

# Chapter 2

## Fundamentals

The following chapter gives an overview of the fundamentals of this work. At first, semiconductor heterostructures are described, building the basis for MODFETs and quantum dots. An overview of the most common semiconductor memories used today, including their advantages and disadvantages, leads the way to the concept of a quantum dot-based memory device which is described at the end of this chapter.

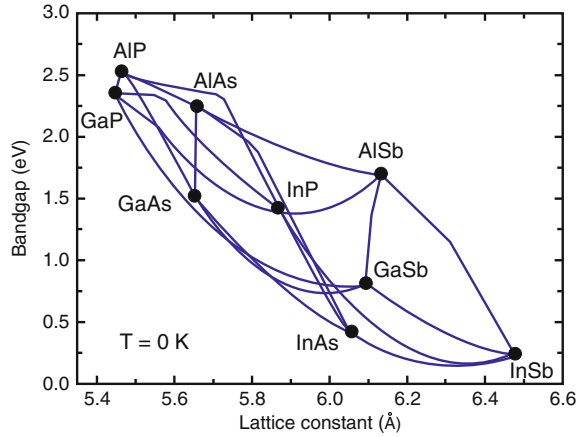
### 2.1 Semiconductor Heterostructures

With the maturing of the quality of semiconductor crystals grown by Molecular Beam Epitaxy (MBE) [1] and Metal-Organic Chemical Vapor Deposition (MOCVD) [2], it became possible to grow almost atomically-abrupt semiconductor layers on top of each other.

When two different semiconductors are brought into contact, a heterostructure is formed. The development of the heterostructure has led to numerous applications, such as the High-Electron Mobility Transistor (HEMT) [3, 4], lasers [5], and amplifiers [6]. This is owed to the fact that the electronic properties of the entire device are determined by the heterostructure, allowing to design devices with specific electronic and optical properties. The most prominent one of such designed heterostructures is the double-heterostructure laser, invented by Zhores I. Alferov and Herbert Kroemer in 1963 [7, 8], which was honored with the Nobel price in the year 2000.

The materials used in heterostructures are predominantly II-VI or III-V compound semiconductors. Besides commonly fabricating binary compounds, it is possible to create a variety of different ternary, quaternary, even quinary compounds, which allow to specifically design the bandgap of the semiconductor. Figure. 2.1 shows the bandgap versus the lattice constant of the most common III-V semiconductors with their ternary alloys [9]. The III-V compounds offer a wide variety of different materials which can be combined in a heterostructure in order to have a precisely controlled band design.

**Fig. 2.1** Energy band gap versus the lattice constant of III-V compound semiconductors. The *black data points* are the binary alloys and the *blue lines* the ternary alloys (after [9])

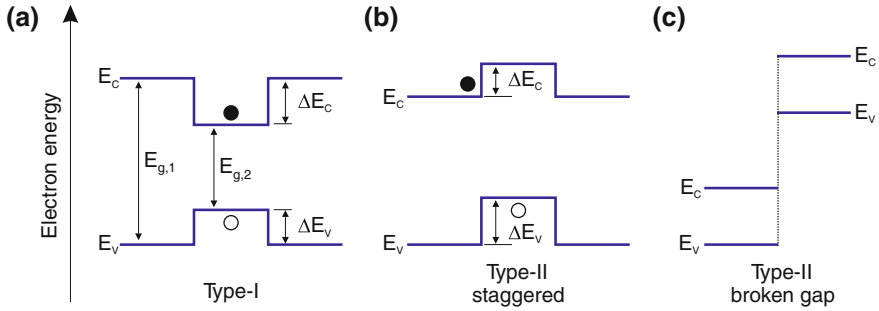


From a device physics point of view, the key property of a heterostructure interface is the band alignment of the two semiconductors. Three different alignments are possible [10, 11] which are shown in Fig. 2.2:

- *Type-I alignment.* The conduction band and the valence band of one semiconductor lie completely within the bandgap of the other semiconductor (Fig. 2.2a). If one semiconductor is surrounded by another semiconductor, charge carriers can be spatially restricted in their ability to move. Material systems of this kind are InAs/GaAs, GaAs/AlGaAs, and GaSb/AlSb.
- *Type-II staggered alignment.* The bandgap of one semiconductor does not lie completely within the bandgap of the other semiconductor, but there is still some overlap (Fig. 2.2b). In such a heterostructure, charge carriers of one type can be confined, while the other type is not confined or even experiences a barrier. A material system with this type of interface is GaSb/GaAs which plays an important role in this work.
- *Type-II broken-gap (also called Type-III) alignment.* The conduction and the valence band of one semiconductor lie completely above or below the conduction and valence band of the other semiconductor (Fig. 2.2c), the bandgaps of the two semiconductors do not overlap. An example for type-III alignment is the material combination InAs/GaSb.

## 2.2 Nanostructures

If the structural size of the functional element in a heterostructure enters the size range of the de Broglie matter wavelength, the electronic and optical properties are governed by size effects described by quantum mechanics. As a result of the reduced effective mass for electrons and holes in semiconductors, the de Broglie wavelength



**Fig. 2.2** Different types of heterostructures:  $E_g$  is the bandgap of the semiconductor,  $\Delta E_C$  and  $\Delta E_V$  are the conduction band and the valence band offsets in the conduction band  $E_C$  and the valence band  $E_V$ , respectively. **a** Type-I: electrons and holes are confined within the same spatial region. **b** Type-II staggered: only one charge carrier type (here holes) is confined. **c** Type-II broken gap: no overlap of the band gaps

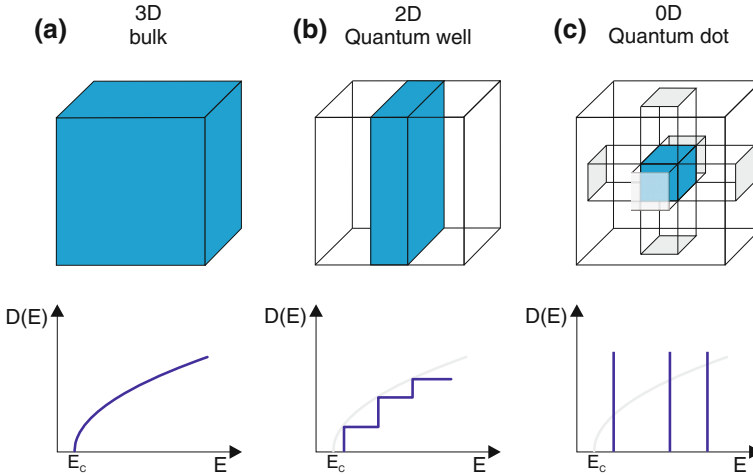
is in the nanometer range. Solving the Schrödinger equation for such nanometer-sized potentials results in quantization energies larger than  $kT$ , which ultimately leads to a completely discrete density of states. Figure 2.3 shows the densities of states for various structures, in which the directions in which charge carriers can freely move are further and further reduced. In a bulk semiconductor the movement of the charge carriers is not restricted and they can freely move in all three spatial dimensions. The density of states is proportional to  $\sqrt{E}$ . Two low-dimensional heterostructures are of importance in this work. In the first, the charge carriers can freely move within two dimensions and are restricted within only one dimension (see Fig. 2.3b). In such a *quantum well* (QW) the density of states for each energy level  $E_i$  is [12]

$$D_{2D}(E_i) = \frac{m^*}{\pi \hbar^2}, \quad (2.1)$$

with the effective mass  $m^*$  and the Planck constant  $\hbar$ . In the second structure, the movement of the carriers is restricted in all three spatial dimensions and the carriers are completely confined. In such a *quantum dot* (QD), complete quantization of the energy levels results in a completely discrete density of states [12], similar to those in an atom

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

with the energy levels at  $E_i$ .



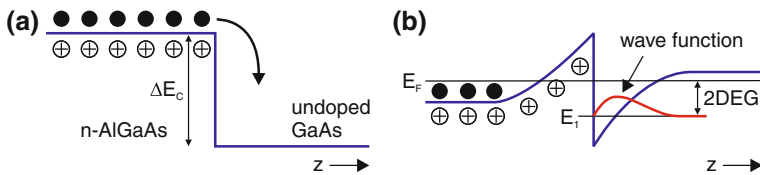
**Fig. 2.3** Density of states  $D(E)$  in semiconductors. **a** Bulk semiconductor with a  $\sqrt{E}$ -shaped density of states. **b** Two-dimensional quantum well with a step-like density of states. **c** Zero-dimensional quantum dot with a completely discrete density of states

## 2.3 Two-Dimensional Carrier Gases and Field-Effect Transistors

Based on the restriction of carrier movement in one spatial dimension, devices containing layers with very high carrier mobility can be built. These layers allow the fabrication of very fast transistors. The basic technique and device structure is described in the following sections.

### 2.3.1 Modulation-Doping

The properties of semiconductors can be altered by doping. By incorporation of donors and acceptors, the conductance of a semiconductor can be enhanced by orders of magnitude. When the doping atoms release their surplus charges, ionized donors or acceptors are left behind and act as Coulomb scatterers (ionized impurity scattering) which can inhibit the motion of free charges [13], blur energy levels, and disrupt the interference of electron waves [12]. The solution of this problem is to separate the active region of the device from the doping atoms, a method which is called remote or modulation doping [14]. Figure 2.4 schematically depicts the formation of a two-dimensional electron gas (2DEG) at the interface of a heterojunction between n-doped AlGaAs and undoped GaAs. Initially, the material is neutral (Fig. 2.4a) and the conduction band is flat in the AlGaAs and in the GaAs. When the electrons of the donors in the AlGaAs are released, they move around and some of them diffuse



**Fig. 2.4** Conduction band around the interface of a heterojunction between n-doped AlGaAs and undoped GaAs. **a** The electrons released by the donors are diffusing to the GaAs and lose potential energy. **b** Field-induced band bending and formation of a two-dimensional electron gas in the triangular potential well

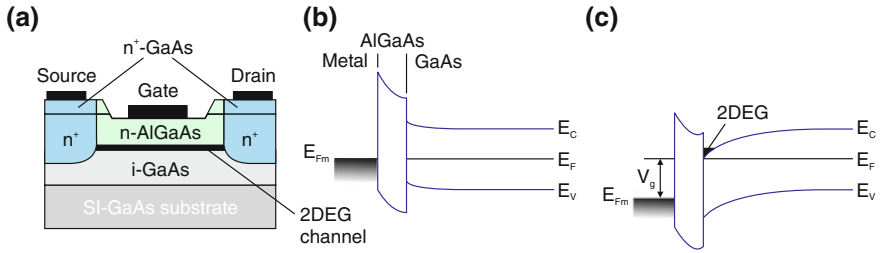
into the GaAs, where they lose potential energy and are trapped in the GaAs because they cannot overcome the barrier  $\Delta E_C$ . The negatively charged electrons are now separated from the positively charged donors leading to an electric field which drives the electrons back towards the interface. The electrons are trapped in a field-induced triangular well with a typical width of  $\sim 10$  nm. As the width of the well is below the de Broglie wavelength of the electrons, the quantization of the energy levels in  $z$ -direction becomes relevant. If only the lowest energy level is occupied, all electrons exhibit the same state in  $z$ -direction, but remain free to move in  $x$ - and  $y$ -direction. This way, a 2DEG is formed with the total energy of an electron sitting on the energy level  $E_n$  [12]

$$E_n(k_x, k_y) = E_n + \frac{\hbar^2 k_x^2}{2m^*} + \frac{\hbar^2 k_y^2}{2m^*}, \quad (2.3)$$

with the wave vectors  $k_x$  and  $k_y$  in  $x$  and  $y$  direction, and the effective mass  $m^*$ . Using the approach of modulation doping, later refined by inserting an undoped spacer layer between the donors and the heterojunction, very high electron mobilities of up to  $> 10^7$  cm<sup>2</sup>/Vs [15] can be achieved. Creating 2DEGs with such high mobilities at low temperature allowed the discovery of the Quantum Hall Effect [16, 17]. If p-doped AlGaAs is used instead of n-doped AlGaAs, a two-dimensional hole gas (2DHG) is formed in a similar way. Here, mobilities of up to  $10^6$  cm<sup>2</sup>/Vs [18] have been achieved. The invention of modulation doping has created high-speed device applications, such as the modulation-doped field-effect transistor (MODFET), also known as high-electron mobility transistor (HEMT) [11].

### 2.3.2 Modulation-Doped Field-Effect Transistor

The principle of modulation doping and the resulting high mobility is the key advantage of a modulation-doped field-effect transistor (MODFET) over conventional MOSFETs. Figure 2.5a depicts the schematic device structure of an enhancement mode AlGaAs/GaAs MODFET. It consists of a layer of undoped GaAs and a top



**Fig. 2.5** Modulation-doped field-effect transistor (MODFET) (after [11]). **a** Schematic device structure. **b** Energy band diagram of an enhancement mode MODFET at equilibrium. **c** Onset of threshold: if  $V_g > V_T$  electrons accumulate at the heterointerface and a 2DEG is formed which acts as channel

layer of n-doped AlGaAs. The electrons from the AlGaAs are transferred to the GaAs, and a 2DEG is formed. The 2DEG is contacted by two Ohmic source and drain contacts, which use highly doped regions fabricated either by ion implantation or introduction of dopants during the alloying step. A Schottky gate contact allows to alter the potential in the 2DEG. The working principle of this enhancement MODFET is shown in Figs. 2.5b, c. In equilibrium ( $V_g = 0$  V), the Fermi level  $E_F = E_{Fm}$  is the same throughout the device, the channel conductance is very low and the transistor is off. If a gate voltage larger than the threshold voltage  $V_T$  is applied, the conduction and the valence band are tilted and the electrons in the GaAs accumulate at the heterointerface, forming a 2DEG inside the triangular well. The channel conductance is now very high and the transistor is on. This way, a controlled switching of the conductance of the channel is achieved.

## 2.4 Quantum Dots

Self-organized quantum dots (QDs) [19, 20] are low-dimensional heterostructures which confine the charge carriers within all three spatial dimensions. The behavior of such a zero-dimensional system is similar to real atoms and has coined the term *artificial atom* for QDs. The qualitatively distinct electronic and optical properties of QDs have led to numerous applications in devices such as field-effect transistors [21–23], memory cells [24–28], lasers [29–34], amplifiers [35], and single photon sources [36–38].

### 2.4.1 Fabrication

There are two fundamentally different approaches to fabricate QDs. In the *top-down* approach QDs are defined by patterning QWs using lithographical techniques (i.e

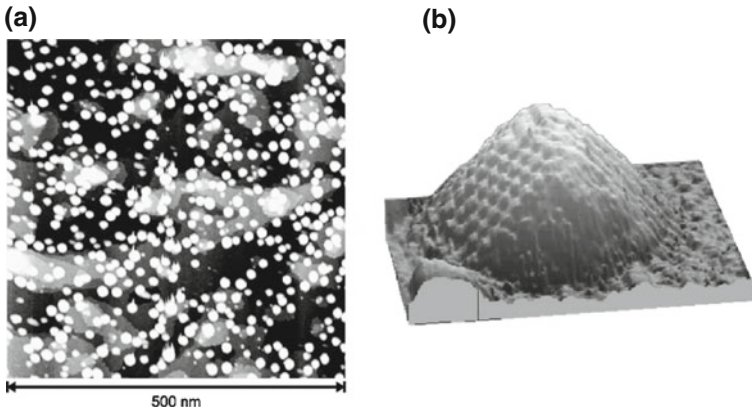
electron beam lithography) [19]. The quantum well material has a smaller bandgap than the surrounding material and confines the charge carriers within the growth direction. A mask is defined on the surface, framing the lateral size of the QD, which is then defined by using standard dry-etching techniques. The definition of these mesas leads to a spatial confinement of charge carriers within the QW layer, now becoming a QD.

In the other approach, which was used for the fabrication of the QDs investigated in this work, the QDs are grown in a self-organized way [19]. In this *bottom-up* approach mainly two different techniques are used: molecular beam epitaxy (MBE) [1] and metal-organic chemical vapor deposition (MOCVD) [2]. During MBE growth, a molecular beam of the source materials is directed onto the substrate and an epitaxial layer is formed on the surface. Under ultra-high vacuum conditions (UHV), a crystalline compound semiconductor can be grown with a very high purity and quality. During MOCVD, precursor gases are transported to the surface of the substrate by a carrier gas. Within the reactor, a chemical reaction cracks the precursors into its constituents, creating a supersaturated gas phase which leads to condensation on top of the substrate surface, where the surface kinetics lead to the formation of an epitaxial layer.

If the epitaxially grown material and the substrate are different semiconductor materials, they usually exhibit a different lattice constant (see Fig. 2.1). Hence, the lattice constant of the deposited layer will adjust to the lattice constant of the layer underneath, leading to coherent strain. If growth proceeds, the successive layer will at some point reach its original lattice constant, as the strain is continuously reduced. If the difference of the lattice constants of the two materials is too large, the strain is relaxed and the deposited layer keeps its original lattice constant, leading to dislocations.

The coherent strain can be utilized to grow self-organized QDs. Essentially, there are three modes of heteroepitaxial growth. If the difference of the lattice constant of the two semiconductors is small, the strain is reduced within a few layers without creating dislocations. In this Frank-van der Merwe growth mode [39], the growth proceeds as a layer-by-layer growth, where the complete surface of the substrate is wetted by the epitaxial layer. If the difference in the lattice constants is larger, the strain can be reduced by a transition to three-dimensional island growth. In this Volmer-Weber [40] growth mode the total energy is minimized by the island formation [41]. In a mixture of the previous two growth modes, called Stranski-Krastanow mode [42], the surface of the substrate is first covered by a wetting layer (WL), followed by the formation of islands in the process of total energy minimization [41]. This growth mode is used to fabricate quantum dots with a few monolayers height and a few ten nanometers in diameter. Depending on the growth conditions, area densities of up to  $1 \times 10^{11} \text{ cm}^{-2}$  can be reached. Figure 2.6 shows a Scanning Tunneling Microscope (STM) image of a surface covered with self-organized QDs and a three-dimensional depiction of a single QD.

All QDs in the samples investigated in this work were fabricated using the Stranski-Krastanow growth mode.



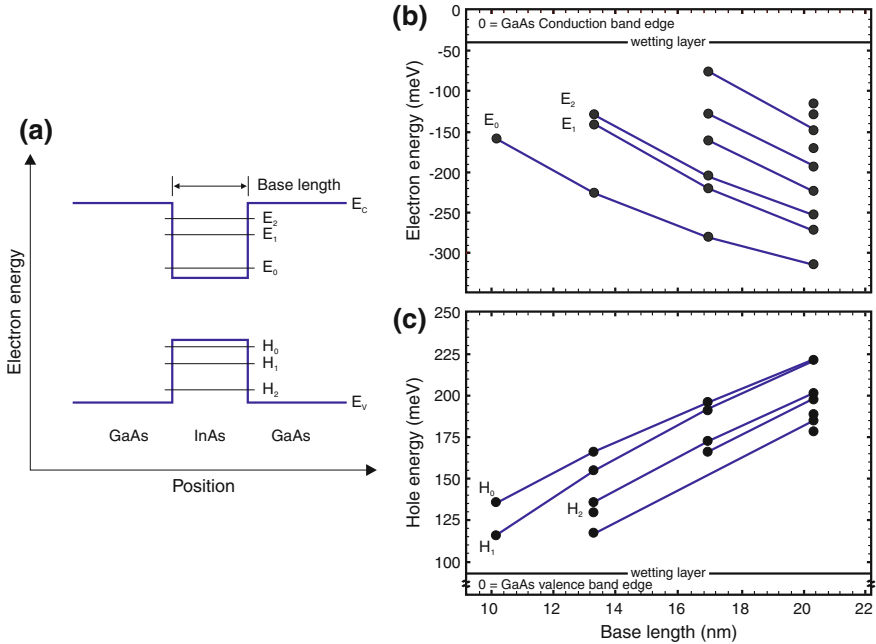
**Fig. 2.6** **a** Surface STM image of self-organized InAs QDs grown on GaAs(001) in the Stranski-Krastanow mode on an area of  $500 \times 500 \text{ nm}^2$ . The QDs (*white*), the wetting layer (*grey*), and the GaAs (*black*) can be clearly distinguished (Reprinted from [43], with permission from Elsevier). **b** High-resolution STM image of a single QD depicted as a three-dimensional model (Reprinted with permission from [44], Copyright 2001, American Institute of Physics)

### 2.4.2 Electronic Structure

As a result of the small structural size of self-organized QDs, they can confine charge carriers and restrict their motion in all three spatial dimensions. In the simplest theoretical model, QDs can be described as a quantum mechanical potential well with infinite barriers. It is well known from quantum mechanics, that the energy levels in such a potential well are quantized [45]. Nevertheless, this simple model neglects many other effects, which are important for the electronic structure. Strain, piezoelectric effects, and the chemical composition within the QDs and the surrounding matrix material have to be taken into account to determine the exact position of the energy levels within the QD.

The exact electronic structures of self-organized QDs can be calculated numerically using a three-dimensional model within the framework of 8-band- $\mathbf{k} \cdot \mathbf{p}$  theory [46–48]. Figure 2.7 shows the results of such a calculation for InAs QDs embedded into a GaAs matrix. It can clearly be seen that the energy levels for electrons and holes shift towards the respective band edge when the size of the QDs is reduced. This shift leads to a reduction of the number of bound states in the QDs with decreasing size. In this particular example QD, there is only one bound state for electrons and two bound states for holes at a base length of 10 nm, whereas the number increases to 9 electron states and 6 hole states at a base length of 20 nm.

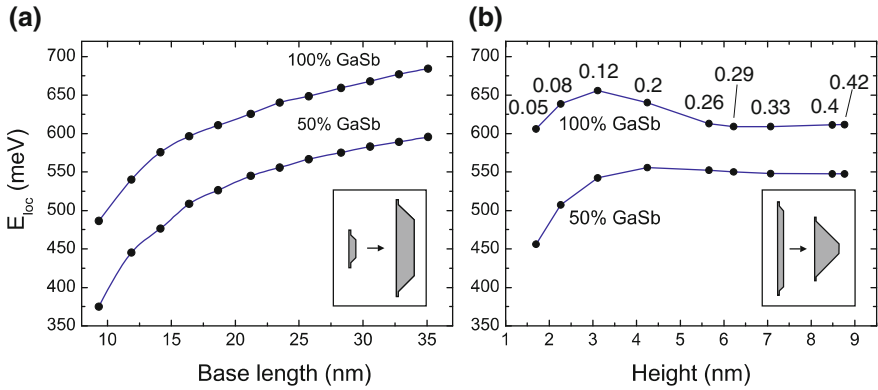
Figure 2.8 shows the hole localization energy (energy difference between the hole ground state and the valence band edge of the surrounding material) for GaSb/GaAs QDs, which are investigated in this work. The calculations have been done by Andrei Schliwa at TU-Berlin. They also include strain and piezoelectric effects and model



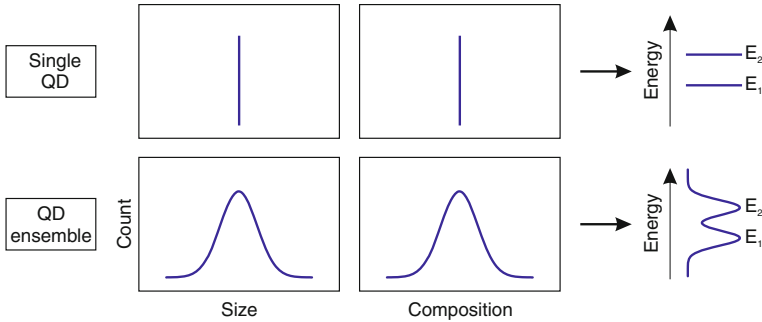
**Fig. 2.7** A quantum dot resembles a three-dimensional potential well. **a** InAs/GaAs QDs can confine electrons and holes due to the discontinuities in the conduction band  $E_C$  and the valence band  $E_V$ . **b** Electron and **c** hole energy levels in InAs/GaAs QDs calculated with 8-band  $\mathbf{k}\cdot\mathbf{p}$ -theory using a continuum mechanical model for the strain (after [47])

the QDs with a pyramidal or truncated-pyramidal shape. Similar to the InAs QDs, it can be seen that an increasing size of the QDs leads to an increasing localization energy for the holes in the GaSb QDs (Fig. 2.8a). If the shape is changed from truncated-pyramidal (with vertical aspect ratio of  $VAR = 0.05$ , defined by height divided by base length) to pyramidal-shaped ( $VAR = 0.45$ ) QDs (Fig. 2.8b), the hole localization energy stays almost constant for the larger heights, while it increases slightly towards the smaller heights. This increase is due to a redistribution of strain from being predominantly hydrostatic for tall QDs to biaxial for flat QDs. If the QDs get flatter, the  $z$ -quantization dominates and the localization energy decreases again. The  $z$ -quantization effect is more pronounced for the QDs with a smaller GaSb content. Here, the overall strain is smaller due to the smaller lattice mismatch to GaAs.

A given population of quantum dots grown by self-organization can be regarded as a statistical ensemble. As their growth process is governed by thermodynamics, their structural and chemical properties will scatter around a mean value, resulting in a Gaussian distribution of size, shape, and chemical composition. This *ensemble broadening* of the structural properties affects the electronic properties and leads to a



**Fig. 2.8** Hole localization energies  $E_{loc}$  in GaSb/GaAs QDs for two different chemical compositions as a function of (a) base length with constant vertical aspect ratio of  $VAR = 0.2$  and (b) shape of the QDs (the numbers indicate the VAR for each data point). The calculations were performed by Andrei Schliwa at TU-Berlin

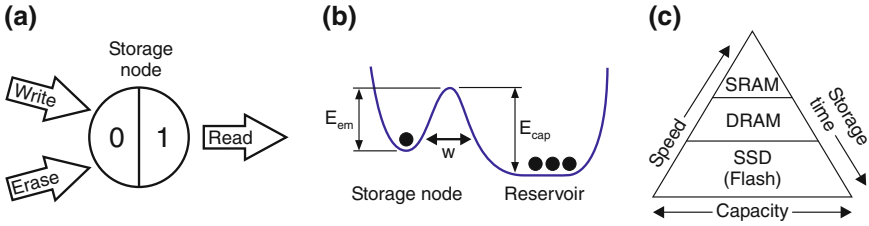


**Fig. 2.9** Inhomogeneous ensemble broadening: Due to the variations in size, shape and composition, the energy levels of the QD ensemble are broadened in comparison to the sharp energy levels of a single QD

similar distribution of the energy levels in the QDs, which is reflected as a Gaussian shaped broadening. The ensemble broadening is illustrated schematically in Fig. 2.9.

## 2.5 Charge-Based Memories

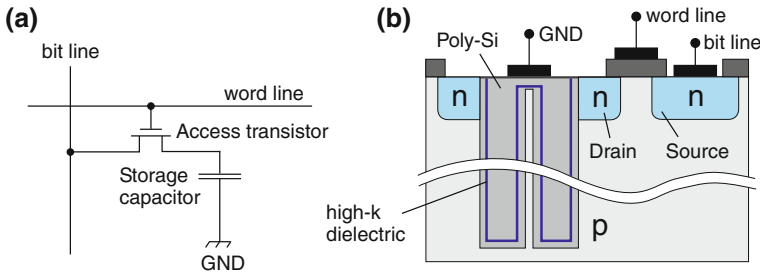
Modern data storage is based on encoding information within the binary system with the two logic states expressed as  $0$  and  $1$  [49, 50]. Digital data storage and data processing systems hence have to distinguish between two states, i.e. a transistor current is switched *on* or *off*. The amount of data stored per storage node is usually one single *bit* (binary digit). A schematic of an abstract binary storage node is depicted



**Fig. 2.10** The constituents of an abstract memory cell. (a) Storage node which stores a single bit with the logic states  $1$  and  $0$ . Write, erase, and read operations can be performed on the storage node. (b) Simple model of a potential minimum used as storage node. The storage node is separated from the electron reservoir by a barrier. Depending on the logic state that is stored, either the emission barrier  $E_{em}$  or the capture barrier  $E_{cap}$  is the limiting factor for the storage time. To write and erase information the electrons have to overcome or tunnel through the barrier. (c) Semiconductor memory hierarchy pyramid

in Fig. 2.10a. The basic operations that have to be performed on a storage node are: writing of information (expressed as writing of a logic  $1$  and erasing the logic  $1$  in order to reach the logic state  $0$ ) and the read-out of the information. A realization of such a storage node is shown in Fig. 2.10b. The storage node is defined by a potential minimum, which is separated from an electron reservoir by a potential barrier. To store the logic  $1$  the potential minimum is occupied by one electron, whereas a logic  $0$  is represented by an unoccupied storage node. The storage of information always involves a non-equilibrium situation. The key properties which limit the performance (storage, write, and erase time) of such a storage node are determined by the energetic height and shape of the barrier which separates the storage node from the reservoir. If an electron is stored inside the storage node (logic  $1$ ), the electron will leave the potential minimum and overcome the barrier with a certain probability. The probability for the electron to tunnel through the barrier is determined by the height  $E_{em}$ , the shape, and width  $w$  of the emission barrier. With thermal energy present, the electron can also thermally overcome the barrier  $E_{em}$ , a process which is determined by the Boltzman statistics and hence proportional to  $exp(-E_{em}/(kT))$ . Hence, the storage time for the  $1$  is limited by the barrier height  $E_{em}$ , the width  $w$ , the shape, and the temperature  $T$ . If the potential minimum is unoccupied (logic  $0$ ), electrons from the reservoir will leak into the storage node with a certain probability. The mechanisms are the same as for the logic  $1$ . The storage time of the logic  $0$  is hence limited by the height, width, and shape of the capture barrier  $E_{cap}$ .

In order to reach a long storage time in the storage node, the barriers need to be designed in such a way that during storage only a very small probability exists for the carriers to transfer between the storage node and the reservoir. In contrast, when information is written or erased from the storage node, the transfer probability should be very high in order to reach fast access times. This dilemma between long storage times and fast access times is a commonly known *trade-off* in memories.



**Fig. 2.11** Schematic cell structure of the storage node in a DRAM. **a** Equivalent circuit of a 1T-1C cell design: the DRAM consists of an access transistor and a storage capacitor connected in an array of *word lines* and *bit lines*. **b** Cross section of an actual DRAM with a deep trench design (after [49])

### 2.5.1 Semiconductor Memories

In order to work around the limitations given by the trade-off between long storage times and fast access times, a hierarchy of different memories has been developed [51], with each individual memory type bridging the gap of a specific combination of speed, capacity, and storage time (see Fig. 2.10c). In the following, only semiconductor memories are considered. On the top of the hierarchy pyramid, at the very fast CPU level, the SRAM (Static Random Access Memory) is used as processor register. It has a total of six transistors for each cell which constitute a flip-flop type behavior. On the bottom of the hierarchy pyramid, the Flash memory in the form of solid-state disks (SSD) is used as data archive, having a slow write time, but a long storage time ( $> 10$  years, non-volatile). In between these two, the DRAM (Dynamic Random Access Memory) is used as a working memory in computers, where the data needed in the upcoming computation cycles of the CPU is stored and cached before it is transferred to the registers.

The semiconductor memory industry focuses mainly on two of these memories: the DRAM and the Flash. Hence, these two play the most important role in the semiconductor memory market. A brief overview of these two memory concepts is given in the following.

#### 2.5.1.1 Dynamic Random Access Memory

The Dynamic Random Access Memory (DRAM) consists of a single access transistor and single storage capacitor per cell, the so-called 1T-1C design [49]. A schematic equivalent circuit of a single cell is shown in Fig. 2.11a. The transistor is needed to access the individual storage node. The capacitor represents the storage node using the electrons on the capacitor plates to encode a single bit per cell. The 1T-1C cells are connected to the control circuits and sense amplifiers by an array of word lines and bit lines. The array matrix allows to address each cell individually.

To initiate the reading or writing operation, the access transistor is switched on by activating the word line. This connects the capacitor with the bit line. In the reading operation the voltage is then sensed via the bit line, whereas in the writing operation electrons are transferred to the capacitor plates via the bit line. To securely detect the logic state stored in the capacitor, a minimum capacitance is required. To reach high bit area densities, today’s DRAMs hence consist of deep trenches buried in a substrate material. The trenches are filled with a high-k dielectric (see Fig. 2.11b).

The main advantages of the DRAM are:

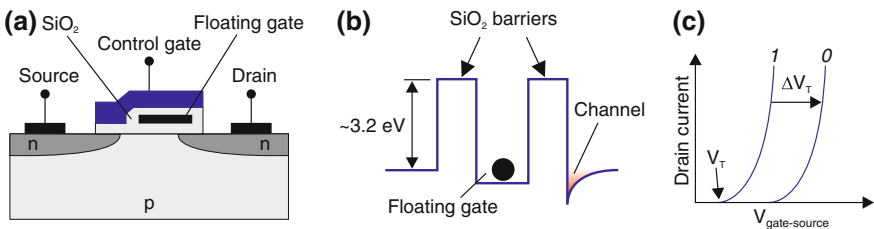
- Fast access time: typically  $\sim 10$  ns,
- Good endurance: typically  $>10^{15}$  write/read cycles.

The disadvantages of the DRAM are:

- Short storage time: typically some 10 ms, requiring cyclical refresh of the information
- Destructive read-out: the read-out destroys the information and a refresh is needed
- Large number of charges per bit: each bit needs  $\sim 10^4$  electrons, which have to be transferred when writing, reading, or refreshing the information. This leads to a large stand-by energy consumption.
- Limited scalability: due to the minimum capacitance required, the scalability is limited. However, a quick end to scaling is not yet in sight [52].

### 2.5.1.2 Flash

The most important non-volatile semiconductor memory is the Flash EEPROM (Electrically erasable read-only memory) [51, 53, 54], which in its modern design was invented in 1984 [55]. It can store information without any power supply. The design consists of a field-effect transistor with an additional gate between the control gate and the channel (see Fig. 2.12a). The additional gate is electrically isolated from its surroundings by a dielectric (e.g. by the native oxide of Silicon  $\text{SiO}_2$ ), and is



**Fig. 2.12** Flash EEPROM. **a** Schematic cross section of the Flash structure. A floating gate surrounded by an insulator is embedded in between the control gate and the channel. **b** Schematic conduction band in a Flash memory. Inside the floating gate a potential well is created due to the band discontinuity of  $\sim 3.2$  eV between Si and the  $\text{SiO}_2$  barriers. **c** Shift of the current-voltage characteristic due to charges inside the floating gate

called *floating gate*. Electrons are trapped inside the potential well, which is formed by the floating gate and the surrounding barriers (Fig. 2.12b). The floating gate is capacitively coupled to the channel underneath. The electrons inside the floating gate affect the electron density in the channel by the field-effect. The change of the electron density in the channel can be directly measured in the drain current, which is illustrated in Fig. 2.12c. If the floating gate is empty (which corresponds to the logic state  $I$ ), the transistor is turned on at the threshold voltage  $V_T$ . In contrast, if electrons are present in the floating gate (logic state  $O$ ), the 2DEG channel underneath is depleted and the threshold voltage is shifted by  $\Delta V_T$ .

The main advantages of the Flash are:

- As a result of the high barriers, the Flash is non-volatile. It can store information for more than 10 years, even without power supply.
- The number of electrons required to store one bit is  $\sim 1000$ , and hence smaller than in the DRAM.
- The read-out of the information is non-destructive.

The disadvantages of the Flash are:

- During writing and erasing of information, the electrons have to overcome or tunnel through the barriers with the fixed height of  $\sim 3.2$  eV. Compared to the thermal energy of the charges in the order of  $\sim kT$ , the barriers are very high, which leads to very slow write and erase times of microseconds to milliseconds.
- The writing and erasing can be done either by hot electron injection or Fowler-Nordheim tunneling [51]. Due to the low injection efficiency and the high electric fields of 8–10 MV/cm [54] which are necessary in these processes, the oxide barriers are successively destroyed by the generation of defects within the barriers. This leads to a low endurance of just  $\sim 10^6$  write/erase cycles.

Two different designs of the Flash exist [51]. Non-volatility and very high densities can be reached in the NAND Flash, and it is used at the basis of the memory hierarchy as data archive, mainly in mobile applications, such as smartphones, USB sticks, and mp3 players. In contrast to the NAND Flash, the NOR Flash has the advantage, that each cell can be randomly accessed individually. Additionally, it works with a very high reliability. It is used mainly in industrial and embedded equipment, where program code (i.e. Firmware or BIOS) and initial data has to be stored with high reliability.

## 2.6 Quantum Dot-Based Flash Memory (QD-Flash)

The DRAM and the Flash memory as the most dominant memory types in the semiconductor memory market both have their advantages and disadvantages. In order to combine the advantages of both memories, the quantum dot-based Flash—the *QD-Flash*—was developed at TU-Berlin [56, 57]. There are three pillars on which the QD-Flash concept is built: the use of III-V compound semiconductors, the use of

nanostructures (i.e. quantum dots) as storage units, and the use of a selectively-doped heterostructure to operate the memory.

The use of III-V compounds instead of the Si/SiO<sub>2</sub> material system with its fixed band offset allows a band structure design, in which the band discontinuities can be specifically tailored. A wide variety of III-V compounds exist (see Fig. 2.1) and can be combined with almost atomically abrupt interfaces which are free of defects. In addition, the III-V compounds can be doped to alter the band structure, which is not possible in SiO<sub>2</sub>.

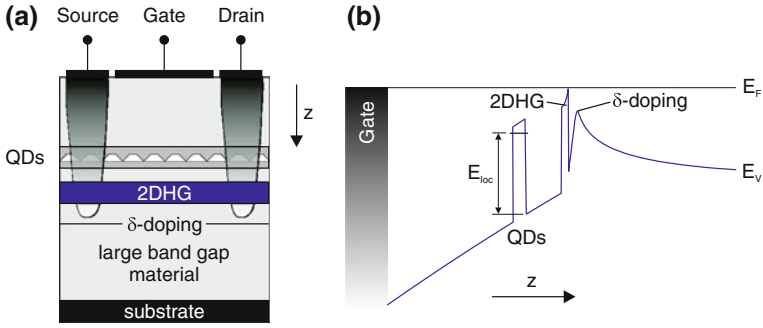
The use of nanostructures, and in particular of self-organized quantum dots, in the QD-Flash concept has many advantages. During the growth of quantum dots, the total energy is minimized (see Sect. 2.4.1) which leads to strain minimization and to dislocation-free interfaces. This results in very high localization energies in the QDs, even if the system is highly strained. It allows the combination of materials which have a large lattice mismatch. Furthermore, the large capture cross section of QDs (typically  $10^{-14}$ – $10^{-12}$  cm<sup>2</sup>), which is orders of magnitude larger than for common traps in semiconductors [58], leads to extremely fast capture times for electrons and holes with time scales in the range of picoseconds [59, 60]. This could facilitate very fast write times in the QD-Flash. Write times in QDs have already been demonstrated to be in the range of the DRAM access time [61], yet limited by the parasitics of the device. QDs also allow the use of holes as information carriers. This allows longer storage times due to the larger effective mass of holes as compared to electrons, and the resulting lower tunneling probability.

Embedding the QDs into a modulation-doped field-effect transistor (MODFET) is necessary to be able to perform the memory operations, i.e. storage, writing, erasing, and reading out. The QDs are hence embedded above the two-dimensional carrier gas of the MODFET. The charges inside the QDs lower the mobility of the 2D gas [62, 14] and reduce its charge carrier density by the field effect [63–65] (see Chap. 4). Hence the logic state stored in the QDs can be detected by measuring the source/drain current of the MODFET.

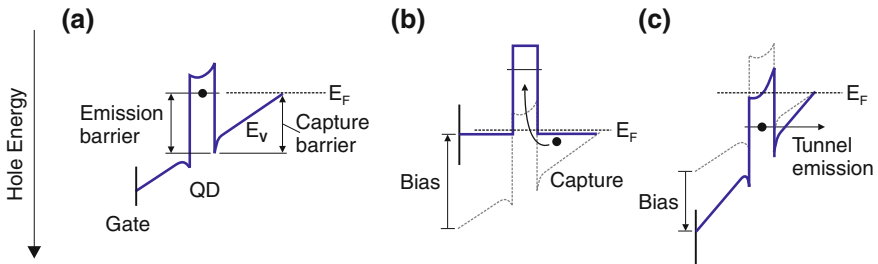
The generic structure of the QD-Flash [27, 28, 66] with hole storage is shown schematically in Fig. 2.13a. A layer of self-organized QDs is embedded into a p-MODFET. The QDs are surrounded by a material with a larger band gap. The 2DHG is formed inside a material with a smaller band gap than the matrix material. The MODFET is nominally undoped, just a thin  $\delta$ -doping (or a thin doping layer) is providing holes for the 2DHG. The 2DHG is contacted via Ohmic source and drain contacts, and a Schottky gate contact allows to manipulate the charge state in the QDs and to change the conductance of the 2DHG channel. Figure 2.13b shows a generic valence band of a QD-Flash with hole storage. The 2DHG pins the Fermi level  $E_F$  throughout the device. By applying a gate bias, the charge state of the QDs and the 2DHG can be controlled.

In the following, the basic memory operations of the QD-Flash are described and are illustrated schematically in Fig. 2.14:

- *Storage.* To store a logic 1 (here defined as holes inside the QDs), an emission barrier is needed to prevent the holes from leaving the QDs by thermal excitation



**Fig. 2.13** Basic structure of the QD-Flash realizing hole storage. **a** Schematic structure of the QD-Flash. The QD layer is embedded into a MODFET structure. The QDs have a smaller band gap than the surrounding material. The charge state in the QDs can be detected via a measurement of the conductance of the 2DHG. **b** Simulated valence band of the QD-Flash



**Fig. 2.14** Schematic depiction of the memory operations in a hole-based QD-Flash. **a** Storage of information. A logic 0 (holes in the QD) requires an emission barrier, while the logic 1 (empty QDs) requires a capture barrier. **b** Writing of information by applying a forward bias. The capture barrier is completely eliminated and holes are captured into the QD. **c** Erasing of information by applying a reverse bias to the gate. This increases the electric field at the position of the QDs leading to a narrower tunnel barrier. The tunneling probability is increased and the holes tunnel out of the QD

(see Fig. 2.10). This emission barrier is formed by the band discontinuity between the QDs and the matrix material. The potential well with the depth of the localization energy  $E_{loc}$  confines the holes inside the QDs and defines the emission probability and the storage time of the logic 1. To store a logic 0 (empty QDs), a capture barrier is needed to prevent holes from outside the QDs to enter the QDs. This capture barrier is formed by the band bending induced by the Schottky diode of the MODFET structure. Hence, the storage time of the 0 is also defined by the position of the QD layer within the device and the magnitude of the band bending. The storage of charges inside the QDs for various material systems has been investigated in [67–69]. The maximum storage time of 1.6 s at room temperature was achieved in InAs/GaAs QDs with an additional  $Al_{0.9}Ga_{0.1}As$  barrier [69].

- **Writing.** In order to charge the QDs with holes (here defined as writing), a gate bias in forward direction of the Schottky diode is applied. Ideally, the device is

designed in such a way, that the forward bias completely eliminates the capture barrier, and holes get captured into the QDs and thermally relax down to the lowest unoccupied state. As mentioned above, the capture time of charge carriers into QDs is extremely fast, on the order of some ps [59, 60]. It was demonstrated, that write times in QDs comparable to the DRAM access time can be achieved [61].

- *Erasing.* To discharge the QDs (here defined as erasing), a gate bias in reverse direction of the Schottky diode is applied. This increases the electric field around the QDs leading to a narrower emission barrier, which leads to an increased tunneling probability of the holes inside the QDs. Eventually, they will tunnel through the barrier, discharging the QDs. The erasing in QDs has been investigated in [70].
- *Read-out.* To read-out the logic state stored in the QDs, a measurement of the conductance of the 2DHG underneath the QDs is performed. Holes inside the QDs decrease the mobility and the charge density in the 2DHG by scattering and the field effect (see Chap. 4). The effect can be directly measured in the source/drain current and can hence be used to detect the logic state in the QDs.

## 2.7 Summary

This chapter is summarized in the following:

- Semiconductor heterostructures and nanostructures allow the controlled confinement of charge carriers within all three spatial dimensions. Within these small dimensions, the electronic and optical properties can be specifically tailored by the size and the geometry of the nanostructures.
- Based on heterostructures, devices with very high mobility can be build. The MODFET is a prominent example for a two-dimensional confinement of charge carriers, which have been separated from their ionized dopant atoms.
- Quantum dots confine charge carriers within all three spatial dimensions. Hence, they are ideally suited to build a memory. Especially self-assembled QDs based on III-V compounds offer many advantages, such as their high-quality growth by self-organization, large confining potentials (large localization energy), and a wide variety of different materials that can be combined.
- The most important semiconductor memories are the DRAM and the Flash. Both have their advantages and disadvantages: the DRAM is fast but volatile, while the Flash is non-volatile, but slow. A way to combine the advantages of both memories is using self-assembled QDs embedded into a MODFET structure. In principle, fast write times and long storage times can be reached.

## References

1. J.R. Arthur, Molecular beam epitaxy. *Surf. Sci.* **500**, 189–217 (2002)
2. G.B. Stringfellow, *Organometallic Vapor-Phase Epitaxy: Theory and Practice*, 2nd edn. (Academic Press, San Diego, 1999)

3. T. Mimura, S. Hiyamizu, T. Fujii, K. Nanbu, A new field-effect transistor with selectively doped GaAs/n-Al<sub>x</sub>Ga<sub>1-x</sub>As heterojunctions. *Jpn. J. Appl. Phys.* **19**, L225–L227 (1980)
4. D. Delagebeaudeuf, P. Delescluse, P. Etienne, M. Laviro, J. Chaplart, N.T. Linh, Two-dimensional electron gas MESFET structure. *Electron. Lett.* **16**(17), 667–668 (1980)
5. I. Hayashi, M.B. Panish, P.W. Foy, S. Sumski, Junction lasers which operate continuously at room temperature. *Appl. Phys. Lett.* **17**, 109 (1970)
6. D. Schicketanz, G. Zeidler, GaAs-Double-Heterostructure Lasers as Optical Amplifiers. *IEEE J. Q. Electron.* **11**(2), 65–69 (1975)
7. Z.I. Alferov, R.F. Kazarinov, Semiconductor laser with electric pumping, Inventor's certificate No. 181737 (1963) (in Russian)
8. H. Kroemer, A proposed class of hetero-junction injection lasers. *Proc. IEEE* **51**(12), 1782 (1963)
9. I. Vurgaftman, J.R. Meyer, L.R. Ram-Mohan, Band parameters for III-V compound semiconductors and their alloys. *J. Appl. Phys.* **89**(11), 5815–5875 (2001)
10. H. Kroemer, Barrier control and measurements: abrupt semiconductor heterojunctions. *J. Vac. Sci. Technol. B* **2**(3), 433–439 (1984)
11. S.M. Sze, K.K. Ng, *Physics of Semiconductor Devices*, 3rd edn. (Wiley, New York, 2006)
12. J.H. Davies, *The Physics of Low-Dimensional Semiconductors* (Cambridge University Press, Cambridge, 1998)
13. E. Conwell, V.F. Weisskopf, Theory of impurity scattering in semiconductors. *Phys. Rev.* **77**(3), 388–390 (1950)
14. R. Dingle, H.L. Störmer, A.C. Gossard, W. Wiegmann, Electron mobilities in modulation-doped semiconductor heterojunction superlattices. *Appl. Phys. Lett.* **33**(7), 665 (1978)
15. C.T. Foxon, Three decades of molecular beam epitaxy. *J. Cryst. Growth* **251**, 1–8 (2003)
16. K.v. Klitzing, G. Dorda, M. Pepper, New method for high-accuracy determination of the fine-structure constant based on quantized hall resistance. *Phys. Rev. Lett.* **45**(6), 494–497 (1980)
17. K.v. Klitzing, The quantized Hall effect, *Rev. Mod. Phys.* **58**(3), 519–531 (1986)
18. M.J. Manfra, L.N. Pfeiffer, K.W. West, R.D. Picciotto, K.W. Baldwin, High mobility two-dimensional hole system in GaAs/AlGaAs quantum wells grown on (100) GaAs substrates. *Appl. Phys. Lett.* **86**, 162106 (2005)
19. D. Bimberg, M. Grundmann, N.N. Ledentsov, *Quantum Dot Heterostructures*. (Wiley, Chichester, 1998)
20. D. Bimberg (ed.), *Semiconductor Nanostructures* (Springer, Berlin, 2008)
21. S. Tiwari, F. Rana, K. Chan, L. Shi, H. Hanafi, Single charge and confinement effects in nano-crystal memories. *Appl. Phys. Lett.* **69**(9), 1232 (1996)
22. K. Koike, K. Saitoh, S. Li, S. Sasa, M. Inoue, M. Yano, Room-temperature operation of a memory-effect AlGaAs/GaAs heterojunction field-effect transistor with self-assembled InAs nanodots. *Appl. Phys. Lett.* **76**(11), 1464 (2000)
23. H. Kim, T. Noda, T. Kawazu, 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)
24. G. Yusa, 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**(3), 345 (1997)
25. C. Balocco, A.M. Song, M. Missous, Room-temperature operations of memory devices based on self-assembled InAs quantum dot structures. *Appl. Phys. Lett.* **85**(24), 5911–5913 (2004)
26. D. Nataraj, N. Ooike, J. Motohisa, T. Fukui, Fabrication of one-dimensional GaAs channel-coupled InAs quantum dot memory device by selective-area metal-organic vapor phase epitaxy. *Appl. Phys. Lett.* **87**, 193103 (2005)
27. A. Marent, T. Nowozin, J. Gelze, F. Luckert, D. Bimberg, Hole-based memory operation in an InAs/GaAs quantum dot heterostructure. *Appl. Phys. Lett.* **95**, 242114 (2009)
28. A. Marent, T. Nowozin, M. Geller, D. Bimberg, The QD-Flash: a quantum dot-based memory device. *Semicond. Sci. Technol.* **26**, 014026 (2011)

29. Y. Arakawa, H. Sakaki, Multidimensional quantum well laser and temperature dependence of its threshold current. *Appl. Phys. Lett.* **40**, 939 (1982)
30. 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. Gösele, J. Heydenreich, Low threshold, large  $T_0$  injection laser emission from (InGa)As quantum dots. *Electron. Lett.* **30**(17), 1416 (1994)
31. F. Heinrichsdorff, C. Ribbat, M. Grundmann, D. Bimberg, High-power quantum-dot lasers at 1100 nm. *Appl. Phys. Lett.* **76**(5), 556–558 (2000)
32. X. Huang, A. Stintz, H. Li, L.F. Lester, J. Cheng, K.J. Malloy, Passive mode-locking in 1.3  $\mu\text{m}$  two-section InAs quantum dot lasers. *Appl. Phys. Lett.* **78**, 2825 (2001)
33. 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, A.R. Kovsh, 35 GHz mode-locking of 1.3  $\mu\text{m}$  quantum dot lasers. *Appl. Phys. Lett.* **85**, 843 (2004)
34. 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, P. Werner, Single-mode submonolayer quantum-dot vertical-cavity surface-emitting lasers with high modulation bandwidth. *Appl. Phys. Lett.* **89**, 141106 (2006)
35. M. Lämmlin, G. Fiol, C. Meuer, M. Kuntz, F. Hopfer, A. R. Kovsh, N.N. Ledentsov, 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)
36. A. Lochmann, E. Stock, O. Schulz, F. Hopfer, D. Bimberg, V. Haisler, A. Toropov, A. Bakarov, A. Kalagin, Electrically driven single quantum dot polarised single photon emitter. *Electron. Lett.* **42**(13), 774–775 (2006)
37. E. Stock, T. Warming, I. Ostapenko, S. Rodt, A. Schliwa, J.A. Töfflinger, A. Lochmann, A.I. Toropov, S.A. Moshchenko, D.V. Dmitriev, V.A. Haisler, D. Bimberg, Single-photon emission from InGaAs quantum dots grown on (111) GaAs. *Appl. Phys. Lett.* **96**, 093112 (2010)
38. W. Unrau, D. Quandt, J.-H. Schulze, T. Heindel, T.D. Germann, O. Hitzemann, A. Strittmatter, S. Reitzenstein, U.W. Pohl, D. Bimberg, Electrically driven single photon source based on a site-controlled quantum dot with self-aligned current injection. *Appl. Phys. Lett.* **101**(21), 211119 (2012)
39. F.C. Frank, J.H.V.D. Merwe, *Proc. Roy. Soc. Lond. A* **98**, 205 (1949)
40. M. Volmer, A. Weber, *Zeitschr. f. phys. Chem.* **119**, 277 (1926)
41. V.A. Shchukin, D. Bimberg, Spontaneous ordering of nanostructures on crystal surfaces. *Rev. Mod. Phys.* **71**, 1125 (1999)
42. I.N. Stranski, L. Krastanow, Zur Theorie der orientierten Ausscheidung von Ionenkristallen aufeinander, *Sitzungsber. Akad. Wiss. Wien, Math.-Naturwiss. Kl., Abt. 2B* **146**, 797 (1938)
43. K. Jacobi, Atomic structure of InAs quantum dots on GaAs. *Prog. Surf. Sci.* **71**, 185–215 (2003)
44. J. Marquez, L. Geelhaar, K. Jacobi, Atomically resolved structure of InAs quantum dots. *Appl. Phys. Lett.* **78**(16), 2309–2311 (2001)
45. C. Cohen-Tannoudji, B. Diu, F. Laloe, *Quantenmechanik — Band I und II*, 2nd edn. (Gruyter, Berlin, 1999)
46. M. Grundmann, O. Stier, D. Bimberg, InAs/GaAs pyramidal quantum dots: strain distribution, optical phonons, and electronic structure. *Phys. Rev. B* **52**, 11969–11981 (1995)
47. O. Stier, M. Grundmann, D. Bimberg, Electronic and optical properties of strained quantum dots modeled by 8-band-k.p theory. *Phys. Rev. B* **59**, 5688 (1999)
48. A. Schliwa, M. Winkelnkemper, D. Bimberg, Impact of size, shape and composition on piezoelectric effects and the electronic properties of InGaAs/GaAs quantum dots. *Phys. Rev. B* **76**, 205324 (2007)
49. R. Waser, *Nanoelectronics and Information Technology* (Wiley-VCH, Berlin, 2003)
50. R. Waser, *Nanotechnology Volume 3: Information Technology I, Nanotechnology* (Wiley-VCH, Weinheim, 2008)
51. J.E. Brewer, M. Gill, *Nonvolatile Memory Technologies with Emphasis on Flash* (Wiley, Hoboken, 2008)

52. International Technology Roadmap for Semiconductors (ITRS), Executive Summary, 2011 edn. (2011), <http://www.itrs.net>
53. P. Pavan, R. Bez, P. Olivo, E. Zanoni, Flash memory cells—an overview. *Proc. IEEE* **85**(8), 1248–1271 (1997)
54. R. Bez, E. Camerlenghi, A. Modeli, A. Visconti, Introduction to flash memory. *Proc. IEEE* **91**(4), 489 (2003)
55. F. Masuoka, M. Asano, H. Iwahashi, T. Komuro, S. Tanaka, A new flash EEPROM cell using triple polysilicon technology. *IEEE IEDM Technical Digest*, pp. 464–467 (1984)
56. M. Geller, A. Marent, D. Bimberg, Speicherzelle und Verfahren zum Speichern von Daten (Memory cell and method for storing data). International patent EP/2097904, 2006
57. A. Marent, M. Geller, T. Nowozin, D. Bimberg, Speicherzelle auf Basis von Nanostrukturen aus Verbindungshalbleitern, International patent application PCT/12/970,744, 2010
58. Landolt-Bornstein—Group III Condensed Matter, in *SpringerMaterials—the landolt-bornstein database*, vol 41A2b. Accessed April 2012. doi:10.1007/b83098
59. T. Müller, F.F. Schrey, G. Strasser, K. Unterrainer, Ultrafast intraband spectroscopy of electron capture and relaxation in InAs/GaAs quantum dots. *Appl. Phys. Lett.* **83**(17), 3572–3574 (2003)
60. M. Geller, A. Marent, E. Stock, D. Bimberg, V.I. Zubkov, I.S. Shulgunova, A.V. Solomonov, Hole capture into self-organized InGaAs quantum dots. *Appl. Phys. Lett.* **89**(23), 232105 (2006)
61. M. Geller, A. Marent, T. Nowozin, D. Bimberg, N. Akçay, N. Öncan, A write time of 6 ns for quantum dot-based memory structures. *Appl. Phys. Lett.* **92**(9), 092108 (2008)
62. M. Russ, C. Meier, B. Marquardt, A. Lorke, D. Reuter, A.D. Wieck, Quantum dot electrons as controllable scattering centers in the vicinity of a two-dimensional electron gas. *Phase Transitions* **79**(9–10), 765–770 (2006)
63. B. Marquardt, M. Geller, A. Lorke, D. Reuter, A.D. Wieck, Using a two-dimensional electron gas to study nonequilibrium tunneling dynamics and charge storage in self-assembled quantum dots. *Appl. Phys. Lett.* **95**, 022113 (2009)
64. M. Russ, C. Meier, A. Lorke, D. Reuter, A.D. Wieck, Role of quantum capacitance in coupled low-dimensional electron systems. *Phys. Rev. B* **73**, 115334 (2006)
65. A. Rack, R. Wetzler, A. Wacker, E. Schöll, Dynamical bistability in quantum-dot structures: role of auger processes. *Phys. Rev. B* **66**(16), 165429 (2002)
66. A. Marent, Entwicklung einer neuartigen Quantenpunkt-Speicherzelle. Technische Universität Berlin, Dissertation, 2010
67. M. Geller, C. Kapteyn, L. Müller-Kirsch, R. Heitz, D. Bimberg, 450 meV hole localization energy in GaSb/GaAs quantum dots. *Appl. Phys. Lett.* **82**(16), 2706–2708 (2003)
68. A. Marent, M. Geller, D. Bimberg, A.P. Vasi'ev, E.S. Semenova, A.E. Zhukov, V.M. Ustinov, Carrier storage time of milliseconds at room temperature in self-organized quantum dots. *Appl. Phys. Lett.* **89**(7), 072103 (2006)
69. A. Marent, M. Geller, A. Schliwa, D. Feise, K. Pötschke, D. Bimberg, N. Akçay, N. Öncan, 10<sup>[sup 6]</sup> years extrapolated hole storage time in GaSb/AlAs quantum dots. *Appl. Phys. Lett.* **91**(24), 242109 (2007)
70. T. Nowozin, A. Marent, M. Geller, D. Bimberg, N. Akçay, N. Öncan, Temperature and electric field dependence of the carrier emission processes in a quantum dot-based memory structure. *Appl. Phys. Lett.* **94**, 042108 (2009)



<http://www.springer.com/978-3-319-01969-7>

Self-Organized Quantum Dots for Memories  
Electronic Properties and Carrier Dynamics  
Nowozin, T.

2014, XVI, 153 p. 91 illus., 9 illus. in color., Hardcover  
ISBN: 978-3-319-01969-7