ^n and σ^p , as well as the emission rates e_n and e_p , are intrinsic properties of the deep level.

In a non-degenerate semiconductor, i. e. $n_C < N_C$, where N_C is the effective density of conduction band states, the free electron concentration n_C is determined by the Boltzmann relation as function a of the temperature T and the conduction band relative to the Fermi energy E_F [82] as:

$$n_C = N_C \exp\left(-\frac{E_C - E_F}{k_B T}\right). \quad (3.8)$$

In the same manner one obtains for a non-degenerate free hole concentration

$$p_V = N_V \exp\left(-\frac{E_F - E_V}{k_B T}\right), \quad (3.9)$$

where N_V is the effective density of valence band states. Substituting Equation (3.1) into Equation (3.6) yields for the electron emission rate

$$e_n = \sigma^n \langle v_n \rangle n_C \frac{g_0}{g_1} \exp\left(\frac{E_t - E_F}{k_B T}\right), \quad (3.10)$$

and with Equation (3.8)

$$\begin{aligned} e_n &= \sigma^n \langle v_n \rangle N_C \frac{g_0}{g_1} \exp\left(\frac{E_t - E_F - (E_C - E_F)}{k_B T}\right) \\ &= \sigma^n \langle v_n \rangle N_C \frac{g_0}{g_1} \exp\left(-\frac{E_C - E_t}{k_B T}\right). \end{aligned} \quad (3.11)$$

Substituting Equation (3.2) and (3.9) into (3.7) yields for the hole emission rate

$$e_p = \sigma^p \langle v_p \rangle N_V \frac{g_1}{g_0} \exp\left(-\frac{E_t - E_V}{k_B T}\right). \quad (3.12)$$

Equation (3.11) and (3.12) describe the temperature-dependent but time-independent emission rate, which is a result from the detailed balance condition in thermal equilibrium. These equations are not an outcome of an evaluation of the fundamental emission processes involved (like Auger or phonon scattering, cf. Section 6.4) but a consequence of the assumption of detailed balance. Therefore, the rate equation depends on the capture cross-section and the thermal velocity: A striking feature for an emission rate. However, for example: A larger capture cross-section implies a faster population of the trap with charge carriers, resulting in a larger emission rate. These “empirical” hole and electron emission rates describe the physical situation very well as they are confirmed by the experiments.

3.1.3 Temperature Dependence of the Emission Rate

As already indicate in Section 3.1.2, the thermal velocity and in the capture cross-section depend on temperature. This section concentrates on the thermal velocity and Section 3.1.4 on the capture cross-section.

For electrons, the root-mean-square (rms) thermal velocity is given by

$$\langle v_n \rangle = \sqrt{\frac{3k_B T}{m_e^*}}, \quad (3.13)$$

where m_e^* is the effective mass of electrons in the conduction band. The conduction band density of states N_C is given by [82]:

$$N_C = 2M_C \left(\frac{2\pi m_e^* k_B T}{h^2}\right)^{\frac{3}{2}}, \quad (3.14)$$

where M_C is the number of conduction band minima, and h is Planck’s constant. Equation (3.11) hence gives

$$e_n(T) = \gamma_n \frac{g_0}{g_1} T^2 \sigma^n \exp\left(-\frac{E_C - E_t}{k_B T}\right), \quad (3.15)$$

where γ_n is a temperature-independent constant

$$\gamma_n = \sqrt{12(2\pi)^3} \frac{k_B^2}{h^3} M_C m_e^*. \quad (3.16)$$

Analogously, the hole emission rate is

$$e_p(T) = \gamma_p \frac{g_1}{g_0} T^2 \sigma^p \exp\left(-\frac{E_t - E_V}{k_B T}\right), \quad (3.17)$$

where γ_p is

$$\gamma_p = \sqrt{12(2\pi)^3} \frac{k_B^2}{h^3} M_V m_p^*. \quad (3.18)$$

Equation (3.15) and (3.17) describe thermally activated emission and it is obvious that a plot of $\ln(e_n/T^2)$ or $\ln(e_p/T^2)$ as function of T^{-1} yields a straight line with a slope of $E_C - E_t$ or $E_t - E_V$, respectively. Experimentally it is found that data for most traps fit an equation of this form over many orders of magnitude of e_n or e_p . However, these values are not the “true” binding/localization energy of the trap level, because the capture cross-section is temperature dependent.

3.1.4 Temperature Dependence of the Capture Cross-Section

The emission rate Equations (3.15) and (3.17) are now expanded in more detail concerning their capture cross-sections. A temperature-dependent capture cross-section due to non-radiative processes (e. g. capture by multi-phonon emission, discussed in Section 6.4 and described in References [52, 53, 83]), could have the form

$$\sigma^n(T) = \sigma_\infty^n \exp\left(-\frac{\Delta E_\sigma^n}{k_B T}\right). \quad (3.19)$$

As a consequence, Equation (3.15) gives the temperature dependence $e_n(T)$ for the electron emission

$$\begin{aligned} e_n(T) &= \gamma_n T^2 \frac{g_0}{g_1} \sigma_\infty^n \exp\left(-\frac{(E_C - E_t) + \Delta E_\sigma^n}{k_B T}\right) \\ &= \gamma_n T^2 \frac{g_0}{g_1} \sigma_\infty^n \exp\left(-\frac{E_A^n}{k_B T}\right) \end{aligned} \quad (3.20)$$

where σ_∞^n is the apparent capture cross-section at infinite temperature $T = \infty$ and E_A^n the apparent activation energy for an electron emission process. Analogously, the hole emission rate is

$$\begin{aligned} e_p(T) &= \gamma_p T^2 \frac{g_1}{g_0} \sigma_\infty^p \exp\left(-\frac{(E_t - E_V) + \Delta E_\sigma^p}{k_B T}\right) \\ &= \gamma_p T^2 \frac{g_1}{g_0} \sigma_\infty^p \exp\left(-\frac{E_A^p}{k_B T}\right) \end{aligned} \quad (3.21)$$

where σ_∞^p is the apparent capture cross-section for holes and E_A^p the corresponding apparent hole activation energy.

Considerable care is required in the physical interpretation of E_A^n and σ_∞^n (or E_A^p and σ_∞^p , respectively). In the formulation here, E_A^n can be identified with $(E_C - E_t) + \Delta E_\sigma^n$, but it does not give the energy level of the trap directly, its “binding energy”. Furthermore this identification holds only if $(E_C - E_t)$ is itself temperature independent. σ_∞^n is the apparent capture cross-section and as derived here, it is not equal to the capture cross-section in the temperature range of measurement but to the extrapolated value at $T = \infty$. This identification is also modified if $(E_C - E_t)$ is temperature dependent. The plot e_n/T^2 versus T^{-1} is called the “trap signature” and, although E_A^n and σ_∞^n cannot be interpreted immediately in terms of energy level and capture cross-section, their values can be used to characterize a trap in terms of its signature defined by Equation (3.20) and (3.21). These values of E_A^n and σ_∞^n are used to catalogue traps found in various materials.

A correct interpretation of E_A^n and σ_∞^n using thermodynamical concepts is not further discussed here, it can be found in [53]. For self-organized QDs, however, the temperature-dependence of the capture cross section in Equation (3.19) can be neglected. The activation energy $E_A^n = E_C - E_{QD}$ is the binding energy of the QD state with respect to the band edge, if no phonon-assisted tunneling occurs. This situation is further discussed in Section 3.2.

3.1.5 Majority and Minority Carrier Traps

From Equations (3.15) and (3.17) it is clear, that trap states in the upper half of the bandgap have $e_n > e_p$, since $(E_C - E_t) < (E_t - E_V)$. For trap states in the lower half of the band gap consequently holds the opposite, $e_p > e_n$. In the former case the deep level is referred to as an *electron trap* and in the latter as a *hole trap*. Without further going into detail (more details in Reference [53]), in an n-type semiconductor, an electron trap captures the predominant carrier type and is therefore commonly referred to as *majority carrier trap*, whereas a hole trap would be a *minority carrier trap*. In a p-type semiconductor, a hole trap must be referred to as a majority carrier trap and an electron trap as a minority carrier trap.

3.1.6 General Solution of the Rate Equation

The interpretation of the differential rate Equation (3.3) has been focused until now on the steady state situation in thermal equilibrium. This section focuses on the general solution in order to derive the time-dependent behavior of the trap population with charge carriers. To describe the transient behavior, determined by Equation (3.3), of the trap occupation from an arbitrary initial state towards the thermal equilibrium, the notation shall be simplified by introducing a and b as being the sum of the rates of electron gain and loss, respectively. Rewriting Equation (3.3) with $a = c_n + e_p$ and $b = e_n + c_p$ yields

$$\frac{dn_n(t)}{dt} = a [N_t - n_n(t)] - b n_n(t). \quad (3.22)$$

With the boundary condition $n_n = n_n(0)$ for $t = 0$, the general solution of Equation (3.22) is

$$n_n(t) = \frac{a}{a+b} N_t - \left(\frac{a}{a+b} N_t - n_n(0) \right) \exp[-(a+b)t]. \quad (3.23)$$

The steady state trap population $n_n(\infty)$ is given by setting $t = \infty$ and Equation (3.22) equal to zero:

$$n_n(\infty) = \frac{a}{a+b} N_t, \quad (3.24)$$

and in general one can write

$$n_n(t) = n_n(\infty) - [n_n(\infty) - n_n(0)] \exp(-t/\tau), \quad (3.25)$$

where τ is the time constant given by

$$\tau^{-1} = a + b = c_n + e_n + c_p + e_p. \quad (3.26)$$

There are two frequently encountered situations, for which Equation (3.23) simplifies to a certain extent. First, when the trap state is initially fully occupied

$$n_n(t) = \frac{a}{a+b} N_t - \frac{b}{a+b} N_t \exp[-(a+b)t], \quad \text{since } n_n(0) = N_t \quad (3.27)$$

and $n_n(t)$ decays, or when the state is initially empty

$$n_n(t) = \frac{a}{a+b} N_t (1 - \exp[-(a+b)t]), \quad \text{since } n_n(0) = 0 \quad (3.28)$$

and $n_n(t)$ increases with time.

Majority Carrier Traps

At first sight, the dynamics of the trap population seems rather complicated, since it generally depends on the four mechanisms of capture and emission of holes and electrons. The situation for majority carrier traps in a n- or p-doped semiconductor is, however, considerably simpler, since usually $e_n \gg e_p$ for an electron trap and $e_p \gg e_n$ for a hole trap (cf. Section 3.1.5), respectively. Therefore the rate equation (3.3) for an electron trap can be simplified to

$$\frac{dn_n(t)}{dt} = c_n [N_t - n_n(t)] - e_n n_n(t). \quad (3.29)$$

The time constant τ_n from the general solution, Equation (3.25), becomes:

$$\frac{1}{\tau_n} = e_n + c_n, \quad (3.30)$$

and the steady state population is reached for $t = \infty$

$$n_n(\infty) = N_t \left(1 + \frac{e_n}{c_n} \right)^{-1}. \quad (3.31)$$

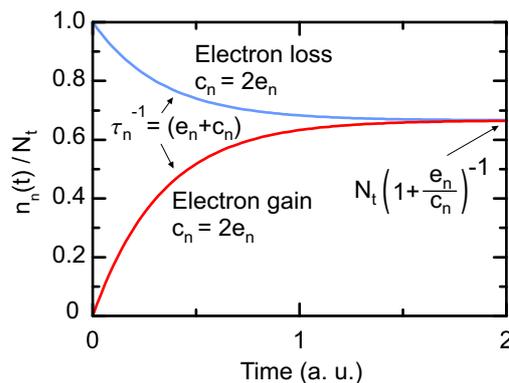


Figure 3.2: Relaxation of the average trap population with electrons $n_n(t)/N_t$ of an electron trap in an n-type material from full (blue line) and empty (red line) initial conditions to the equilibrium population. The time constant is the same in both cases.

The general solution of the rate equation shows, that if the population is momentarily perturbed from $n_n(0)$ to $n_n(\infty)$ it relaxes mono-exponentially to the steady state population with time constant $\tau_n = (e_n + c_n)^{-1}$. The time constant is the sum of the rates of all operative processes, both emission and capture. Whether the relaxation is an emission or capture process depends on whether $n_n(0)$ is greater (emission) or less (capture) than $n_n(\infty)$. The initial population of the trap is usually set by the experimental conditions, though there are two particular situations which are frequently encountered.

In Fig. 3.2, the relaxation process is depicted for the case that (i) all trap states are initially filled with electrons (blue line) and (ii) all trap states are initially empty (red line). In both cases $c_n = 2e_n$ was chosen and the average population $n_n(t)/N_t$ is shown time-dependent. For the initially filled situation, electron emission occurs with the time constant $\tau_n = (e_n + c_n)^{-1}$, whereas electron capture is observed with the same time constant for the initially empty trap situation. The average electron population strives in both cases to $2/3$ for $t = \infty$.

Initially Filled Majority Carrier Traps

The situation for majority carrier traps in the depletion region is even more simpler, as in addition the free carrier density vanishes in the depletion region leading to $c_n = 0$ or $c_p = 0$, respectively. It is furthermore assumed that the emitted carriers are swept out of the depletion region and no re-capturing can take place. Therefore, the transient population of an initially completely filled electron trap, $n_n(0) = 1$, in the depletion region of an n-type semiconductor is given by

$$\frac{dn_n(t)}{dt} = -e_n n_n(t) \quad (3.32)$$

and integration leads to

$$n_n(t) \sim \exp(-e_n t) \quad (3.33)$$

When thermal emission of electrons is the only operative process, the emission rate of electrons from the trap is given by Equation (3.33), which is assumed to be valid for deep levels [53]. The same holds for an analogous equation for hole emission.

3.2 Quantum Dots in Semiconductors

The carrier emission processes in QDs have certain similarities with the ones from deep levels. A localized state in a QD can be regarded as a “giant trap” and the thermal activation energy of the emission rate can be used to measure (with some limitations, which will be discussed later) the thermal energy separation between the confined state and the band edge, i. e. the binding or localization energy.² The formalism as previously developed for a bulk density of single-level traps, however, has to undergo some modifications to account for QDs in semiconductors.

The energy of a trap level E_t is now replaced by the energy of the QD state E_{QD} , either electron or hole state. It is assumed, that the carrier dynamics of a number of QDs with density N_{QD} is probed, where all QDs have the same energy spectrum, that means, no ensemble broadening due to fluctuations in size, shape and composition exist. Furthermore, only one QD state should be present and all figures are depicted for electrons in the conduction band.

3.2.1 Thermal emission rate

Coulomb Charging

The charge carrier accumulation in a QD leads to a rather complicated three-dimensional potential distribution. This results from Coulomb interaction between the charges inside the QD and in the bulk material. This causes a local band bending surrounding the dot, schematically depicted in Fig. 3.3(b) and labeled with E_{CC}^1 . Coulomb interaction of charge carriers inside the same QD leads to an intra-QD Coulomb energy term, which lifts the QD energy level with respect to the band edge. This term is labeled with E_{CC}^2 in Fig. 3.3(b). It is not obvious that the interaction strength between the charge located inside the QD and the carriers in the matrix E_{CC}^1 and the interaction strength between charge carriers inside the same dot E_{CC}^2 is identical. For a point defect this may be a valid assumption. In the case of QDs, the intra-QD carrier Coulomb interaction may be stronger. However, the Coulomb charging energy E_{CC}^2 inside the dot and the local band bending energy E_{CC}^1 are assumed to be equal, that means $E_{CC}^1 = E_{CC}^2 := E_{CC}$.

An estimate of the Coulomb charging energy E_{CC} can be obtained from the capacitance of a QD C_{QD} by

$$E_{CC} = \frac{e^2}{C_{QD}}. \quad (3.34)$$

²The name “localization energy” will now only be used for the binding energy of the confined electron/hole ground state with respect to the conduction/valence band edge.

The capacitance of a QD can be approximated by the well-known capacitance of a disk or sphere. Assuming a disk with radius r in a matrix material with the relative permittivity ϵ_r , the capacitance is given by

$$C_{QD}^{disk} = 8 \epsilon_0 \epsilon_r r, \quad (3.35)$$

and for a sphere with the same radius by

$$C_{QD}^{sphere} = 4\pi \epsilon_0 \epsilon_r r. \quad (3.36)$$

More realistic values for the Coulomb charging energies can be determined from self-consistent numerical calculations of many-particle effects in QDs [48, 84].

Assuming a disk with radius $r = 10$ nm a Coulomb Charging energy of 17 meV is obtained. For a sphere of the same radius one obtains 11 meV. This estimation is in agreement with experimentally observed Coulomb charging energies, e. g. Reference [85].

The local band bending E_{CC} decreases the density of majority charge carriers in the matrix material around the QD. In first approximation, the total band bending is linearly increased by the number of charge carriers inside the dot f : $E_{CC} = fE_\alpha$. Accordingly, the charge carrier density at the conduction band edge n_C is reduced by the factor resulting from the Boltzmann relation, since the Coulomb interaction provides a capture barrier, see Fig. 3.3(b):

$$n_{QD} = n_C \cdot \exp\left(-\frac{E_{CC}}{k_B T}\right), \quad (3.37)$$

where n_{QD} is here the density of electrons in the conduction band directly at the QD position. Therefore, Equation (3.1) for the capture rate has to be rewritten

$$c_n(T) = \sigma^n \langle v_n \rangle n_C \cdot \exp\left(-\frac{E_{CC}}{k_B T}\right). \quad (3.38)$$

The additional contribution E_{CC} to the electrochemical potential also modifies the Fermi-Dirac distribution in thermal equilibrium, Equation (3.5), to [54]

$$\frac{n_n^{QD}(\infty)}{N_{QD}} = \frac{1}{1 + \frac{g_0}{g_1} \exp\left(\frac{E_{QD} - E_F + E_{CC}}{k_B T}\right)}, \quad (3.39)$$

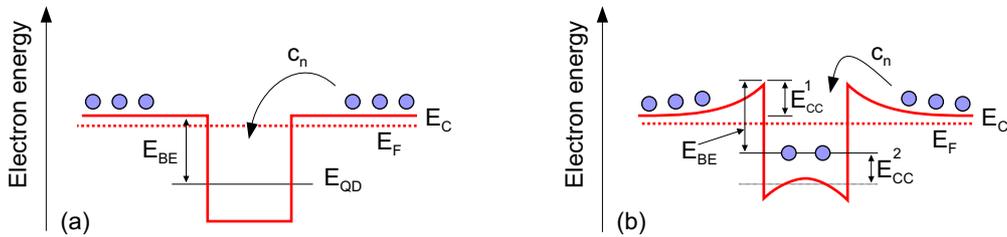


Figure 3.3: The occupation with charge carriers of a QD state with binding energy E_{BE} (a) leads to a local band bending E_{CC}^1 outside the dot (b). The QD states are also lifted by the Coulomb charging energy E_{CC}^2 .

where $n_n^{QD}(\infty)$ is the electron density inside the dot in thermal equilibrium, N_{QD} the QD sheet density and E_{QD} the energy of a certain QD state.

Substituting Equation (3.39) into Equation (3.4) yields:

$$e_n(T) = c_n(T) \frac{g_0}{g_1} \exp\left(\frac{E_{QD} + E_{CC} - E_F}{k_B T}\right). \quad (3.40)$$

Substituting Equation (3.38) for the capture rate $e_n(T)$, Equation (3.8) for the number of electrons in the conduction band n_C and Equation (3.13) for the thermal velocity into Equation (3.40) yields the electron emission rate for QDs:

$$e_n(T) = \gamma_n \frac{g_0}{g_1} \sigma^n T^2 \exp\left(-\frac{E_{BE}^n}{k_B T}\right), \quad (3.41)$$

where γ_n is the temperature-independent constant as defined in Equation (3.16) and $E_{BE}^n = E_C - E_{QD}$ the binding energy of the QD electron state E_{QD} with respect to the conduction band edge E_C , Fig. 3.3(a). The Coulomb charging energy E_{CC} equals zero, since it occurs in the Fermi distribution, Equation (3.39), and in the emission rate, Equation (3.40).

In contrast, the expression for the emission rate given by Equation (3.41) remains unchanged in comparison to the emission rate for deep levels. Only the capture rate c_n is really changed by the presence of a capture barrier due to the local band bending.

Analogously, the hole emission rate from a QD hole level E_{QD} with the binding energy $E_{BE}^p = E_{QD} - E_V$ is

$$e_p(T) = \gamma_p \frac{g_1}{g_0} \sigma^p T^2 \exp\left(-\frac{E_{BE}^p}{k_B T}\right), \quad (3.42)$$

where γ_p is the temperature-independent constant as defined in Equation (3.18).

Temperature-Dependent Capture Cross-Section

As already described for deep levels in Section 3.1.4, the capture cross-section for QDs $\sigma^n(T)$ is also temperature-dependent, since the underlying physical mechanism for charge carrier capture into QD states is a temperature dependent Auger [86–89] and/or phonon scattering process [19,90–94]. The capture cross-section can be described by an exponential dependence in a multiphonon capture model [83], yielding Equation (3.19). For QDs, substituting Equation (3.19) into Equation (3.41) results in

$$e_n(T) = \gamma_n T^2 \frac{g_0}{g_1} \sigma_\infty^n \exp\left(-\frac{E_{BE}^n + \Delta E_\sigma^n}{k_B T}\right), \quad (3.43)$$

where σ_∞^n is now the the apparent electron capture cross-section for $T = \infty$, which is temperature-independent. An analogous Equation can be obtained for the hole emission rate.

For deep levels, ΔE_σ can exhibit values up to several hundred meV [83], leading to an overestimation of the trap binding energy. For self-organized QDs, the temperature-dependent capture and relaxation time of electrons into InGaAs/GaAs QDs was determined by Müller et al. [19] by using interband-pump intraband-probe spectroscopy.

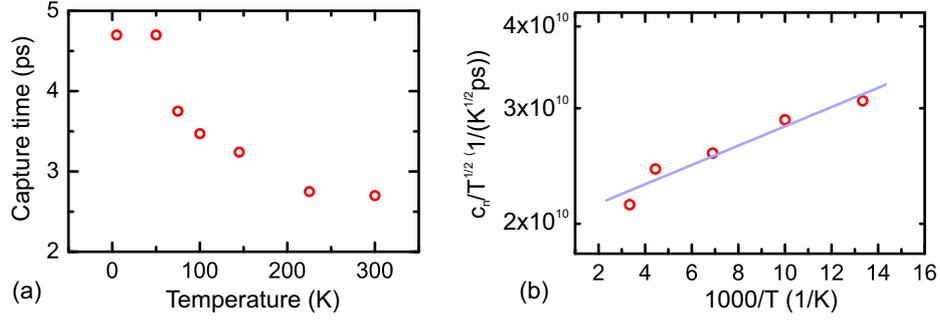


Figure 3.4: (a) Temperature dependence of the electron capture time, obtained by using interband-pump intraband-probe spectroscopy. After [19]. (b) Measured capture times divided by \sqrt{T} on a semilogarithmic scale versus inverse temperature. A linear fit to the data points (blue line) yields the temperature-dependent part of the capture cross-section: ΔE_σ^n .

The measured values are shown in Fig. 3.4(a). For an excitation density of 25 W/cm^2 , corresponding to an electron concentration of about 10^{16} cm^{-3} (cf. Engström et al. [95]), the capture time constant c_n was found to be 4.7 ps at 5 K (decreasing down to 2.7 ps at 300 K). This measurement can be used here to estimate the temperature-dependence of the capture cross-section in QDs.

Substituting the temperature-dependent capture cross-section $\sigma^n(T)$ in Equation (3.19) into the capture rate Equation (3.1) and using a constant electron concentration at the conduction band edge n_C^{const} yields

$$c_n(T) = \sigma_\infty^n \sqrt{\frac{3k_B T}{m_e^*}} n_C^{const} \cdot \exp\left(-\frac{\Delta E_\sigma^n}{k_B T}\right). \quad (3.44)$$

A logarithmic plot of the capture rate divided by the square-root of the temperature $\ln(c_n/\sqrt{T})$ as function of the inverse temperature T^{-1} is shown in Fig. 3.4(b). A linear fit to the data for $T > 70$ K can be used to estimate roughly the temperature-dependence of the capture process into QDs and, hence, the value for the temperature-dependence of the capture cross-section. The fit yields a value for ΔE_σ of about 3 meV. For $T < 70$ K no linear dependence is observed anymore, however, the slope is even smaller, that means the temperature-dependence of the capture process is also smaller for temperatures below 70 K. Therefore, ΔE_σ can be neglected in Equation (3.43) for the determination of the binding energy E_{BE}^n in QDs and has only a minor influence on the temperature-dependence of the capture cross-section. In the following, the apparent capture cross-section σ_∞ at infinite temperature will be used.

3.2.2 Electric Field Enhanced Emission

If the QDs are placed inside a depletion region of a p-n diode structure, as it is the case for capacitance spectroscopy measurements, the effect of the electric field across the depletion region has to be taken into account. The field induced band tilt provides an effective capture barrier with energy E_{Cap} . Figure 3.5 displays the situation

for electrons in a n-type semiconductor with the Coulomb charging energy E_{CC} . The energy levels of the QDs are lifted by the sum of the capture barrier and the Coulomb interaction: $E_{Cap} + E_{CC}$. The capture barrier is also enhanced by the Coulomb interaction E_{CC} . Both emission and capture barrier are reduced equally by the Poole-Frenkel effect δE , already described in Section 2.2.3.

The influence of the total capture barrier $E_{Cap} + E_{CC} - \delta E$ on the electron density at the QDs is given by the Boltzmann relation:

$$n_{QD} = n_C \cdot \exp\left(-\frac{(E_{Cap} + E_{CC} - \delta E)}{k_B T}\right). \quad (3.45)$$

Therefore, the capture rate for electrons is

$$c_n = \sigma_\infty^n \langle v \rangle n_C \cdot \exp\left(-\frac{(E_{Cap} + E_{CC} - \delta E)}{k_B T}\right). \quad (3.46)$$

The energy terms related to the band bending also change the Fermi-Dirac function to

$$\frac{n_n^{QD}(\infty)}{N_{QD}} = \frac{1}{1 + \frac{g_0}{g_1} \exp\left(\frac{E_{QD} + E_{CC} + E_{Cap} - E_F}{k_B T}\right)}, \quad (3.47)$$

where $n_n^{QD}(\infty)$ is the density of QD states filled with electrons in thermal equilibrium and N_{QD} the QD sheet density.

Substituting Equation (3.47) into Equation (3.4) yields the thermal emission rate for electrons in an applied electric field F :

$$e_n^F(T) = \gamma_n \frac{g_0}{g_1} \sigma_\infty^n T^2 \exp\left(-\frac{E_A^n}{k_B T}\right), \quad (3.48)$$

where E_A^n is the thermal activation energy of a certain QD electron state with energy E_{QD} . Which emission process from which QD state is probed depends on the experimental situation in the deep level transient spectroscopy (DLTS) measurements,

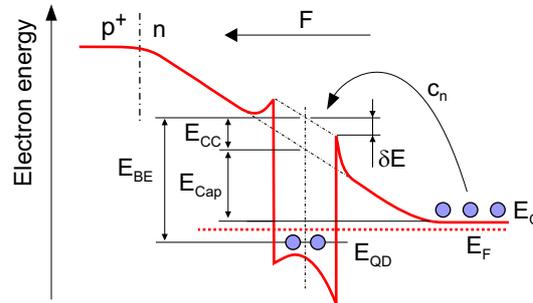


Figure 3.5: Schematic bandstructure for the situation that the QDs are inside a depletion region of a p-n diode. The QD states are energetically lifted by the Coulomb charging energy and the band bending: $E_{CC} + E_{Cap}$. In addition, the Coulomb charging and the depletion region provide a potential barrier for charge carrier capture c_n , which is reduced by the value of the Poole-Frenkel energy: $E_{CC} + E_B - \delta E$.

cf. Chapter 4. E_A^n is, hence, the binding energy of the QD state ($E_C - E_{QD}$) reduced by the Poole-Frenkel energy δE : $E_A^n = (E_C - E_{QD}) - \delta E$.

Analogously, the emission rate for holes as majority charge carriers yields:

$$e_p^F(T) = \gamma_p \frac{g_1}{g_0} \sigma_\infty^p T^2 \exp\left(-\frac{E_A^p}{k_B T}\right), \quad (3.49)$$

where E_A^p is the thermal activation energy of a certain QD hole state with energy E_{QD} . Here, E_A^p is the binding energy of the hole state reduced by the Poole-Frenkel energy δE : $E_A^p = (E_{QD} - E_V) - \delta E$.

In capacitance spectroscopy measurements of structures with a band bending due to the p-n interface, the energy obtained by an Arrhenius plot can be interpreted as the binding energy of the charge carrier minus the Poole-Frenkel energy. Since the Poole-Frenkel energy has only a minor impact, the charge carrier binding energy of a certain QD state with respect to the conduction/valence band edge can be obtained. However, it is very important to notice here, that this conclusion is only valid for the case, that tunneling emission is negligible. Otherwise phonon-assisted tunneling occurs, which reduces the observed activation energy (see next section).

Pure Tunneling and Phonon-Assisted Tunneling

The pure tunneling rate e^T in an electric field F for a QD with an elongation in z -direction L and the height E_B of the triangular barrier, Fig. 3.6(a), can be described by an exponential dependence, cf. Equation (2.8):

$$e^T = \frac{\hbar\pi}{2m^*L^2} \exp\left(-\frac{4}{3}\sqrt{\frac{2m^*}{\hbar^2}} \frac{\Delta E_B^{3/2}}{eF}\right). \quad (3.50)$$

However, for QDs in an electric field the charge carrier emission process is in most cases a combination of thermal activation and tunneling. Therefore, the activation energies E_A obtained from capacitance spectroscopy measurements do not represent the binding energies of the QD states with respect to the corresponding band edge, as the underlying physical mechanism is a phonon-assisted tunneling process [62,80]: Thermal activation from a QD state into an intermediate state and subsequent tunneling through the remaining triangular barrier E_B , schematically depicted in Fig. 3.6(b). The total emission rate depends on the applied electric field [62] and the rate equation has to be multiplied by the transparency factor of a triangular tunneling barrier [56]. The phonon-assisted tunneling rate for electrons is then given by:³

$$e^P(T, F) = \gamma_n \frac{g_0}{g_1} T^2 \sigma_{T,F=\infty}^n \cdot \exp\left(-\frac{E_A^n}{k_B T}\right) \cdot \exp\left(\frac{-4}{3\hbar F e} \sqrt{2m^* E_B^3}\right), \quad (3.51)$$

where m^* is the effective mass in the matrix and F the applied electric field. $\sigma_{T,F=\infty}^n$ is the apparent capture cross-section for infinite high temperature and infinite high electric field (see below).

³Here, the capture cross-section for QDs has no temperature-dependence anymore. This is in contrast to Equation (2.10), where the same Equation is shown for the general situation of a temperature-dependent capture cross-section σ^n .

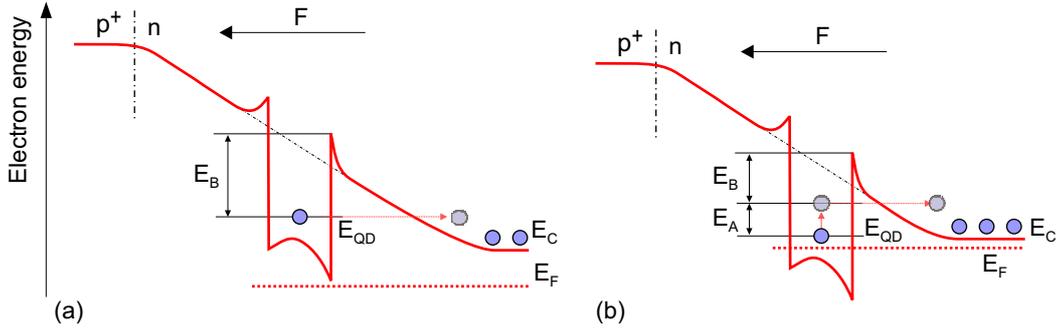


Figure 3.6: Carrier emission due to tunneling (a) and phonon-assisted tunneling (b). Phonon-assisted tunneling is a combination process of thermal activation E_A from initial state E_{QD} into a higher QD state and subsequent tunneling through the remaining triangular barrier E_B .

In capacitance spectroscopy measurements, the activation energy E_A can be separated from the phonon-assisted tunneling process by an Arrhenius plot, where $\ln(e^P/T^2)$ is plotted as function of T^{-1} . The slope yields the activation energy E_A . From the intersection with the y-axis the apparent capture cross-section $\sigma_{T=\infty}$ for infinite high temperature can be deduced, which still depends on the applied electric field F , cf. Equation (3.51):

$$\sigma_{\infty}^n(F) = \sigma_{T,F=\infty}^n \cdot \exp\left(\frac{-4}{3\hbar F e} \sqrt{2m^* E_B^3}\right). \quad (3.52)$$

That means, the capture process is changed by an electric field leading to an electric-field dependent capture cross-section. Since the electric field in a p-n diode structure is related to a depletion region with a band bending (see Fig. 3.5) the capture cross-section σ_{∞}^n describes a capture process with a capture barrier.

As reported by Chang et al. [62], plotting the apparent capture cross section σ_{∞} as function of $1/F$ can be used to determine the tunneling barrier height of the intermediate state E_B . In addition, an electric field independent apparent capture cross-section $\sigma_{T,F=\infty}$ is obtained. This value describes the capture process for infinite high temperature and infinite high electric field – that means, with no temperature and electric-field dependence. This equals to a carrier capture process for flat-band condition without a capture barrier, like in Fig. 3.3(a). More details can be found in Chapter 5, where the hole capture into QDs is investigated by using capacitance spectroscopy.

3.2.3 Influence of a Barrier Below the Quantum Dots

In this section, the influence of a barrier below the QD layer is discussed. Figure 3.7 shows schematically the conduction band structure including Coulomb interaction and the band bending. The QD energy levels are lifted by $E_{Cap} + E_{CC}$, while the emission barrier is enhanced by the barrier height and lowered by the Poole-Frenkel energy: $E_{Bar} - \delta E$. The capture barrier has the energy: $E_{Cap} + E_{CC} + E_{Bar} - \delta E$. The energy

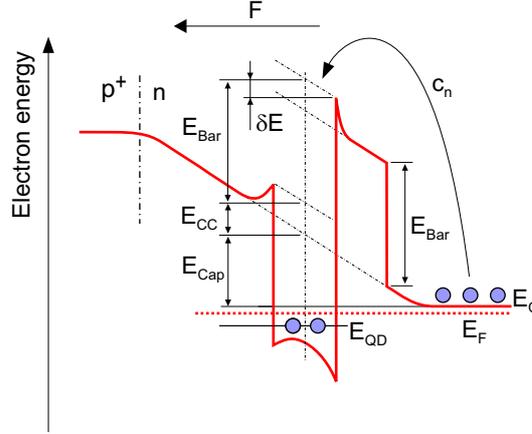


Figure 3.7: Schematic bandstructure for the situation, that the QDs are inside a depletion region of a p-n diode and a barrier is situated directly below the QD layer. The QD states are energetically lifted by the Coulomb charging energy and the band bending: $E_{CC} + E_{Cap}$. The emission barrier is enhanced by the barrier height E_{Bar} and slightly reduced by the Poole-Frenkel energy δE . The entire capture barrier is $E_{Cap} + E_{CC} + E_{Bar} - \delta E$.

of the QD state E_{QD} is not influenced by the barrier E_{Bar} , hence, the Fermi-Dirac distribution, given in Equation (3.47), can be used in the following.

Analogously to the previous Section 3.2.2 the capture rate is:

$$c_n = \sigma_{\infty}^n \langle v \rangle n_C \cdot \exp \left(- \frac{(E_{Cap} + E_{CC} + E_{Bar} - \delta E)}{k_B T} \right). \quad (3.53)$$

and the electron emission rate e_n with a barrier is given by:

$$e_n^{Bar}(T) = \gamma_n \frac{g_0}{g_1} \sigma_{\infty}^n T^2 \exp \left(- \frac{E_A^n}{k_B T} \right). \quad (3.54)$$

where E_A^n is again the thermal activation energy for electron emission from a certain QD electron state with energy E_{QD} over the barrier E_{Bar} . E_A^n is the binding energy of the QD electron state ($E_C - E_{QD}$) plus the barrier height E_{Bar} and reduced by the Poole-Frenkel energy δE : $E_A^n = (E_C - E_{QD}) + E_{Bar} - \delta E$.

Analogously, the emission rate for holes as majority charge carriers yields:

$$e_p^{Bar}(T) = \gamma_p \frac{g_1}{g_0} \sigma_{\infty}^p T^2 \exp \left(- \frac{E_A^p}{k_B T} \right), \quad (3.55)$$

where E_A^p is the thermal activation energy for thermal hole emission from a certain QD hole state E_{QD} over the barrier E_{Bar} . E_A^p is analogously to the electron situation given by: $E_A^p = (E_{QD}^p - E_V) + E_{Bar} - \delta E$.

Therefore, from an Arrhenius plot one obtains a thermal activation energy of the QD charge carriers that represents the binding energy E_{BE} plus the barrier height E_{Bar} (and minus the Poole-Frenkel energy). In addition, it is possible to determine the barrier height E_{Bar} only, if the QDs are completely charged up to the band edge and the emission over the barrier is observed, see Section 5.4.

3.2.4 Charge Carrier Capture

To analyze the capture process into QDs by using capacitance spectroscopy, presented in Chapter 6, a temperature-dependent rate equation for carrier capture is needed. The electron capture rate, Equation (3.1), for a temperature-independent capture cross section σ_∞^n and an electron density at the QD position n_{QD} is:

$$c_n = \sigma_\infty^n \langle v \rangle n_{QD}. \quad (3.56)$$

In a capacitance measurement of the capture process, the electron density at the QD position is reduced by a thermal activation barrier $E_A^c = E_{Cap} + E_{CC} - \delta E$, schematically displayed in Fig. 3.5. Therefore, the electron density at the QD position is given by:

$$n_{QD} = n_C \cdot \exp\left(-\frac{E_A^n}{k_B T}\right), \quad (3.57)$$

where $n_C = N_C \cdot \exp[(E_C - E_F)/k_B T]$ is the charge carrier density at the conduction band edge. An equation for the thermal capture rate with an activation barrier $E_{A,Capt}^n$ can be obtained using Equation (3.14):

$$c_n(T) = \gamma_n T^2 \frac{g_0}{g_1} \sigma_\infty^n \cdot \exp\left(-\frac{E_{A,Capt}^n}{k_B T}\right), \quad (3.58)$$

where $E_{A,Capt}^n$ is the thermal activation energy for electron capture over a capture barrier into a certain QD electron state: $E_{A,Capt}^n = (E_C - E_F) + E_{Cap} + E_{CC} - \delta E$. The capture barrier height is given by the experimental condition and equals the emission barrier inside the QD except the term $(E_C - E_F)$, which can be neglected (see below).

For holes as majority charge carriers, an analogous equation can be obtained:

$$c_p(T) = \gamma_n T^2 \frac{g_1}{g_0} \sigma_\infty^p \cdot \exp\left(-\frac{E_{A,Capt}^p}{k_B T}\right). \quad (3.59)$$

where $E_{A,Capt}^p$ is the thermal activation energy for hole capture: $E_{A,Capt}^p = (E_F - E_V) + E_{Cap} + E_{CC} - \delta E$.

The difference between the rate equation for emission and capture lies within the additional term for the energy difference between the Fermi level and the valence or conduction band edge, $(E_F - E_V)$ or $(E_C - E_F)$, respectively.⁴

⁴However, in the temperature range of the capacitance measurements (~ 50 K) on hole capture in Chapter 6, it is in the order of 10 meV for a doping concentration of $3 \times 10^{16} \text{ cm}^{-3}$ [82]. Therefore, this term is neglected and from Arrhenius plots activation energies $E_{A,Capt}$ and apparent capture cross-sections σ_∞ can also be obtained from capacitance measurements of carrier capture for this experimental situation.

“IT DOESN’T MATTER HOW BEAUTIFUL YOUR THEORY IS, IT DOESN’T MATTER HOW SMART YOU ARE. IF IT DOESN’T AGREE WITH EXPERIMENT, IT’S WRONG”
RICHARD FEYNMAN

Chapter 4

Capacitance Spectroscopy

Within this thesis, the charge carrier emission and capture are investigated with capacitance spectroscopy methods, like static capacitance-voltage profiling and time-resolved capacitance spectroscopy. Section 4.1 summarizes the basic principles of a capacitance associated with the depletion region (the depletion capacitance) of a Schottky and p-n junction. The capacitance-voltage (C-V) profiling and the deep level transient spectroscopy (DLTS) are presented in Section 4.2 and 4.3, respectively. Two improved methods to study QDs in particular were developed during this work and will be presented in Section 4.3: The charge-selective DLTS (Section 4.3.2) and the time-resolved tunneling capacitance measurement (TRTCM) (Section 4.3.3).

4.1 Depletion Region

The Fermi energy with respect to the vacuum level in a metal or doped semiconductor is not the same for different metals or semiconductors having different doping concentrations. If a metal and a semiconductor (or two differently doped semiconductors) are in electric contact, the free charge carriers are exchanged between the different materials until a thermodynamic equilibrium is reached and the Fermi energy is equal throughout the entire structure. As a consequence, ionized donors and acceptors are present in the vicinity of the interface while all free charge carriers are moved away, Fig. 4.1. This layer depleted from free charge carriers is usually referred to as “depletion region” [53,82]. The width of the depletion region depends on the doping concentration and the potential difference between the materials, of which the latter can easily be modified by an externally applied bias.

A metal-semiconductor contact can be described by the Schottky model and is referred to as “Schottky contact”. A junction of a p-doped and n-doped semiconductor is a “p-n junction”. Both types of contacts provide a depletion region and are briefly described in the following.

4.1.1 Schottky Contact

A metal-semiconductor contact is usually described in the framework of the Schottky model, which is an acceptable approach to construct the band diagram of the contact,

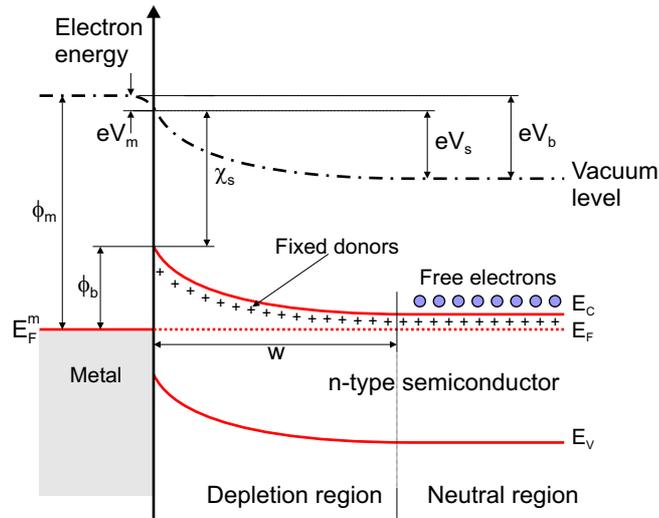


Figure 4.1: Energy band diagram of a metal-semiconductor Schottky contact. The width of the depletion region w depends on the doping concentration in the semiconductor. A built-in voltage V_b is present without external bias due to the band bending of the depletion region.

Fig. 4.1. Within this model, the barrier height ϕ_b inside the metal is independent or just weakly dependent on the applied external bias.

According to the Schottky model, the energy band diagram is constructed by reference to the vacuum level, defined as the energy of a free electron to rest outside the material. The work function of the metal ϕ_m and the electron affinity of the semiconductor χ_s are defined as the energies required to remove an electron from the metal's Fermi level or semiconductor conduction band edge, respectively, to the vacuum level. These properties are assumed to be constant in a given material and it is further assumed that the vacuum level is continuous across the interface. The Fermi levels in the metal and semiconductor must be equal in thermal equilibrium. These conditions result in a band diagram for the interface as shown in Fig. 4.1.

Since the vacuum level is the same at the interface of the metal and the semiconductor, a step between the Fermi level in the metal and the conduction band edge E_C of the semiconductor occurs. This is the barrier height ϕ_b given by¹

$$\phi_b = \phi_m - \chi_s - eV_m. \quad (4.1)$$

The band bending in the metal is very small due to the large density of electron states, hence, eV_m can be neglected. Therefore, the Schottky barrier height is usually written as

$$\phi_b = \phi_m - \chi_s. \quad (4.2)$$

For increasing distance from the interface, the conduction band energy decreases. At the end of the depletion region it has the same value as in the neutral semiconductor

¹For a Schottky barrier ϕ_m must exceed χ_s otherwise the bands bend in the opposite direction.

with respect to the Fermi level. The resulting band bending is an effect of the removed free charge carriers, leaving behind a distribution of fixed positive charges from ionized donors. The depletion region ends at that position, where the bands become flat and the associated electric field is zero. The width of the depletion region w is determined by the net ionized charge density according to Poisson's equation (see Section 4.1.3).

Since the density of states in the metal is much greater than the doping density in the semiconductor, the depletion width in the metal is much smaller. Therefore, it can be assumed that the potential difference across the metal near the contact (V_m) is negligibly small compared to that in the semiconductor (V_s). The total zero bias band bending of the Schottky contact, also referred to as "build-in potential" or "build-in voltage" V_b can be written as

$$eV_b \approx eV_s = \phi_m - \chi_s - (E_C - E_F) = \phi_b - (E_C - E_F) \quad (4.3)$$

Experimental values for various metal Schottky contact on GaAs can be found in [96].

4.1.2 p-n Junction

The band diagram of an abrupt p-n junction, shown in Fig. 4.2, is considered in a similar manner. Again, two rules are used to construct the diagram: (i) The vacuum level is continuous and (ii) the Fermi energy is constant across the junction in thermal equilibrium. For the same p- and n-doped semiconductor the electron affinity χ_s is the same on both sides of the junction. Therefore, the band bending is caused entirely by the difference in the Fermi level with respect to the conduction band of the two differently doped materials. From Fig. 4.2, using subscripts to denote the n and p side, one obtains for zero external bias

$$(E_C^p - E_F) + \chi_s = eV_b + \chi_s + (E_C^n - E_F), \quad (4.4)$$

hence, the build-in voltage is

$$eV_b = E_g - (E_F - E_V^p) - (E_C^n - E_F), \quad (4.5)$$

where E_g is the energy gap of the semiconductor. The build-in voltage as the total electrostatic potential difference between the p-side and the n-side is temperature-dependent and a function of the fixed donor and acceptor charges of density N_d and N_a [82]:

$$V_b = \frac{k_B T}{e} \ln \left(\frac{N_a N_d}{n_i^2} \right), \quad (4.6)$$

where n_i is the intrinsic carrier density

$$n_i = \sqrt{N_C N_V} \exp \left(-\frac{E_g}{2k_B T} \right), \quad (4.7)$$

with the effective density of states in the valence N_V and conduction band N_C , respectively, defined in Equation (3.14). A typical doping density of 10^{17} cm^{-3} yields for GaAs a build-in voltage of 1.3 V at 300 K.

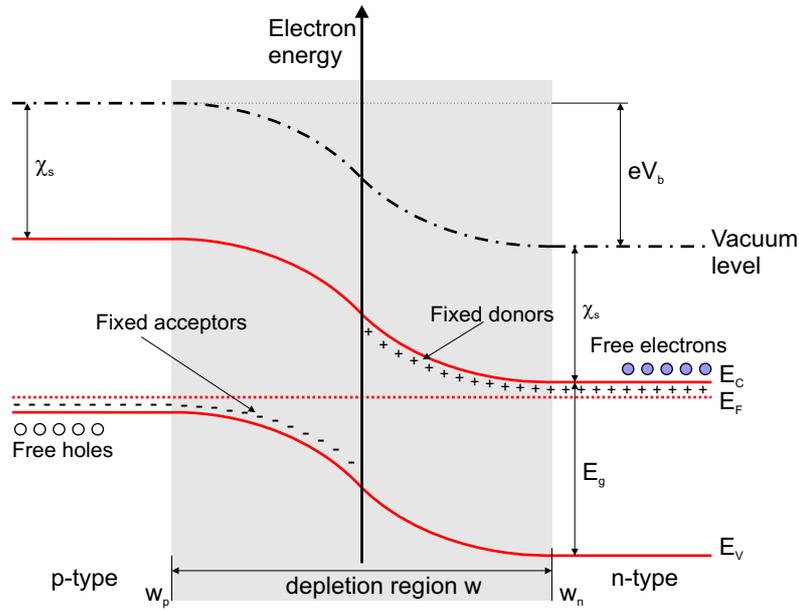


Figure 4.2: Energy band diagram of a p-n junction.

In a p-n junction, depletion regions on each side of the contact exist, where the fixed donor and acceptor charges lead to the band bending. Since the total charge in the p-n diode is zero (charge neutrality), in the depletion approximation [53] of an abrupt depletion layer edge

$$N_a w_p = N_d w_n \quad (4.8)$$

must hold. For similar doping concentrations, the depletion width on the p- and n-side of the semiconductor will be comparable. However, for the purpose of material characterization (such as capacitance experiments) the doping concentrations are often chosen, such that the depletion region is situated almost entirely on one side of the junction. The depletion region of such an asymmetrical p-n junction resembles almost the depletion region of a Schottky contact. Such asymmetrical junctions are briefly denoted as p⁺-n or n⁺-p junction with $N_a \gg N_d$ or $N_d \gg N_a$, respectively.

4.1.3 Width of the Depletion Region

The entire band bending across the depletion region is defined by the sum of the built-in voltage of the contact V_b , Equation (4.5) or (4.6), and the applied external bias V_a in reverse direction: $V = V_b + V_a$. The width of the depletion region and the electric field can be calculated using Poisson's equation. The electrostatic potential ψ at any point is given by

$$-\frac{\partial^2 \psi}{\partial x^2} = \frac{\partial F}{\partial x} = \frac{\rho(x)}{\epsilon \epsilon_0}, \quad (4.9)$$

where F is the electric field, ϵ_0 the vacuum permittivity, and ϵ the dielectric constant of the semiconductor material. If the donors or acceptors are entirely ionized, the charge density is eN_d or eN_a , respectively:

$$-\frac{\partial^2\psi}{\partial x^2} = \frac{eN_d}{\epsilon\epsilon_0} \quad \text{for } 0 \leq x \leq w_n. \quad (4.10)$$

For a Schottky contact or an abrupt asymmetric p⁺-n junction, integration of Equation (4.9) yields

$$F(x) = F_0 + \frac{eN_dx}{\epsilon\epsilon_0} \quad \text{for } 0 \leq x \leq w_n, \quad (4.11)$$

while $F(x) = 0$ for $x < 0$ and $x > w_n$. Hence, using the approximation $F(x) = 0$ at the edge of the depletion region, gives the boundary condition for the integration constant:

$$F_0 = -\frac{eN_dw_n}{\epsilon\epsilon_0}, \quad (4.12)$$

and represents the electric field at the interface $F(0)$, where it has its maximum. Therefore, the electric field F across the depletion region in the n-doped semiconductor with the donor concentration N_d is

$$F(x) = \frac{eN_d}{\epsilon\epsilon_0}(x - w_n) \quad \text{for } 0 \leq x \leq w_n. \quad (4.13)$$

Analogously, the electric field distribution across the depletion region on the p-doped side is:

$$F(x) = -\frac{eN_a}{\epsilon\epsilon_0}(x + w_p) \quad \text{for } -w_p \leq x \leq 0. \quad (4.14)$$

The potential distribution across the depletion region for a p⁺-n junction inside the n-doped semiconductor is now obtained by integration of Equation (4.13):

$$\psi(x) = -\int_0^x \frac{eN_d}{\epsilon\epsilon_0}(x - w_n) dx = \frac{eN_dw_n}{\epsilon\epsilon_0} \left(x - \frac{x^2}{2w_n} \right) + \psi(0), \quad (4.15)$$

with $\psi(0) = 0$ as a reference for the potential distribution. For an asymmetric p⁺-n junction, one notices that the band bending in the present approximation only occurs in the n-doped region of the junction. The contact potential is equal to the total band bending, that means, to the built-in voltage plus the external bias $\psi(0) = -V = -(V_b + V_a)$. Since the potential on the edge of the depletion region is zero, $\psi(w_n) = 0$, one obtains

$$V = V_b + V_a = \frac{eN_d}{2\epsilon\epsilon_0} w_n^2, \quad (4.16)$$

or the expression for the depletion width

$$w_n = \sqrt{\frac{2\epsilon\epsilon_0}{eN_d} V}. \quad (4.17)$$

If the p-n junction is symmetric and the depletion region extends into the p- and n-doped semiconductor, Equation (4.16) is modified to

$$V = V_b + V_a = \frac{eN_d}{2\epsilon\epsilon_0}w_n^2 + \frac{eN_a}{2\epsilon\epsilon_0}w_p^2, \quad (4.18)$$

with the net acceptor doping concentration N_a and the width of the depletion region w_p in the p-doped side. The total depletion width w is given by:

$$w = (w_n + w_p) = \sqrt{\frac{2\epsilon\epsilon_0}{e} \left(\frac{N_a + N_d}{N_a N_d} \right) V}. \quad (4.19)$$

The assumption of an abrupt depletion layer edge, the so-called “depletion approximation”, is not generally true. A more general description of the depletion region can be found in Reference [53]; it is not further discussed here.

4.1.4 Depletion Layer Capacitance

The capacitance which arises from the depletion region is the so-called “depletion capacitance”. In comparison to a plate capacitor, here, the charge is accumulated inside the depletion region of the p-n diode. The number of ionized donors depends linearly on the width w_n , while again the width is a function of the square root of the voltage, Equation (4.17) and (4.19). Hence, the capacitance does not depend linearly on the voltage and has to be defined differentially for a small bias signal ΔV :

$$C = \lim_{\Delta V \rightarrow 0} \frac{\Delta Q}{\Delta V} = \frac{dQ}{dV}. \quad (4.20)$$

Using Equation (4.17), the entire stored charge inside the depletion region of a Schottky diode or an abrupt asymmetric p⁺-n junction with area A is:

$$Q = eN_d A \cdot w_n(V) = A\sqrt{2e\epsilon\epsilon_0 N_d V}, \quad (4.21)$$

and the capacitance is

$$C(V) = \left. \frac{dQ}{dV} \right|_V = A\sqrt{\frac{\epsilon\epsilon_0 e N_d}{2V}} = \frac{\epsilon\epsilon_0 A}{w_n}. \quad (4.22)$$

The depletion capacitance resembles the capacitance of a plate capacitor with a distance of w_n between the plates and a dielectric with relative permittivity ϵ , although the charge is actually stored in the volume rather than on the edges of the space charge region.

4.2 Static Capacitance-Voltage Profiling

In the previous section, it was shown that the capacitance depends on the width of the depletion region, which itself depends on the doping concentration and the applied

bias. Therefore, capacitance-voltage measurements (C-V) can be used to measure depth profiles of the doping concentration in semiconductor devices.

Rewriting Equation (4.22) and using Equation (4.17) yields:

$$\frac{1}{C^2} = \frac{2V}{A^2 \epsilon \epsilon_0 e N_d}. \quad (4.23)$$

It is obvious, that plotting $1/C^2$ versus V yields a straight line for constant doping concentration. Therefore, the doping concentration N_d can be obtained from a C-V measurement:

$$N_d = \frac{2}{e \epsilon \epsilon_0 A^2} \left(\frac{d(1/C^2)}{dV} \right)^{-1}. \quad (4.24)$$

The applied bias for which $1/C^2 = 0$ corresponds to the build-in voltage.

Even for a varying doping density Equation (4.22) remains valid, since in the depletion layer approximation the charge fluctuations occur only at the depletion layer edge. With help of Equation (4.22) a doping-voltage profile $N_d(V)$ and with

$$w_n = \frac{\epsilon \epsilon_0 A}{C(V)}, \quad (4.25)$$

a depth profile of the doping density $N_d(w_n)$ can be determined from measuring the capacitance as function of the applied voltage $V = V_b + V_a$.

4.3 Capacitance Transient Spectroscopy

As already mentioned, the depletion width depends on the applied voltage and the doping concentration, i. e. on the charge stored inside the space charge region. As a consequence, the measured capacitance is sensitive to charge carrier population inside deep levels or QDs situated at some position inside the depletion region [97]. The carrier dynamics of QDs or deep levels can be studied by means of a time-resolved measurement of the capacitance of a Schottky or p-n diode.

Historically, this time-resolved capacitance measurement was initially used to study and characterize deep levels caused by impurities or defects. In this context it is usually referred to as ‘‘Deep Level Transient Spectroscopy’’ (DLTS) or ‘‘Capacitance Transient Spectroscopy’’ [52, 98–102]. Besides the determination of activation energies and capture cross-sections, DLTS also permits to obtain a depth profile of the trap concentration density and the investigation of the influence of the electric field on the emission process. As QDs act more or less as deep levels, the DLTS has been used successfully to study thermal emission processes, activation energies, and capture cross-sections of various QD systems [66, 67, 79, 80, 103–105].

4.3.1 Measurement Principle

In the following, the DLTS measurement principle will be described for QDs in a p-n or Schottky diode structure. A detailed description concerning deep levels can be found in [53] and references therein.

First of all, a single layer of QDs with density N_{QD} per area in an n-doped material (with doping concentration N_d) shall be considered. The work cycle of an DLTS experiment of a p^+ -n structure with QDs embedded is depicted in Fig. 4.3. The upper row displays schematically the p-n diode while the lower row depicts the corresponding potential distribution of the conduction band.

In an initial step, Fig. 4.3(a), the reverse bias V_r is chosen, such that the depletion region extends well over the QDs. The QDs are completely depleted from charge carriers and the Fermi level is below the QD levels in this situation. During the pulse V_p , the depletion region is shorter than the distance of the QD layer from the p-n interface and the QDs are consequently filled with carriers – electrons in Fig. 4.3(b). The Fermi level is now above the QD states. After switching back to the reverse bias situation, Fig. 4.3(c), the QDs are again situated inside the depletion region but still filled with electrons. The emission of carriers from QD states is usually slow, whereas the free carriers in the matrix material are considered to follow the change in the external bias instantaneously at the time-scale of the experiment. The depletion width is larger for QDs filled with charge carriers, hence, the capacitance after the pulse V_p is smaller and increases again when electron/holes are emitted due to thermal activation or tunneling emission. By recording time-resolved the depletion capacitance, transients are observed, which represent the carrier emission processes from the QD states. The pulse sequence during the experiment and the measured capacitance is schematically depicted in Fig. 4.4(a).

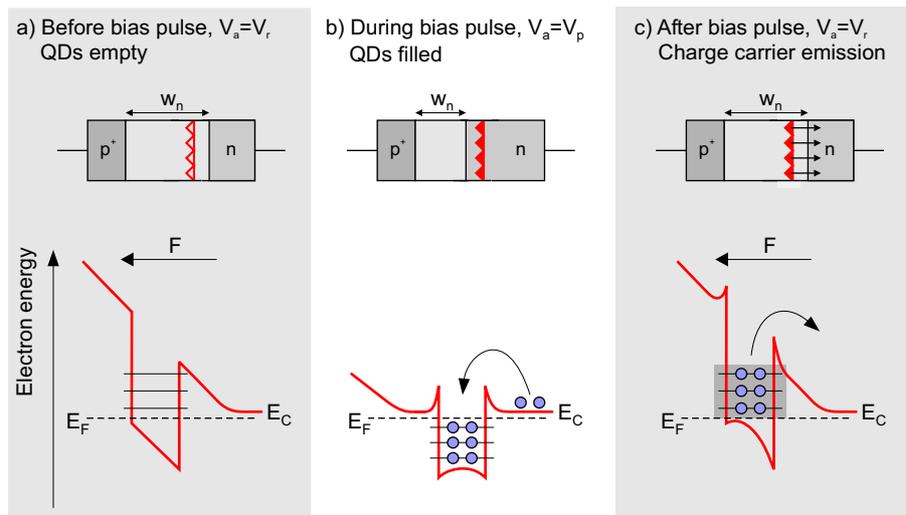


Figure 4.3: DLTS work cycle of a p^+ -n diode with embedded QDs. The upper row shows schematically sketches of the devices for three different bias situations: (a) Before, (b) during and (c) after the bias pulse, respectively. The lower row shows the corresponding band diagrams.

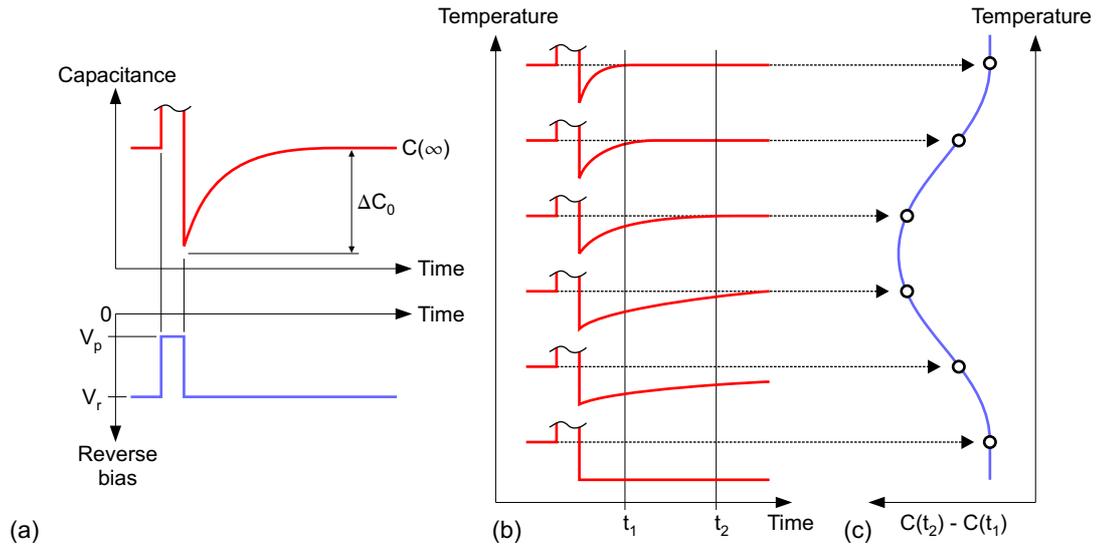


Figure 4.4: DLTS work cycle during a DLTS experiment. The lower part of (a) displays the external bias on the device as function of time, the upper part the corresponding capacitance. The evaluation of capacitance transients (b) for increasing temperature by a rate window, defined by t_1 and t_2 , leads to a DLTS plot of a thermally activated emission process (c).

Rate Window and Double-Boxcar Method

The emission time constants e_n or e_p are obtained by evaluation of the capacitance transient. The time constant is needed to determine activation energies and capture cross-sections, described in Section 3.2.

The capacitance transient, recorded from a single emission process is usually mono-exponential having the time-constant τ . Hence, the capacitance transient is given by [53]

$$C(t) = C(\infty) - \Delta C_0 \exp\left(-\frac{t}{\tau}\right), \quad (4.26)$$

where $C(\infty)$ is the steady state capacitance at V_r and ΔC_0 the entire change in the capacitance $C(t)$ for $t = \infty$, see Fig. 4.4. The emission time constant could be determined by using Equation (4.26) and plotting the transient data on a semi-logarithmic scale. However, in general the capacitance transient for deep levels and QDs are multi-exponential due to the ensemble broadening [106] and a linear fit is consequently in most cases impossible.

In order to obtain the emission time constant, the activation energies and capture cross-section from multi-exponential transients, the rate window concept is commonly applied. One basically investigates the contribution to the observed emission process at a chosen reference time constant τ_{ref} . By plotting the boxcar amplitude $C(t_2) - C(t_1)$ for that reference time constant as function of temperature, the relation between temperature and emission rate can be evaluated for a thermal activated process.

In the rate window concept, the selection of the contribution for a chosen reference time constant is done by a simple technique: The DLTS signal at a certain temperature $S(T, t_1, t_2)$ is given by the difference of the capacitance at two times t_1 and t_2 by

$$S(T, t_1, t_2) = C(T, t_2) - C(T, t_1), \quad (4.27)$$

or with Equation (4.26)

$$S(T, t_1, t_2) = \Delta C_0 \left[\exp\left(-\frac{t_2}{\tau(T)}\right) - \exp\left(-\frac{t_1}{\tau(T)}\right) \right]. \quad (4.28)$$

The two times t_1 and t_2 define the rate window, which has the reference time constant

$$\tau_{\text{ref}} = \frac{t_2 - t_1}{\ln(t_2/t_1)}. \quad (4.29)$$

Plotting $S(T, t_1, t_2)$ as function of temperature yields the DLTS spectrum, schematically depicted in Fig. 4.4(b). A maximum appears at that temperature, where the emission time constant of the thermally activated process equals almost the applied reference time constant: $\tau(T) = \tau_{\text{ref}}$. A maximum appears only for a thermally activated process, a temperature independent tunneling process leads to a constant DLTS signal [54]. This relation is also valid for inhomogeneous broadened DLTS spectra of self-organized QDs. The measured emission time constant is the average time constant at the maximum of the Gaussian ensemble distribution [106].

For varying τ_{ref} different peak positions T_{max} are obtained. If the capacitance transient is given by an expression like Equation (4.26), i. e. $\tau = e_n^{-1}$, and $e_n(T)$ as in Equation (3.48), the activation energy can be determined from the slope of an Arrhenius plot of $\ln(T_{\text{max}}^2 \tau_{\text{ref}})$ as function of T^{-1} . From the y-axis intersection of the extrapolated data the apparent capture cross-section σ_{∞} at infinite temperature can be obtained.

In order to improve the signal-to-noise ratio (SNR), the capacitance transients near t_1 and t_2 can be averaged over an interval t_{av} [107]. The SNR in this case is found to scale as $\sim \sqrt{t_{av}}$ [100], and the reference time constant is in good approximation

$$\tau_{\text{ref}} = \frac{t_2 - t_1}{\ln\left(\frac{t_2 + \frac{1}{2}t_{av}}{t_1 + \frac{1}{2}t_{av}}\right)}. \quad (4.30)$$

This method to improve the SNR is referred to as “Double-Boxcar”.

4.3.2 Charge-Selective Deep Level Transient Spectroscopy

An improved method to analyze the emission from QDs has been established within this thesis: the charge-selective DLTS. As already described in the previous section, usually the QDs are completely filled with charge carriers during the pulse bias V_p , Fig. 4.3(b). As a consequence, after the pulse bias the emission of many charge carriers from multiply charged QDs is probed. The activation energy of each emitted charge carrier depends on the actual charge state in such conventional DLTS experiments

on QDs. The DLTS spectrum of is broadened due to many emission processes from different QD states having different emission time constants, previously nicely observed in DLTS experiments on Ge/Si QDs [105].

In order to study the charge states in more detail, the charge-selective DLTS probes the emission of approximately one charge carrier per QD. The principle of this method is schematically illustrated in Fig. 4.5. Before the filling pulse, the QDs might be already filled with charge carriers up to the Fermi level. During the filling pulse, Fig. 4.5(b), the Fermi level is adjusted by the applied pulse bias, that approximately one carrier per QD will be captured. The pulse bias with respect to the reverse bias is always set to a constant pulse bias height, so that one carrier is captured into different QD states. After the bias pulse, the reverse bias is set again to the initial condition. Now, the previously capture carrier is emitted again and the emission process from approximately one charge carrier per QD is observed. Narrow peaks appear in the DLTS spectra, which are due to differently charged QDs, see Section 5.1.3.

Moreover, even for an energy-broadened ensemble of QDs (normally measured in a DLTS experiments) the charge-selective DLTS offers a simple method to probe the emission from many charge carriers in different QDs having all the same activation energy. Increasing the reverse bias and keeping the pulses bias height fixed (to ensure the emission of only one charge carrier per QDs) gives the activation energy starting at the ground states of the ensemble to the excited states and finishing at completely charged QDs.

Carrier Capture in Charge-Selective DLTS

In conventional DLTS experiments the electron/hole capture is normally not observed for two reasons: Firstly, during the pulse bias the QDs are outside the depletion region of the p-n diode, cf. 4.3(b) and no change in the capacitance due to charge carrier capture can be observed [53]. Secondly, the carrier capture time is expected to be in the picosecond range for QDs [19,95]. Therefore, carrier capture is not observable in DLTS measurements having a typical time resolution of milliseconds. A common approach

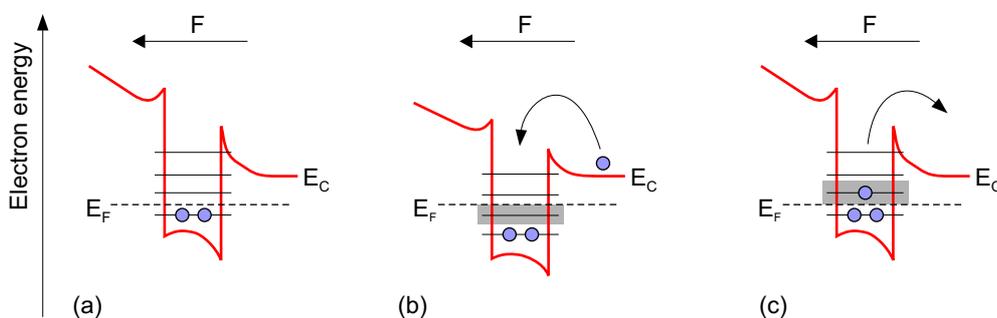


Figure 4.5: Work cycle during a charge-selective DLTS experiment (a) before, (b) during and (c) after the bias pulse, respectively. The filling pulse V_p is chosen, that the emission or capture of approximately one charge carrier per QD is probed.

to study carrier capture into deep levels is the electrical filling pulse method [108,109], which has a better time resolution in the order of nanoseconds. However, even this improved time-resolution is not sufficient to study carrier capture into QDs in the picosecond range.

Both problems can be solved by using charge-selective DLTS, where the QDs are inside the depletion region during the capture process and the band bending induced by the p-n junction provides an activation barrier, Fig. 4.5(b). The capture time is increased up to milliseconds and transients for emission *and* capture can be observed (see Chapter 6). The capture process can be easily recorded during the pulse bias.

4.3.3 Time-Resolved Tunneling Capacitance Measurement

Space charge transient methods as DLTS are very powerful tools to study thermal emission processes and thermal activation energies of QD systems. The thermal activation energy, however, does not represent in general the binding or localization energy of the QD state, as the underlying physical mechanism is phonon-assisted tunneling, cf. Section 3.2.2. Since the DLTS is only sensitive to a thermal activated process, the observed activation energy equals the separation energy between the two energy levels involved. The correct QD localization/binding energy cannot easily be observed in conventional DLTS measurements, due to this competing tunneling process in an electric field.

A method to study the tunneling emission of electrons/holes and determine the “true” localization energy of a QD state (which has been defined as the binding energy of the electron/hole ground state with respect to the conduction/valence band edge) has been developed within this thesis: The time-resolved tunneling capacitance measurement (TRTCM). This method is schematically illustrated in Fig. 4.6. At low temperature, where thermal activation of electron/holes from the ground states of the QDs is negligible, the QDs are initially empty, Fig. 4.6(a). During the pulse bias, they are filled with approximately one charge carrier per QD in the ground states, comparable to the pulse bias situation of the charge-selective DLTS. For high reverse biases, the

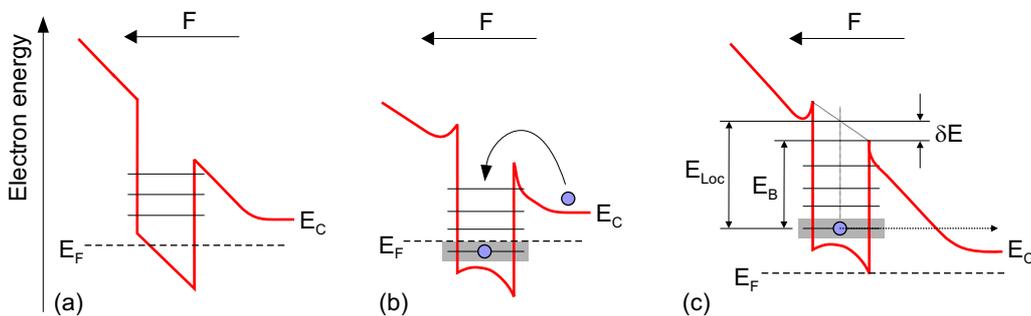


Figure 4.6: Work cycle during a time-resolved tunneling capacitance measurement. Before the filling pulse (a) the QDs are empty, and are charged with approximately one carrier in the ground state during the bias pulse (b). At low temperature, tunneling emission occurs in a high electric field (c) and can be observed in the time-resolved capacitance measurements.

charge carriers will be emitted at low temperatures due to tunneling through the triangular barrier. This method allows to obtain tunneling times by recording capacitance transients of the tunneling emission.

When measuring the tunneling time as function of the electric field strength, Equation (3.50) can be used to obtain the height of the triangular barrier E_B , depicted in Fig. 4.6(c). A logarithmic plot of the tunneling rate $\ln(e^T)$ as function of the inverse electric field $1/F$ yields a straight line with a slope that is determined by the entire barrier height E_B . The electron or hole ground state localization energy E_{Loc} is the sum of the barrier height E_B and a small energy correction due to the Poole-Frenkel effect δE .

“A THEORY IS SOMETHING NOBODY BELIEVES, EXCEPT THE PERSON WHO MADE IT. AN EXPERIMENT IS SOMETHING EVERYBODY BELIEVES, EXCEPT THE PERSON WHO MADE IT.”

ALBERT EINSTEIN

Chapter 5

Storage and Emission of Charge Carriers

The experimental results of charge carrier storage and emission for different QD systems are presented in this chapter. In Section 5.1, thermally activated emission of electrons and holes from InGaAs/GaAs QDs is investigated by using charge-selective DLTS and an average hole/electron storage time in an applied electric field is extrapolated to room temperature. Then, in Section 5.2, the tunneling emission of electrons and holes from InGaAs/GaAs QDs is studied using the TRTCM method and a storage time at room temperature without an applied electric field is derived. In order to increase the retention time, two different material combinations are investigated: GaSb/GaAs and InAs/GaAs QDs with an additional AlGaAs barrier, reported in Section 5.3 and 5.4, respectively. The latter yields an average hole storage time of milliseconds at room temperature – the DRAM milestone. Type-II GaSb-based QDs, like (In)(Ga)Sb, are on the other hand the most promising material system to achieve a hole retention time in the order of years at 300 K – the important prerequisite to use QDs in future Flash memory applications; everything discussed in Section 5.5. Information on the experimental setup can be found in Appendix A.1.

5.1 Thermal Emission from InGaAs/GaAs QDs

The conventional DLTS signal from electron and hole emission from InAs QDs has already been intensively studied previously by C. Kapteyn [54, 66, 67]. Here, the electron/hole emission from InGaAs QDs is investigated in more detail by using the charge-selective DLTS method.

In order to study both charge carrier types, C-V and DLTS measurements of two different samples were performed. The samples were grown on a GaAs(001) substrate by using MOCVD. The QD layer was incorporated in either a p^+-n or a n^+-p diode structure, i. e. the surrounding material is n-doped/p-doped to allow controlled charging of the electron and hole states, respectively. The emission of one charge carrier type

can be studied individually.

After a general characterization of the two samples in Section 5.1.1, the results of the capacitance-voltage and the charge-selective DLTS measurements are presented in Section 5.1.2 and 5.1.3, respectively. The outcome of the investigations concerning storage times at room temperature in an applied electric field is discussed in Section 5.2.5.

5.1.1 Sample Structures

The first sample, referred to as H1 (which is used to study the hole emission) is a n^+ -p structure containing a single $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ QD layer, Fig. 5.1(a), formed by deposition of ~ 3 monolayer (ML) $\text{In}_{0.8}\text{Ga}_{0.2}\text{As}$ at 500°C . The QD sheet-density amounts to $N_D^{\text{H1}} \approx 3 \times 10^{10} \text{cm}^{-2}$. The QDs exhibit a ground-state transition at 1.12 eV as observed by photoluminescence (PL) at He temperature [110], and an average QD base-length of about 16 nm.

The layer sequence of sample H1 is as follows: On top of a GaAs(001) substrate, a 500 nm thick highly p-doped GaAs ($\sim 1 \times 10^{18} \text{cm}^{-3}$) layer followed by 700 nm slightly p-doped ($\sim 3 \times 10^{16} \text{cm}^{-3}$) GaAs were deposited. Subsequently, on top of a 10 nm thick undoped GaAs layer, ~ 3 ML $\text{In}_{0.8}\text{Ga}_{0.2}\text{As}$ were deposited forming the QDs. The QDs were eventually overgrown with 7 nm undoped GaAs and a 500 nm layer of slightly p-doped ($\sim 3 \times 10^{16} \text{cm}^{-3}$) GaAs. Finally, a 400 nm highly n-doped ($\sim 7 \times 10^{18} \text{cm}^{-3}$) GaAs cap layer formed the n^+ contact.

The second sample, referred to as E1, is a p^+ -n structure with $\text{In}_x\text{Ga}_{1-x}\text{As}$ QDs embedded in the n-region of GaAs, to study electron emission, Fig. 5.1(b). The $\text{In}_x\text{Ga}_{1-x}\text{As}$ QDs were grown at 500°C by deposition of ~ 7 ML $\text{In}_{0.7}\text{Ga}_{0.3}\text{As}$ on top of a QD seed layer, which was formed by the deposition of ~ 3.5 ML $\text{In}_{0.7}\text{Ga}_{0.3}\text{As}$. The GaAs spacer

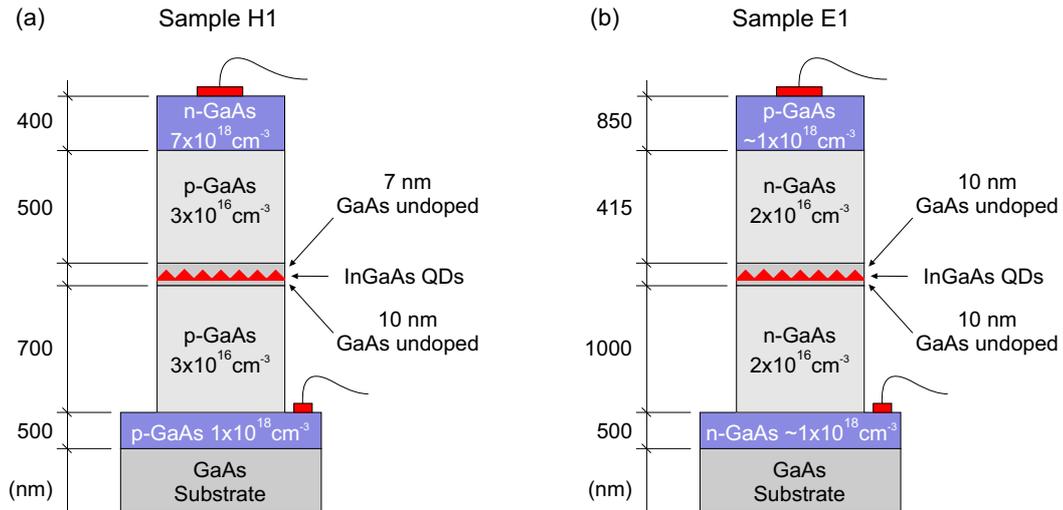


Figure 5.1: Schematic layer structure of (a) sample H1 (to study the hole emission) and (b) sample E1 (to study the electron emission).

of only 3 nm thickness ensures vertical growth correlation between the two QD layers. The QDs of the second layer nucleate directly above the QDs of the seed layer and have also a bigger size than the seed QDs. This method was used to achieve a high QD density of large In_xGa_{1-x}As QDs, see Reference [111]. Transmission electron microscopy (TEM) images of comparable samples [112] show a QD area density of $\sim 5 \times 10^{10} \text{cm}^{-2}$, a truncated pyramidal shape with an average base width of about 18 – 20 nm and a height of ~ 4 nm. The ground-state transition energy in PL is ~ 1.04 eV at He temperatures [110]. The smaller In_xGa_{1-x}As seed QDs have only a minor impact on the electronic and optical properties of the In_xGa_{1-x}As QDs due to the large mismatch of the ground-state transition energies of the two QD types [113].

The layer sequence of sample E1 is as follows: On top of a GaAs(001) substrate, a 500 nm thick highly n-doped GaAs ($\sim 1 \times 10^{18} \text{cm}^{-3}$) layer followed by 1000 nm slightly n-doped ($\sim 2 \times 10^{16} \text{cm}^{-3}$) GaAs were deposited. Subsequently, on top of a 10 nm thick undoped GaAs layer, the seed layer and the QD layer were deposited. The QDs were eventually overgrown with 10 nm undoped GaAs and a 415 nm layer of slightly n-doped ($\sim 2 \times 10^{16} \text{cm}^{-3}$) GaAs. Finally, a 850 nm highly p-doped ($\sim 1 \times 10^{18} \text{cm}^{-3}$) GaAs cap layer formed the p⁺ contact.

For both samples, photolithography and chemical wet-etching were employed to form mesa structures with a diameter of 800 μm . Ohmic contacts were formed on top and back of the structures by metal evaporation and alloying. For details of the sample processing see Appendix A.2. With a diameter of 800 μm of the mesa structures and an average QD sheet density of $3 \times 10^{10} \text{cm}^{-2}$, charge carrier emission from an energy broadened ensemble of millions of QDs is studied in the following. The energy broadening is due to fluctuations in size, shape, and composition of the single QDs in the ensemble.

5.1.2 Capacitance-Voltage Measurements

The current-voltage (I-V) dependence of the processed diodes were measured firstly to check for proper device operation. The break-through reverse bias was typically between 10 – 20 V, and the build-in voltages V_b between 1.2 and 1.4 V, in good agreement with the expected value of about 1.3 V for the given doping concentration.

Thereafter, static C-V measurements at varying temperatures (between 20 K and 350 K) and measurement frequencies (between 220 Hz and 1 MHz) of both samples were made. The AC signal amplitude of the capacitance meter was usually fixed at 100 mV for an acceptable signal to noise ratio (SNR) and measurement integration times. The C-V curves showed a significant dependence on the frequency and the temperature, since at high frequencies and low temperatures the localized charge carriers can not follow the applied AC measurement frequency. As a consequence, the plateau structure due to the QD states is not visible anymore in the data. For temperatures above $T \approx 150$ K (depending on the localization energy of the charge carriers) the plateaus tend to disappear, as the charge carrier are thermally exited into the valence/conduction band and not localized inside the QD layer anymore. The optimum conditions in order

to resolve the full structure were determined to be at about 1 kHz for a temperature around 100 K for both samples. Two representative curves are shown in Fig. 5.2, where the static C-V trace (solid red line) of samples H1 (a) and E1 (b) is displayed, recorded at $T=100$ K at a modulation frequency of 1 kHz. In both C-V curves a pronounced plateau is visible, which reflects the charge accumulation in the QD layer.

For sample H1, the depletion region of the n^+p diode reaches the QD layer at a reverse bias of about 7.0 V and is pinned there, precisely marked by a maximum of the second derivative (dashed line) in Fig. 5.2(a). As a consequence, the capacitance remains roughly constant for increasing reverse bias until all holes are removed from the QDs at 9.4 V. The small quantization energies of the hole levels do not allow to distinguish between the population of ground and excited states in the C-V measurement.

For sample E1 in Fig. 5.2(b), the onset of the plateau at a reverse bias of about 6.5 V, as marked by the peaks of the second derivative (dashed line), denotes the beginning of the population of the electron ground states. The larger level splitting of the electron states allows here to distinguish between ground and excited state population. Between 6.5 V and 5.5 V the ground state is populated with two electrons. The excited states are populated with electrons below 5.5 V, whereas below 3.0 V the QDs leave the space-charge region and are completely filled.

The C-V results allow to estimate the number of charge carriers accumulated in the QD layer from: $Q \approx C_p \Delta V$, where C_p is the capacitance of the plateau region, and ΔV the width of the plateau. Furthermore, the knowledge about the QD area density is required to evaluate the average population of completely charged QD. For sample H1, a maximum number of approximately seven holes per QD is found. For sample E1, the evaluation of the C-V data yields two charge carriers per QD for the reverse bias region from 5.5 V to 6.5 V, as expected for a spin-degenerated ground state. Five more electrons can be charged into the excited state for the voltage region between 3.0 V and

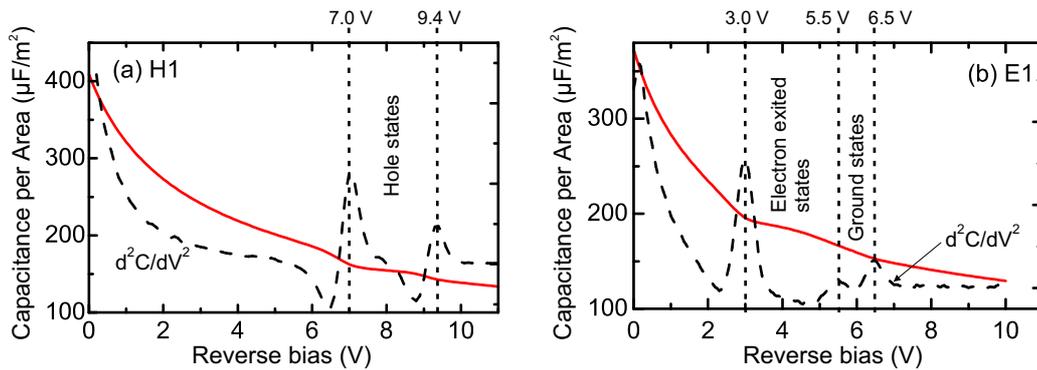


Figure 5.2: Static capacitance-voltage (C-V) characteristics of sample H1 (a) and E1 (b) at $T=100$ K for a measurement frequency of $f=1$ kHz (solid red line). The plateaus in the C-V traces are due to hole/electron accumulation, respectively, in the QD layer. The maxima of the second derivative d^2C/dV^2 (dashed line) indicate the beginning and end of the plateaus.

5.5 V. In total, sample E1 can be charged with up to seven electrons per QD.

The C-V characterization reveals the bias needed to charge the QDs in the following charge-selective DLTS (Section 5.1.3) and TRTCM experiments (Section 5.2). Especially for the TRTCM method, precise bias values are required to populate only the electron/hole ground states with charge carriers.

5.1.3 DLTS Measurements

In this section, the time-resolved capacitance measurements, the conventional and the charge-selective DLTS, of both samples are summarized. Typically, transients of 300 ms duration were sampled with a resolution of 2000 points or more (yielding a minimum time resolution of 150 μ s) for temperature steps of about 1.5 K. The data was subsequently evaluated by the double-boxcar method (see Section 4.3.1). More information on the experimental setup can be found in Appendix A.1.

Hole Emission

Figure 5.3(a) depicts a conventional DLTS measurement of sample H1 for a pulse bias of $V_p = 7.0$ V and a detection reverse bias of $V_r = 9.4$ V. For a reverse bias of 9.4 V, the QDs are inside the depletion region, i. e. they are totally depleted from charge carriers, cf. the C-V measurement in Fig. 5.2(a). At a pulse bias of $V_p = 7.0$ V, the QDs are completely filled with holes during a 10 ms electrical pulse. After the bias is set back to the initial situation, the QDs are again in the depletion region in a non-equilibrium state with the reservoir and all trapped holes are emitted, see schematic inset in Fig. 5.3(a). The DLTS spectrum is a result of thermal emission of completely charged QDs, where the thermal activation energy of each emitted hole depends on the actual charge state. As a consequence, the DLTS peak is broadened and extends towards temperatures below 30 K.

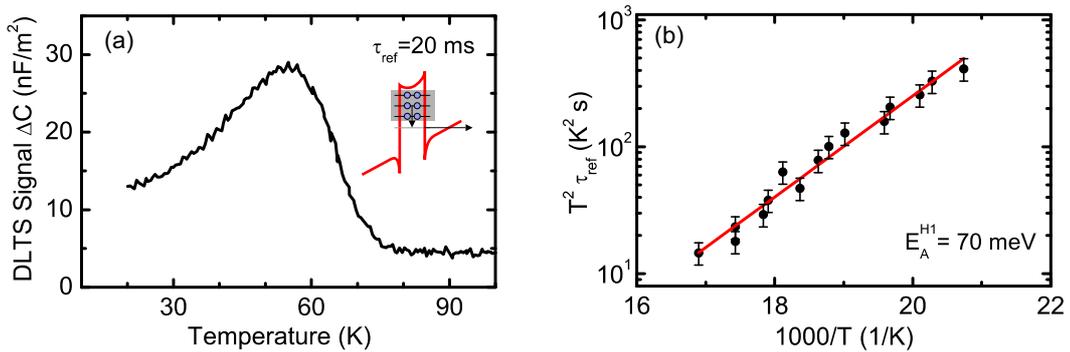


Figure 5.3: (a) DLTS spectra of thermally activated hole emission at a reference time constant of $\tau_{\text{ref}} = 20$ ms, a reverse bias of $V_r = 9.4$ V and a pulse length of $t_{\text{pulse}} = 10$ ms. For this pulse bias situation of $V_p = 7.0$ V the QDs are completely filled with holes and parallel emission of multiple charged QDs is observed (schematic inset). Panel (b) displays the Arrhenius plot, yielding an activation energy of $E_A^{H1} = 70$ meV.

From a standard Arrhenius plot of the DLTS peak position for varying reference time constants τ_{ref} , Fig. 5.3(b), an activation energy of about (70 ± 10) meV is obtained. This value represents the average activation energy of hole emission from completely charged QDs, where many QD states are involved. In addition, in an applied electric field phonon-assisted tunneling occurs, inset in Fig. 5.3(a), and the DLTS measurement separates only the activation energy E_A^{H1} from this two step process, which is not equal to the localization energy E_{Loc}^{H1} .

Charge-Selective DLTS

In order to study the activation energies with more precision, the charge-selective DLTS method is used. As mentioned before, this technique always sets the pulse bias in relation to the reverse bias, such that on average the thermal emission of about one charge carrier per QD is probed. By changing the reverse bias, this method allows to control the QDs' charge state, see Section 4.3.2.

The pulse bias was always set to $V_p = V_r - 0.2$ V, while the reverse bias was decreased from 9.6 up to 7.2 V in Fig. 5.4. The reference time constant was always set to 20 ms. For higher reverse biases than 9.4 V no distinct DLTS signal is visible, as the QDs are now completely inside the depletion region of the p-n diode and not

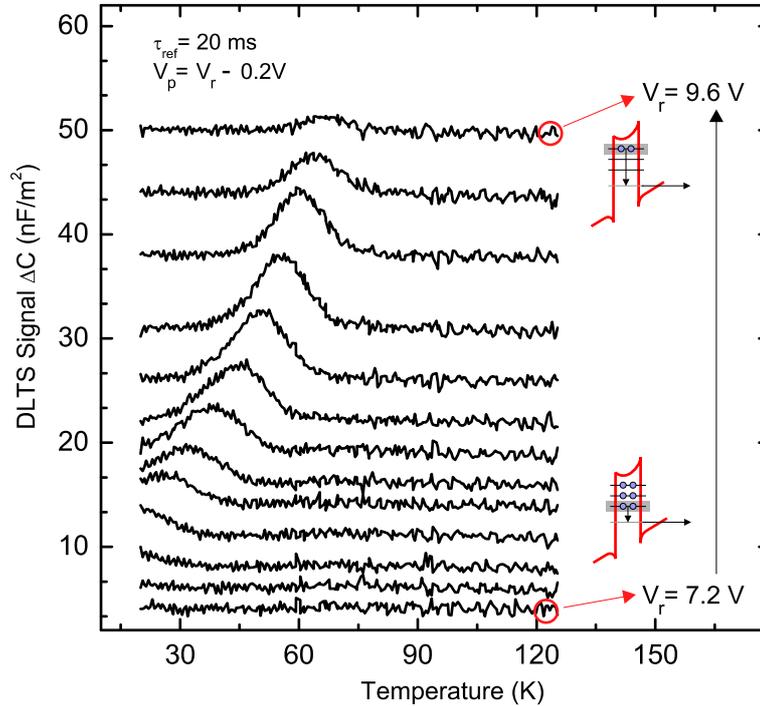


Figure 5.4: (a) Charge-selective DLTS spectra of sample H1 at a reference time constant $\tau_{\text{ref}} = 20$ ms for increasing reverse bias V_r between 7.2 and 9.6 V. The pulse height is 0.2 V for all spectra – that means $V_p = V_r - 0.2$ V – and the pulse width 10 ms. The data is displayed vertically shifted for clarity.

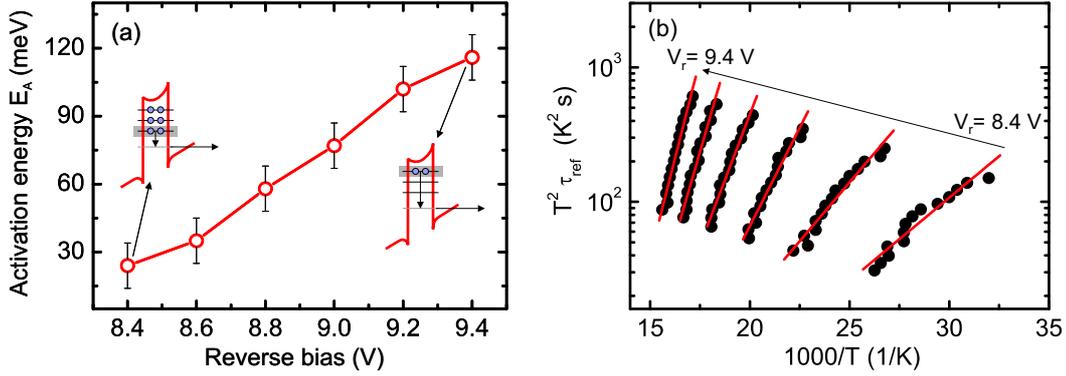


Figure 5.5: (a) Activation energies in dependence of the reverse bias V_r for the charge-selective DLTS spectra in Fig. 5.4. Panel (b) displays the corresponding Arrhenius plots.

charged with holes anymore during the pulse bias. For a reverse bias of $V_r = 9.4$ V only the ground states are filled with approximately one hole per QD. As a consequence, a narrow peak at about $T = 60$ K appears in the DLTS spectrum, Fig. 5.4, while no DLTS signal below 50 K exists. From an Arrhenius plot in Fig. 5.5(b), an activation energy of (120 ± 10) meV with a capture cross-section of about $\sigma_\infty^{H1} \approx 7 \times 10^{-14}$ cm² is obtained.

The DLTS peak shifts to lower temperature for decreasing reverse bias, Fig. 5.4. The activation energy decreases accordingly from ~ 120 meV at $V_r = 9.4$ V down to ~ 30 meV at $V_r = 8.4$ V in Fig. 5.5(a). The DLTS spectra for a reverse bias between 8.4 and 9.4 V are attributed to hole emission from differently charged QDs. All these spectra exhibit a maximum in the temperature range between 20 and 60 K, the range covered by the broadened peak in Fig. 5.3(a). The maximum activation energy of about 120 meV represents the thermal activation energy E_A^{H1} of the phonon-assisted tunneling process from the QD hole ground states into an intermediate state. The decrease in the activation energy from 120 meV down to 30 meV corresponds to an increase in the average occupation of the QDs. With increasing amount of charge in the QDs, state filling lowers the thermal activation barrier and less confined QD states are probed.

As expected, the activation energy of 120 meV from the charge-selective DLTS is much larger than the value of 70 meV from multiple hole emission, where a number of weaker bound states are involved. However, it is evident that this value still underestimates the ground state localization energy, if compared with the theoretical prediction from 8-band k-p theory in Fig. 2.7(a). For QDs with a base-length of ~ 16 nm a hole ground state localization of about 200 meV is predicted. Previous admittance spectroscopy measurements from Chang et al. [62] confirm this DLTS measurements and explain also the significant underestimation of the hole localization energy caused by the phonon-assisted tunneling process.

Ensemble Effects on the DLTS Measurements

Even the DLTS curves for the charge-selective DLTS measurements in Fig. 5.4 do not represent a single emission process, as always an energy broadened ensemble of QDs is probed. For comparison, the black dashed line in Fig. 5.6 displays a simulated DLTS spectrum by using Equation (2.3), that stems from thermal emission from a single energy level. This simulation shows clearly a smaller peak width than the experimental data (black solid line for $V_r = 9.4$ V). The solid red line in Fig. 5.6 displays a simulation that is fitted to the experimental data, assuming thermal emission from an energy broadened ensemble of QDs. A good agreement with the experimental data is obtained for a Gaussian shape with a full width at half maximum (FWHM) of 30 meV.

Electron emission

The black solid line in Figure 5.7(a) shows a conventional DLTS measurement of the electron sample E1 for a pulse bias of $V_p = 3.0$ V and a detection reverse bias of $V_r = 6.8$ V. The solid red line displays a pulse bias situation ($V_r = 6.0$ V), where only the electron ground states are filled with charge carriers, cf. the C-V measurement of sample E1 in Fig. 5.2(b). Both DLTS spectra were recorded after a pulse duration of 10 ms and evaluated with a double-boxcar having a reference time of $\tau_{\text{ref}} = 30$ ms.

Clearly visible are two peaks in the black line at a temperature of 50 K and 30 K, respectively. Even for this bias condition, where the QDs are completely charged with electron, two different thermal activation processes are distinguishable in the DLTS spectrum. This is in contrast to the hole sample H1, since the electron states have a larger level splitting in the order of 80 meV [66].

The red line in Fig. 5.7(a) shows the charge-selective DLTS measurement of sample E1, for a reverse/pulse bias of 6.8 V/6.0 V, such that only an energy broadened

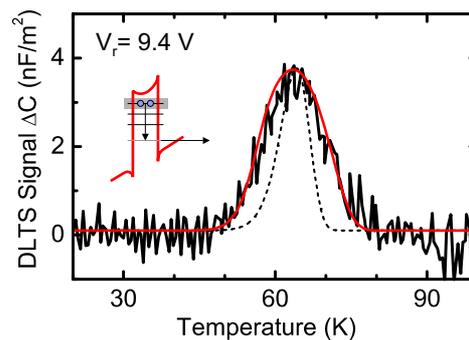


Figure 5.6: Charge-selective DLTS spectrum of sample H1 at a reference time constant of $\tau_{\text{ref}} = 20$ ms, a reverse/pulse bias of $V_r = 9.4$ V/ $V_p = 9.2$ V, respectively. The signal is due to thermal emission of holes from the ground states of the QD ensemble (see schematic inset). The black dashed line represents a simulated DLTS spectrum, that stems from thermal emission from a single energy level. The solid red line displays a simulation, assuming an energy broadened ensemble of QDs with a Gaussian shape and a FWHM of 30 meV.

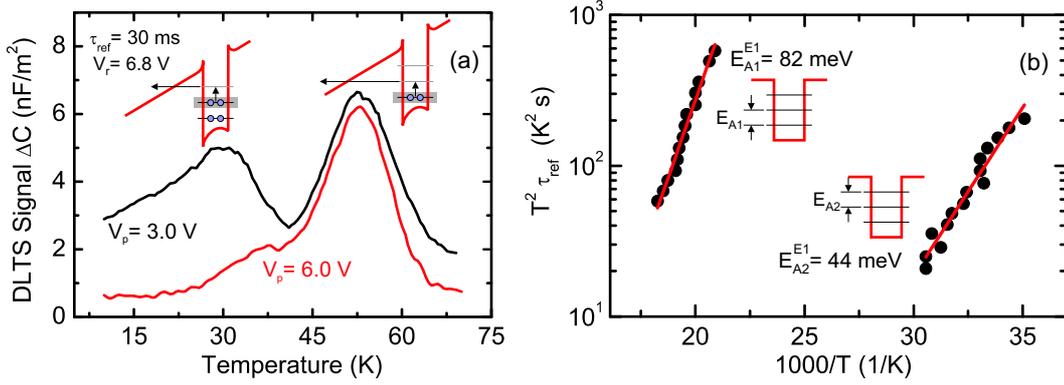


Figure 5.7: (a) DLTS spectra of sample E1 at a reference time constant of $\tau_{\text{ref}} = 30$ ms, a reverse bias of $V_r = 6.8$ V, and a pulse duration of 10 ms. The black solid line displays a DLTS spectrum, where the QDs are completely filled with electrons for a pulse bias of $V_p = 3.0$ V. The solid red line shows a pulse bias situation, $V_p = 6.0$ V, where only the electron ground states are filled with charge carriers. Panel (b) displays the corresponding Arrhenius plots.

ensemble of electron ground states are charged with approximately one electron per QD. Now, only a single DLTS peak occurs at about 50 K, while the second peak at 30 K is vanished. The corresponding thermal activation energy from an Arrhenius plot is $E_{A1}^{E1} = (82 \pm 10)$ meV and displayed in Fig. 5.7(b), representing the ground to first excited state level splitting (previously demonstrated in Reference [67]). Since electrons have a larger tunneling probability than holes for the same barrier height, thermal activation excites electrons mainly to the first excited state before tunneling through the remaining triangular barrier takes place, see inset in Fig. 5.7(a). The thermal activation energy is, thus, more reduced for electrons than for holes. In addition, from the Arrhenius plot a capture cross-section of about $\sigma_{\infty, A1}^{E1} \approx 2 \times 10^{-14}$ cm² is obtained for this thermal activation process.

The second peak at 30 K can be attributed to the level splitting between the first and second excited QD electron state. Here, the phonon-assisted tunneling between these two states in the energy-broadened ensemble of QDs is probed. The thermal activation energy from an Arrhenius plot in Fig. 5.7(b) is $E_{A2}^{E1} = (44 \pm 10)$ meV, and the corresponding capture cross-section has a value of $\sigma_{\infty, A2}^{E1} \approx 1 \times 10^{-15}$ cm².

5.1.4 Carrier Storage Times in an Electric Field

A peak in the DLTS signal appears for that temperature, at which the reference time constant τ_{ref} equals almost the average emission time of the charge carriers, cf. Equation (4.29). That means, the average hole storage time in an applied electric field is 20 ms at $T = 65$ K for the emission process from the ground state with an activation energy of ~ 120 meV in sample H1. Analogously, the average electron storage time is 30 ms at $T = 55$ K for the ground to first excited state emission process having an activation energy of 82 meV in sample E1.

The important value for QDs in future memory applications is the storage time at room temperature. This value can be extrapolated by using the thermal rate Equation (2.3) or (2.4) for electrons/holes with the values for the activation energy and capture cross-section obtained from the DLTS experiment. The average storage time in an applied electric field at $T = 300$ K is approximately 50 ps/10 ps for electrons/holes, respectively, in the ground state of InGaAs/GaAs QDs. It is obvious that the storage time in InGaAs/GaAs QDs is not sufficiently long enough for future memory applications. However, these values are related to carrier storage in an electric field, where phonon-assisted tunneling reduces the retention time. A longer storage time is expected for a situation without an electric field. Then, the storage time will be limited purely by thermal activation of the charge carriers, which is controlled by the entire localization energies. The knowledge about the localization energy in QDs is needed to estimate the storage time without an electric field, discussed in the next section.

5.2 Tunneling Emission from InGaAs Quantum Dots

In this section, the observation of tunneling emission of electrons and holes from InGaAs/GaAs QDs in time resolved capacitance measurements (TRTCM) is presented. The time-dependent decrease of the electron or hole ground state population is a result of pure tunneling escape at low temperatures, where thermally activated emission is negligible. In addition, by measuring the tunneling time as function of the electric field strength, the electron and hole localization energy is determined. Both samples, investigated in this section, are the same InGaAs QD samples, already presented in Section 5.1.1 and investigated by using charge-selective DLTS in Section 5.1.3.

After a general introduction with a check for a possible temperature-dependence of the tunneling measurement in Section 5.2.1, the tunneling emission from the hole and electron ground states into the valence/conduction band are presented in Section 5.2.2 and 5.2.3, respectively. Then, in Section 5.2.4 the electron and hole localization energies are derived by using the dependence of the tunneling time on the applied electric field. Finally, the electronic properties and storage times in InGaAs/GaAs QDs are discussed in Section 5.2.5.

5.2.1 Check for Temperature Dependence

Firstly, capacitance transients for increasing temperature are recorded, to ensure for the subsequent TRTCM investigations, that the observed change in the capacitance is only due to tunneling.

In general, the capacitance reflects the population of QD states with majority charge carriers. In Fig. 5.8, for instance, the minimum capacitance at $t = 0$ represents the maximum population of the QD ground states with holes. For $t > 0$ an increasing capacitance is caused by a decreasing hole population due to tunneling emission. The filling pulse is not shown in the following spectra, i. e. the capacitance transients begin at that time, when the tunneling emission process starts.

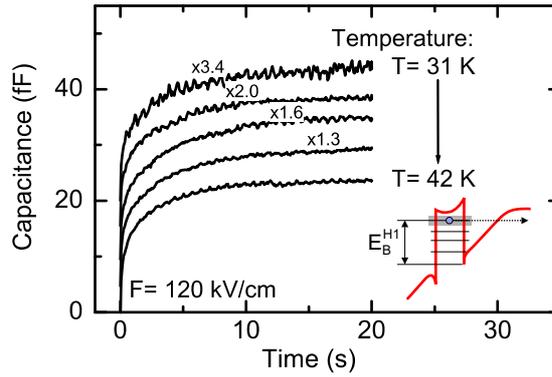


Figure 5.8: Capacitance transients for increasing temperature from $T= 31$ K up to 42 K at a reverse bias of $V_r= 10.4$ V. The reverse bias corresponds to an electric field of 120 kV/cm in the QD layer of sample H1. The almost constant emission time proves that no thermal emission is present and pure hole tunneling is observed. The data is vertically shifted for clarity.

Figure 5.8 displays capacitance transients at temperatures between 31 K and 42 K. The reverse bias was set to $V_r= 10.4$ V, which corresponds to an electric field of about 120 kV/cm in the QD layer. The electric field was determined based on a simple approximation for p-n junctions (see below). The pulse bias V_p was adjusted to 9.2 V, such that on average only the hole ground states are filled. An almost constant emission time with an average value of (2.1 ± 0.3) s is obtained, which proves that pure hole tunneling is observed in this temperature range. In the corresponding charge-selective DLTS measurement (Fig. 5.4 for a reverse bias of $V_r= 9.4$ V), the peak starts at about $T= 45$ K for identical reverse/pulse bias situation, due to the initiation of the thermal hole emission on a 20 ms time scale. Thermally activated emission also occurs in the TRTCM spectra for $T > 45$ K, leading to a decreasing emission time with increasing temperature. That means, up to $T= 45$ K, thermally activated emission can be neglected and pure hole tunneling is observed in this temperature range.¹

Determination of the Electric Field Strength

The electric field at the position of the QD layer x has been determined by using Equation (4.13) and (4.14), respectively, for an electric field F across the depletion region of an abrupt p-n junction. The expression for the width of the depletion region w is given by Equation (4.19).

As already mentioned, Equation (4.13) and (4.19) describe an idealized model in which the boundary at the edge of the depletion region is abrupt and beyond the depletion region the GaAs is electrically neutral. It is a good practical approximation to obtain the width of the depletion region and the strength of the electric field. This

¹Note the minor effect of an increasing amplitude of the transients with increasing temperature in Fig. 5.8. The QD charging process is slightly temperature-dependent, due to the fact that the charge carriers have to overcome a capture barrier during the capture process, if the pulse bias is not set to flat-band condition, cf. Section 4.3.2 and Fig. 4.5(b).

simple approximation characterizes a p-n junction without an embedded QD layer, which pins the depletion region, as clearly visible in the static C-V measurements in Fig. 5.2. Nevertheless, during the tunneling emission of the charge carriers the QDs are completely inside the space-charge region for all applied reverse biases and only charged with up to one charge carrier per QD. Therefore, it can be assumed that Equation (4.13) is reasonable to determine the electric field, since almost uncharged QDs have only a minor effect on the width of the depletion region and, hence, on the strength of the electric field.²

This assumption can be easily confirmed by a simple estimate of the influence of singly charged QDs on the depletion region: For a QD sheet density of about $3 \times 10^{10} \text{ cm}^{-2}$ and a depletion width always larger than 700 nm for all applied reverse biases, the QD charge is equivalent to a space-charge concentration less than $N \approx 4 \times 10^{14} \text{ cm}^{-3}$. This value is almost two orders of magnitude smaller than the doping concentration, which determines in addition to the applied reverse bias V_r the width of the depletion region and the magnitude of the electric field. The error in the determination of the electric field due to the assumption of singly charged QDs is estimated to be always smaller than 0.5 kV/cm. Therefore, the influence of the charged QDs can be neglected.

The accuracy of the determination of the applied electric field is, hence, mainly due to the uncertainty of the doping concentration N_a and N_d , respectively. The doping concentration was determined from static C-V measurements to $N_a^{H1} = (1.6 \pm 0.3) \times 10^{16} \text{ cm}^{-3}$ and $N_d^{E1} = (2.0 \pm 0.3) \times 10^{16} \text{ cm}^{-3}$ at the tunneling measurement temperature of $T = 40 \text{ K}/20 \text{ K}$ for sample H1/E1, respectively. This differs slightly from the nominal doping concentration, described in Section 5.1.1. Note, that even with an uncertainty of about 20 percent for the doping concentration the total accuracy of the electric field is better than five percent, due to the error propagation theorem.

5.2.2 Hole Tunneling

If the temperature is kept fixed at $T = 40 \text{ K}$, a strong dependence of the tunneling emission time on the applied electric field is visible in Fig 5.9(a), which additionally confirms the observed pure hole tunneling. The electric field is increased in small steps of 2.25 kV/cm by increasing the reverse bias from $V_r = 10.4 \text{ V}$ up to $V_r = 12.5 \text{ V}$. The capacitance transients are recorded for a duration of 5 s up to 30 s.

From the recorded capacitance transients, the hole tunneling times from the QD ground states through the GaAs barrier E_B^{H1} can be obtained. As mentioned before, in capacitance experiments the carrier emission from an energy broadened ensemble of QDs is observed (Section 5.1.3), hence, the capacitance transients cannot be described by a single exponential function. Different single QDs with varying size, shape and composition are expected to have different tunneling rates. As a consequence, tunneling emission with different time constants is observed, leading to multi-exponential transients.

²For a better prediction of the electric field, a self-consistent calculation according to the model described in [54] is necessary.

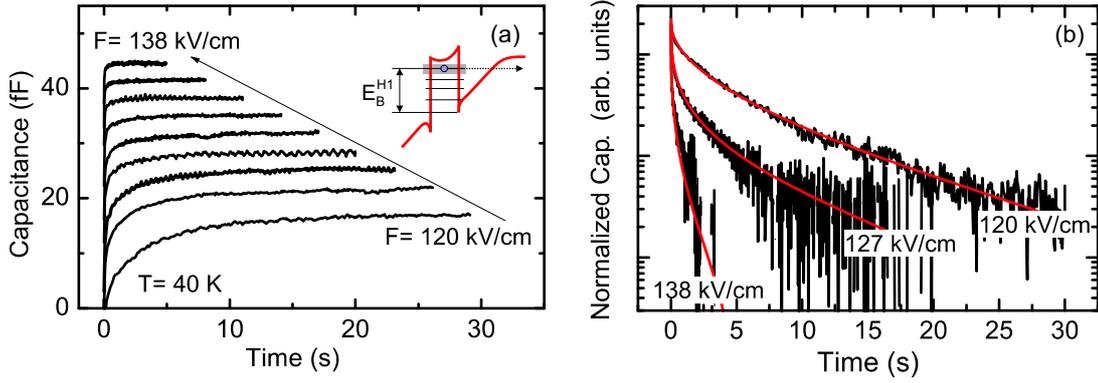


Figure 5.9: (a) Capacitance transients for increasing applied electric field from 120 kV/cm up to 138 kV/cm. The pulse bias $V_p = 9.2$ V was set such that on average only the hole ground states are occupied and tunneling through the entire barrier E_B^{H1} is observed (see schematic inset). The pulse width was set to $t_{\text{pulse}} = 900$ ms and the transients are vertically shifted for clarity. (b) Capacitance transients of hole tunneling, displayed on a semi-logarithmic scale. Only three transients from (a) are shown for clarity. Red solid lines are fits with stretched exponential functions (see text).

Figure 5.9(b) depicts three selected capacitance transients from Fig. 5.9(a) as black lines on a semi-logarithmic scale. Clearly visible is a non-exponential decay, similar to DLTS experiments on deep levels in semiconductor alloys, see Omling et al. [106]. The capacitance transients $C(t)$ are fitted with a so-called stretched exponential function, see Scher et al. [114] and references therein:

$$C(t) = C(t_0) \cdot \exp[-(t/\tau^*)^\beta]. \quad (5.1)$$

Equation (5.1) describes a superposition of different exponential tunneling processes, with τ^* being an effective time constant and β the stretching parameter. The latter describes the spread of time constants and is unity for a mono-exponential decay. A stretching parameter smaller than one indicates a variation from a mono-exponential transient. Probing always an ensemble of QDs with a FWHM of about 30 meV, see Section 5.1.3, the capacitance transients of the tunneling emission can be fitted well by Equation (5.1) with a stretching parameter of about 0.6. The multi-exponential fits are displayed as red lines in Fig. 5.9(b).

The tunneling times of sample H1 in dependence of the applied electric field are shown as red dots in Fig. 5.10. In addition, the $1/e$ -decay time of the transients as a measure for the average emission time of the ensemble (black dots in Fig. 5.10) has been evaluated, to obtain a better accuracy in the determination of the localization energy. The tunneling times obtained from the stretched-exponential fit is also in very good agreement with the values from the $1/e$ -decay time evaluation. The tunneling time decreases for hole emission in sample H1 from about 2 s down to 30 ms for an increasing electric field from 120 kV/cm (at $V_r = 10.4$ V) to 138 kV/cm (at $V_r = 12.5$ V). The solid lines are fits to the data, as predicted from Equation (2.8), that means, the tunneling time on a logarithmic scale is fitted with a function that depends

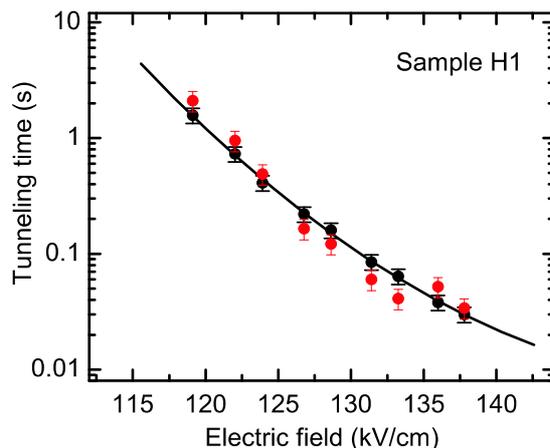


Figure 5.10: Hole tunneling time of sample H1 in dependence of the applied electric field. The tunneling times were obtained by evaluation of the 1/e decay time (black dots) and stretched exponential fits (red dots). The solid line is a fit to the data (see text).

on the inverse electric field.

5.2.3 Electron Tunneling

The electron tunneling emission from the ground states through the GaAs barrier E_B^{E1} is shown in Fig. 5.11(a) on a linear scale. The pulse bias was set to $V_p = 6.0$ V. Hence, again only the ground states are occupied with electrons, cf. C-V curve in Fig. 5.2(b). At $T = 20$ K, the thermal electron emission can be neglected. The electric field is increased in small steps of 1.3 kV/cm by increasing the reverse bias from $V_r = 7.0$ V up to $V_r = 8.0$ V. The decrease in the population of the QD ground states was recorded time-resolved for a duration between 10 s and 60 s.

Figure 5.11(b) depicts three selected capacitance transients from Fig. 5.11(a) as black lines on a semi-logarithmic scale. Clearly visible is again a non-exponential decay, which is fitted with a stretched exponential function, red lines in Fig. 5.11(b). The effective time constants τ^* are shown as red dots in Fig. 5.12. The 1/e-decay times are displayed as black dots. Increasing the electric field from 79 kV/cm (at $V_r = 7.0$ V) up to 92 kV/cm (at $V_r = 8.0$ V) reduces the tunneling time from about 5 s down to 80 ms. The solid lines are again fits to the data, as predicted from Equation (2.8).

5.2.4 Localization Energies

The GaAs barrier height for electrons E_B^{E1} and holes E_B^{H1} can be obtained from the dependence of the tunneling time on the applied electric field. The barrier height is almost equal to the entire localization energy E_{Loc} . The difference is due to the Poole-Frenkel effect, which slightly lowers the barrier height in an applied electric field. The localization energy E_{Loc} is the sum of the tunneling barrier height and an estimated Poole-Frenkel energy δE , explained in Section 3.2.

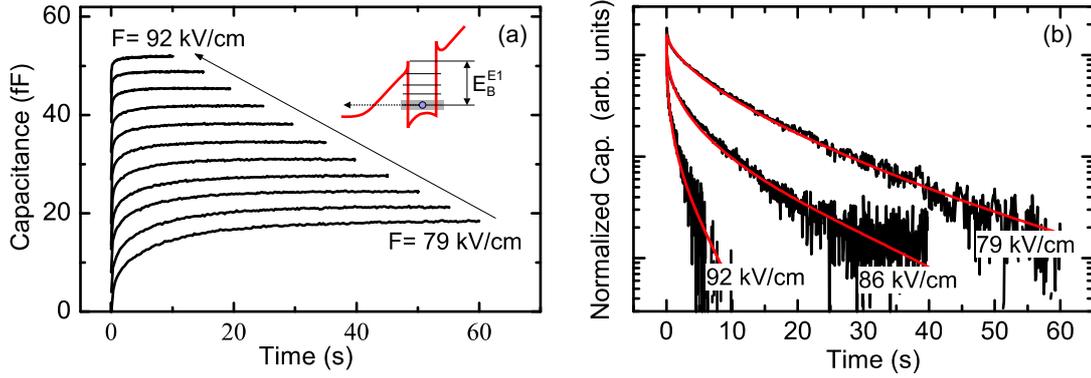


Figure 5.11: (a) Electron tunneling from the ground states through the GaAs barrier E_B^{E1} for sample E1, observed in capacitance transients at low temperature ($T=20$ K). The pulse bias V_p was set to 5.8 V to occupy only the electron ground states (see inset), while the reverse bias was increased to apply an electric field from 79 kV/cm up to 92 kV/cm. The pulse width was set to $t_{\text{pulse}}=900$ ms and the transients are vertically shifted for clarity. (b) Capacitance transients of electron tunneling, displayed on a semi-logarithmic scale. Only three transients from (a) are shown for clarity. Red solid lines are fits with stretched exponential functions (see text).

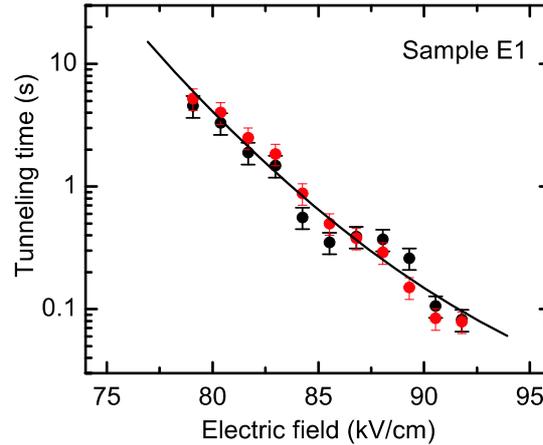


Figure 5.12: Electron tunneling time of sample E1 in dependence of the applied electric field. The tunneling times were obtained by evaluation of the $1/e$ decay time (black dots) and stretched exponential fits (red dots). The solid line is a fit to the data (see text).

The Poole-Frenkel effect reduces the hole barrier for QDs [104] with an average height of 4 nm by approximately 24 meV for an electric field of ~ 120 kV/cm. For electrons, the barrier height is reduced by approximately 16 meV for an electric field of about 80 kV/cm. The accuracy of δE is limited by the knowledge of the QD height. The Poole-Frenkel effect, however, has only a minor impact on the determination of the localization energy.

As predicted by Equation (2.8), a linear dependence of the tunneling rate (on a logarithmic scale) via the inverse electric field occurs in Fig. 5.13. A hole barrier height

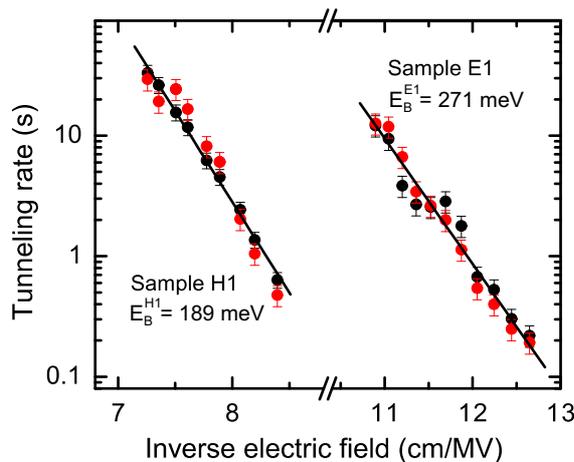


Figure 5.13: Tunneling rate for holes (sample H1) and electrons (E1) in dependence of the inverse electric field. From the linear dependence, as predicted in Equation (2.8), the hole ($E_B^{H1} = 189$ meV) and electron barrier height ($E_B^{E1} = 271$ meV) is obtained. Adding an approximate value for the Poole-Frenkel energy δE leads to the hole ($E_{Loc}^{H1} = 210$ meV) and electron ($E_{Loc}^{E1} = 290$ meV) localization energy.

of ($E_B^{H1} = 189 \pm 11$) meV can be determined from a linear fit to the data. The accuracy is mainly limited by the uncertainty of the heavy hole effective mass. For the [100] direction, which is the growth and tunneling direction, the GaAs heavy hole effective mass is $m^* = 0.40 m_0$ [115], with a standard deviation of $0.06 m_0$. A minor limitation of the accuracy is given by the determination of the electric field strength (see Section 5.2.1). From a linear fit of the electron sample data, Fig. 5.13, an electron barrier height of $E_B^{E1} = (271 \pm 24)$ meV is determined. Here, the recommended value for the GaAs electron effective mass [28, 116] of $0.0635 m_0$ was used.³

The hole and electron ground state localization energy is the sum of the barrier height E_B and the estimated energy lowering due to the Poole-Frenkel effect δE : $E_{Loc} = E_B + \delta E$. Hence, the hole and electron localization energies are: $E_{Loc}^{H1} = (210 \pm 20)$ meV and $E_{Loc}^{E1} = (290 \pm 30)$ meV, respectively.⁴

5.2.5 Discussion: Electronic Properties and Storage Times in InAs QDs

The result of the measurements and the derived or estimated values presented are summarized in Table 5.14. It turns out, that the charge carrier emission process from

³The effective mass of the GaAs barrier is slightly reduced, due to the seed layer below the QDs in sample E1. A simple estimation of the fraction of InAs in the GaAs barrier leads to a reduction of the effective mass to approximately $m_e^* \approx 0.058$, an uncertainty of about nine percent. However, the total accuracy of the tunneling barrier height (including the linearisation uncertainty) is only five percent, due to the error propagation theorem.

⁴If a very poor accuracy of the Poole-Frenkel energy of about 50 percent is assumed (due to the limited knowledge of the QD height), the total accuracy of the localization energy is in any case better than eight percent.

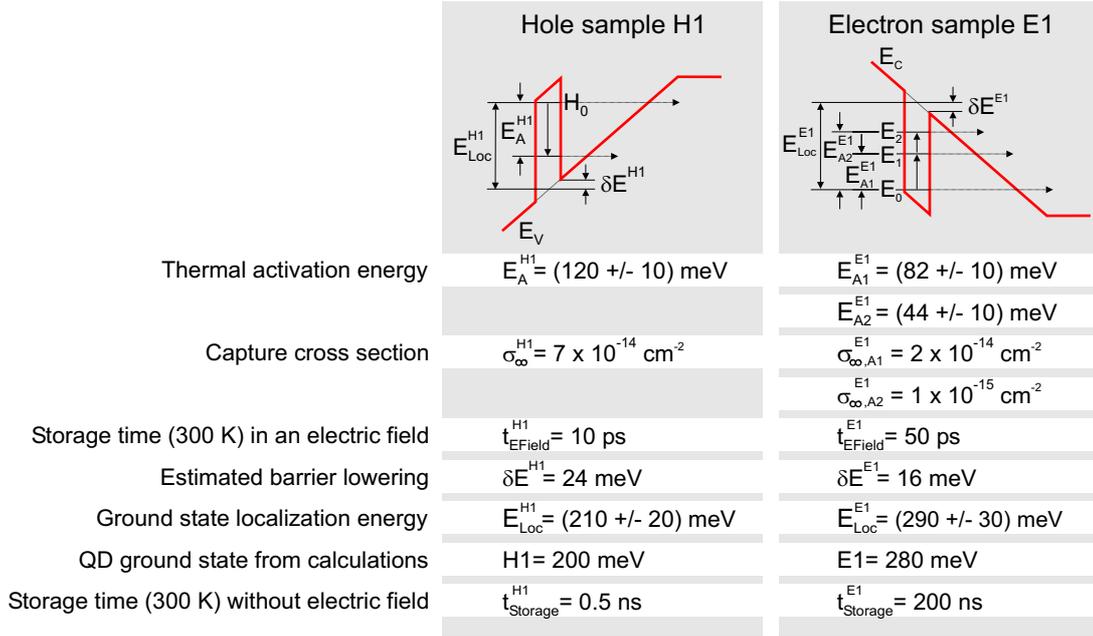


Figure 5.14: Summary of the results from the charge-selective DLTS and Tunneling emission experiments. The calculated values for the electron and hole ground state are taken from [48], see Fig. 2.7.

self-organized QDs in an electric field is always controlled by phonon-assisted tunneling. The influence of the tunneling part, however, depends strongly on the effective mass and the strength of the electric field. Again, the electric field strength is mainly attributed to the applied bias and the doping concentration. Previous DLTS measurements on hole emission from InAs QDs [54] showed a negligible tunneling contribution for a low doping concentration in the order of $6 \times 10^{15} \text{ cm}^{-3}$. For the electron emission, a strong tunneling influence was observed previously [66] and is confirmed here, due to the smaller effective mass. For InAs and GaAs the heavy hole effective mass is typically by a factor of 10 larger than the electron effective mass.

In this work, the tunneling contribution has a significant influence on the observed activation energies for both electron and hole emission. The hole tunneling is strongly increased due to the higher doping concentration in the order of $3 \times 10^{16} \text{ cm}^{-3}$, leading to a larger electric field in the QD layer. The DLTS experiments on a InAs/GaAs QD sample with very low doping concentration of about $6 \times 10^{15} \text{ cm}^{-3}$ (Section 5.4) show also a very small influence of the tunneling emission.

The hole ground state activation energy $E_A^{H1} = (120 \pm 10) \text{ meV}$ is only an underestimated value for the localization energy. The hole localization energy was determined by using the TRTCM method to $E_{\text{Loc}}^{H1} = (210 \pm 20) \text{ meV}$. The 8-band k-p theory [48] predicts a hole ground state localization energy of $\sim 200 \text{ meV}$ for QDs having a base-length of about 16 nm (as in sample H1). Therefore, the experimental value for the hole localization energy is in very good agreement with the theoretical prediction.

For the electrons, two contributions to the DLTS signal were observed: A DLTS signal having an activation energy of $E_{A1}^{E1} = (82 \pm 10)$ meV, which is in good agreement with the theoretically predicted value for the ground to first excited state energy splitting (70 meV). This a well-known value, already obtained previously from DLTS measurements [67] and admittance spectroscopy measurements [117]. Furthermore, a DLTS signal with a smaller activation energy of $E_{A2}^{E1} = (44 \pm 10)$ meV is observed. This value is attributed to the first/second excited state energy splitting and in satisfying agreement with the theoretically predicted value of 50 meV.

The electron ground state localization energy was not accessible with conventional DLTS experiments, due to the higher tunneling probability. This value is determined here for the first time by the TRTCM method to $E_{Loc}^{E1} = (290 \pm 30)$ meV. The theoretical value for the electron localization energy from 8-band k·p calculations is in the order of 280 meV (for QDs with a base length of about 18 nm). Therefore, the electron ground state localization energy obtained experimentally from the TRTCM method is also in very good agreement with the theoretical prediction.

The important storage time for *thermal excitation* from the electron/hole ground states can be estimated by using Equation (2.3), the capture cross-sections and the localization energies for sample E1/H1, respectively (Table 5.14). For zero electric field, an average storage time at room temperature for electrons/holes of about 200 ns/ 0.5 ns is obtained, respectively. That means, even without an applied electric field, where only thermally activated emission occurs, InGaAs/GaAs QDs do not exhibit sufficiently long storage times at $T = 300$ K to act as building blocks in future semiconductor memories. Different materials for the QDs and surrounding matrix must be taken into consideration.

5.3 Hole Storage in GaSb/GaAs Quantum Dots

Type-II heterostructures (e. g. GaSb/GaAs or InSb/GaAs) are expected to have large band discontinuities. QDs consisting of these material combinations are supposed to have strong confinement and large localization energies [118–122], therefore being very attractive for future memory applications. This section confirms the expectation of a strong hole localization in type-II GaSb/GaAs QDs by charge-selective DLTS measurements.

After a characterization of the samples in Section 5.3.1, the results of the C-V and the charge-selective DLTS measurements are presented in Section 5.3.2 and 5.3.3, respectively. The outcome of the investigations concerning localization energy and storage time at room temperature is discussed in Section 5.3.4.

5.3.1 Sample Structures

Two different samples, both n⁺-p diode structures, are investigated. The first, referred to as the "QD sample", contains a layer of GaSb QDs created by deposition of nominally 2.8 ML of GaSb. The second sample has an identical structure, except for the amount

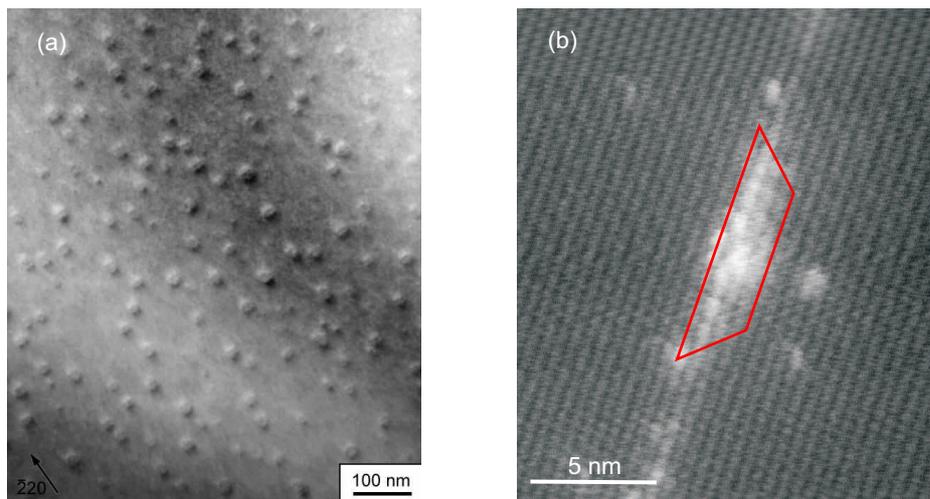


Figure 5.15: (a) Plan-view TEM dark-field image of GaSb QDs after a growth interruption time of 5 s [123]. (b) Cross-sectional STM image of an GaSb QD sample [124].

of GaSb deposited (nominally 2.0 ML), and only contains the wetting layer. The second structure is referred to as "WL sample".

The diode structures were grown in the following layer sequence, see Fig. 5.16: On top of a semi-insulating GaAs substrate, a 500 nm thick highly p-doped contact layer ($p = 1 \times 10^{18} \text{ cm}^{-3}$) and a 700 nm thick p-doped GaAs layer ($p = 3 \times 10^{16} \text{ cm}^{-3}$) were deposited. Subsequently, the GaSb layer embedded between two 10 nm undoped GaAs spacers was grown, followed by 500 nm p-doped GaAs ($p = 3 \times 10^{16} \text{ cm}^{-3}$). Finally, 400 nm highly n-doped GaAs ($n = 7 \times 10^{17} \text{ cm}^{-3}$) was deposited. Employing standard optical lithography and chemical wet-etching mesa structures with a diameter of $800 \mu\text{m}$ were created. Ohmic contacts were formed on the top and the back of the structures by metal evaporation and subsequent alloying. For details of the sample processing see Appendix A.2.

GaSb/GaAs QD samples, grown on identical conditions, were structural characterized by plan-view dark-field TEM [123], see Fig. 5.15(a). In addition, cross-sectional STM [124] was performed on different MOCVD-grown GaSb/GaAs QDs for comparison, displayed in Fig. 5.15(b). From the TEM images, an area density of about $3 \times 10^{10} \text{ cm}^{-2}$, an average QD height of about 3.5 nm, and an average base width of about 26 nm, is determined. The cross-sectional STM image shows a significantly smaller QDs size. The base width in these images is around 8 nm, and the shape of the QD is rather flat with the largest lateral extension just at the base. In [001] growth direction, the bright contrast extends over four atomic chains, indicating that the height of the dot amounts to about 2 nm.

A standard photoluminescence (PL) measurement at low temperature $T = 10 \text{ K}$, using the 514.5 nm line of a Ar^+ -laser as excitation source with a power density of about 500 W/cm^2 on the sample surface, was used to measure spectra of the QD and the WL sample, displayed in Fig. 5.17. In both samples, the peaks of the GaAs bulk

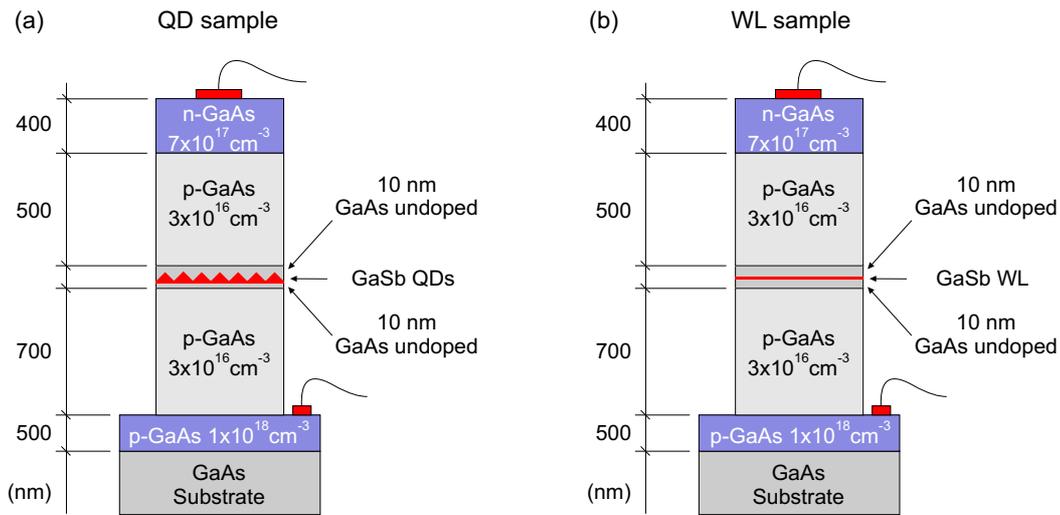


Figure 5.16: Schematic layer structure of the QD sample (a) and the WL reference sample (b).

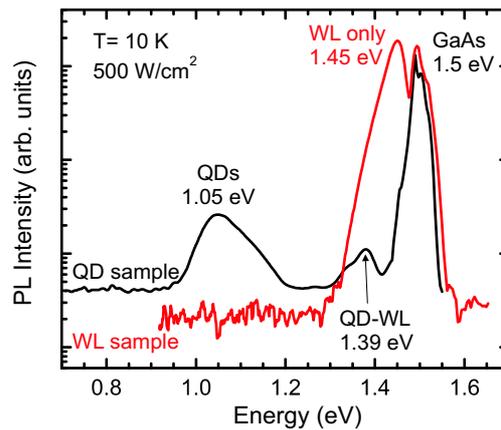


Figure 5.17: Low-temperature photoluminescence spectra with an excitation density of 500 W/cm² of the GaSb/GaAs QD sample (black line) and the WL reference sample (red line). The data is displayed with offset for clarity.

exciton and its phonon replica appear at about 1.5 eV. In the WL sample, a strong peak at 1.45 eV is due to the recombination of electrons with the two-dimensionally confined holes in the GaSb quantum well. The reference WL sample does not show any further luminescence.

In the QD sample, a similar but weaker peak is present at 1.39 eV, which again is due to the recombination in the wetting layer. Furthermore, a rather broad but pronounced signal from the recombination of electrons in the conduction band with confined holes in the GaSb QDs appears at 1.05 eV. This peak has a full width at half maximum (FWHM) of about 80 meV. For a type-II QD system the electron localization can be neglected (see Section 2.2.2), therefore, the hole ground state energy is approximately

given by the energy difference between the GaAs bulk exciton and the QD ground state transition. A hole ground state energy with respect to the GaAs valence band of about 450 meV can be estimated from the QD maximum at 1.05 eV.

5.3.2 Capacitance-Voltage Measurements

In order to determine the appropriate bias conditions for the DLTS experiments, static C-V characteristics were performed at both samples at different temperatures and measurement frequencies, as already described in Section 5.1.2 for the InGaAs/GaAs QD samples. The C-V curve of the QD sample shows also a significant dependence on frequency and temperature, since at high frequencies and low temperatures the localized charge carriers can not follow the applied AC measurement frequency. The optimum condition in order to resolve the full structure were determined to be at about 10 kHz for a temperature of about 300 K. Two representative C-V curves for the QD and the WL sample are displayed in Fig. 5.18.

The C-V characteristics of the QD sample reveals a plateau for a reverse bias between 4.5 V and 10 V (solid line in Fig. 5.18) due to the additional hole states in the sample. At a reverse bias of about 4.5 V the depletion region of the n^+ -p diode reaches the QD layer and is pinned there. As a consequence the capacitance remains roughly constant for increasing reverse bias until all charge is completely removed from the QDs at about 10 V [125]. The voltage range ΔV of the plateau in the C-V curve allows to estimate the charge Q accumulated in the insertion layer. Taking the QD area density into account, a charge accumulation of about 15 holes for completely filled QDs is derived. In contrast, the C-V data of the WL sample (red dashed line in Fig. 5.18) follows almost the common bulk dependence for constant doping concentration $C(V) \sim (V_b + V_a)^{-1/2}$ on the external reverse bias V_a , cf. Equation (4.25).

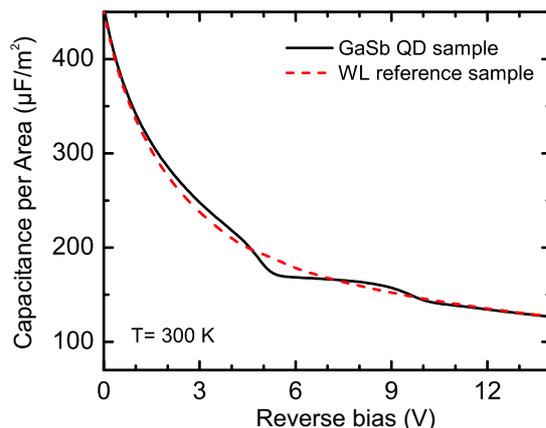


Figure 5.18: C-V characteristics at $T= 300$ K of the QD (black line) and WL sample (red dashed line) for a measurement frequency of 10 kHz. The plateau structure in the C-V of the QD sample is due to hole accumulation in the QDs. The plateau in the C-V starts at a reverse bias of 4.5 V and ends at about 10 V.

5.3.3 DLTS Measurements

In this section, the time-resolved capacitance measurements (the conventional and the charge-selective DLTS) of the GaSb QD sample is summarized. Typical transients of 300 ms duration were sampled with a resolution of 2000 points or more, yielding a minimum time resolution of 150 μs , at temperature steps of about 1.5 K. The data was subsequently evaluated by the double-boxcar method, see Section 4.3.1.

Multiple Hole Emission

In Fig. 5.19, conventional DLTS measurements of the QD and WL sample for a reverse bias $V_r = 10$ V and a pulse bias $V_p = 4$ V are displayed. The pulse length was 10 ms. For these conditions, the QDs are completely filled during the bias pulse and release all trapped holes after the pulse.

The DLTS spectrum of the QD sample, Fig. 5.19(a), shows three maxima, in the following denoted by "A", "B" and "C". Peak C appears in the DLTS spectrum of the WL and the QD sample. From an Arrhenius plot, an activation energy $E_A^C = (530 \pm 20)$ meV and a capture cross-section $\sigma_\infty^C \approx 6 \times 10^{-16}$ cm² is obtained for peak C. This signature can be identified as the bulk hole trap H3 in p-type GaAs [126].

Peak A and B, observed at about $T = 140$ K and 230 K in the QD sample, are obviously related to QD formation. An activation energy of about 600 meV is found for peak B. This relatively high activation energy is comparable to the unstrained GaSb/GaAs valence band offset [118,127] and could be explained by the existence of relaxed GaSb islands in the QD sample.

Peak A in the DLTS spectrum of the QD sample in Fig. 5.19(a) is attributed to hole

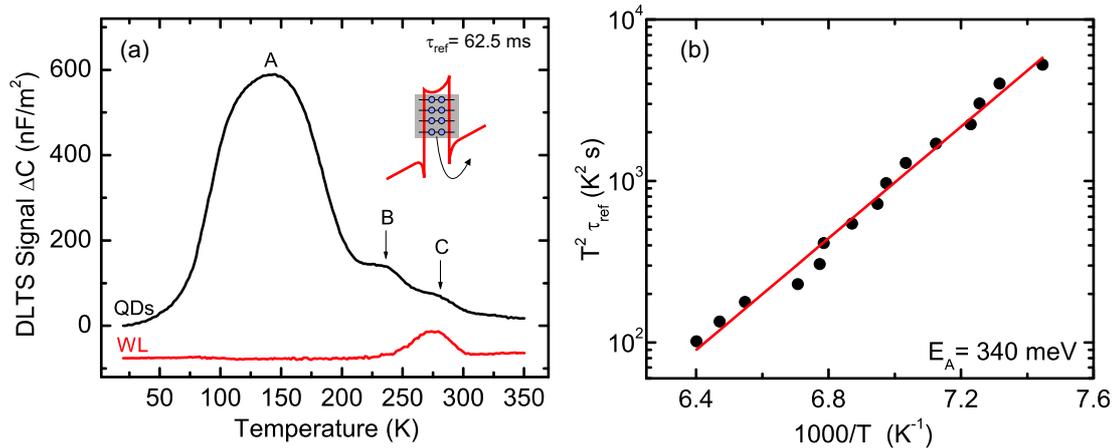


Figure 5.19: (a) DLTS signal of the QD and WL sample, measured at a reverse bias $V_r = 10$ V and a pulse bias $V_p = 4$ V with a filling pulse width of 10 ms and a reference time constant $\tau_{\text{ref}} = 62.5$ ms. The data of the WL sample is displayed with a negative offset. (b) The corresponding Arrhenius plot for peak A, which is related to multiple hole emission from the GaSb QD states. An average activation energy of $E_A = 340$ meV is obtained.

emission from the GaSb QD states, where the activation energy of each successively emitted hole depends on the actual charge state. The DLTS peak is broadened as previously observed for multiply charged Ge/Si QDs [105]. From a standard Arrhenius plot in Fig. 5.19(b), an activation energy of $E_A = 340$ meV is obtained. Consequently, this activation energy of peak A represents only an average value for hole emission from completely charged QDs, where 15 holes from different states with different confinement are involved.

Charge-Selective DLTS

In order to study the charge states in more detail, the QD sample was studied with the charge-selective DLTS method. The pulse bias was always set to $V_p = V_r - 0.5$ V, while the reverse bias was increased from 4.5 V up to 9.5 V. Therefore, narrow peaks appear in the spectra in Fig. 5.20. At $V_r = 4.5$ V the DLTS spectrum shows a maximum at about 80 K while for increasing reverse bias, the DLTS peak shifts to higher temperature. The activation energy increases accordingly from 150 meV at 4.5 V to 450 meV for 9.5 V in Fig. 5.21(a), obtained by standard Arrhenius plots.

The DLTS spectra for a reverse bias between 4.5 and 9.5 V in Fig. 5.20 are attributed to the hole emission from differently charged QDs. All these spectra exhibit a maximum

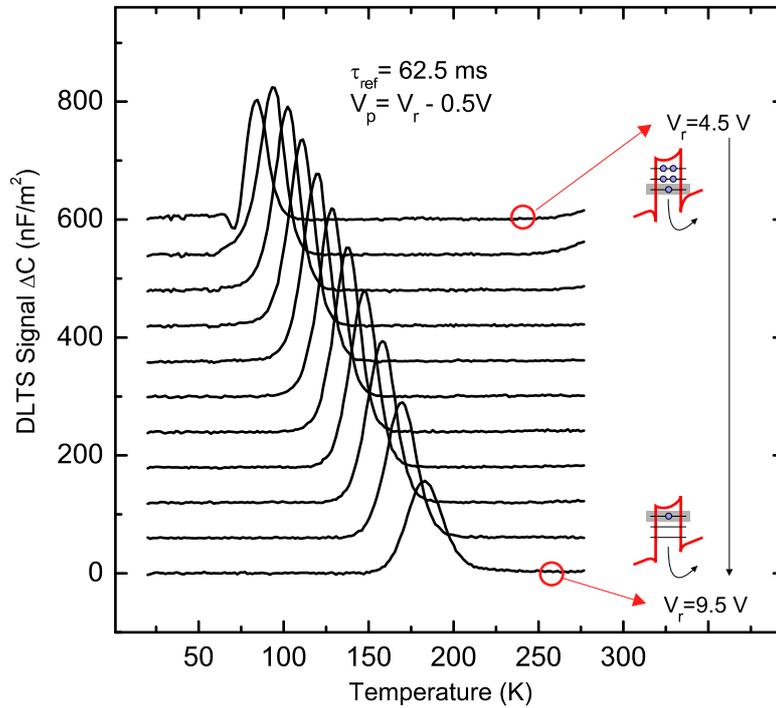


Figure 5.20: Charge-selective DLTS spectra of hole emission from GaSb/GaAs QDs for a reference time constant of $\tau_{\text{ref}} = 62.5$ ms. The reverse bias V_r is increased from 4.5 V up to 9.5 V while the pulse height is fixed at 0.5 V for all spectra. The pulse width is 10 ms and the data is displayed vertically shifted for clarity.

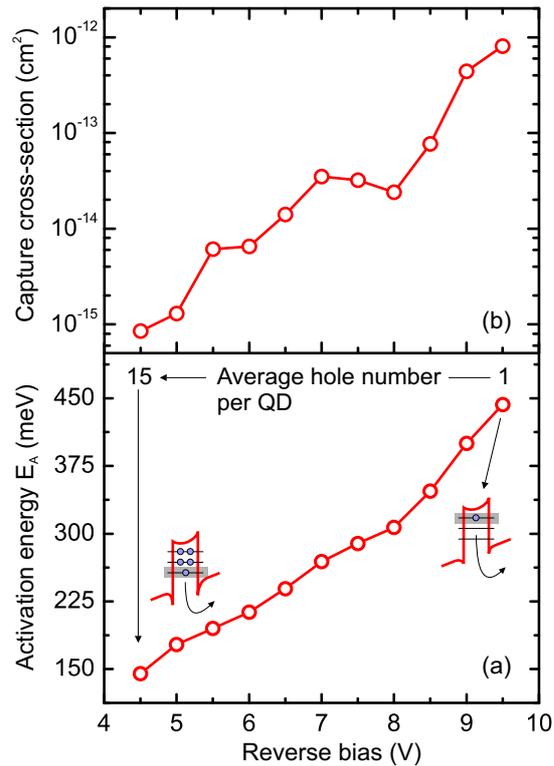


Figure 5.21: Dependence of the activation energy E_A and the capture cross-section σ_∞ on the reverse bias are displayed in panel (a) and (b), respectively. For a reverse bias of $V_r = 9.5$ V hole emission from the QD ground states is probed.

in the temperature range between 80 K and 180 K, the range covered by peak A in Fig. 5.19(a). The maximum activation energy of 450 meV represents the QD hole ground state energy and is in very good agreement with the ground state transition obtained from the PL measurement in Fig. 5.17. The results indicate that the vertical confinement by the GaAs matrix controls the hole escape from the QD layer. The decrease in the activation energy from 450 meV down to 150 meV corresponds to an increase in the average occupation of the QDs. With increasing amount of charge in the QDs, state filling lowers the thermal activation barrier and different QD states are probed, starting at the ground state and ending at that state, where the QD are completely charged.

The capture cross-section σ_∞ increases from $\sigma_\infty \approx 1 \times 10^{-15} \text{ cm}^2$ at $V_r = 4.5$ V up to $\sigma_\infty \approx 1 \times 10^{-12} \text{ cm}^2$ at $V_r = 9.5$ V, see Fig. 5.21(b). This behaviour can be qualitatively explained by an increase of a repulsive Coulomb potential for an increasing confined charge in the QDs. At a reverse bias of 4.5 V the hole emission of completely charged QDs with an average amount of about 15 holes is investigated. A strong Coulomb potential with a capture barrier for other free holes is present and the effective capture cross-section is quenched. For increasing reverse bias, the amount of charge per QD and, hence, the strength of the Coulomb potential decreases and consequently the

magnitude of the capture cross-section increases.

5.3.4 Summary and Discussion

Numerical calculations of the QD states and the according wavefunctions were performed by A. Schliwa using 8-band k-p theory and considering the effects of strain and piezoelectricity. QDs with a geometry of a truncated pyramid were calculated, shown in Fig. 5.22. The width and the height of the pyramid was 22 nm/4.4 nm (shape A), 25 nm/2.2 nm (shape B), and 13 nm/2.2 nm (shape C), respectively. The optimal agreement with a ground state localization energy of 450 meV is obtained for shape C in Fig. 5.22(a). A base width of only 13 nm and a QD height of 2.2 nm predicts a hole ground state energy H0 of 445 meV. The structural parameters are in sufficient agreement with the cross-sectional STM measurements on similar GaSb QDs, showing a base width of about 8 nm and a height of 2 nm in Fig. 5.15(b). In addition, the calculations predict a quantization energy of about 55 meV. The probability density of the hole states for shape B and C are depicted in Fig 5.22(b). The H0 state has a s-like structure, since no nodal planes are visible. The excited states H1 and H2 resembles p-orbitals having one nodal plane perpendicular to each other. The degeneracy between these two states is lifted due to piezo-electricity. The excited state H3 has two nodal planes.

The results obtained from the experiments and the numerical calculations are summarized in Fig. 5.23. The barrier lowering due to the Poole-Frenkel effect is neglected, as it plays a minor role for a localization energy of 450 meV.⁵ Phonon-assisted tunnel-

⁵For an estimated electric field in the order of 100 kV/cm and a QD base width of about 13 nm an

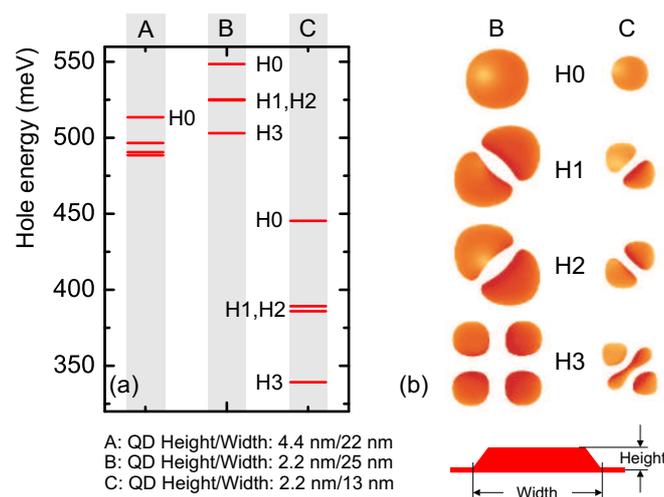


Figure 5.22: (a) Scheme of the hole levels for GaSb/GaAs QDs. The lowest four hole states were calculated by eight-band k-p theory from A. Schliwa. The calculation was done for QDs with different height and base width. (b) Corresponding isosurface plots for 65% of the total probability density for the energy scheme B and C.

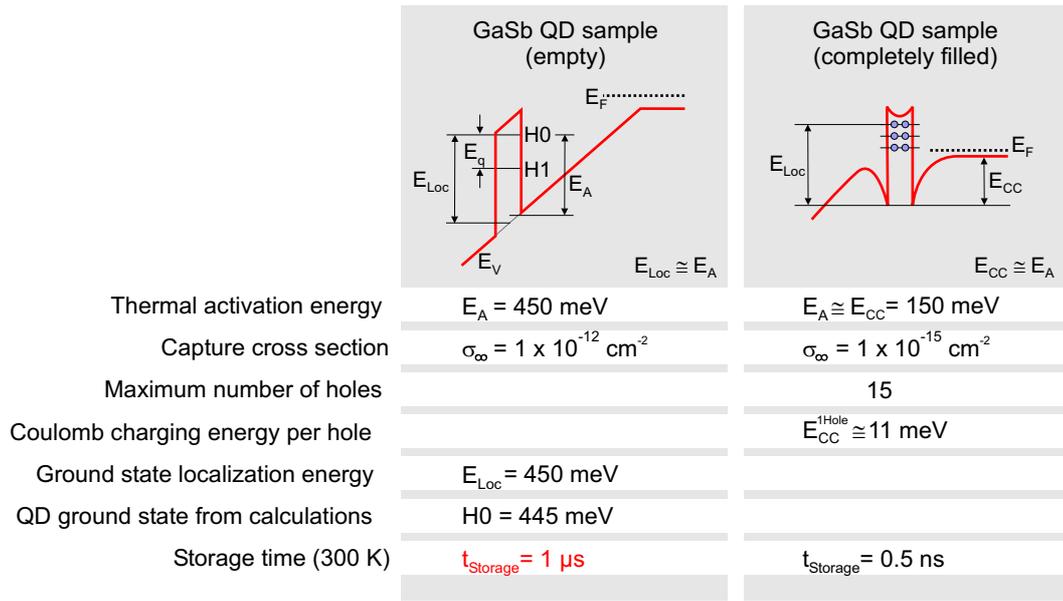


Figure 5.23: Summary of the results from the charge-selective DLTS experiments on GaSb/GaAs QDs and numerical calculations done by A. Schliwa.

ing has not been taken into account for hole emission from GaSb QDs, i. e. the thermal activation energy E_A is assumed to be almost equal to the localization energy E_{Loc} .⁶

The Coulomb charging energy per hole $E_{CC}^{1\text{Hole}}$ can be estimated from the activation energy of completely charged QDs $E_A = 150 \text{ meV}$. The completely charged QDs are filled with 15 holes up to the Fermi level at the valence band edge, whereas the barrier height is generated by the entire Coulomb charging energy of 150 meV, see schematic inset on the upper-right corner in Fig. 5.23. The Coulomb charging energy per hole is therefore approximately $E_{CC}^{1\text{Hole}} \approx 11 \text{ meV}$.

In order to compare the emission time constants for GaSb/GaAs with InGaAs QDs, the observed emission rates were extrapolated to room temperature. A time constant for the thermally activated hole emission process of about $1 \mu\text{s}$ for the localized holes with the localization energy of 450 meV can be estimated. GaSb QDs exhibit a hole retention time at room temperature, that is about three orders of magnitude longer than InGaAs/GaAs QDs. With increasing amount of charge carriers inside the QDs, state filling lowers the thermal activation barrier down to 150 meV and accordingly a storage time of about 0.5 ns at room temperature is estimated.

In conclusion, the carrier emission from multiply charged GaSb QDs was studied. It was found, that these QDs can be charged with up to 15 holes. A ground state activation energy of 450 meV was determined, which accounts for a room temperature emission time in the order of a microsecond. The ground state localization is about

Pool-Frenkel energy of only $\sim 7 \text{ meV}$ is obtained.

⁶The PL measurement confirms this assumption, as a localization energy of 450 meV is in very good agreement with this optical measurement in Fig. 5.17.

twice as large and the retention time at room temperature is about three orders of magnitude longer than in InGaAs QDs. With increasing number of holes per QD, a significant decrease of the activation energy and capture cross-section is observed. This behaviour reflects the strong influence of many-particle effects on the electronic properties of the QDs. Based on the experimental results, GaSb/GaAs QDs are very promising for future research directed towards memory applications.

5.4 InAs/GaAs Quantum Dots with AlGaAs Barrier: Hole storage

To increase the storage time further, besides type-II systems, material combinations of QD and matrix having a large difference in the bandgap energy are promising. Hole storage in InAs/GaAs QDs with an additional AlGaAs barrier is presented in this section. The additional AlGaAs barrier increases the activation energies sufficiently to reach a retention time of milliseconds at room temperature. Instead of using it as a barrier, doped AlGaAs could consequently be used as the surrounding matrix material.

Two samples were investigated, both characterized in Section 5.4.1 and studied with C-V measurements in Section 5.4.2. Sample B contains InAs QDs embedded in a GaAs matrix and an additional AlGaAs barrier. Sample A is a reference sample with only InAs QDs and without AlGaAs barrier. The DLTS measurements of sample A are discussed in Section 5.4.3, while the results of the sample B (with the AlGaAs barrier) are presented in the following Section 5.4.4. The outcome of the investigations concerning the storage times at room temperature and localization energies are summarized and discussed in Section 5.4.5.

5.4.1 Sample Structures

Before the sample growth, the following considerations have been made in order to optimize the charge-selective DLTS measurements. Since the DLTS is only sensitive to the thermal emission processes, the tunneling emission should be reduced. A small electric field in the p-n diode structure causes a small tunneling probability. A very low doping concentration in the matrix material surrounding the QDs is used ($2 \times 10^{15} \text{ cm}^{-3}$) to obtain a significant reduction of the tunneling emission. To increase the storage time at room temperature, an additional $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}$ barrier below the QD layer is incorporated and a sample without an AlGaAs barrier is used as reference. In addition, the influence of DX centers in AlGaAs on the carrier dynamic can be neglected for the investigated sample with nominally undoped AlGaAs in a p-doped GaAs matrix [128,129]. This enables to unambiguously identify the hole emission from the QD states in sample B.

Both samples are n^+p diode structures, grown by molecular beam epitaxy (MBE) at the Ioffe Institute in St. Petersburg. The schematic layer structure of both samples is shown in Fig. 5.24. The first, referred to as "Sample A" is the reference sample without the AlGaAs barrier and consists of a semi-insulating GaAs substrate, followed

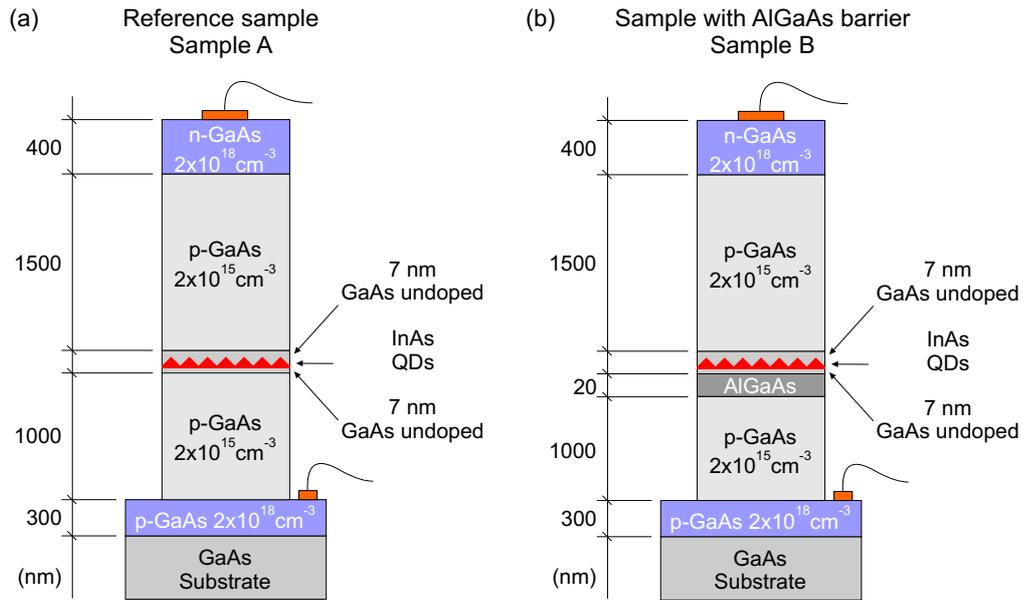


Figure 5.24: (a) Schematic layer structure of reference sample A without the barrier. The layer structure of the reference sample B containing a layer of InAs/GaAs QDs and an additional $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}$ barrier is shown in panel (b).

by a 300 nm thick highly p -doped GaAs contact layer ($p = 2 \times 10^{18} \text{ cm}^{-3}$) and a 1000 nm thick p -doped GaAs layer ($p = 2 \times 10^{15} \text{ cm}^{-3}$). Subsequently, a layer of InAs QDs (nominally ~ 2.2 ML) was deposited with a growth rate of about 0.1 ML/sec at 485 °C. The QD layer is sandwiched between two 7 nm thick undoped GaAs spacers. After deposition of 1500 nm p -doped GaAs ($p = 2 \times 10^{15} \text{ cm}^{-3}$), 400 nm of highly n -doped GaAs ($n = 2 \times 10^{18} \text{ cm}^{-3}$) was grown to form the n^+p diode structure.

The second sample, referred to as "Sample B" is the QD sample with an additional AlGaAs barrier. It has an identical layer structure besides a nominally undoped $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}$ barrier of 20 nm thickness below the QD layer, depicted in Fig 5.24(b). Mesa structures with diameters of 800 μm were created by employing standard optical lithography and chemical wet-etching. Finally, Ohmic contacts were formed by metal evaporation and subsequent alloying, see Appendix A.2.

5.4.2 Capacitance-Voltage Measurements

In order to determine the appropriate bias condition for the DLTS experiments, static C-V characteristics were performed at both samples at different temperatures and measurement frequencies, as already described in Section 5.1.2 and 5.3.2. The optimum condition to resolve the full structure were determined to be at about 5 kHz at $T = 160$ K for sample A and 5 kHz at $T = 350$ K for sample B. Two representative C-V curves for sample A and B are displayed in Fig. 5.25.

A plateau due to hole accumulation in the InAs QDs is clearly visible in both samples. The maxima of the second derivative d^2C/dV^2 , dashed line in Fig.5.25, precisely

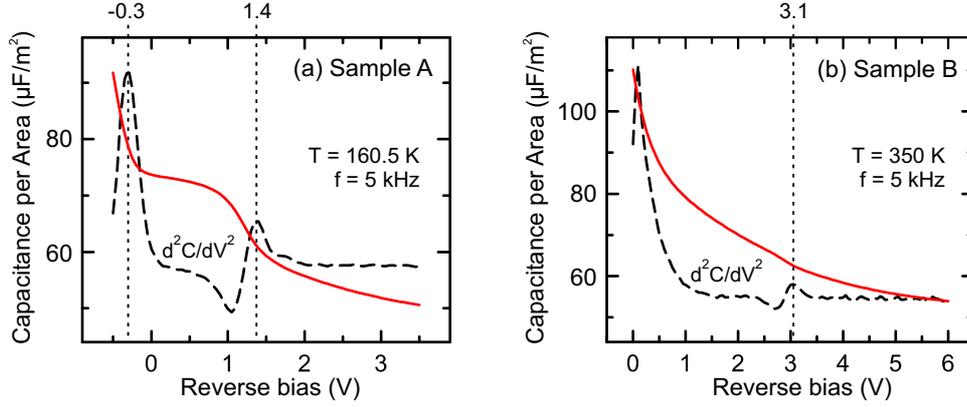


Figure 5.25: (a) Static C-V characteristics of sample A, containing InAs QDs and (b) sample B containing InAs QDs with an additional AlGaAs barrier. The C-V measurements are taken at different frequencies and temperatures (see labels). The plateau in the C-V traces are due to hole accumulation in the QD layer. The maxima of the second derivative d^2C/dV^2 (dashed line) precisely indicate the beginning and end of the plateaus.

indicate the beginning and end of the plateaus, where the QDs are completely charged with holes or completely empty. The end of the plateau for sample B exists at a reverse bias $V_r = 3.1$ V, whereas the beginning is not clearly visible. The reference sample B, Fig. 5.25(b), exhibits a plateau from $V_r = -0.3$ V⁷ up to 1.4 V.

5.4.3 DLTS Measurements: Reference Sample A

Firstly, the DLTS measurements of the reference sample A are discussed in this section, which will be compared with the outcome of sample B in the following Section 5.4.4.

Using the results from the C-V measurement of sample A, the reverse and filling pulse bias were chosen, such that hole emission from completely charged QDs is observed in a conventional DLTS experiment. Figure 5.26(a) shows a DLTS spectrum for a pulse bias of $V_p = -0.5$ V (QDs completely filled) and a reverse bias of $V_r = 1.6$ V (QDs are completely empty). The reference time constant was set to 5 ms and the pulse duration to 1 s. A broadened maximum at 80 K represents the hole emission from completely filled QDs of the entire ensemble. An average activation energy of 116 meV is obtained from an Arrhenius plot. It is worth noticing here, that no further DLTS signal related to unwanted deep levels or defects appears in the spectrum up to 350 K, that means, the crystal quality of the samples seems to be almost perfect.

Figure 5.26(b) shows charge-selective DLTS spectra of the reference sample A. The pulse bias height was fixed to 0.2 V for all spectra (i.e. $V_p = V_r - 0.2$ V) in order to observe on average the emission of less than one hole per QD. As already described, the emission from differently charged QDs, i.e. different energy levels is observed by changing the reverse bias. A pronounced DLTS peak at 90 K for a reference time

⁷Negative signs for the reverse/pulse bias means positive voltage in forward direction of the p-n diode.

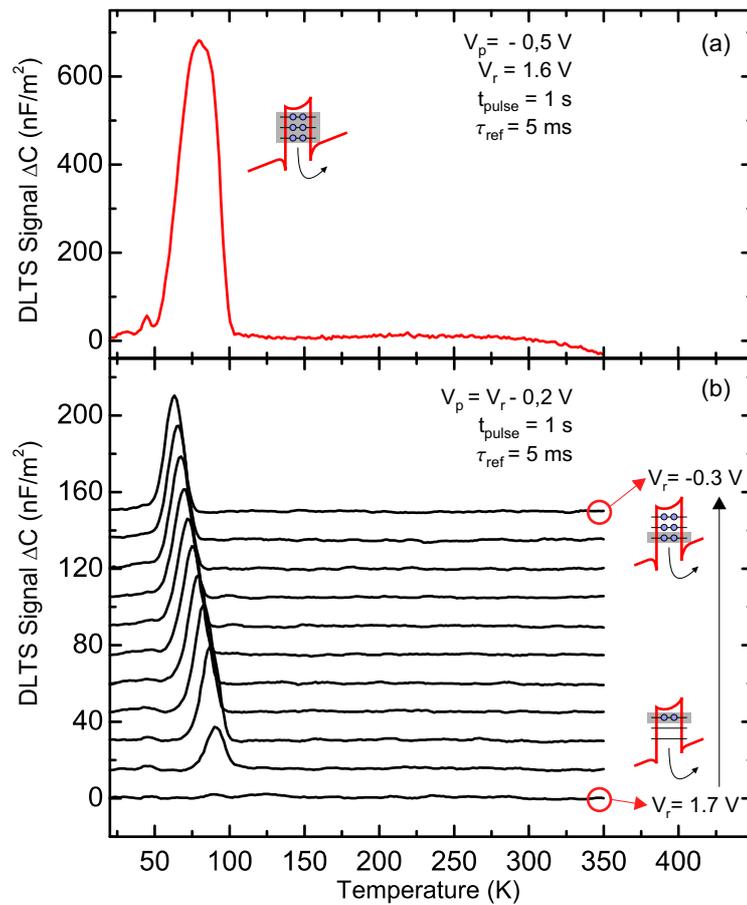


Figure 5.26: (a) Conventional DLTS signal of the reference sample A, measured at a reverse bias $V_r = 1.6$ V and a pulse bias $V_p = -0.5$ V with a filling pulse width of 1 s and a reference time constant $\tau_{\text{ref}} = 5$ ms. The QDs are completely charged during the filling pulse and multiple hole emission is observed. (b) Charge-selective DLTS spectra for increasing reverse bias V_r from -0.3 V up to 1.7 V. The pulse height is fixed at 0.2 V for all spectra. The pulse width is 1 s and the data is displayed vertically shifted for clarity.

constant of 5 ms and a reverse bias of $V_r = 1.5$ V is visible, which represents hole escape from the QD ground states. A decrease of the reverse bias leads to emission from successively higher QD states and the DLTS signal appears at lower temperatures.

Activation Energies

From Arrhenius plots of the DLTS peak positions, Fig. 5.27(b), thermal activation energies for hole emission from differently charged QDs are obtained. Figure 5.27(a) shows the dependence of the activation energy on the reverse bias for sample A. The highest value of (180 ± 10) meV at $V_r = 1.5$ V is related to thermal activation from the QD hole ground states. For decreasing reverse bias a decreasing activation energy due to QD state-filling is visible. At $V_r = 0.1$ V the QDs are completely filled and between $V_r = 0.1$ V and $V_r = -0.3$ V the activation energy remains roughly constant with a mean

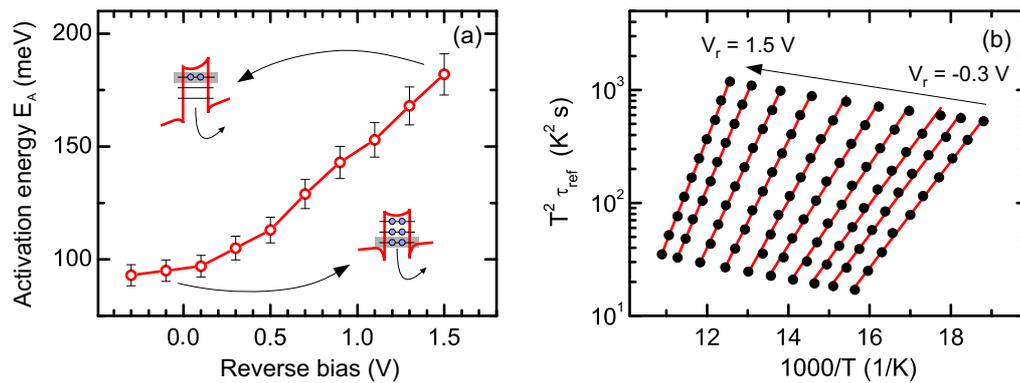


Figure 5.27: (a) Activation energies in dependence of the reverse bias V_r for the charge-selective DLTS spectra in Fig. 5.26 (Sample A). The highest value of 180 meV at $V_r = 1.5$ V is related to the thermal activation from QD hole ground states (inset in the upper left corner). The mean value of 95 meV between 0.1 and -0.3 V represents the Coulomb barrier height for completely charged QDs (inset in the lower right corner). Panel (b) displays the corresponding Arrhenius plots.

value of 95 meV. This energy represents the Coulomb barrier height for completely charged QDs [79, 130], as already described for completely charged GaSb/GaAs QDs in Section 5.3.

Influence of the Electric Field

As previously shown, carrier emission in DLTS experiments always occurs in presence of an electric field where phonon-assisted tunneling occurs. To obtain the localization energy, the barrier height for tunneling through the triangular barrier must be added to the thermal activation energy. However, for this sample the tunneling probability should be strongly reduced due to the low doping concentration. Nevertheless, the barrier height should be estimated using a simple approximation to confirm the assumption.

This tunneling barrier height can be estimated by comparison of the tunneling emission rate, using Equation for the tunneling rate (3.50), and the thermal emission rate, Equation (2.3), for a given barrier height. To first order, tunneling dominates the thermal emission for barrier heights at which the tunneling emission rate is larger than the thermal emission rate.

For the detection of the hole emission from the QD ground states, a reverse bias of $V_r = 1.5$ V was applied at the p-n diode; for the detection of the emission from completely charged dots a reverse bias of $V_r = -0.3$ V. Equation (4.14) for the electric field across the depletion width yields an electric field at the QD layer of 9 kV/cm and 2 kV/cm, respectively. The electric field is strongly reduced in this InAs QD sample in comparison to the samples in Section 5.1, since the doping concentration in the QD surrounding GaAs matrix is more than one order of magnitude smaller. The influence of the electric field on the activation energy is smaller and the tunneling barrier height for the phonon-assisted emission processes is reduced.

A tunneling barrier height of about 25 meV at $V_r = 1.5$ V is obtained from this approximation and the hole localization energy is therefore 205 meV. This value is in good agreement with the theoretical prediction of ~ 200 meV from 8-band k-p theory for QDs with a base-length of about 16-18 nm [48] and tunneling measurements, Section 5.2, yielding a value of 210 meV. A tunneling barrier height of about 7 meV is estimated for the emission from completely charged QDs. The activation energy of 95 meV for a reverse bias of -0.3 V represents the Coulomb charging energy for completely charged QDs, i. e. the Coulomb charging energy can be estimated in a similar way to ~ 100 meV, taking into account the tunneling barrier height.

The Poole-Frenkel energy, which leads to a lowering of the barrier height in an applied electric field, can be neglected in the present case. For a electric field strength of about 9 kV/cm a Poole-Frenkel energy of 3 meV can be estimated for QDs with a height of 6 nm. This value is well within the measurement accuracy.

The storage time for InAs QDs without an AlGaAs barrier can again be extrapolated to room temperature using Equation (2.3), the ground state activation energy and the capture cross section, which are obtained from the Arrhenius plots in Fig. 5.27(b). The storage time at $T = 300$ K for the hole ground state is in the order of nanoseconds, in good agreement with the experiments in Section 3.1.3.

Influence of the Pulse Width

The activation energy and the capture cross-section can significantly depend on the filling pulse duration t_{pulse} , if the QD occupation with charge carriers is temperature-dependent. The QD occupation is controlled by the capture process, cf. Equation (3.1) and (3.2), which again can strongly depend on temperature due to two reasons: (i) The presence of a capture barrier, which leads to a temperature dependence of the capture cross-section (Section 3.1.4) and (ii) a low doping concentration in the matrix material.

A capture barrier is always present in a charge-selective DLTS experiment. A doping concentration in the order of 10^{16} cm^{-3} is sufficiently high, that a filling pulse

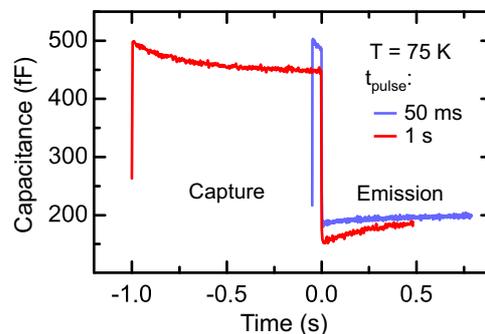


Figure 5.28: Capacitance transients for hole capture and emission at $T = 75$ K and a pulse width of 50 ms (blue line) and 1 s (red line). The reverse bias V_r and pulse bias V_p were set to 1.5 V/1.4 V, respectively.

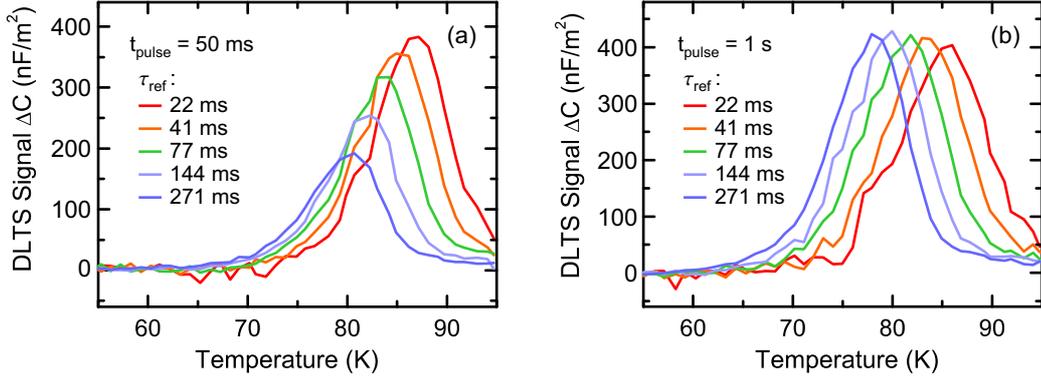


Figure 5.29: Charge-selective DLTS spectra for different reference time constants and a pulse duration of 50 ms (a) and 1 s (b). The reverse bias V_r and pulse bias V_p were always set to 1.5 V/1.4 V, respectively.

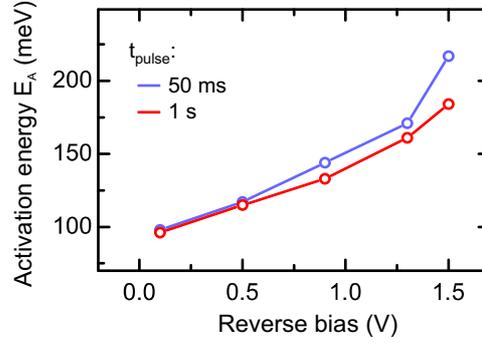


Figure 5.30: Activation energies of charge-selective DLTS for increasing reverse bias between 0.1 and 1.5 V and two different pulse durations. The activation energies for a pulse duration of 50 ms are slightly overestimated.

duration t_{pulse} of 10 ms is enough to reach thermal equilibrium between the QD levels and the surrounding matrix for the temperature range of measurement. As a consequence, the charge-selective DLTS measurements on GaSb/GaAs QDs (Section 5.3) and InGaAs/GaAs (Section 5.1) showed no dependence on the pulse duration. Here, in the experiments on InAs/GaAs QDs the doping concentration is more than one order of magnitude smaller and a strong influence of the pulse width on the activation energies and capture cross-sections is observed.

Figure 5.28 displays two capacitance transients of hole capture and emission from charge-selective DLTS measurements on sample A. The reverse bias V_r and pulse bias V_p were set to 1.5 V/1.4 V, respectively, for both transients. At a temperature $T = 75$ K a pulse duration of 50 ms (blue line) is not sufficient to reach equilibrium and the amplitude of the capacitance transients depends on temperature. A decrease in the DLTS amplitude is clearly visible at lower temperature for an increasing reference time constant, see Fig. 5.29(a).

The activation energies for a filling pulse width of 50 ms are overestimated, see Fig. 5.30, since the maxima of the DLTS spectra in Fig. 5.29(a) for longer reference time constants appear at higher temperatures as expected. To obtain correct activation energies and capture cross-sections, a saturation situation must be ensured in the capacitance transients. A pulse duration of about 1 s (red line in Fig. 5.28) is sufficient here to reach an equilibrium state in the QDs for the considered temperature range. Hence, the amplitude in the DLTS spectra is constant, even at lower temperatures for increasing reference time constants, shown in Fig. 5.29(b). The activation energies are smaller than for the 50 ms pulse duration situation (Fig. 5.30) and match the correct values.

5.4.4 DLTS Measurements: AlGaAs Barrier Sample B

In this section, the results from DLTS measurements on InAs QDs with an additional AlGaAs barrier (sample B) are discussed. Figure 5.31(a) shows the conventional DLTS spectrum, where the QDs were completely filled with holes. Appropriate reverse and pulse bias conditions were obtained from the C-V measurements (Section 5.4.2). The pulse bias was set to $V_p = 0$ V (QDs completely filled) and the reverse bias to $V_r = 3.2$ V (QDs are completely empty). The reference time constant was set to 5 ms and the pulse duration to 1 s. A peak at $T = 150$ K and a broadened maximum from 150 K up to 300 K are visible. The broadened maximum represents the hole emission from completely filled QDs, while the peak is attributed to hole emission over the AlGaAs barrier (see below).⁸ Due to the broadened maximum in a range of more than 150 K an Arrhenius plot cannot be created and thus, it is not possible to determine an average activation energy.

Figure 5.31(b) shows charge-selective DLTS spectra of sample B for increasing reverse bias V_r between 0.2 and 3.4 V. The pulse bias height was fixed to 0.2 V for all spectra (i. e. $V_p = V_r - 0.2$ V) and the pulse width t_{pulse} was set to 1 s. For reverse biases above $V_r = 3.2$ V, no DLTS signal is visible, as the Fermi level is energetically above the QD states. Therefore, no QD states are occupied. By decreasing the reverse bias the Fermi level reaches the QD ground state at $V_r = 3.2$ V and a peak in the DLTS spectrum appears at 300 K. This peak is related to thermally activated hole emission from the QD ground state over the AlGaAs barrier, schematically depicted in the inset in the upper right corner of Fig. 5.31(b). A further decrease of the reverse bias leads to QD state-filling. Successive emission from higher QD states is observed and the maxima in the DLTS spectra appear at lower temperatures. As mentioned before, a peak in the DLTS spectra is observed at that temperature, at which the averaged time constant of the thermally activated emission almost equals the applied reference time constant τ_{ref} . Therefore, the peak at 300 K for $\tau_{\text{ref}} = 5$ ms ($V_r = 3.2$ V in Fig. 5.31(b)) represents an averaged emission time constant (storage time) of 5 ms for hole emission from the QD ground states over the AlGaAs barrier. In this situation, the storage time at

⁸No further DLTS signal related to unwanted deep levels or defects appear in the spectrum below 150 K, that means, the crystal quality of this sample is also extremely good.

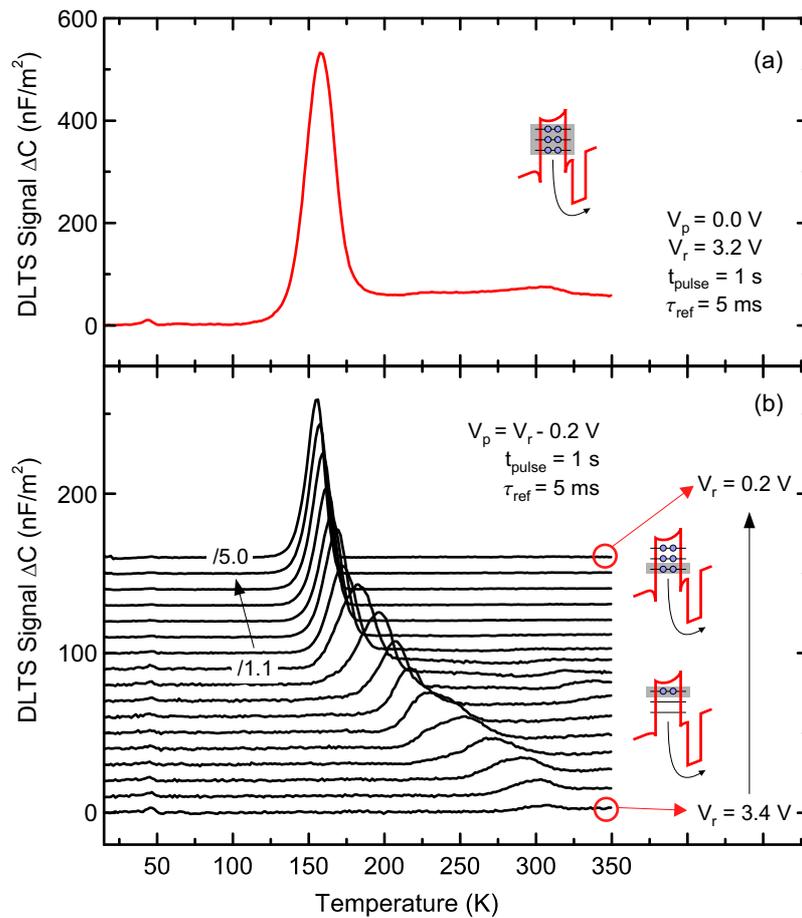


Figure 5.31: (a) Conventional DLTS signal of sample B, measured at a reverse bias $V_r = 3.2$ V and a pulse bias $V_p = 0$ V with a filling pulse width of 1 s and a reference time constant $\tau_{\text{ref}} = 5$ ms. The QDs are completely charged during the filling pulse and multiple hole emission is observed. (b) Charge-selective DLTS spectra for increasing reverse bias V_r from 0.2 V up to 3.4 V. The pulse height is fixed at 0.2 V for all spectra. The pulse width is 1 s and the data is displayed vertically shifted for clarity. Spectra below $V_r = 1.6$ V are divided by a factor from 1.1 up to 5.

room temperature is directly measured in the DLTS experiment and not extrapolated as before.

Activation Energies

From Arrhenius plots of the DLTS peak position for varying reference time constants τ_{ref} (Fig. 5.32), thermal activation energies for hole emission from differently charged QDs are obtained.

The thermal activation energies for sample B are shown in Fig. 5.33.⁹ The highest

⁹The missing three points in Fig. 5.33 are considered as experimental artifacts. An increase of the activation energy with decreasing reverse bias cannot lead to a consistent interpretation of the experimental data.

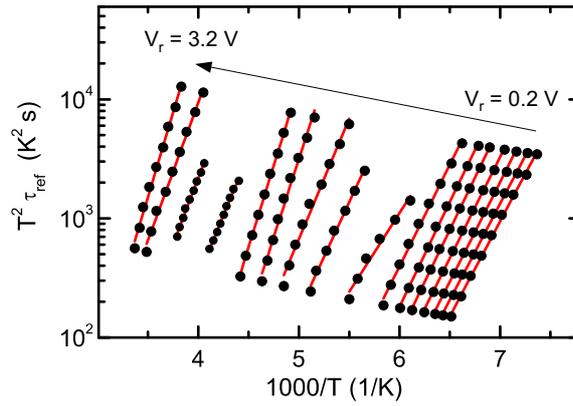


Figure 5.32: Arrhenius plot for the charge-selective DLTS measurements of sample B (InAs QDs with AlGaAs barrier) in Fig. 5.31.

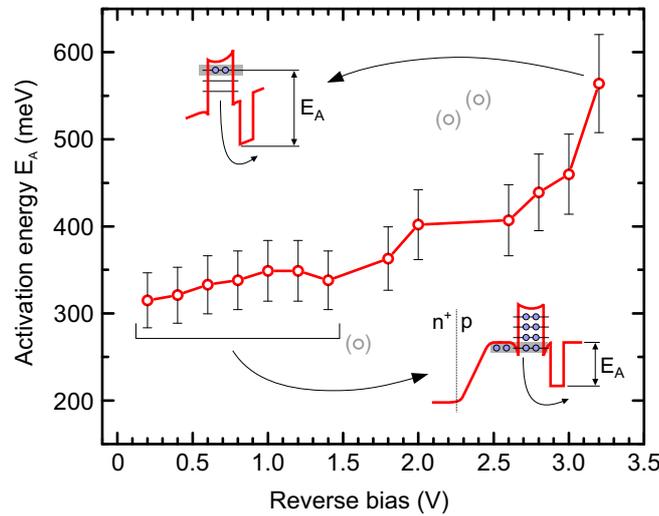


Figure 5.33: Dependence of the thermal activation energy on the reverse bias V_r for sample B (with an AlGaAs barrier). The observed maximum activation energy of 560 meV at $V_r = 3.2$ V corresponds to the thermal activation of holes from the QD ground state over the entire AlGaAs barrier (inset in the upper left corner). The mean value of 340 meV between 1.4 and 0.2 V represents the AlGaAs barrier height (inset in the lower right corner).

value of (560 ± 60) meV at $V_r = 3.2$ V is related to thermal activation from QD hole ground states over the AlGaAs barrier. This value represent in good approximation the hole localization energy plus the barrier height, as the influence of the electric field on the activation energy can be neglected.¹⁰ As in sample A, for decreasing reverse bias a decreasing activation energy due to QD state-filling is observed. At $V_r = 1.4$ V the energetic position of the Fermi level is at the valence band edge and no further

¹⁰Within the accuracy of 60 meV, the influence of tunneling through the AlGaAs barrier on the ground state activation energy of 560 meV can be neglected for an electric field strength of about 10 kV/cm, cf. Section 5.4.3.

QD state-filling is possible. As a consequence, between $V_r = 1.4$ V and $V_r = 0.2$ V carrier emission from the valence band edge is probed and the activation energy remains roughly constant with a mean value of 340 meV.¹¹ This energy represents the AlGaAs barrier height, see schematic inset in the lower right corner of Fig. 5.33.

Comparing the charge-selective DLTS measurements on sample A (without AlGaAs barrier) and sample B (with AlGaAs barrier), the AlGaAs barrier leads to a shift of the DLTS spectra to higher temperatures. The hole emission for the same reference time constant appears at higher temperature as a higher thermal activation energy is needed to overcome the AlGaAs barrier. Consequently, the observed thermal activation energy increases from 205 meV up to 560 meV, where the difference is due to the barrier height.

Influence of the AlGaAs Barrier

The thermal activation energy for hole emission from the ground states of the InAs QDs over the AlGaAs barrier was determined to 560 meV. However, the barrier has a strong impact on the emission and capture processes and also on the determination of the activation energies. Mainly two different influences must be taken into account, otherwise the activation energies will be overestimated, described in the following.

1. Influence of The Pulse Width: The dependence of the activation energy on the pulse duration t_{pulse} has already been discussed in detail in Section 5.4.3. The influence of the pulse width on the activation energy is even enhanced for sample B with an additional AlGaAs barrier and a low doping concentration in the order of 10^{15} cm⁻³. For a short pulse width in the order of milliseconds the activation energy is overestimated up to 200 meV. Therefore, a pulse duration of 20 s was used for the Arrhenius plots in Fig. 5.32 to obtain correct activation energies. The DLTS spectra in Fig. 5.31 were measured with 1 s pulses for a large temperature range between 15 K and 350 K. A measurement with 20 s pulse duration was not done for the DLTS spectra, as it would be too time-consuming.

2. Temperature Dependence of the Equilibrium Occupation: A saturation in the capacitance transients is obtained, if the pulse duration is sufficiently long. Figure 5.34(a) shows the capacitance transient at $T = 200$ K for a pulse width of 20 s of sample B (with AlGaAs barrier). Clearly visible is the saturation of the capacitance during the capture process. However, even if a saturation of the capacitance transients is achieved for the entire temperature range of measurement, the observed amplitude is still temperature-dependent. As a consequence, the activation energies are again overestimated.

¹¹The influence of the Coulomb charging energy on the determination of the AlGaAs barrier height can be neglected. A simple estimation yields 9 meV for completely charged QDs, less than three percent of the total barrier height of 340 meV.

This effect has already been discussed for deep levels in semiconductors. Mainly two reasons are given for the temperature-dependent capacitance amplitude in experiments on deep levels: Lee et al. [131] attributed it to a minority-carrier emission process (maybe due to leakage currents), while Cavalcoli et al. [132] relates it to the presence of a repulsive Coulomb barrier. It is assumed here, that on one hand, the additional AlGaAs barrier leads to a strongly temperature-dependent capture cross-section. On the other hand, even for long pulse durations, the QD states are not in thermal equilibrium with the surrounding matrix and even if the capacitance transient saturates, the QDs are only partly charged with holes. The AlGaAs barrier decouples the QD states from the surrounding matrix and the Fermi level is not constant throughout the entire diode. The Fermi level inside the QDs seems to have a strong temperature dependence in this situation. A more detailed analysis of the “local thermal equilibrium” in the QDs is required to understand the observation in more detail, which is not done here. A microstate model using a master equation approach was already applied to Ge/Si QDs [54] and discussed for photoluminescence experiments [133] to understand the charge carrier population inside the QDs.

The dependence of the capacitance amplitude ΔC_0 on temperature is visible in the amplitude of the DLTS spectra for different reference time constants, Fig. 5.34(b). The solid yellow line displays ΔC_0 as function of temperature. The values are obtained from the measured capacitance transients. The temperature dependence is even present for long pulse durations of 20 s and, hence, not related to the saturation effect of the transients, cf. Fig. 5.34(a).

A simulation of the DLTS spectra has been performed to estimate the influence of this effect on the determination of the activation energies. Figure 5.35(a) shows simulated DLTS spectra for different reference time constants for a constant amplitude

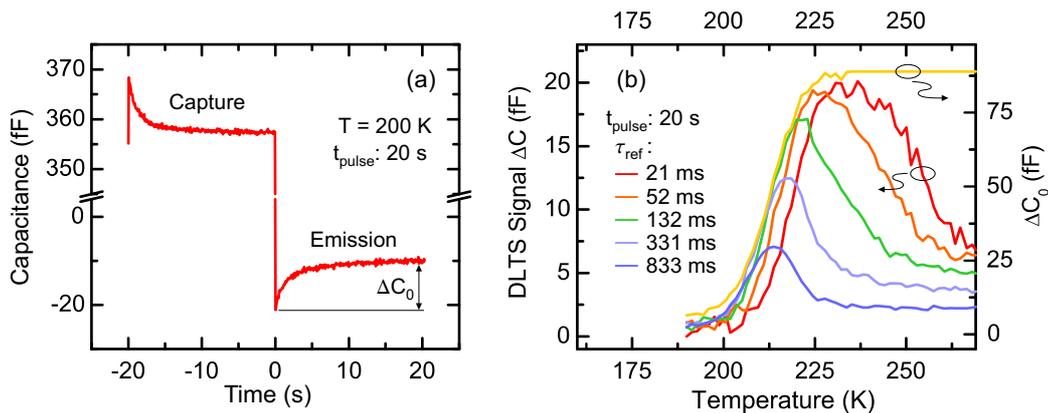


Figure 5.34: (a) Capacitance transient at $T=200$ K for a pulse width t_{pulse} of 20 s. The pulse width is not sufficient to drive the QDs into equilibrium state with the surrounding matrix. (b) Charge-selective DLTS spectra for different reference time constants. The reverse/pulse bias was set to $V_r=2.6$ V and $V_p=2.4$ V, respectively. The decrease of the DLTS amplitude at lower temperatures indicates a temperature dependence of the amplitude of the capacitance transients ΔC_0 .

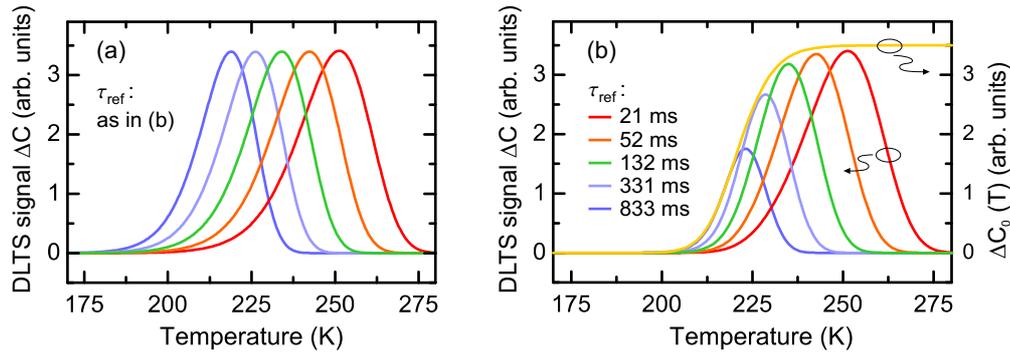


Figure 5.35: (a) Simulation of DLTS Spectra for different reference time constants τ_{ref} and a temperature-independent amplitude of the capacitance transients. (b) Simulation of the DLTS spectra for a temperature-dependent amplitude of the capacitance transients ΔC_0 , solid yellow line.

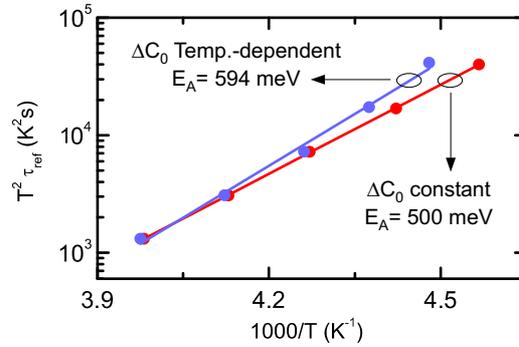


Figure 5.36: Arrhenius plots, obtained from the simulated DLTS spectra in Fig. 5.35. The red solid line yields the correct activation energy of $E_A = 500$ meV for a constant ΔC_0 , while the blue line overestimates the activation energy due to a temperature-dependent ΔC_0 .

ΔC_0 . The constant ΔC_0 leads to a constant DLTS amplitude. DLTS spectra for a temperature-dependent ΔC_0 are displayed in Fig. 5.35(b). The yellow line shows the ΔC_0 function, while the DLTS spectra for different reference time constants reveal almost the experimentally observed behavior. The DLTS amplitude in the simulation decreases for lower temperatures, similar to the measured DLTS spectra in Fig. 5.34.

The DLTS spectra were simulated for the emission from an energy level having an activation energy of 500 meV. The evaluation of the simulated DLTS spectra with an Arrhenius plot yields an activation energy of about 600 meV for a temperature-dependent ΔC_0 (blue solid line in Fig. 5.36). The Arrhenius plot shows no significant deviation from linearity. The Arrhenius plot for a constant ΔC_0 yields, of course, the correct value of 500 meV (solid red line in Fig. 5.36).

The temperature dependence of the amplitude of the capacitance transients $\Delta C_0(T)$ has been taken into account during the evaluation of the DLTS spectra of sample B with the AlGaAs barrier. The amplitude ΔC_0 was obtained experimentally from the transients and the transients were normalized to this amplitude, such that a constant

DLTS peak amplitude results for all reference time constants.

5.4.5 Summary

The experimental results from both samples can now be used to obtain a consistent picture of the energy diagram for the valence band of sample B. Figure 5.37 shows the energy diagram for the valence band.

The ground state hole localization energy is 205 meV. The entire Coulomb charging energy was estimated to ~ 100 meV. Within an accuracy of 60 meV, the influence of tunneling through the AlGaAs barrier on the ground state activation energy of 560 meV is neglected. The barrier height is determined as 340 meV, which is in good agreement with the valence band offset between GaAs and $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}$, shown in Fig. 2.3. By adding the barrier height of 340 meV, the ground state activation energy for the emission over the AlGaAs barrier yields 545 meV. This value is in very good agreement with the measured activation energy of about 560 meV, obtained from the DLTS measurements on sample B. The activation energy for the QD ground states increases from 205 meV up to 560 meV, due to the additional AlGaAs barrier. Accordingly, the hole storage time at room temperature increases by six orders of magnitude from nanoseconds up to milliseconds.

5.5 Storage Time in Quantum Dots

In Chapter 5, carrier emission from different QD systems was studied in order to determine the carrier storage time at room temperature. In addition, the localization energy was obtained and related to the hole/electron retention time. The results of the experiments are summarized in Fig. 5.38. The value for Ge/Si QDs are taken from Kapteyn et al. [105].

An electron/hole localization energy of 290/210 meV, respectively, was obtained for InGaAs QDs embedded in a GaAs matrix. Based on these values, the storage time at

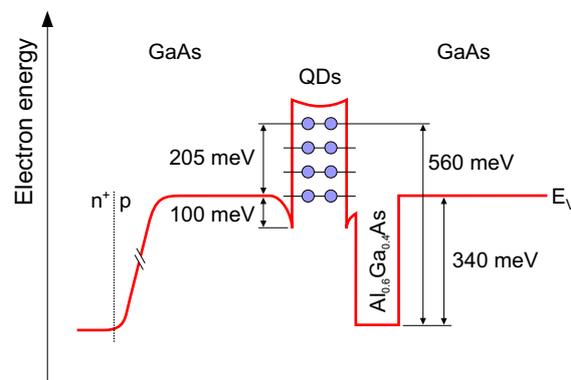


Figure 5.37: Energy diagram for the valence band of sample B with InAs QDs and an additional $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}$ barrier.

Material system	Charge carrier type	Localization energy	Storage time at 300 K
InAs / GaAs [1]	Hole Electron	210 meV 290 meV	~0.5 ns ~200 ns
Ge / Si [2]	Hole	350 meV	~0.1 μ s
GaSb / GaAs [3]	Hole	450 meV	~1 μ s
InAs / GaAs with $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}$ barrier [4]	Hole	560 meV	~5 ms

[1] M. Geller et al., PRB **73**, 205331 (2006) [3] M. Geller et al., Appl. Phys. Lett. **82**, 2706 (2003)
 [2] C. M. A. Kapteyn et al., Appl. Phys. Lett. **77**, 4169 (2000) [4] A. Marent et al., Appl. Phys. Lett. **89**, 072103 (2006)

Figure 5.38: Summary of the electron/hole storage time (at $T=300$ K) and localization energy for different QD systems.

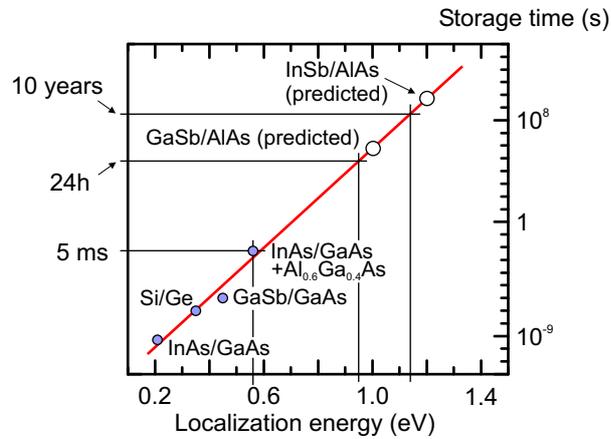


Figure 5.39: Blue points display the hole measured/extrapolated storage time at room temperature versus the localization energy (from Fig. 5.38). The red line is a fit to the data and corresponds to Equation (2.3). The open circles are predicted values for GaSb and InSb QDs in an AlAs matrix, respectively.

$T=300$ K was estimated to be ~ 200 ns for electrons and ~ 0.5 ns for holes. Furthermore, the more promising type-II GaSb/GaAs system has been investigated in detail. It was found, that these QDs can be charged with up to 15 holes. A ground state localization energy of 450 meV was determined, which accounts for a room temperature emission time in the order of one microsecond. The hole localization energy is about twice as large and the retention time at room temperature is about three orders of magnitude longer than in InGaAs/GaAs QDs. Finally, hole emission from InGaAs/GaAs QDs with an additional $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}$ barrier has been studied. The activation energy for the QD ground states increases from 210 meV to 560 meV due to the additional AlGaAs barrier. Accordingly, the hole storage time at room temperature increases by six orders of magnitude, reaching the important DRAM refresh time in the millisecond range.

From the experimental results, a material combination can be predicted to obtain a storage time of more than ten years at room temperature. Figure 5.39 displays the hole

storage time in dependence of the localization energy. The red line is a fit to the data and corresponds to Equation (2.3). This rough estimation gives a retention time of 24 hours for a localization energy of about 0.95 eV. The hole ground state localization energy in GaSb/GaAs QDs is 450 meV and the AlAs/GaAs valence band offset is known to be 550 meV [129], yielding a possible localization energy in GaSb/AlAs QDs of about 1.0 eV. Hence, a retention time of several days at room temperature is possible for this material combination.

(In)(Ga)Sb QDs in an Al(Ga)As matrix may provide a hole localization of more than 1.2 eV. This localization energy could maintain a storage time in the order of more than 200 years, fulfilling an important prerequisite for a non-volatile memory. Therefore, based on the experimental results within this thesis, type-II (In)(Ga)Sb QDs in a Al(Ga)As matrix are considered to be a very promising material combination for a future non-volatile QD memory.

“THE MOST EXCITING PHRASE TO HEAR IN SCIENCE,
THE ONE THAT HERALDS NEW DISCOVERIES, IS NOT
“HEUREKA!” (I FOUND IT!) BUT “THATS FUNNY ...!”
ISAAC ASIMOV

Chapter 6

Carrier Capture into Quantum Dots

Normally, carrier capture into QDs is studied by interband-absorption optical techniques, like time-resolved photoluminescence (PL) spectroscopy [91, 134]. However, such experiments probe the exciton dynamics, that means, the electron-hole capture and relaxation into the QD states is measured simultaneously. A detailed knowledge of either electron or hole capture is also of importance, especially for future QD-based memory applications. The average electron/hole capture time limits the possible write time in such memories. There are just a few works that investigated the electron or hole capture only, like the interband-pump and intraband-probe experiment of Müller et al. [19], sensitive to the capture¹ and relaxation only of electrons. This chapter reports on a new approach to investigate the carrier capture process by using the charge-selective DLTS.

The chapter is organized as follows: Section 6.1 reports on DLTS spectra and results for hole emission *and* capture. Activation energies and capture cross-sections obtained from Arrhenius plots are presented in Section 6.2. Phonon-assisted tunneling is assumed to occur also for the capture process into QDs. In Section 6.3, a temperature and electric field independent capture cross-section is discussed, which describes hole capture in a flat-band situation. Using this capture cross-section, a hole capture time is estimated in dependence on the temperature (from 5 K up to 300 K). The chapter closes with Section 6.4 with a discussion on carrier relaxation, uncertainties in DLTS measurements and a conclusion of the results.

6.1 DLTS Measurements on Hole Capture

The charge-selective DLTS measurements of hole capture into InGaAs/GaAs QDs are presented in this section. The investigated sample H1 is a n^+p diode structure and has

¹Capture means in the following carrier capture *and* relaxation from the valence band into a certain QD state. Relaxation is only the charge carrier transfer inside the dots from an excited initial to a certain final state.

been described already in Section 5.1.1. The QD layer is incorporated in the slightly p-doped ($\sim 3 \times 10^{16} \text{ cm}^{-3}$) region of the diode and situated 500 nm below the p-n junction. Static capacitance measurements of the sample are shown in Fig. 5.2(a), Section 5.1.2.

In conventional DLTS experiments electron/hole capture into QDs is not directly observed for two reasons: Firstly, during the filling pulse the QDs are outside the depletion region of the p-n diode and no change in the capacitance due to charge carrier capture is observed. Secondly, the carrier capture time is expected to be in the picosecond range [19,95] for QDs, due to a capture cross-section of QDs which is orders of magnitude larger than that of deep levels. Therefore, carrier capture is not observable in DLTS measurements having a typical time resolution of milliseconds. A common approach to study capture into deep levels with conventional DLTS is the electrical filling pulse method [108,109], which has a better time resolution in the order of nanoseconds. However, even this time resolution is not sufficient to study carrier capture into QDs.

Both problems can be solved by using charge-selective DLTS, where the QDs are inside the depletion region during the capture process and the band bending induced by the p-n junction provides an activation barrier, schematically depicted in Fig. 6.6. The capture time is increased up to milliseconds and transients for emission *and* capture can be observed for this situation.

Capacitance transients were recorded during temperature scans from 15 K to 100 K

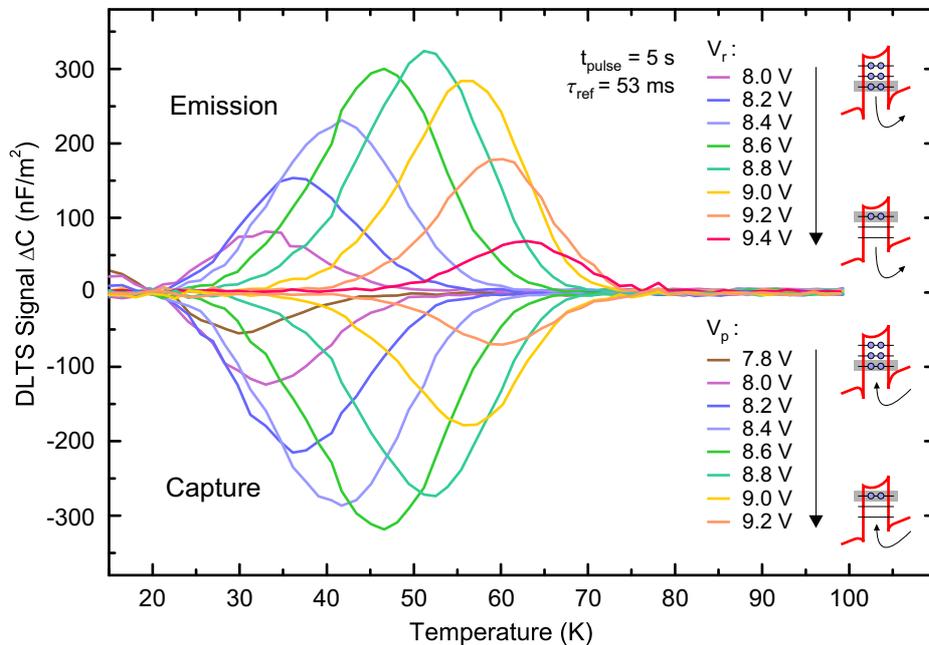


Figure 6.1: Charge-selective DLTS spectra of hole emission and capture. The emission process shows a positive signal while the capture has a negative one. The pulse length and reference time constant are set to 5.0 s and 53 ms, respectively.

and converted into DLTS plots using a double-boxcar correlator with a reference time constant τ_{ref} . Figure 6.1 displays capture (negative signal) and emission (positive signal) spectra after the boxcar conversion. The pulse bias was set to $V_p = V_r - 0.2$ V for the emission DLTS spectra. The reverse bias was set to $V_r = V_p + 0.2$ V for capture spectra. As a consequence, the emission and capture of about one hole per QDs is probed.

Two interesting features are now visible in the DLTS spectra: Firstly, the maxima for hole capture at a given pulse bias value V_p appear at almost the same temperature as the maxima for hole emission for a given reverse bias V_r (for $V_p = V_r$), indicated by identical colours in Fig. 6.1. That means, almost identical capture and thermal emission rates are observed. Secondly, the DLTS amplitude for hole emission at a reverse bias V_r is almost equal to the amplitude for hole capture at a pulse bias $V_p = V_r - 0.2$ V. To explain these two observations, the emission and capture transient are discussed in the following.

6.1.1 Emission and Capture Transients

The capture and emission process in charge-selective DLTS experiments seems to be the following, discussed for two different pulse/reverse bias situations in Fig. 6.2: The capture process for a pulse bias of $V_p = 8.4$ V in certain QD levels in the broadened ensemble is probed in Fig. 6.2(a). This capture process has a capture rate c_1 , defined by the pulse bias.² After switching to a reverse bias of $V_r = 8.6$ V, hole emission from the same energy levels is measured (see schematic pictures in the bottom row of Fig. 6.2), having a different rate, here, the emission rate e_2 . As the same QD levels in the broadened ensemble are probed during capture and emission, the amplitudes of the transients are equal. This explains the equal amplitudes in the DLTS spectra in Fig. 6.1. The DLTS amplitude of a capture process with a certain pulse bias V_p is almost identical to the amplitude of the emission process for $V_r = V_p + 0.2$ V.

In Fig. 6.2(b), the transients for a pulse/reverse bias of $V_p = 8.6$ V and $V_r = 8.8$ V are shown. The capture process for $V_p = 8.6$ V has now the same rate c_2 as the emission process e_2 for an equal reverse bias of $V_r = 8.6$ V in Fig. 6.2(a). The same applied

²As described in Section 3.1.6 for electron emission from deep levels, it is assumed here, that the hole capture rate is much larger than the emission rate ($c_p \gg e_p$) during the capture process. Analogously, the hole emission rate should be much larger than the capture rate during the emission process ($e_p \gg c_p$). This assumption seems to be valid for deep levels [53] and QDs [54] in conventional DLTS experiments, where the deep levels/QDs are well inside the depletion region during emission and the emitted carriers are swept out by the electric field and no re-capturing can take place. During capture, the deep levels/QDs are outside the depletion region and the charge carriers are captured directly from the band edge.

In the charge-selective DLTS experiments, it seems to be more complicated. The QDs are always inside the depletion region and a capture and emission barrier ($E_A^c + E_B$ and $E_A^e + E_B$ in Fig. 6.6) exist, both almost equal. However, a simple approximation of the capture and emission rates c_p and e_p using Equation (3.49) leads to the same conclusion. If a QD level is just 10 meV below the Fermi level in the matrix (i. e. the capture barrier is 10 meV smaller), the capture rate is already one order of magnitude larger than the emission rate (for an activation energy of 120 meV at a temperature of 50 K). The same is valid for the emission situation, where the emission barrier is reduced in comparison to the capture barrier.

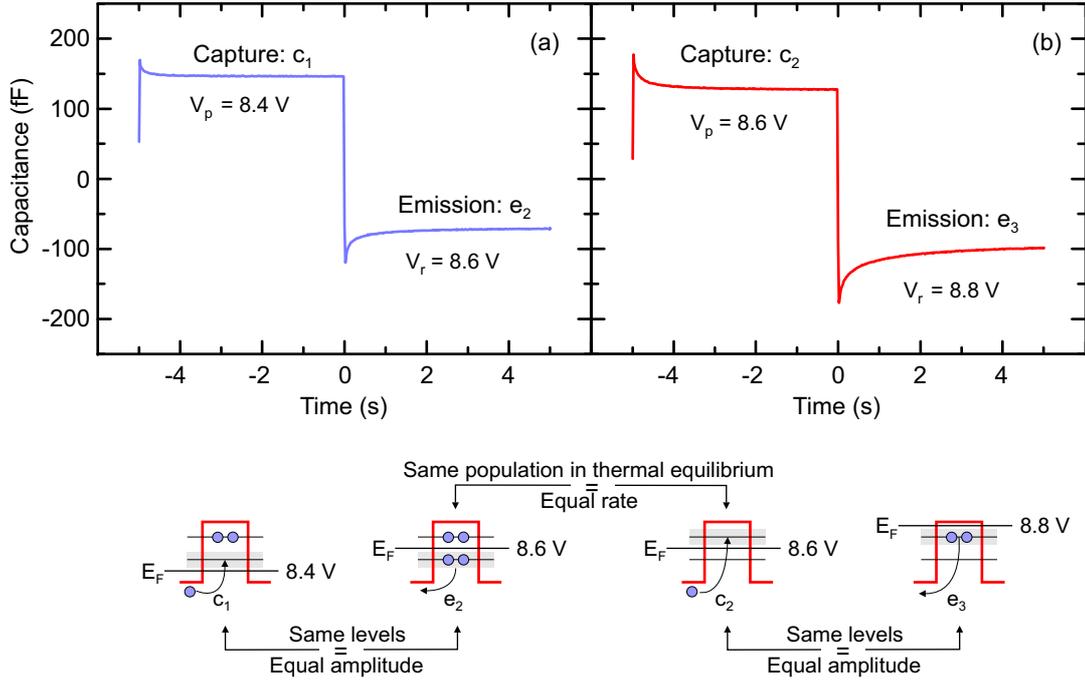


Figure 6.2: Capacitance transients from charge-selective DLTS of InGaAs/GaAs QDs. The measurements were made at a temperature of $T = 49.5$ K for a reverse/pulse bias 8.6 V/8.4 V in panel (a). Panel (b) shows transients for a pulse/reverse bias of 8.6 V/ 8.8 V, respectively, for the same temperature. Clearly visible are transients for hole capture and emission. The bottom row explains schematically the capture and emission process into the different levels having different capture and emission rates.

bias of 8.6 V (V_p or V_r) implies the same carrier population (local Fermi level) inside the QDs, that means, the same population with charge carriers in thermal equilibrium $n_{QD}(\infty)/N_{QD}$. As already described in Section 3.1.6 and schematically depicted in Fig. 3.2, this leads to identical rates for emission and capture. In other words, the charge carrier population in thermal equilibrium, which is defined by the applied voltage at the diode, determines the emission and capture rate, i. e. the same applied voltage leads to equal rates for emission and capture. This explains why the maxima in the DLTS spectra in Fig. 6.1 appear at identical temperatures for the same voltage.

6.2 Activation Energies and Capture Cross-Sections

To elucidate this observation of equal capture and emission rates, the activation energies and capture cross-sections from the emission *and* capture DLTS are evaluated.

To obtain activation energies E_A^e and apparent capture cross-sections σ_∞^e from the emission DLTS spectra, the thermal rate Equation (3.42) in combination with standard Arrhenius plots, shown in Fig. 6.3(b), is used. Analogously, the capture DLTS spectra were analyzed by Arrhenius plots, Fig. 6.3(a), using Equation (3.59), to obtain activation energies E_A^c and capture cross-section σ_∞^c from the capture measurements. The

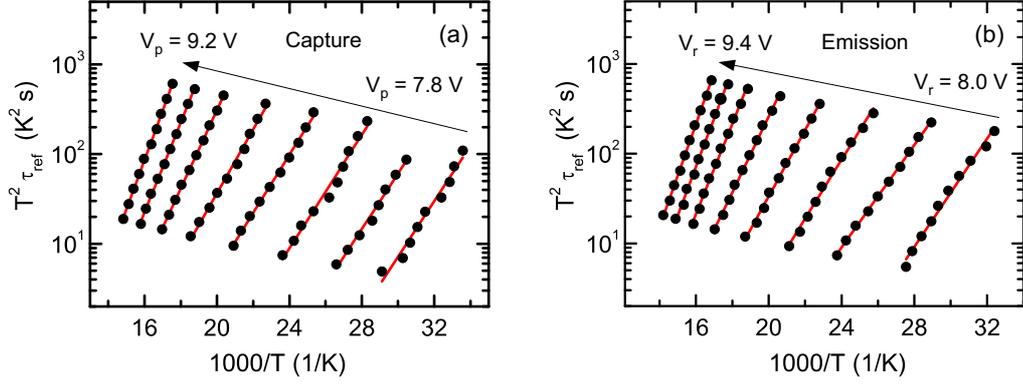


Figure 6.3: Arrhenius plots for hole capture (a) and emission (b) for different reverse/pulse biases.

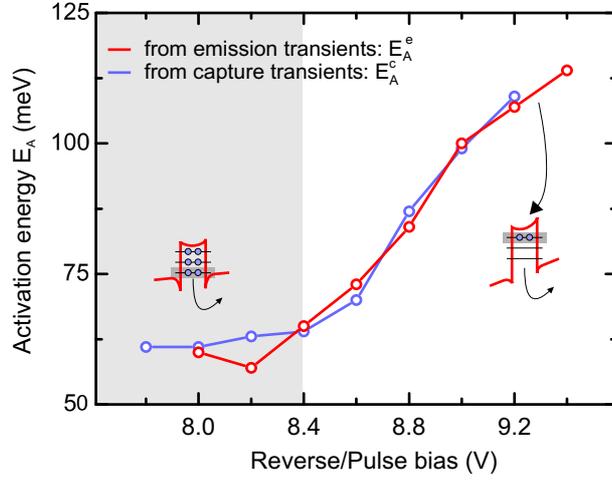


Figure 6.4: Dependence of the activation energy E_A on the reverse bias for emission (red circles) and capture (blue circles). The x-axis displays the reverse/pulse bias situation for the emission/capture transients, respectively. The pulse bias is $V_p = V_r - 0.2$ V for the emission transients and the reverse bias is $V_r = V_p + 0.2$ V for the capture transients.

activation energies and capture cross-sections for hole emission and capture are shown in Fig. 6.4 and 6.5, respectively.

As already described for thermally activated emission for QDs in an electric field, the activation energies E_A^e do not represent the localization energies of the QD states with respect to the corresponding band edge, as the underlying physical mechanism is a phonon-assisted tunneling process. Thermal activation from a QD state into an intermediate state and subsequent tunneling through the remaining triangular barrier E_B occurs, schematically depicted in Fig. 6.6. The total emission rate, hence, depends on the applied electric field and the rate equation has to be multiplied by the transparency factor of a triangular tunneling barrier, Equation (3.51). However, from the slope of the Arrhenius plot in Fig. 6.3(b), the thermal activation energy E_A^e is separated from the phonon-assisted tunneling process. From the intersection with the y-axis the apparent

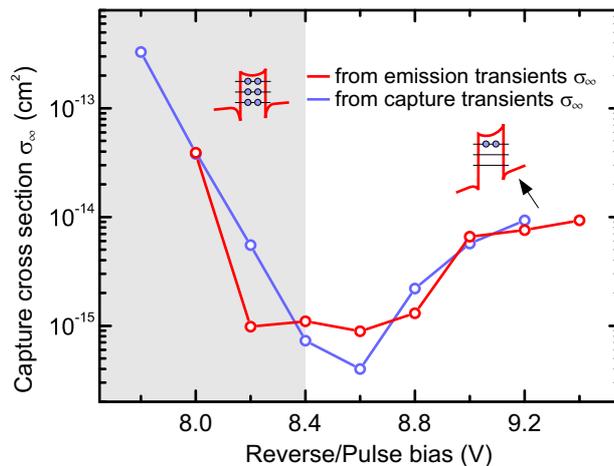


Figure 6.5: Dependence of the apparent capture cross-section σ_∞ for infinite temperature on the reverse/pulse bias for emission (red circles) and capture (blue circles), respectively. The x-axis displays the reverse/pulse bias for the emission/capture transients, respectively. The pulse bias is $V_p = V_r - 0.2$ V for the emission transients and the reverse bias is $V_r = V_p + 0.2$ V for the capture transients.

capture cross-section σ_∞ is deduced, which still depends on the applied electric field F , cf. Equation (3.52).

Figure 6.4 displays activation energies E_A^e and E_A^c for different reverse biases, obtained from capture (blue circles) and emission (red circles) transients. The activation energy of 120 meV for a reverse bias of $V_r = 9.4$ V is attributed to hole emission from the QD ground states. With increasing reverse bias a decrease in the activation energy E_A^e is observed due to an increase in the average occupation of the QDs and thermal emission from successively higher QD states. Accordingly, the decrease in the activation energy E_A^c for the capture process is due to thermally activated *capture* into successively higher QD states.

The apparent capture cross-section decreases in Fig. 6.5 from $\sigma_\infty \approx 1 \times 10^{-14}$ cm² at $V_r = 9.4$ V down to $\sigma_\infty \approx 1 \times 10^{-15}$ cm² at $V_r = 8.4$ V for both, emission and capture. Below a bias of 8.4 V, the capture cross section increases again, while the activation energy remains almost constant at 60 meV (gray shaded area in Fig. 6.4 and 6.5). This energy represents the barrier height resulting from Coulomb interaction of completely filled QDs.

6.2.1 Phonon-Assisted Tunneling During Hole Capture

Due to the almost identical activation energies (E_A^e and E_A^c) and apparent capture cross-sections (σ_∞^e and σ_∞^c), it is assumed here, that carrier capture into QDs in an electric field has the same underlying physical mechanism as the emission: phonon-assisted tunneling.

During the capture process, the holes may firstly be thermally activated with an activation energy E_A^c and tunnel afterwards through the triangular barrier E_B into

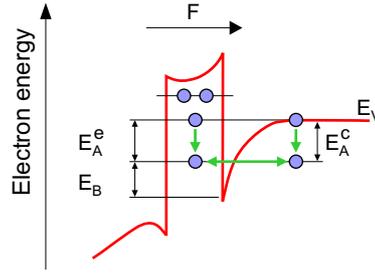


Figure 6.6: Schematic picture of the valence band structure for phonon-assisted tunneling during hole emission and capture. The thermal activation energies for emission E_A^e and capture E_A^c are equal in a charge-selective DLTS measurement.

the QDs, see Fig. 6.6. Remarkably, this phonon-assisted capture process has to occur in opposite direction to the electric field. However, the experimental findings of equal rates for emission and capture, and the according activation energies and capture cross-sections justify this interpretation.

As a consequence of the phonon-assisted capture process, the total capture rate should also depend on the electric field and Equation (3.59) for the capture process has to be multiplied by the transparency factor of a triangular tunneling barrier. In the following evaluation, the apparent capture cross-section for emission and capture can be treated in the same manner – since they are equal – and both will be used to obtain an capture cross-section $\sigma_{T,F=\infty}$ at infinite electric field, described in Section 3.2.4.

6.3 Hole Capture Without a Barrier

The apparent capture cross-section σ_∞ for a phonon-assisted tunneling process still depends on the electric field and the barrier height E_B , cf. Equation (3.52). Plotting the capture cross-section σ_∞ for $V_r > 8.4$ V as function of $1/F$, the slope can be used to determine the tunneling barrier height³ of the intermediate state [62], shown in Fig. 6.7. Using the GaAs hole effective mass of about $m^* = 0.40 m_0$ in the [001] direction, the tunneling barrier height can be estimated to about 30 meV. The apparent capture cross-section $\sigma_{T,F=\infty}$ for an infinitely large electric field is estimated from the intersection to be $\sim 7 \times 10^{-12}$ cm², comparable to the result of Engström et al. [95]. This value represents the capture cross-section for hole capture into the QD ground state in a flat-band situation *without* a capture barrier.

The capture rate Equation (3.56) for holes:⁴

$$c_p = \sigma_{T,F=\infty}^p \langle v \rangle p = \sigma_{T,F=\infty}^p \sqrt{\frac{3kT}{m^*}} p \quad (6.1)$$

³The estimation of the electric field is based on a simple approximation for an abrupt p-n junction, already presented in Section 5.2.1

⁴The temperature dependence of the capture cross-section has been neglected here for the estimation of the capture time. This dependence has only a minor influence on the capture time, discussed in Section 3.2.4.

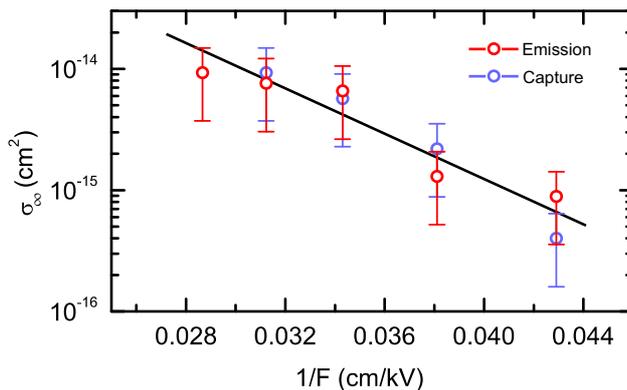


Figure 6.7: The temperature-independent apparent capture cross-section σ_∞ as function of the inverse electric field (for $V_r > 8.4$ V in Fig. 6.5). The solid black line is a fit to the data. The intersection with the y-axis gives an apparent capture cross-section $\sigma_{T,F=\infty}$ at infinite electric field, which is related to hole capture in a flat-band condition.

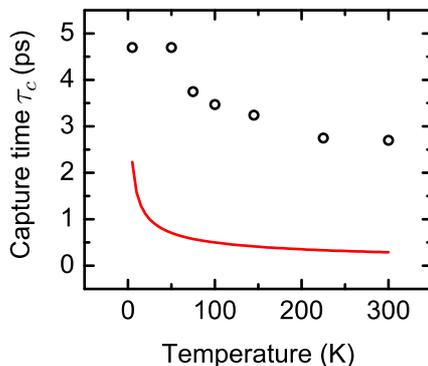


Figure 6.8: The red solid line displays the hole capture time versus temperature, according to Equation (3.2), for an apparent capture cross-section at infinite temperature and electric field of $\sigma_{T,F=\infty} = 7 \times 10^{12}$ cm². This rough estimation for the hole capture time τ_c is compared with the results from Müller et al. [19] for electron capture into InAs QDs (black open circles).

can now be used to obtain an average capture time for the present doping concentration of $p = 3 \times 10^{16}$ cm⁻³. It is in the order of 2 ps at 5 K and decreases down to 0.3 ps at 300 K, red solid line in Fig. 6.8. The results suggest that not only the capture, but also the relaxation of holes within the QDs occurs on a picosecond time scale.

Figure 6.8 compares the capture time versus temperature from the DLTS measurement with the results from Müller et al. [19] for electron capture into InAs QDs (open black circles).⁵ The electron capture time was found to be about 4.7 ps at 5 K and decreasing down to 2.7 ps at 300 K. The red curve for the hole capture time has roughly the same temperature dependence as the data for the electron capture. The accuracy of the capture cross-section is limited and the capture time is only extrapolated via Equation

⁵The measurement were performed for an excitation density of 25 W/cm², corresponding to an electron concentration of about 10^{16} cm⁻³, cf. Engström et al. [95].

(6.1). The comparison suggests nevertheless, that the \sqrt{T} -temperature-dependence of Equation (6.1) describes the hole capture process sufficiently well.

6.4 Discussion and Summary

The main observations in this chapter can be summarized as follows: The charge-selective DLTS method allows to observe the capture process in the capacitance transients. The capture time is increased by a thermal activation barrier during capture and, hence, only the capture process over this barrier can be directly observed. The experimentally observed equal rates for emission and capture lead to the assumption of phonon-assisted tunneling also for carrier capture into QDs. From a linearisation of the apparent capture cross-section at infinite temperature versus the inverse electric field, see Fig. 6.7, an apparent capture cross-section for hole capture in a flat-band condition – i. e. carrier capture from the valence band edge without a capture barrier – can be extrapolated. Using this extrapolated value and the empirical capture rate equation allows to roughly estimate a hole capture time for different temperatures, Fig. 6.8. However, these values for the capture cross-section and hole capture time have their limitations in accuracy, as there are obtained only by extrapolation and evaluation of Equation (6.1). These uncertainties and limitations and some aspects of the underlying physical mechanisms for carrier relaxation within the QDs are discussed in this section.

6.4.1 Discussion: Carrier Capture and Relaxation into QD States

To obtain the apparent capture cross-section at infinite temperature σ_∞ , the well-known rate Equation (3.42) has been used to analyze the DLTS data from the emission measurements. An analogous Equation (3.59) was derived for the capture process. Both equations exhibit a T^2 -temperature-dependence in the pre-exponential factor. This temperature dependence is mostly a result of the temperature dependence of the capture process, here, given by Equation (3.1).

The actual temperature dependence of the capture cross-section is not obvious and was only assumed to have an exponential character in Section 3.2.1. In addition, the carrier relaxation process within the QD influences the capture cross-section and its temperature dependence, too. This is of relevance for the choice of a suitable set of parameters to be displayed in the Arrhenius plot for determining the thermal activation energies and capture cross-sections. The temperature dependence of the relaxation rate results from multi-phonon processes, discussed in the following.

Multi-phonon scattering

To estimate the temperature dependence of the relaxation process, the physical mechanism which mainly controls the carrier relaxation processes in QDs is needed. In the early days of self-organized QDs, relaxation in QDs was predicted to be rather slow [135], since phonon-scattering was assumed to be the relaxation process. The

confinement in all three spatial directions leads to a δ -function-like density of states and a so-called "phonon-bottleneck" should suppress the phonon-emission efficiency from the ground states, since the carriers can not relax sufficiently fast from the valence/conduction band into these states. In higher-dimensional systems, e. g. bulk or quantum wells, the dominant relaxation process is longitudinal-optical (LO) phonon emission with relaxation times in the sub-picosecond range. In a QD system, the discrete spectrum of energy levels should suppress the energy relaxation via LO phonons, unless the level separation equals the LO phonon energy $\hbar\omega_{LO} \approx 30$ meV, which is of course seldom the case.

However, no experimental evidence for a particularly weak relaxation probability has been found. As a consequence, a number of different processes have been proposed to account for carrier relaxation in QDs, including Auger scattering [87, 89, 136, 137], multi-phonon processes [90, 91, 138–140], and defect-assisted relaxation [141, 142]. One of the most promising relaxation processes is the multi-phonon relaxation, where one or more low frequency LA phonon and one or more LO phonon or a combination of both are predicted to be involved.

This multi-phonon relaxation process can be used to estimate the possible temperature-dependence of the capture rate. A more detailed analysis can be found in Reference [54]. The multi-phonon scattering rate (for LO or LA) can be written as (see e. g. [19, 134, 143]):

$$c = \Gamma^0 (N(T) + 1)^n, \quad (6.2)$$

where Γ^0 is the scattering rate at $T = 0$ K and $N(T) = [\exp(\hbar\omega_q/k_B T) - 1]^{-1}$ the Bose Einstein distribution and n accounts for the number of phonons involved in the scattering process.

For a single-phonon scattering process, two cases of asymptotic behavior can be derived from this expression. If the phonon energy is much greater than the thermal energy, $\hbar\omega_q \gg k_B T$, Equation (6.2) turns into $c \sim 1$. A temperature-independent carrier relaxation due to purely spontaneous phonon emission is observed. If $\hbar\omega_q \ll k_B T$, one obtains $c \sim T$. Two-phonon emission exhibits, according to Equation (6.2), a $(N(T) + 1)^2$ -dependence with the two asymptotic cases of $c \sim 1$ and $c \sim T^2$. The T^2 -dependence may serve as an upper limit of the capture rate for a two-phonon assisted relaxation process. Higher order phonon scattering will have significantly smaller probabilities.

Activation Energies in DLTS Measurements

A \sqrt{T} -dependence is used in Equation (6.1), which is well within the expected range of T^x , $x \in [0, 2]$ for a multi-phonon scattering process. This temperature-dependence yields a T^2 -dependence in the pre-factor of Equation (3.59), used for the evaluation of the capture transients. Analogously, the assumption of a detailed balance condition, Section 3.1.2, gives a T^2 -dependence in the pre-factor for the emission rate (Equation (3.17)).

In standard Arrhenius plots ($\ln(T^2 \tau_{\text{ref}})$ versus T^{-1}), a T^2 -dependence of the pre-factor in the expression for the emission/capture rate is assumed. If the relaxation process in the QDs has (as part of the carrier capture process) a different temperature-dependence, different Arrhenius plots, such as $\ln(\sqrt{T} \tau_{\text{ref}})$ versus T^{-1} , would be required. For the limiting case of a $\ln(\tau_{\text{ref}})$ versus T^{-1} Arrhenius plot, a linear fit in such a graph yields an activation energy decreased by less than 10%. Without knowing the exact dependence of the temperature dependence, the results presented in this work for the activation energies can still be assumed to be correct within an accuracy of about 10%.

Capture Cross-Sections

The uncertainties concerning the temperature dependence of the emission rate hamper a straightforward interpretation of the experimentally determined capture cross-section in all DLTS experiments. For the limiting case of a temperature-independent relaxation rate, a $\ln(\tau_{\text{ref}})$ versus T^{-1} Arrhenius plot would yield a two orders of magnitude larger capture cross-section. The hole capture time would be up to two orders of magnitude faster, too. That means, the presented capture cross-section of about $7 \times 10^{-12} \text{ cm}^2$ is just a lower limit for an assumed T^2 -temperature dependence. Analogously, the estimated capture time of about 0.3 ps at room temperature serves as an upper limit for hole capture and relaxation in a flat-band condition. However, the good agreement of the capture time versus temperature from the DLTS measurement with the results from Müller et al. [19] for electron capture into InAs QDs confirms the T^2 -temperature dependence and the assumptions made.

6.4.2 Summary

In this section, measurements of the carrier capture of holes into InGaAs QDs were presented using the charge-selective DLTS method. From the capacitance transients thermal activation energies and apparent capture cross sections are obtained. The apparent capture cross-section for hole capture into the QD ground states in a flat-band situation is estimated to about $7 \times 10^{-12} \text{ cm}^2$, leading to fast hole capture in the order of 2 ps at 5 K (0.3 ps at 300 K). Therefore, the write process in a future QD-based memory device via carrier capture from the band-edge into the QD states could be more than four orders of magnitude faster than the write access time of a DRAM cell ($\sim 20 \text{ ns}$).

“PHYSICS IS LIKE SEX. SURE, IT MAY GIVE SOME PRACTICAL RESULTS, BUT THAT’S NOT WHY WE DO IT.”

RICHARD FEYNMAN

Chapter 7

Conclusion

In the present thesis, charge carrier storage and dynamics of self-organized quantum dots (QDs) were investigated with the aim of using QDs as building blocks in future memory devices. This work has been divided in two parts: carrier emission (which determines the storage time) and carrier capture (which determines the write time), both crucial for memory applications. Time-resolved capacitance spectroscopy (also known as “deep level transient spectroscopy” – DLTS) has been used to study different QD systems concerning their electron/hole localization energy and storage time at room temperature. Furthermore, hole capture into InGaAs/GaAs QDs has been examined and a capture time was estimated. A concept of a QD-based memory has also been developed, combining the advantages of DRAM and Flash to an universal memory. Such a novel memory type should have a storage time in the order of years (non-volatility), fast write/erase and read access time (<10 ns) in combination with good endurance ($>10^{15}$ write/erase cycles).

In order to study different QD systems concerning their storage time and localization energy as well as the emission and capture dynamics, the capacitance spectroscopy has been refined within this thesis. An improved DLTS method has been developed, the charge-selective DLTS, which allows to observe thermally activated charge carrier emission from differently charged QDs and probe carrier escape from different QD states in an energy-broadened ensemble. Furthermore, the time-resolved tunneling capacitance measurement (TRTCM) has been established to study electron and hole tunneling and determine localization energies. This method enables the direct observation of tunneling emission in capacitance measurements. The time-dependent decrease of the ground state population (due to pure tunneling escape) is observed at low temperatures, where thermally activated emission is negligible. By measuring the tunneling time as function of the electric field strength, the electron and hole localization energy can be determined.

In the first set of experiments, the electron and hole emission from a single layer of InGaAs/GaAs QDs was studied by using the charge-selective DLTS. It turns out that the charge carrier emission process from QDs in an electric field is always controlled by phonon-assisted tunneling. The influence of the tunneling part strongly depends

on the effective mass and the strength of the electric field. For electron emission, a strong tunneling influence was observed previously [66] and is confirmed here. As a result, activation energies of 82 meV for the transition from the ground to the first excited state and 44 meV for the transition between the first and second excited state were obtained. For the hole emission an energy of 120 meV was obtained, which corresponds to thermal activation from the hole ground state to an intermediate state. The electron/hole activation energy from the ground state to the conduction/valence band (the localization energy) was not accessible in charge-selective DLTS measurements for InGaAs/GaAs QDs. A storage time at room temperature in an InGaAs/GaAs system with an applied electric field was estimated to 50 ps /10 ps for electrons/holes.

In the next part, the newly developed capacitance measurement of the tunneling emission was used to determine the localization energy of the electron/hole ground state in InGaAs/GaAs to 290 meV/210 meV, respectively. The average storage time in this QD system is estimated to be about 200 ns/ 0.5 ns, respectively, for a field-free situation. Obviously, the retention time in InGaAs/GaAs QDs at room temperature is not sufficiently long for memory applications.

Two possibilities were then studied to increase the storage time at room temperature: (i) Changing the QD material to a system that exhibits a type-II band alignment and (ii) changing the matrix material, such that a larger difference in the bandgap occurs. Firstly, type-II GaSb/GaAs QDs were investigated with the charge-selective DLTS method. These QDs exhibit a band alignment such that a hole localization appears in the valence band while in the conduction band a repulsive barrier for electrons exists. It was found that these QDs can be charged with up to 15 holes. A ground state localization energy of 450 meV was determined, which accounts for a room temperature emission time in the order of one microsecond. The ground state localization is about twice as large and the hole retention time at room temperature is about three orders of magnitude longer than in InGaAs/GaAs QDs. Secondly, InAs/GaAs QDs in combination with an $\text{Al}_{0.6}\text{Ga}_{0.4}\text{As}$ barrier were used, providing a larger difference in the bandgaps. The barrier increases the hole activation energy from 210 meV up to 560 meV and a signal for hole emission over the barrier is observed for a reference time of 5 ms in the DLTS spectra at 300 K. That means, a storage time of about 5 ms was measured at room temperature, which is in the order of the important DRAM refresh time.

In the next part, the charge-selective DLTS was used to investigate carrier capture into InGaAs/GaAs QDs. In capacitance transients the hole capture has been observed and activation energies and capture cross-sections for emission *and* capture were determined. The experimental findings lead to the assumption that carrier capture into QDs in an applied electric field is also controlled by phonon-assisted tunneling – similar to carrier emission. A large capture cross-section of about $7 \times 10^{-12} \text{ cm}^2$ is obtained and an average capture time of $\sim 0.3 \text{ ps}$ at room temperature is estimated, which is more than four orders of magnitude faster than the write time in a DRAM ($\sim 20 \text{ ns}$). Hence, a very fast write time seems to be possible in QD-based memory devices.

In this work, self-organized QDs were investigated concerning their suitability to act

as information storage units in future memories. A storage time of milliseconds at room temperature was demonstrated and the experimental results allow to predict a retention time of more than 10 years for (In)(Ga)Sb QDs in an AlAs matrix. Furthermore, the estimated hole capture time in the sub-picosecond range confirms the possibility of a QD-based memory with improved performance in comparison to DRAM and Flash. Even an universal memory with long storage time in the order of years in combination with fast write/erase and read times seems to be feasible.

“SCIENCE IS ALWAYS WRONG. IT NEVER SOLVES A PROBLEM WITHOUT CREATING TEN MORE.”

GEORGE BERNARD SHAW

Chapter 8

Outlook

The first realizations of self-organized quantum dots (QDs) by modern epitaxial growth techniques were reported 15 years ago [7, 8] and optical applications, like lasers and amplifiers were the earliest devices based on these QDs [10–14]. The first attempts to use QDs in memory structures [32, 36], however, were mainly hampered by the coexistence of deep levels in the device structures [144], which causes the difficulty of a well-defined carrier storage in the QDs, i. e. carrier storage in defect levels could not be excluded. Within this thesis, the capacitance spectroscopy methods allowed to unambiguously identify the carrier emission from the QD states and determine their storage times: an important prerequisite for further investigations directed towards a QD-based memory. Some possible next steps are briefly outlined in the following.

Storage times at room temperature: To improve the retention time, QD structures based on different III-V materials (InGaAl)(AsSb) should be developed and tested by capacitance spectroscopy methods and/or detection of the resistance of a 2DEG below the QD layer (Section 2.3.3). An average storage time of more than ten years at room temperature, as predicted for an In(Ga)Sb/AlAs material combination, is the ultimate goal.

Testing of the write/erase and read-out time: Within this thesis, the capture time into the QD states was investigated. The write/erase and read times in a single-cell QD memory device are of importance, too. Especially, the relation between storage and erase time is crucial, as a trade-off between both exists. A larger localization energy increases the storage time and the erase time slows down simultaneously, as the charge carriers have to tunnel through a thicker triangular barrier.

Some solutions to increase the erase speed in a memory cell with strong localization can be imagined: (i) Using quantum well layers below the QDs for resonant tunneling or (ii) using a barrier similar to the InAs/GaAs QD sample with the AlGaAs barrier (Section 5.4). On one hand, a barrier below the QD layer instead of a doped matrix material will reduce the tunneling distance from the QD states into the valence/conduction band, hence, increasing the erase speed again. On the other hand, the write speed will slow down, since the charge carriers can not directly relax into the QD states from the band edge anymore. Therefore, a QD memory cell has to be optimized to ensure fast

write/erase and read-out times below 10 ns in combination with long storage times at room temperature. In contrast to today's CMOS technology, III-V compound semiconductors and their ternary alloys offer the possibility of a controlled band structure engineering with adjustable barrier heights to meet the challenge.

After designing, processing and testing single memory cells concerning their storage and access times, three different QD-based memories with a better performance than today's DRAM or Flash should be possible: (i) A novel DRAM, (ii) an improved non-volatile memory and (iii) an universal memory.

A novel DRAM based on QDs should improve its access time below nanoseconds and will extend the storage time by orders of magnitude from milliseconds to more than seconds at room temperature. The improved non-volatile memory should have a storage time of more than 10 years and an optimized write access time below 1 μ s, which is the standard value in nowadays Flash memories. The ultimate goal would be an universal memory, which combines the novel DRAM and the improved non-volatile memory. Such a QD memory will be non-volatile and provides fast access times well below 10 ns in combination with an endurance of more than 10^{15} write/erase cycles (like in a DRAM cell).

Appendix A

Experimental Details

A.1 DLTS Setup

A schematic layout of the experimental setup is shown in Fig. A.1. The center of the experimental setup is a Helium constant flow cryostat (CryoVac) with a CryoVac TIC 303M temperature controller. This cryostat/controller combination has a temperature range from 10 K up to 450 K with an accuracy of about ± 0.3 K. For static capacitance measurements (C-V), a HP4284A LCR meter was used. This capacitance bridge has an internal voltage source and provides a measurement frequency between 20 Hz and 1 MHz with an AC signal amplitude ranging from 5 mV up to 1 V. The accuracy of the C-V measurement depends on the measurement frequency, the signal amplitude and the integration time. For more details see Reference [145].

For the time-resolved capacitance measurements, two different capacitance meter with an operation frequency of 1 MHz were used: An analog Boonton 72B and a digital Boonton 7200, both operating with an AC signal amplitude of 100 meV. The digital Boonton 7200 can also supply signal amplitudes down to 10 meV. The accuracy of the capacitance measurements is typically better than 0.25% of the determined value. The rise time of both devices is about 50 μ s. In order to allow operation in the most sensitive 2 pF range (Model 7200) or 1 pF range (Model 72B), a step-motor driven external capacitance correction (with a variable capacitor of 3 pF to 240 pF) is used to compensate the static sample capacitance after each change of temperature and before transient recording. The next sensitive capacitance range – 20 pF (Model 7200) or 10 pF (Model 72B) – was often used for samples with higher serial or lower parallel resistance. The ranges with lower sensitivity are less sensitive to leakage currents or high serial resistances (more details can be found in Reference [146,147]). The samples have been checked and the optimal capacitance range were used.

Reverse and pulse biases were externally supplied by a HP-8115A pulse generator and a Agilent 33250A arbitrary waveform generator. The HP-8115A allows a maximum pulse amplitude of 16 V for a minimum pulse width of 9 ns, while the Agilent 33250 A provides an amplitude of 10 V for a minimum pulse width of 8 ns. The main advantage of the Agilent waveform generator is a pulse period of up to 2000 s for the DLTS

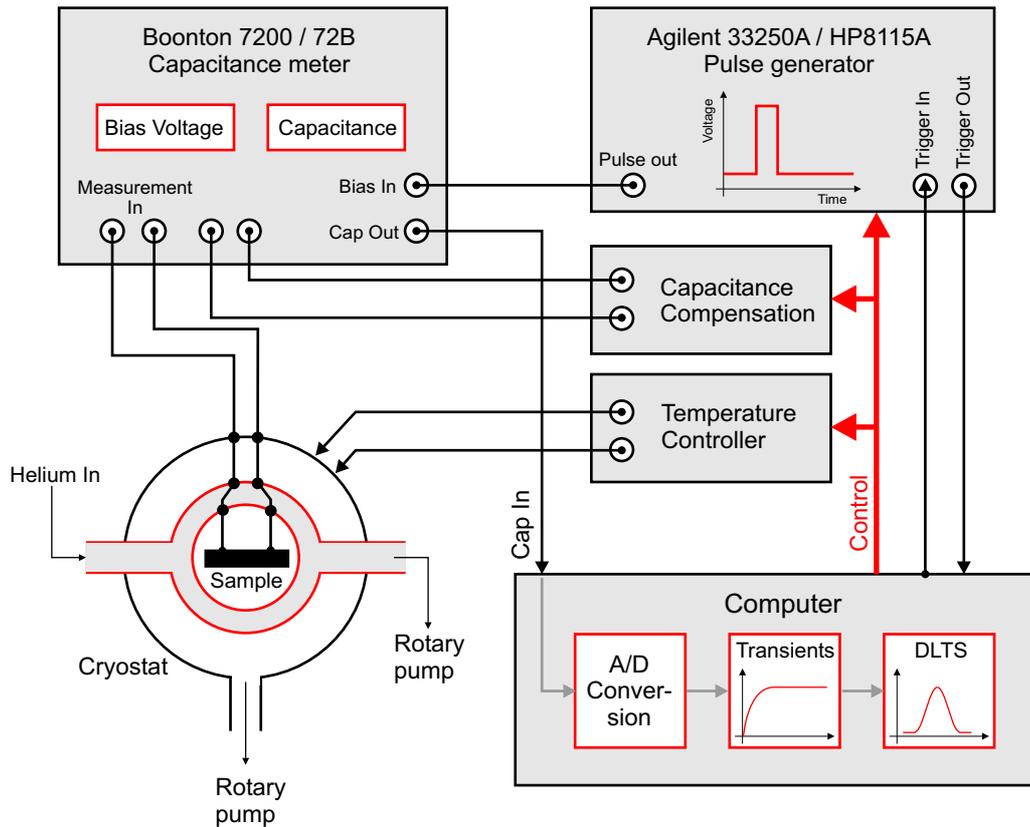


Figure A.1: Experimental setup for the DLTS measurements.

measurement on the InAs QDs with the additional AlGaAs barrier. The HP pulse generator has only a maximum period of 1 s. Of importance is the pulse edge time for the DLTS measurements. A short edge time in combination with high pulse amplitude leads to an overshoot of the pulse voltage at the sample, since the capacitance meter acts as a low-pass filter. The edge time has been adjusted from $1 \mu\text{s}$ up to $300 \mu\text{s}$, depending on the pulse amplitude. The pulse form has to be carefully checked after passing the Boonton. In addition, the analog Boonton 72B stays in an overload, if the measured capacitance is out of range. Therefore, the amplifier of the analog Boonton 72B has to be switched off during and switch on again after the pulse. A custom-made switch was installed, which is described in detail in Reference [148].

The whole setup is controlled by a computer, which records the capacitance transients from the analog output of the capacitance meter after 12 bit A/D conversion with up to 50.000 points per transient. Typically 2000 to 3000 points per transient with $105\text{-}150 \mu\text{s}$ resolution were used. Averaging of up to 30 transients was applied to increase the signal to noise ratio. Details on the data format used can be found in Reference [149, 150]. Typically 1.5 K temperature steps were used.

A.2 Sample Processing

The samples were partly processed in the class 10 cleanroom ("Center of NanoPhotonics") and a class 100 cleanroom of the "Institut für Festkörperphysik" of the "Technische Universität Berlin". Standard optical lithography with a Süss mask aligner, wet chemical etching, and a Veeco metal evaporation system were used. The latest optimized recipes for processing of the p-n diodes are presented in the following.

A.2.1 p-n and n-p Diodes

Both diodes are processed in almost the same manner, except the metals used for the Ohmic back and top contacts. The entire process consists of three steps: 1. Lithography and chemical wet etching of mesa structures. 2. Lithography and formation of Ohmic top contacts. 3. Lithography and formation of Ohmic back contacts.

1. Lithography and Wet Etching

- **Sample Cleaning**

- Place the sample on a plastic sifter, not directly on the bottom of a beaker
- Rinse the sample for about 2 min in beakers filled with Aceton ($\times 2$) and Isopropanol ($\times 2$) on a 76°C hotplate
- Dry with N₂
- Never clean with H₂O when using MA-photoresist!

- **Lithography of the mesa structures**

- Spin photoresist MA-P1215 (3000 rnds/min for 40 s and 6000 rnds/min for 15 s)
- Hotplate 90°C, 5 min
- Let sample cool down, 5 min
- Exposure with mask, 6 mW/cm², 10 s
- Develop with MA-D331, 60–90 s, stir slowly
- H₂O dip, 10 s ($\times 2$)
- Dry with N₂

- **Wet chemical etching**

- Place the sample on a plastic sifter
- Etch sample with 100 ml H₂O, 1 ml H₂SO₄, 8 ml H₂O₂ (100:1:8), ≈ 365 nm/min in a beaker
- Stop with H₂O, 10 s ($\times 2$)
- Check with profiler

- **Sample Cleaning** – as above

2. Lithography of Top Contacts

- **Sample Cleaning** – as above
- **Lithography**
 - Spin photoresist MA-N-440 (3000 rnds/min for 40 s and 6000 rnds/min for 15 s)
 - Hotplate 90°C, 5 min
 - Let sample cool down, 5 min
 - Exposure with mask, 6 mW/cm², 10 s
 - Develop with MA-D532, 60–90 s, stir slowly
 - H₂O dip, 10 s (×2)
 - Dry with N₂
 - Use plasma-asher to remove remaining photoresist
- **Evaporation of top contacts**
 - HCl- dip, 30 s
 - H₂O cascade
 - Dry with N₂
 - Metal evaporation (Section A.2.2)
- **Liftoff**
 - Liftoff in remover (Hotplate 76°C)
 - Aceton, 2 min (×2)
 - Isopropanol 2 min (×2)
 - Dry with N₂
- **Sample Cleaning** – as above
- **Annealing** – 120 s at 390°C in N₂ atmosphere

3. Lithography of Back Contacts

As Lithography of top contacts – see above.

A.2.2 Ohmic Contacts

For the formation of metal-semiconductor Ohmic contacts, a standard metal-evaporation procedure with subsequent alloying is used. Depending on whether the material is n- or p-type, layers of different metals are used. In addition, two different p-type contacts (Ni/Zn/Au and Ti/Pt/Au) are possible, both having their advantages and

disadvantages. The Ni/Zn/Au contact has a very good adherence to the GaAs semiconductor. However, Zn diffuses easily into the active layer and can produce defects or additional doping. On the other hand, Ti/Pt/Au has no diffusion problem but a worse adherence to the semiconductor.

The layer evaporation sequence and the according thickness is given below. After the evaporation, the samples were annealed in N₂ atmosphere to allow for diffusion of the metal into the semiconductor. The conditions used were: 120 s at 390°C.

- **n-type GaAs:** 7.5 nm Ni – 100 nm Au/Ge (88:12) – 200 nm Au
- **p-type GaAs:** 5 nm Ni – 50 nm Zn – 200 nm Au
- **p-type GaAs:** 30 nm Ti – 50 nm Pt – 350 nm Au

A.2.3 Sulfur Passivation

It is well-known, that the surface of III-V compounds exposed to air exhibits a high density of states and hence, very high surface recombination velocity, resulting in leakage currents and poor device performance. Various passivation methods were developed to improve and stabilize the surface properties. The sulfide passivation method with sulfide solutions has widely studied and used (see for example References [151–156]). This treatment significantly reduces the surface state density and results in improved device performance. The passivation method used works as follows:

- Place the sample on a plastic sifter
- Etch sample with H₂O:H₂SO₄ (100:1), 10 s
- H₂O dip, 10 s (×2)
- Warm up 60 ml H₂O to 60°C, stir slowly, place sample inside
- Add 3.2 ml (NH₄)₂S_x
- Cover with thermoplastic foil, wait 30 min
- H₂O dip, 10 s (×2), no ultrasonic bath
- Hotplate 125°C, 5 min

Bibliography

- [1] G. E. Moore, Cramming more components onto integrated circuits, *Electronics* **38** (1965).
- [2] B. Hoeneisen and C. A. Mead, Fundamental limitations in microelectronics: MOS technology, *Solid-State Electronics* **15**, 819 (1972).
- [3] T. Sugano, Physical and Technological Limits in Size of Semiconductor Devices, *Jpn. J. Appl. Phys.* **15**, 329 (1976).
- [4] International roadmap for semiconductors, (2005 Edition, January 2006).
- [5] I. N. Stranski and L. Krastanow, Zur Theorie der orientierten Ausscheidung von Ionenkristallen aufeinander, *Sitzungsber. Akad. Wiss. Wien, Math.-Naturwiss. K1, Abt. 2B* **146**, 797 (1938).
- [6] D. Bimberg, M. Grundmann, and N. N. Ledentsov, *Quantum Dot Heterostructures*, John Wiley & Sons, Chichester, 1998.
- [7] D. J. Eaglesham and M. Cerullo, Dislocation-Free Stranski-Krastanow Growth of Ge on Si (100), *Phys. Rev. Lett.* **64**, 1943 (1990).
- [8] Y.-W. Mo, D. E. Savage, B. S. Swartzentruber, and M. G. Lagally, Kinetic pathway in Stranski-Krastanov growth of Ge on Si(001), *Phys. Rev. Lett.* **65**, 1020 (1990).
- [9] Y. Arakawa and H. Sakaki, Multidimensional quantum well laser and temperature dependence of its threshold current, *Appl. Phys. Lett.* **40**, 939 (1982).
- [10] N. Kirstaedter, N. N. Ledentsov, M. Grundmann, D. Bimberg, V. M. Ustinov, S. S. Ruvimov, M. V. Maximov, P. S. Kop'ev, Z. I. Alferov, U. Richter, P. Werner, U. Gosele, and J. Heydenreich, Low threshold, large T_0 injection laser emission from (InGa)As quantum dots, *Electron. Lett.* **30**, 1416 (1994).
- [11] M. Grundmann, The present status of quantum dot lasers, *Physica E* **5**, 167 (2000).
- [12] N. N. Ledentsov, M. Grundmann, F. Heinrichsdorff, D. Bimberg, V. M. Ustinov, A. E. Zhukov, M. V. Maximov, Z. I. Alferov, and J. A. Lott, Quantum-Dot

- Heterostructure Lasers, *IEEE Journal of Selected Topics in Quantum Electronics* **6**, 439 (2000).
- [13] F. Heinrichsdorff, C. Ribbat, M. Grundmann, and D. Bimberg, High-power quantum-dot lasers at 1100 nm, *Appl. Phys. Lett.* **76**, 556 (2000).
- [14] M. Kuntz, G. Fiol, M. Lämmlin, D. Bimberg, M. G. Thompson, K. T. Tan, C. Marinelli, R. V. Penty, I. H. White, V. M. Ustinov, A. E. Zhukov, Y. M. Shernyakov, and A. R. Kovsh, 35 GHz mode-locking of 1.3 μm quantum dot lasers, *Appl. Phys. Lett.* **85**, 843 (2004).
- [15] P. Michler, A. Kiraz, C. Becher, W. V. Schoenfeld, P. M. Petroff, L. Zhang, E. Hu, and A. Imamoglu, A quantum dot single-photon turnstile device, *Science* **290**, 2282 (2000).
- [16] C. Santori, D. Fattal, J. Vučković, G. S. Solomon, and Y. Yamamoto, Indistinguishable photons from a single-photon device, *Nature* **419**, 594 (2002).
- [17] L. Geppert, The new indelible memories, *IEEE Spectrum*, March , 49–54 (2003).
- [18] R. Waser, *Microelectronics and Information Technology*, Wiley-VCH, Berlin, 2003.
- [19] T. Müller, F. F. Schrey, G. Strasser, and K. Unterrainer, Ultrafast intraband spectroscopy of electron capture and relaxation in InAs/GaAs quantum dots, *Appl. Phys. Lett.* **83**, 3572 (2003).
- [20] M. Geller, A. Marent, E. Stock, A. E. Zubkov, I. S. Shulgunova, A. V. Solomonov, and D. Bimberg, Hole capture into self-organized InGaAs quantum dots, *Appl. Phys. Lett.* **89**, 232105 (2006).
- [21] H. Kroemer, Barrier Control and Measurements - Abrupt Semiconductor Heterojunctions, *J. of Vac. Sci. & Technol. B* **2**, 433 (1984).
- [22] Z. I. Alferov and R. F. Kazarinov, Semiconductor laser with electric pumping, Inventor's Certificate No. 181737 [in Russian], Application No. 950840, (1963).
- [23] H. Kroemer, Semiconductor laser with optical pumping, US Patent No. 3309553, (1963).
- [24] I. Hayashi, M. B. Panish, P. W. Foy, and S. Sumski, Junction Lasers Which Operate Continuously at Room Temperature, *Appl. Phys. Lett.* **17**, 109 (1970).
- [25] F. K. Reinhart, I. Hayashi, and M. B. Panish, Mode Reflectivity And Waveguide Properties of Double-Heterostructure Injection Lasers, *J. Appl. Phys.* **42**, 4466 (1971).

-
- [26] R. A. Milano, T. H. Windhorn, E. R. Anderson, G. E. Stillman, R. D. Dupuis, and P. D. Dapkus, Al_{0.5}Ga_{0.5}As-GaAs Heterojunction Phototransistors Grown by Metalorganic Chemical Vapor-Deposition, *Appl. Phys. Lett.* **34**, 562 (1979).
- [27] C. Y. Chen, A. Y. Cho, C. G. Bethea, P. A. Garbinski, Y. M. Pang, B. F. Levine, and K. Ogawa, Ultrahigh Speed Modulation-Doped Heterostructure Field-Effect Photodetectors, *Appl. Phys. Lett.* **42**, 1040 (1983).
- [28] I. Vurgaftman, J. R. Meyer, and L. R. Ram-Mohan, Band parameters for III-V compound semiconductors and their alloys, *J. Appl. Phys.* **89**, 5815 (2001).
- [29] E. T. Yu, J. O. Mccaldin, and T. C. McGill, Band Offsets in Semiconductor Heterojunctions, *Solid State Physics* **46**, 1 (1992).
- [30] *Landolt-Börnstein — Semiconductors: Physics of Group IV Elements and III-V Compounds*, volume III/17a, Springer, Berlin, 1982.
- [31] S. Tiwari, F. Rana, H. Hanafi, A. Hartstein, E. F. Crabbé, and K. Chan, A silicon nanocrystal based memory, *Appl. Phys. Lett.* **68**, 1377 (1996).
- [32] K. Koike, K. Saitoh, S. Li, S. Sasa, M. Inoue, and M. Yano, Room-temperature operation of a memory-effect AlGaAs/GaAs heterojunction field-effect transistor with self-assembled InAs nanodots, *Appl. Phys. Lett.* **76**, 1464 (2000).
- [33] H. Kim, T. Noda, T. Kawazu, and H. Sakaki, Control of Current Hysteresis Effects in a GaAs/n-AlGaAs Quantum Trap Field Effect Transistor with Embedded InAs Quantum Dots, *Jpn. J. Appl. Phys.* **39**, 7100 (2000).
- [34] G. Yusa and H. Sakaki, Trapping of photogenerated carriers by InAs quantum dots and persistent photoconductivity in novel GaAs/n-AlAs field-effect transistor structures, *Appl. Phys. Lett.* **70**, 345 (1997).
- [35] J.-E. Oh and J.-W. Kim, Room-temperature Memory Operation of AlGaAs/GaAs High Electron Mobility Transistors with InAs Quantum Dots embedded in the Channel, in *Proceedings of the International Electron Device Meeting 2000, IEDM*, page 106, 2000.
- [36] C. Balocco, A. M. Song, and M. Missous, Room-temperature operations of memory devices based on self-assembled InAs quantum dot structures, *Appl. Phys. Lett.* **85**, 5911 (2004).
- [37] J. C. Campbell, D. L. Huffaker, H. Deng, and D. G. Deppe, Quantum dot resonant cavity photodiode with operation near 1.3 μm wavelength, *Electr. Lett.* **33**, 1337 (1997).
- [38] L. Chu, A. Zrenner, M. Bichler, and G. Abstreiter, Quantum-dot infrared photodetector with lateral carrier transport, *Appl. Phys. Lett.* **79**, 2249 (2001).

-
- [39] M. Lämmlin, G. Fiol, C. Meuer, M. Kuntz, F. Hopfer, A. R. Kovsh, N. N. Ledentsov, and D. Bimberg, Distortion-free optical amplification of 20-80 GHz modelocked laser pulses at $1.3 \mu\text{m}$ using quantum dots, *Electronics Lett.* **42**, 697 (2006).
- [40] X. D. Huang, A. Stintz, H. Li, L. F. Lester, J. Cheng, and K. J. Malloy, Passive mode-locking in $1.3 \mu\text{m}$ two-section InAs quantum dot lasers, *Appl. Phys. Lett.* **78**, 2825 (2001).
- [41] F. Hopfer, A. Mutig, M. Kuntz, G. Fiol, D. Bimberg, N. N. Ledentsov, V. A. Shchukin, S. S. Mikhlin, D. L. Livshits, I. L. Krestnikov, A. R. Kovsh, N. D. Zakharov, and P. Werner, Single-mode submonolayer quantum-dot vertical-cavity surface-emitting lasers with high modulation bandwidth, *Appl. Phys. Lett.* **89**, 141106 (2006).
- [42] A. Lochmann, E. Stock, O. Schulz, R. Hopfer, D. Bimberg, V. A. Haisler, A. I. Toropov, A. K. Bakarov, and A. K. Kalagin, Electrically driven single quantum dot polarised single photon emitter, *Electronics Lett.* **42**, 774 (2006).
- [43] H. Eisele, O. Flebbe, T. Kalka, C. Preinesberger, F. Heinrichsdorff, A. Krost, D. Bimberg, and M. Dähne-Prietsch, Cross-sectional scanning-tunneling microscopy of stacked InAs quantum dots, *Appl. Phys. Lett.* **75**, 106 (1999).
- [44] J. Marquez, L. Geelhaar, and K. Jacobi, Atomically resolved structure of InAs quantum dots, *Appl. Phys. Lett.* **78**, 2309 (2001).
- [45] Y. Heiner, O. Stier, V. Türck, J. Vaschull, B. Sumpf, and A. Ostermeier, Evolution strategies applied to least squares curve fitting of spectroscopic data, *Journal of Quantitative Spectroscopy and Radiative Transfer* **56**, 769 (1996).
- [46] W. D. Sheng and J. P. Leburton, Electron-hole alignment in InAs/GaAs self-assembled quantum dots: Effects of chemical composition and dot shape, *Phys. Rev. B* **63**, 161301 (2001).
- [47] M. Grundmann, O. Stier, and D. Bimberg, InAs/GaAs Quantum Pyramids: Strain Distribution, Optical Phonons and Electronic Structure, *Phys. Rev. B* **52**, 11969 (1995).
- [48] O. Stier, M. Grundmann, and D. Bimberg, Electronic and optical properties of strained quantum dots modeled by 8-band-k.p theory, *Phys. Rev. B* **59**, 5688 (1999).
- [49] O. Stier, *Electronic and Optical Properties of Quantum Dots and Wires*, volume 7 of *Berlin Studies in Solid State Physics*, Wissenschaft und Technik Verlag, Berlin, 2001, [Dissertation, Technische Universität Berlin].
- [50] M. Grundmann, *Nano-Optoelectronics*, Springer, Berlin, 2002.

-
- [51] L.-W. Wang, J. Kim, and A. Zunger, Electronic structures of [110]-faceted self-assembled pyramidal InAs/GaAs quantum dots, *Phys. Rev. B* **59**, 5678 (1999).
- [52] J. Bourgoin and M. Lannoo, *Point Defects in Semiconductors II - Experimental Aspects*, volume 35 of *Springer Series in Solid-State Sciences*, Springer, Berlin, 1983.
- [53] P. Blood and J. W. Orton, *The Electrical Characterization of Semiconductors: Majority Carriers and Electron States*, Academic Press, London, 1992.
- [54] C. Kapteyn, *Carrier Emission and Electronic Properties of Self-Organized Semiconductor Quantum Dots*, Mensch & Buch Verlag, Berlin, 2001, [Dissertation, Technische Universität Berlin].
- [55] J. Frenkel, On pre-breakdown phenomena in insulators and electronic semiconductors, *Phys. Rev.* **54**, 647 (1938).
- [56] G. Vincent, A. Chantre, and D. Bois, Electric field effect on the thermal emission of traps in semiconductor junctions, *J. Appl. Phys.* **50**, 5484 (1979).
- [57] J. L. Hartke, The Three-Dimensional Poole-Frenkel Effect, *J. Appl. Phys.* **39**, 4871 (1968).
- [58] P. A. Martin, B. G. Streetman, and K. Hess, Electric field enhanced emission from non-Coulombic traps in semiconductors, *J. Appl. Phys.* **52**, 7409 (1981).
- [59] S. Makram-Ebeid and M. Lannoo, Electric-Field-Induced Phonon-Assisted Tunnel Ionization from Deep Levels in Semiconductors, *Phys. Rev. Lett.* **48**, 1281 (1982).
- [60] P. W. Fry, J. J. Finley, L. R. Wilson, A. Lemaître, D. J. Mowbray, and M. S. Skolnick, Electric-field-dependent carrier capture and escape in self-assembled InAs/GaAs quantum dots, *Appl. Phys. Lett.* **77**, 4344 (2000).
- [61] S. Makram-Ebeid and M. Lannoo, Quantum model for phonon-assisted tunnel ionization of deep levels in semiconductors, *Phys. Rev. B* **25**, 6406 (1982).
- [62] W. H. Chang, W. Y. Chen, T. M. Hsu, N. T. Yeh, and J. I. Chyi, Hole emission processes in InAs/GaAs self-assembled quantum dots, *Phys. Rev. B* **66**, 195337 (2002).
- [63] K. Irmscher, H. Klose, and K. Maas, Electric Field Enhanced Electron Emission from Gold Acceptor Level and A-Centre in Silicon, *Phys. stat. sol. (a)* **7**, 25 (1983).
- [64] U. S. Qurashi, M. Z. Iqbal, C. Delerue, and M. Lannoo, Electric-field dependence of electron emission from the deep-level oxygen defect in GaP, *Phys. Rev. B* **45**, 13331 (1992).

- [65] A. Dadgar, R. Engelhardt, M. Kuttler, and D. Bimberg, Capacitance transient study of the deep Fe acceptor in indium phosphide, *Phys. Rev. B* **56**, 10241 (1997).
- [66] C. M. A. Kapteyn, F. Heinrichsdorff, O. Stier, R. Heitz, M. Grundmann, N. D. Zakharov, D. Bimberg, and P. Werner, Electron Escape from InAs Quantum Dots, *Phys. Rev. B* **60**, 14265 (1999).
- [67] C. M. A. Kapteyn, M. Lion, R. Heitz, D. Bimberg, P. N. Brunkov, B. V. Volovik, S. G. Konnikov, A. R. Kovsh, and V. M. Ustinov, Hole and electron emission from InAs quantum dots, *Appl. Phys. Lett.* **76**, 1573 (2000).
- [68] W.-H. Chang, W. Chen, M. Cheng, C. Lai, T. Hsu, N.-T. Yeh, and J.-I. Chyi, Charging of embedded InAs Self assembled quantum dots by space-charge techniques, *Phys. Rev. B* **64**, 125315 (2001).
- [69] J. Z. Zhang and I. Galbraith, Intraband absorption for InAs/GaAs quantum dot infrared photodetectors, *Appl. Phys. Lett.* **84**, 1934 (2004).
- [70] O. G. Schmidt, N. Y. Jin-Phillipp, C. Lange, U. Denker, K. Eberl, R. Schreiner, H. Grbeldinger, and H. Schweizer, Long-range ordered lines of self-assembled Ge islands on a flat Si (001) surface, *Appl. Phys. Lett.* **77**, 4139 (2000).
- [71] S. M. Sze, *Semiconductor Devices*, John Wiley & Sons, 2nd edition, 2002.
- [72] H. Mizuta, Single- and few-electron memories, in *12th Japanese-German IT Forum*, 1998.
- [73] S. M. Sze, Evolution of Nonvolatile Semiconductor Memory: From Floating-Gate Concept to Single-Electron Memory Cell, in *Future Trends in Microelectronics*, edited by S. Luryi, J. Xu, and A. Zaslavsky, page 291, John Wiley & Sons, Inc., 1999.
- [74] N. R. Mahapatra and b. Venkatrao, The Processor-Memory bottleneck: Problems and Solutions., *ACM Crossroads* - (<http://www.acm.org/crossroads/xrds5-3/pmgap.html>) (1999).
- [75] Evolution of Parallel Computing, <http://www.intel.com/platforms/parallel.htm>, Intel Platforms 2006.
- [76] M. Geller, A. Marent, and D. Bimberg, Speicherzelle und Verfahren zum Speichern von Daten, Deutsche Patentanmeldung Nr. 10 2006 059 110.0, (27.10.2006).
- [77] M. Lannoo and J. Bourgoin, *Point Defects in Semiconductors I - Theoretical Aspects*, volume 22 of *Springer Series in Solid-State Sciences*, Springer, Berlin, 1981.

- [78] C. M. A. Kapteyn, M. Lion, R. Heitz, D. Bimberg, P. Brunkov, B. Volovik, S. G. Konnikov, A. R. Kovsh, and V. M. Ustinov, Time-Resolved Capacitance Spectroscopy of Hole and Electron Levels in InAs/GaAs Quantum Dots, *Phys. stat. sol. (b)* **224**, 57 (2001).
- [79] O. Engström, M. Malmkvist, Y. Fu, H. O. Olafsson, and E. O. Sveinbjörnsson, Thermal emission of electrons from selected s-shell configurations in InAs/GaAs quantum dots, *Appl. Phys. Lett.* **83**, 3578 (2003).
- [80] S. Schulz, S. Schnüll, C. Heyn, and W. Hansen, Charge-state dependence of InAs quantum-dot emission energies, *Phys. Rev. B* **69**, 195317 (2004).
- [81] C. Kittel and H. Kroemer, *Thermal Physics*, W. H. Freeman, San Francisco, 1980.
- [82] S. M. Sze, *Semiconductor Devices - Physics and Technology*, John Wiley & Sons, New York, 1985.
- [83] D. V. Lang and C. H. Henry, Nonradiative Recombination at Deep Levels in GaAs and GaP by Lattice-Relaxation Multiphonon Emission, *Phys. Rev. Lett.* **35**, 1525 (1975).
- [84] A. Wojs and P. Hawrylak, Charging and infrared spectroscopy of self-assembled quantum dots in a magnetic field, *Phys. Rev. B* **53**, 10841 (1996).
- [85] B. T. Miller, W. Hansen, S. Manus, R. J. Luyken, A. Lorke, J. P. Kotthaus, S. Huan, G. Medeiros-Ribeiro, and P. M. Petroff, Few-electron ground states of charge-tunable self-assembled quantum dots, *Phys. Rev. B* **56**, 6764 (1997).
- [86] B. Ohnesorge, M. Albrecht, J. Oshinowo, A. Forchel, and Y. Arakawa, Rapid carrier relaxation in self-assembled $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ quantum dots, *Phys. Rev. B* **54**, 11532 (1996).
- [87] S. Sanguinetti, K. Watanabe, T. Tateno, M. Wakaki, N. Koguchi, T. Kuroda, F. Minami, and M. Gurioli, Role of the wetting layer in the carrier relaxation in quantum dots, *Appl. Phys. Lett.* **81**, 613 (2002).
- [88] I. Magnúsdóttir, S. Bischoff, A. V. Uskov, and J. Mork, Geometry dependence of Auger carrier capture rates into cone-shaped self-assembled quantum dots, *Phys. Rev. B* **67**, 205326 (2003).
- [89] G. A. Narvaez, G. Bester, and A. Zunger, Carrier relaxation mechanisms in self-assembled (In,Ga)As/GaAs quantum dots: Efficient P \rightarrow S Auger relaxation of electrons, *Phys. Rev. B* **74**, 075403 (2006).
- [90] R. Heitz, M. Grundmann, N. N. Ledentsov, L. Eeckey, M. Veit, D. Bimberg, V. M. Ustinov, A. Y. Egorov, A. E. Zhukov, P. S. Kopev, and Z. I. Alferov, Multiphonon-relaxation processes in self-organized InAs/GaAs quantum dots, *Appl. Phys. Lett.* **68**, 361 (1996).

-
- [91] R. Heitz, M. Veit, N. N. Ledentsov, A. Hoffmann, D. Bimberg, V. M. Ustinov, P. S. Kop'ev, and Z. I. Alferov, Energy relaxation by multiphonon processes in InAs/GaAs quantum dots, *Phys. Rev. B* **56**, 10435 (1997).
- [92] I. Magnusdottir, A. V. Uskov, S. Bischoff, B. Tromborg, and J. Mork, One- and two-phonon capture processes in quantum dots, *J. Appl. Phys.* **92**, 5982 (2002).
- [93] I. Magnusdottir, A. V. Uskov, S. Bischoff, B. Tromborg, and J. Mork, Two-phonon capture processes into quantum dots: the role of intermediate states, *Physica E* **17**, 111 (2003).
- [94] J. Siegert, S. Marcinkevicius, and Q. X. Zhao, Carrier dynamics in modulation-doped InAs/GaAs quantum dots, *Phys. Rev. B* **72**, 085316 (2005).
- [95] O. Engström, M. Kaniewska, Y. Fu, J. Piscator, and M. Malmkvist, Electron capture cross sections of InAs/GaAs quantum dots, *Appl Phys. Lett.* **85**, 2908 (2004).
- [96] G. Myburg, F. D. Auret, W. E. Meyer, C. W. Louw, and M. J. v. Staden, Summary of Schottky barrier height data on epitaxially grown n- and p-GaAs, *Thin Solid Films* **325**, 181 (1998).
- [97] L. C. Kimerling, Influence of deep traps on the measurement of free-carrier distributions in semiconductors by junction capacitance techniques, *J. Appl. Phys.* **45**, 1839 (1974).
- [98] C. T. Sah, L. Forbes, L. L. Rosier, and A. F. Tasch, Jr., Thermal and optical emission and capture rates and cross sections of electrons and holes at imperfection centers in semiconductors from photo and dark junction current and capacitance experiments, *Solid-State Electronics* **13**, 759 (1970).
- [99] D. V. Lang, Deep-level transient spectroscopy: A new method to characterize traps in semiconductors, *J. Appl. Phys.* **45**, 3023 (1974).
- [100] G. L. Miller, D. V. Lang, and L. C. Kimerling, Capacitance Transient Spectroscopy, *Ann. Rev. Mat. Sci.* **7**, 377 (1977).
- [101] E. H. Rhoderick and R. H. Williams, *Metal Semiconductor Contacts*, Oxford Science Publications, 2nd edition, 1988.
- [102] H. G. Grimmeiss and C. Ovrén, Fundamentals of junction measurements in the study of deep energy levels in semiconductors, *J. Phys. E: Sci. Instrum.* **14**, 1032 (1981).
- [103] S. Anand, N. Carlsson, M.-E. Pistol, L. Samuelson, and W. Seifert, Deep level transient spectroscopy of InP quantum dots, *Appl. Phys. Lett.* **67**, 3016 (1995).

-
- [104] S. Anand, N. Carlsson, M.-E. Pistol, L. Samuelson, and W. Seifert, Electrical characterization of InP/GaInP quantum dots by space charge spectroscopy, *J. Appl. Phys.* **84**, 3747 (1998).
- [105] C. M. A. Kapteyn, M. Lion, R. Heitz, D. Bimberg, C. Miesner, T. Asperger, and G. Abstreiter, Many-particle effects in Ge quantum dots investigated by time-resolved capacitance spectroscopy, *Appl. Phys. Lett.* **77**, 4169 (2000).
- [106] P. Omling, L. Samuelson, and H. G. Grimmeis, Deep level transient spectroscopy evaluation of nonexponential transients in semiconductor alloys, *J. Appl. Phys.* **54**, 5117 (1983).
- [107] D. S. Day, M. Y. Tsai, B. G. Streetman, and D. V. Lang, Deep-level-transient spectroscopy: System effects and data analysis, *J. Appl. Phys.* **50**, 5093 (1979).
- [108] C. H. Henry, H. Kukimoto, G. L. Miller, and F. R. Merritt, Photocapacitance Studies of the Oxygen Donor in GaP. II. Capture Cross Sections, *Phys. Rev. B* **7**, 2499 (1973).
- [109] A. Majid, M. Z. Iqbal, A. Dadgar, and D. Bimberg, Osmium impurity-related deep levels in n-type GaAs, *J. Appl. Phys.* **98**, 083709 (2005).
- [110] F. Guffarth, R. Heitz, A. Schliwa, O. Stier, M. Geller, C. Kapteyn, R. Sellin, and D. Bimberg, Few-particle interactions in charged $\text{In}_x\text{Ga}_{1-x}\text{As}/\text{GaAs}$ quantum dots, *Phys. Rev. B* **67**, 235304 (2003).
- [111] I. Mukhametzhanov, R. Heitz, J. Zeng, P. Chen, and A. Madhukar, Independent manipulation of density and size of stress-driven self-assembled quantum dots, *Appl. Phys. Lett.* **73**, 1841 (1998).
- [112] F. Heinrichsdorff, *MOCVD growth and laser applications of In(Ga)As/GaAs quantum dots*, Mensch & Buch, Berlin, 1998, [Dissertation, Technische Universität Berlin].
- [113] R. Heitz, I. Mukhametzhanov, P. Chen, and A. Madhukar, Excitation transfer in self-organized asymmetric quantum dot pairs, *Phys. Rev. B* **58**, 10151 (1998).
- [114] H. Scher, M. F. Shlesinger, and J. T. Bendler, Time-scale invariance in transport and relaxation, *Phys. Today* **44**, 26 (1991).
- [115] *Gallium arsenide (GaAs), valence bands, effective masses*, volume 41/A1A of *Landolt-Börnstein - Group III Condensed Matter*, Springer, Berlin, 2001.
- [116] *Gallium arsenide (GaAs), conduction band, effective masses and related parameters*, volume 41/A1A of *Landolt-Börnstein - Group III Condensed Matter*, Springer, Berlin, 2001.

-
- [117] P. N. Brunkov, A. R. Kovsh, V. M. Ustinov, Y. G. Musikhin, N. N. Ledentsov, S. G. Konnikov, A. Polimeni, A. Patan, P. C. Main, L. Eaves, and C. M. A. Kapteyn, Emission of electrons from the ground and first excited states of self-organized InAs/GaAs quantum dot structures, *J. Electr. Mat.* **28**, 486 (1999).
- [118] F. Hatami, N. N. Ledentsov, M. Grundmann, J. Böhrer, F. Heinrichsdorff, M. Beer, D. Bimberg, S. S. Ruvimov, P. Werner, U. Gösele, J. Heydenreich, U. Richter, S. V. Ivanov, B. Y. Meltser, P. S. Kop'ev, and Z. I. Alferov, Radiative recombination in type-II GaSb/GaAs quantum dots, *Appl. Phys. Lett.* **67**, 656 (1995).
- [119] F. Hatami, M. Grundmann, N. N. Ledentsov, F. Heinrichsdorff, R. Heitz, J. Böhrer, D. Bimberg, S. S. Ruvimov, P. Werner, V. M. Ustinov, P. S. Kop'ev, and Z. I. Alferov, Carrier dynamics in type-II GaSb/GaAs quantum dots, *Phys. Rev. B* **57**, 4635 (1998).
- [120] B. R. Bennet, B. V. Shannabrook, and R. Magno, Phonons in self-assembled (In,Ga,Al)Sb quantum dots, *Appl. Phys. Lett.* **68**, 958 (1996).
- [121] E. R. Glaser, B. R. Bennet, B. V. Shannabrook, and R. Magno, Photoluminescence studies of self-assembled InSb, GaSb, and AlSb quantum dot heterostructures, *Appl. Phys. Lett.* **68**, 3614 (1996).
- [122] R. Magno, B. R. Bennett, and E. R. Glaser, Deep level transient capacitance measurements of GaSb self-assembled quantum dots, *J. Appl. Phys.* **88**, 5843 (2000).
- [123] L. Müller-Kirsch, R. Heitz, U. W. Pohl, and D. Bimberg, Temporal evolution of GaSb/GaAs quantum dot formation, *Appl. Phys. Lett.* **79**, 1027 (2001).
- [124] R. Timm, H. Eisele, A. Lenz, S. K. Becker, J. Grabowski, T. Y. Kim, L. Müller-Kirsch, K. Potschke, U. W. Pohl, D. Bimberg, and M. Dahne, Structure and intermixing of GaSb/GaAs quantum dots, *Appl. Phys. Lett.* **85**, 5890 (2004).
- [125] R. Wetzler, A. Wacker, E. Schöll, C. M. A. Kapteyn, R. Heitz, and D. Bimberg, Capacitance-voltage characteristics of InAs/GaAs quantum dots embedded in a pn structure, *Appl. Phys. Lett.* **77**, 1671 (2000).
- [126] D. Stievenard, X. Boddaert, and J. Bourgoin, Irradiation-induced defects in p-type GaAs, *Phys. Rev. B* **34**, 4048 (1986).
- [127] S. M. North, P. R. Briddon, M. A. Cusack, and M. Jaros, Electronic structure of GaSb/GaAs quantum dots, *Phys. Rev. B* **58**, 12601 (1998).
- [128] P. M. Mooney, Deep donor levels (DX centers) in III-V semiconductors, *J. Appl. Phys.* **67**, R1 (1990).

-
- [129] J. Batey and S. L. Wright, Energy band alignment in GaAs:(Al,Ga)As heterostructures: The dependence on alloy composition, *J. Appl. Phys.* **59**, 200 (1986).
- [130] M. Geller, C. Kapteyn, L. Müller-Kirsch, R. Heitz, and D. Bimberg, 450 meV hole localization energy in GaSb/GaAs quantum dots, *Appl. Phys. Lett.* **82**, 2706 (2003).
- [131] H. H. Lee, R. J. Racicot, and S. H. Lee, Surface passivation of GaAs, *Appl. Phys. Lett.* **54**, 724 (1988).
- [132] D. Cavalcoli, A. Cavallini, and E. Gombia, Anomalous temperature dependence of deep-level-transient-spectroscopy peak amplitude, *Phys. Rev. B* **56**, 14890 (1997).
- [133] M. Grundmann and D. Bimberg, Theory of random population for quantum dots, *Phys. Rev. B* **55**, 9740 (1997).
- [134] M. D. Giorgi, C. Lingk, G. v. Plessen, J. Feldmann, S. D. Rinaldis, A. Passaseo, M. D. Vittorio, R. Cingolani, and M. Lomascolo, Capture and thermal re-emission of carriers in long-wavelength InGaAs/GaAs quantum dots, *Appl. Phys. Lett.* **79**, 3968 (2001).
- [135] H. Benisty, C. M. Sotomayor-Torres, and C. Weisbuch, Intrinsic mechanism for the poor luminescence properties of quantum-box systems, *Phys. Rev. B* **44**, 10945 (1991).
- [136] U. Bockelmann and T. Egeler, Electron relaxation in quantum dots by means of Auger processes, *Phys. Rev. B* **46**, 15574 (1992).
- [137] R. Ferreira and G. Bastard, Phonon-assisted capture and intradot Auger relaxation in quantum dots, *Appl. Phys. Lett.* **74**, 2818 (1999).
- [138] T. Inoshita and H. Sakaki, Electron relaxation in a quantum dot: Significance of multiphonon processes, *Phys. Rev. B* **46**, 7260 (1992).
- [139] M. J. Steer, D. J. Mowbray, W. R. Tribe, M. S. Skolnick, M. D. Sturge, M. Hopkinson, A. G. Cullis, C. R. Whitehouse, and R. Murray, Electronic energy levels and energy relaxation mechanisms in self-organized InAs/GaAs quantum dots, *Phys. Rev. B* **54**, 17738 (1996).
- [140] T. R. Nielsen, P. Gartner, and F. Jahnke, Many-body theory of carrier capture and relaxation in semiconductor quantum-dot lasers, *Phys. Rev. B* **69**, 235314 (2004).
- [141] D. F. Schoeter, D. J. Griffiths, and P. C. Sercel, Defect-assisted relaxation in quantum dots at low temperature, *Phys. Rev. B* **54**, 1486 (1996).

-
- [142] X. Q. Li and Y. Arakawa, Ultrafast energy relaxation in quantum dots through defect states: A lattice-relaxation approach, *Phys. Rev. B* **56**, 10423 (1997).
- [143] P. Y. Yu and M. Cardona, *Fundamentals of Semiconductors - Physics and Materials Properties*, Springer, Berlin, 1996.
- [144] S. W. Lin, C. Balocco, M. Missous, A. R. Peaker, and A. M. Song, Coexistence of deep levels with optically active InAs quantum dots, *Phys. Rev. B* **72**, 165302 (2005).
- [145] HP 4284A Precision LCR Meter Operation Manual, 1996.
- [146] Instruction Manual Model 72B Capacitance Meter, 1983.
- [147] Instruction Manual Model 7200 Capacitance Meter, 1993.
- [148] T. I. Chappell and C. M. Ransom, Modifications to the Boonton 72BD capacitance meter for deep-level transient spectroscopy applications, *Rev. Sci. Instrum.* **55**, 200 (1983).
- [149] M. Lion, *Deep Level Transient Spectroscopy an selbstorganisierten InAs/GaAs Quantenpunkten*, Studienarbeit, Technische Universität, 1999.
- [150] M. Lion, *Optisch angeregte Kapazitätsspektroskopie an selbstorganisierten Quantenpunkten*, Diplomarbeit, Technische Universität, 2000.
- [151] M. S. Carpenter, M. R. Melloch, M. S. Lundstrom, and S. P. Tobin, Effects of Na_2S and $(\text{NH}_4)_2\text{S}$ Edge Passivation Treatments on the Dark Current-Voltage Characteristics of GaAs pn Diodes, *Appl. Phys. Lett.* **52**, 2157 (1988).
- [152] V. N. Bessolov, M. V. Lebedev, E. B. Novikov, and B. V. Tsarenkov, Sulfide Passivation of III-V semiconductors: Kinetics of the photoelectrochemical reaction, *J. Vac. Sci. Technol. B* **11**, 10 (1992).
- [153] V. Fischer, T.-J. Kim, P. H. Holloway, E. Ristolainen, and D. Schoenfeld, Formation of ohmic contacts to n-GaAs using $(\text{NH}_4)_2\text{S}$ surface passivation, *J. Vac. Sci. Technol. B* **12**, 1419 (1993).
- [154] V. N. Bessolov, E. V. Konenkova, and M. V. Lebedev, Solvent effect on the properties of sulfur passivated GaAs, *J. Vac. Sci. Technol. B* **14**, 2761 (1996).
- [155] V. N. Bessolov, E. V. Konenkova, and M. V. Lebedev, Passivation of GaAs in alcohol solutions of ammonium sulfide, *Semiconductors* **31**, 1164 (1997).
- [156] Y. Dong, X. M. Ding, X. Y. Hou, Y. Li, and X. B. Li, Sulfur passivation of GaAs metal-semiconductor field-effect transistor, *Appl. Phys. Lett.* **77**, 3839 (2000).

Publications

Parts of this work have already been published:

- M. Geller, C. Kapteyn, L. Müller-Kirsch, R. Heitz, and D. Bimberg
450 meV hole localization in GaSb/GaAs quantum dots
Appl. Phys. Lett. **82**, 2706 (2003)
- M. Geller, C. Kapteyn, L. Müller-Kirsch, R. Heitz, and D. Bimberg
Strong hole localization in type-II GaSb/GaAs quantum dots
Proc. 26th Int. Conf. on the Physics of Semiconductors, Edinburgh, UK, A. R. Long and J. H. Davies eds., Institute of Physics Publishing, H184 (2003).
- M. Geller, C. Kapteyn, L. Müller-Kirsch, R. Heitz, and D. Bimberg
Hole storage in GaSb/GaAs quantum dots for memory devices
Phys. stat. sol. (b) **238**, 258 (2003).
- M. Geller, C. Kapteyn, E. Stock, L. Müller-Kirsch, R. Heitz, and D. Bimberg
Energy-selective charging of type-II GaSb/GaAs quantum dots
Physica E **21**, 474-478 (2004).
- M. Geller
Deep level transient spectroscopy of self-organized quantum dots
Semiconductor News **13**, 235 (2004).
- M. Geller, E. Stock, C. Kapteyn, R. L. Sellin, and D. Bimberg
Tunneling emission from self-organized In(Ga)As/GaAs quantum dots observed via time-resolved capacitance measurements
Phys. Rev. B **73**, 205331 (2006).
- M. Geller, E. Stock, R. L. Sellin, and D. Bimberg
Direct observation of tunneling emission to determine localization energies in self-organized In(Ga)As quantum dots
Physica E **32**, 171 (2006).
- M. Geller, E. Stock, A. Marent, A. P. Vasi'ev, E. S. Semenova, A. E. Zhukov, V. M. Ustinov, and D. Bimberg
Carrier storage and confinement in GaAs-based self-organized quantum dots
Phys. stat. sol. (c) **3**, 504 (2006).

- A. Marent, M. Geller, A. P. Vasi'ev, E. S. Semenova, A. E. Zhukov, V. M. Ustinov, and D. Bimberg
Carrier storage time of milliseconds at room temperature in self-organized In(Ga)As quantum dots
Appl. Phys. Lett. **89**, 072103 (2006)
- M. Geller, A. Marent, E. Stock, A. E. Zubkov, I. S. Shulgunova, A. V. Solomonov, and D. Bimberg
Hole capture into self-organized InGaAs quantum dots
Appl. Phys. Lett. **89**, 232105 (2006).
- M. Geller, A. Marent, A. P. Vasi'ev, E. S. Semenova, A. E. Zhukov, V. M. Ustinov, and D. Bimberg
5 ms hole storage time at room temperature in self-organized InGaAs QDs
Proceeding of 14th Int. Symp. "Nanostructures: Physics and Technology", page 125-126, St. Petersburg, Russia June 26-30, (2006).
- A. Marent, M. Geller, A. E. Zubkov, I. S. Shulgunova, A. V. Solomonov, and D. Bimberg
Carrier capture into self-organized InGaAs quantum dots
Proceeding of 14th Int. Symp. "Nanostructures: Physics and Technology", page 152-153, St. Petersburg, Russia June 26-30, (2006).

Patent:

- M. Geller, A. Marent, and D. Bimberg
Speicherzelle und Verfahren zum Speichern von Daten
Deutsche Patentanmeldung Nr. 10 2006 059 110.0, (27. 10. 2006)

Other Publications:

- M. Geller, I. Manke, K. Hodeck, R. Heitz, and M. Dähne
Ringlike emission profiles in scanning near-field photoluminescence images of single InGaAs quantum dots
Phys. Rev. B **64**, 233312 (2001).
- F. Guffarth, R. Heitz, M. Geller, C. Kapteyn, H. Born, R. Sellin, A. Hoffmann, D. Bimberg, N. A. Sobolev, and M. C. Carmo
Radiation hardness of InGaAs/GaAs quantum dots
Appl. Phys. Lett. **82**, 1941 (2003).
- F. Guffarth, R. Heitz, A. Schliwa, O. Stier, C. M. A. Kapteyn, M. Geller, R. Sellin, and D. Bimberg
Excited state transitions in negatively and positively charged InGaAs/GaAs quantum dots

-
- Proc. 26th Int. Conf. on the Physics of Semiconductors, Edinburgh, UK, A. R. Long and J. H. Davies eds., Institute of Physics Publishing, D165 (2003).
- K. Hodeck, I. Manke, M. Geller, J. L. Spithoven, J. Lorbacher, R. Heitz, F. Heinrichsdorff, A. Krost, D. Bimberg, and M. Dähne
Photoluminescence of individual InGaAs quantum dots
Proc. 26th Int. Conf. on the Physics of Semiconductors, Edinburgh, UK, A. R. Long and J. H. Davies eds., Institute of Physics Publishing, D165 (2003).
 - K. Hodeck, I. Manke, M. Geller, R. Heitz, F. Heinrichsdorff, A. Krost, D. Bimberg, H. Eisele, M. Dähne
Multiline photoluminescence of single InGaAs quantum dots
Phys. stat. sol. (c) **4**, 1209 (2003).
 - F. Guffarth, R. Heitz, A. Schliwa, O. Stier, M. Geller, C. M. A. Kapteyn, R. Sellin, and D. Bimberg
Few-particle interactions in charged InGaAs/GaAs quantum dots
Phys. Rev. B **67**, 235304 (2003).
 - M. Hayne, O. Razinkova, S. Bersier, R. Heitz, L. Müller-Kirsch, M. Geller, D. Bimberg and V. V. Moshchalkov
Optically induced charging effects in self-assembled GaSb/GaAs quantum dots
Phys. Rev. B **70**, 081302 (2004).
 - T. Warming, W. Wiczorek, M. Geller, A. Zhukov, V. M. Ustinov, and D. Bimberg
Spectral hole burning by storage of electrons or holes
AIP Conf. Proc. **772**, 1555 (2005).
 - W. Wiczorek, T. Warming, M. Geller, and D. Bimberg
Charge and spin storage in self-organized quantum dots
Appl. Phys. Lett. **88**, 182107 (2006).
 - W. Wiczorek, T. Warming, M. Geller, and D. Bimberg
Charge and spin storage in self-organized quantum dots
Proc. 27th Int. Conf. on the Physics of Semiconductors, Wien (2007), in print.
 - A. Marent, M. Geller, A. P. Vasi'ev, E. S. Semenova, A. E. Zhukov, V. M. Ustinov, and D. Bimberg
DRAM storage time of milliseconds demonstrated in self-organized quantum dots
Proc. 27th Int. Conf. on the Physics of Semiconductors, Wien (2007), in print.

Book chapter:

- M. Geller and A. Marent
Quantum dots for memories

in Book "Semiconductor Nanostructures", in preparation.
Springer (2007), editor: D. Bimberg.

Danksagung

An dieser Stelle möchte ich mich allen bedanken, die zum Gelingen dieser Arbeit beigetragen haben. Alle, die ich vergessen haben sollte, mögen mir verzeihen.

Mein ganz besonderer Dank gilt Herrn Prof. Bimberg, der mir die Möglichkeit gegeben hat, in seiner Gruppe zu promovieren. Neben der exzellenten und vielgelobten Ausstattung möchte ich mich vor allem für das nahezu grenzenlose Vertrauen in meine Arbeit und die großartige finanzielle und fachliche Unterstützung bedanken. Ich habe viel gelernt, danke!

Zwei Personen haben meine wissenschaftliche Arbeit und mein „physikalisches Denken“ entscheidend geprägt: Christian Kapteyn und Robert Heitz. Beiden möchte ich ganz herzlich für die Unterstützung danken. Robert Heitz gilt nicht nur mein besonderer Dank, sondern auch mein Versprechen, ihn nicht zu vergessen.

Einen entscheidenden Anteil am Zustandekommen dieser Arbeit haben Erik Stock und Andreas Marent, deren Diplomarbeiten ich betreuen durfte. Beiden möchte ich ganz besonders danken und mich gleichzeitig bei Till Warming für die endlosen Diskussionen über DLTS-Probleme während der Mittagspause entschuldigen. Bei Andreas Marent möchte ich mich ganz besonders für die vielen kritischen Fragen und ausgiebigen Diskussionen bedanken, die uns immer wieder vorangebracht haben, auch wenn sich nicht alle Fragen klären ließen. Außerdem möchte ich mich bei Andreas für die sehr genaue Korrektur dieser Arbeit bedanken. Ich wünsche ihm viel Erfolg bei der Realisierung eines ersten „Quantenpunktspeichers“.

Sven Rodt sei ganz herzlich für seine unerschöpfliche Hilfsbereitschaft gedankt, die nicht nur Computerfragen betrifft, sondern auch z. B. das Korrigieren einzelner Passagen dieser Arbeit. Genauso möchte ich mich bei Robert Seguin für die Korrektur von Teilen dieser Arbeit bedanken. Andrei Schliwa danke ich für die 8-band k - p Berechnungen der GaSb-Proben. Bei Lutz Müller-Kirsch möchte ich mich für die MOCVD-gewachsenen GaSb-Proben bedanken. Konstantin Pötschke danke ich für die Hilfe bei Computerproblemen und die eine oder andere politische Diskussion. For the MBE-grown InAs samples I have to thank Alexey Vasi'ev, Elisaveta Semenova, Alexey Zhukov, and Victor Ustinov from the Ioffe Institute in St. Petersburg. Ich möchte allen derzeitigen und auch früheren Mitglieder der Arbeitsgruppe Bimberg für die gute Arbeitsatmosphäre, die vielen Doktorfeiern und die permanente Unterstützung danken.

Herrn Prof. Marius Grundmann danke ich für die Zweitkorrektur dieser Arbeit, Herrn Prof. Thomas Möller für den Prüfungsvorsitz.

Bei Christian Kristukat möchte ich mich ganz besonders dafür bedanken, dass er ans andere Ende der Welt gezogen ist, um meine Arbeit in aller Ruhe korrigieren zu können.

Mein Dank gilt meinen Eltern, für die stete Unterstützung in allem, was ich je getan habe. Bei meiner heißgeliebten Nachteule, meiner Tochter Mia, möchte ich mich für die vielen wachen Nächte bedanken. Der größte Dank gebührt Nele dafür, dass sie immer da war, wenn ich sie brauchte.