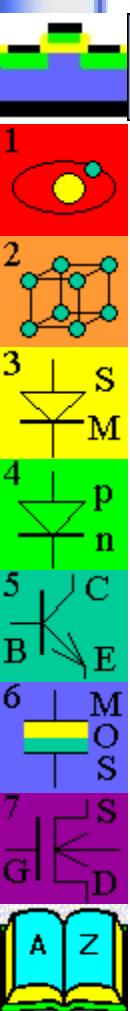


Principles of Semiconductor Devices



- [Help](#)
- [Table of Contents](#)
- [Slideshow](#)
- [Movie](#)



Principles of Semiconductor Devices



- [Help](#)
- [Table of Contents](#)
- [Slideshow](#)
- [Movie](#)

Appendix:



Appendix 1: List of Symbols

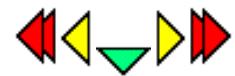
Symbol	Description	MKS Units
A	Area	m^2
c	Speed of light in vacuum	m/s
C	Capacitance per unit area	F/m^2
C_{FB}	Flatband capacitance per unit area of a MOS structure	F/m^2
C_j	Junction capacitance per unit area	F/m^2
C_{ox}	Oxide capacitance per unit area	F/m^2
D_n	Electron diffusion constant	m^2/s
D_p	Hole diffusion constant	m^2/s
E	Energy	Joule
E	Electric field	V/m
E_a	Acceptor energy	Joule
E_c	Conduction band energy of a semiconductor	Joule
E_d	Donor energy	Joule
E_F	Fermi energy (thermal equilibrium)	Joule
E_g	Energy bandgap of a semiconductor	Joule
E_i	Intrinsic Fermi energy	Joule
E_v	Valence band energy of a semiconductor	Joule
E_{vacuum}	Electron energy in vacuum	Joule
$f(E)$	Distribution function (probability density function)	
F_n	Quasi-Fermi energy of electrons	Joule
F_p	Quasi-Fermi energy of holes	Joule
$g_c(E)$	Density of states in the conduction band per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
$g_v(E)$	Density of states in the valence band per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
G_n	Electron generation rate	$\text{m}^{-3}\text{s}^{-1}$
G_p	Hole generation rate	$\text{m}^{-3}\text{s}^{-1}$
h	Plank's constant	Js
\hbar	Reduced Plank's ($= h / 2\pi$)	Js
I	Current	A
J	Current density	A/m^2
J_n	Electron current density	A/m^2
J_p	Hole current density	A/m^2
k	Boltzmann's constant	J/K
l	Mean free path	m

<i>L</i>	Length	m
<i>L_n</i>	Electron diffusion length	m
<i>L_p</i>	Hole diffusion length	m
<i>m</i>	Mass	kg
<i>m₀</i>	Free electron mass	kg
<i>m_e[*]</i>	Effective mass of electrons	kg
<i>m_h[*]</i>	Effective mass of holes	kg
<i>n</i>	Electron density	m ⁻³
<i>n_i</i>	Intrinsic carrier density	m ⁻³
<i>n(E)</i>	Electron density per unit energy and per unit volume	m ⁻³
<i>n₀</i>	Electron density in thermal equilibrium	m ⁻³
<i>n_i</i>	Intrinsic carrier density	m ⁻³
<i>N</i>	Doping density	
<i>N_a</i>	Acceptor doping density	m ⁻³
<i>N_a⁻</i>	Ionized acceptor density	m ⁻³
<i>N_B</i>	Base doping density	m ⁻³
<i>N_c</i>	Effective density of states in the conduction band	m ⁻³
<i>N_C</i>	Collector doping density	m ⁻³
<i>N_d</i>	Donor doping density	m ⁻³
<i>N_d⁺</i>	Ionized donor density	m ⁻³
<i>N_E</i>	Emitter doping density	m ⁻³
<i>N_v</i>	Effective density of states in the valence band	m ⁻³
<i>p</i>	Hole density	m ⁻³
<i>p(E)</i>	Hole density per unit energy	m ⁻³
<i>p₀</i>	Hole density in thermal equilibrium	m ⁻³
<i>p_n</i>	Hole density in an n-type semiconductor	m ⁻³
<i>q</i>	electronic charge	C
<i>Q</i>	Charge	C
<i>Q_d</i>	Charge density per unit area in the depletion layer of an MOS structure	C/m ²
<i>Q_{d,T}</i>	Charge density per unit area at threshold in the depletion layer of an MOS structure	C/m ²
<i>Q_i</i>	Interface charge density per unit area	C/m ²
<i>R</i>	Resistance	Ohm
<i>R_n</i>	Electron recombination rate	m ⁻³ s ⁻¹
<i>R_p</i>	Hole recombination rate	m ⁻³ s ⁻¹
<i>t</i>	Thickness	m
<i>t_{ox}</i>	Oxide thickness	m
<i>T</i>	Temperature	Kelvin
<i>U_n</i>	Net recombination rate of electrons	m ⁻³ s ⁻¹
<i>U_p</i>	Net recombination rate of holes	m ⁻³ s ⁻¹

v	Velocity	m/s
v_{th}	Thermal velocity	m/s
V_a	Applied voltage	V
V_B	Base voltage	V
V_C	Collector voltage	V
V_D	Drain voltage	V
V_E	Emitter voltage	V
V_{FB}	Flatband voltage	V
V_G	Gate voltage	V
V_t	Thermal voltage	V
V_T	Threshold voltage of an MOS structure	V
w	Depletion layer width	m
w_B	Base width	m
w_C	Collector width	m
w_E	Emitter width	m
w_n	Width of an n-type region	m
w_p	Width of a p-type region	m
x	Position	m
x_d	Depletion layer width in an MOS structure	m
$x_{d,T}$	Depletion layer width in an MOS structure at threshold	m
x_j	Junction depth	m
x_n	Depletion layer width in an n-type semiconductor	m
x_p	Depletion layer width in a p-type semiconductor	m
α	Transport factor	
β	Current gain	
γ	Body effect parameter	V ^{1/2}
γ_E	Emitter efficiency	
δ_n	Excess electron density	m ⁻³
δ_p	Excess hole density	m ⁻³
$\Delta Q_{n,B}$	Excess electron charge density in the base	C/m ²
ϵ_{ox}	Dielectric constant of the oxide	F/m
ϵ_s	Dielectric constant of the semiconductor	F/m
μ_n	Electron mobility	m ² /V-s
μ_p	Hole mobility	m ² /V-s
ρ	Charge density per unit volume	C/m ³
	Resistivity	Ω m
ρ_{ox}	Charge density per unit volume in the oxide	C/m ³
σ	Conductivity	Ω^{-1} m ⁻¹
τ_n	Electron lifetime	s

τ_p	Hole lifetime	s
ϕ	Potential	V
ϕ_B	Barrier height	V
ϕ_F	Bulk potential	V
ϕ_i	Built-in potential of a p-n diode or Schottky diode	V
ϕ_s	Potential at the semiconductor surface	V
Φ_M	Workfunction of the metal	V
Φ_{MS}	Workfunction difference between the metal and the semiconductor	V
Φ_S	Workfunction of the semiconductor	V
χ	Electron affinity of the semiconductor	V

Chapter 7: MOS Field-Effect-Transistors



7.1. Introduction

The n-type Metal-Oxide-Semiconductor Field-Effect-Transistor (nMOSFET) consists of a source and a drain, two highly conducting n-type semiconductor regions, which are isolated from the p-type substrate by reversed-biased p-n diodes. A metal or poly-crystalline gate covers the region between source and drain. The gate is separated from the semiconductor by the gate oxide. The basic structure of an n-type MOSFET and the corresponding circuit symbol are shown in Figure 7.1.1.

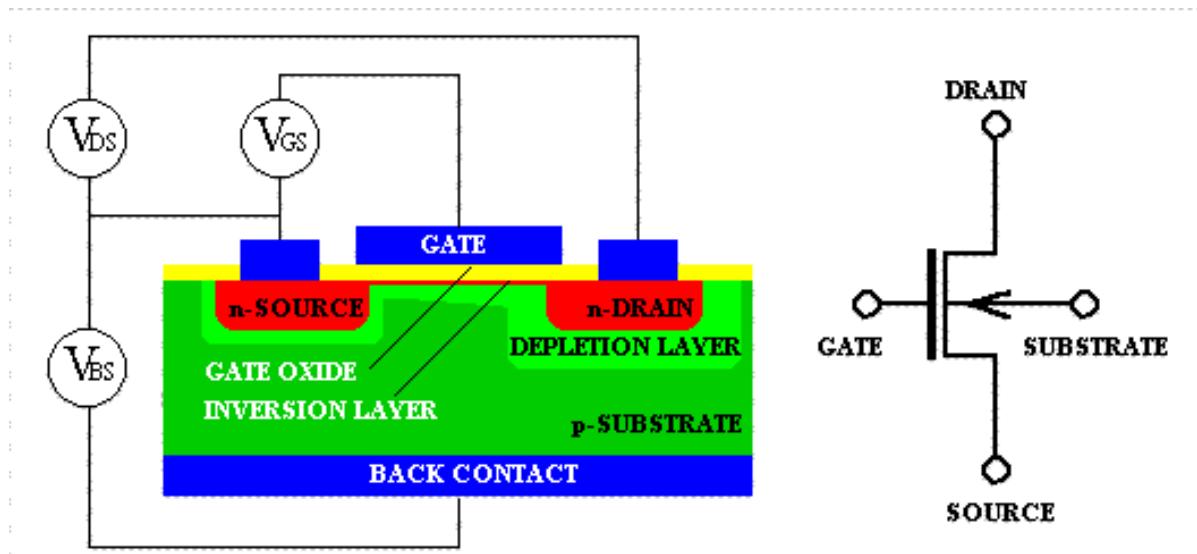


Figure 7.1.1 : Cross-section and circuit symbol of an n-type Metal-Oxide-Semiconductor-Field-Effect-Transistor (MOSFET)

As can be seen on the figure the source and drain regions are identical. It is the applied voltages, which determine which n-type region provides the electrons and becomes the source, while the other n-type region collects the electrons and becomes the drain. The voltages applied to the drain and gate electrode as well as to the substrate by means of a back contact are referred to the source potential, as also indicated Figure 7.1.1.

A conceptually similar structure was proposed and patented independently by Lilienfeld and Heil in 1930, but was not successfully demonstrated until 1960. The main technological problem was the control and reduction of the surface states at the interface between the oxide and the semiconductor.

Initially it was only possible to deplete an existing n-type channel by applying a negative voltage to the gate. Such devices have a conducting channel between source and drain even when no gate voltage is applied and are called "depletion-mode" devices.

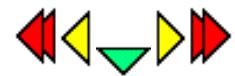
A reduction of the surface states enabled the fabrication of devices, which do not have a conducting channel unless a positive voltage is applied. Such devices are referred to as "enhancement-mode" devices. The electrons at the oxide-semiconductor interface are concentrated in a thin (~10 nm thick) "inversion" layer. By now, most MOSFETs are "enhancement-mode" devices.

While a minimum requirement for amplification of electrical signals is power gain, one finds that a device with both voltage and current gain is a highly desirable circuit element. The MOSFET provides current and voltage gain yielding an output current into an external load which exceeds the input current and an output voltage across that external load which exceeds the input voltage.

The current gain capability of a Field-Effect-Transistor (FET) is easily explained by the fact that no gate current is required to maintain the inversion layer and the resulting current between drain and source. The device has therefore an infinite current gain in DC. The current gain is inversely proportional to the signal frequency, reaching unity current gain at the transit frequency.

The voltage gain of the MOSFET is caused by the current saturation at higher drain-source voltages, so that a small drain-current variation can cause a large drain voltage variation.

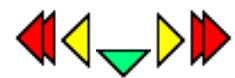
Chapter 6: MOS Capacitors



6.1. Introduction

The primary reason to study the Metal-Oxide-Silicon (MOS) capacitor is to understand the principle of operation as well as the detailed analysis of the Metal-Oxide-Silicon Field Effect Transistor (MOSFET). In this chapter, we introduce the MOS structure and its four different modes of operation, namely accumulation, flatband, depletion and inversion. We then consider the flatband voltage in more detail and present the MOS analysis based on the full depletion approximation. Finally, we analyze and discuss the MOS capacitance.

Chapter 5: Bipolar Junction Transistors



5.1. Introduction

The bipolar junction transistor was the first solid-state amplifier element and started the solid-state electronics revolution. Bardeen, Brattain and Shockley at the Bell Laboratories invented it in 1948 as part of a post-war effort to replace vacuum tubes with solid-state devices. Solid-state rectifiers were already in use at the time and were preferred over vacuum diodes because of their smaller size, lower weight and higher reliability. A solid-state replacement for a vacuum triode was expected to yield similar advantages. The work at Bell Laboratories was highly successful and culminated in Bardeen, Brattain and Shockley receiving the Nobel Prize in 1956.

Their work led them first to the point-contact transistor and then to the bipolar junction transistor. They used germanium as the semiconductor of choice because it was possible to obtain high purity material. The extraordinarily large diffusion length of minority carriers in germanium provided functional structures despite the large dimensions of the early devices.

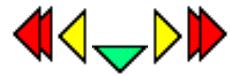
Since then, the technology has progressed rapidly. The development of a planar process yielded the first circuits on a chip and for a decade, bipolar transistor operational amplifiers, like the 741, and digital TTL circuits were the workhorses of any circuit designer.

The spectacular rise of the MOSFET market share during the last decade has completely removed the bipolar transistor from center stage. Almost all logic circuits, microprocessor and memory chips contain exclusively MOSFETs.

Nevertheless, bipolar transistors remain important devices for ultra-high-speed discrete logic circuits such as emitter coupled logic (ECL), power-switching applications and in microwave power amplifiers.

In this chapter we first present the structure of the bipolar transistor and show how a three-layer structure with alternating n-type and p-type regions can provide current and voltage amplification. We then present the ideal transistor model and derive an expression for the current gain in the forward active mode of operation. Next, we discuss the non-ideal effects, the modulation of the base width and recombination in the depletion region of the base-emitter junction.

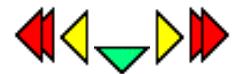
Chapter 4: p-n Junctions



4.1. Introduction

P-n junctions consist of two semiconductor regions of opposite type. Such junctions show a pronounced rectifying behavior. They are also called p-n diodes.

The p-n junction is a versatile element, which can be used as a rectifier, as an isolation structure and as a voltage-dependent capacitor. In addition, they can be used as solar cells, photodiodes, light emitting diodes and even laser diodes. They are also an essential part of Metal-Oxide-Silicon Field-Effects-Transistors (MOSFETs) and Bipolar Junction Transistors (BJTs).

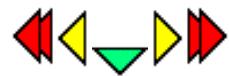


Chapter 3: Metal-Semiconductor Junctions

3.1 Introduction

Metal-to-semiconductor junctions are of great importance since they are present in every semiconductor device. They can behave either as a Schottky  barrier or as an ohmic contact dependent on the characteristics of the interface. We will focus primarily on the Schottky barriers. This chapter contains an analysis of the electrostatics of the M-S junction. Calculated are the charge, field and potential distribution within the device. This chapter also contains a derivation of the current voltage characteristics due to diffusion, thermionic emission and tunneling in Metal-Semiconductor junctions.

Chapter 2: Semiconductor Fundamentals



2.1 Introduction

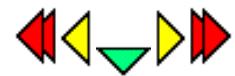
To understand the fundamental concepts of semiconductors, one must apply modern physics to solid materials. More specifically, we are interested in semiconductor crystals. Crystals are solid materials consisting of atoms, which are placed in a highly ordered structure called a lattice. Such a structure yields a periodic potential throughout the material.

Two properties of crystals are of particular interest, since they are needed to calculate the current in a semiconductor. First, we need to know how many fixed and mobile charges are present in the material. Second, we need to understand the transport of the mobile carriers through the semiconductor.

In this chapter we start from the atomic structure of semiconductors and explain the concepts of energy band gaps, energy bands and the density of states in an energy band. We also show how the current in an almost filled band can more easily be analyzed using the concept of holes. Next, we discuss the probability that energy levels within an energy band are occupied. We will use this probability density to find the density of electrons and holes in a band.

Two transport mechanisms will be considered. The drift of carriers in an electric field and the diffusion of carriers due to a carrier density gradient will be discussed. Recombination mechanisms and the continuity equations are then combined into the diffusion equation. Finally, we present the drift-diffusion model, which combines all the essential elements discussed in this chapter.

Chapter 1: Review of Modern Physics



1.1 Introduction

The fundamentals of semiconductors are typically found in textbooks discussing quantum mechanics, electro-magnetics, solid-state physics and statistical thermodynamics. The purpose of this chapter is to review the physical concepts, which are needed to understand the semiconductor fundamentals of semiconductor devices. While an attempt was made to make this section comprehensible even to readers with a minimal background in the different areas of physics, readers are still referred to the bibliography for a more thorough treatment of this material. Readers with sufficient background in modern physics can skip this chapter without loss of continuity.

Introduction



The Semiconductor Industry

Semiconductor devices such as diodes, transistors and integrated circuits can be found everywhere in our daily lives, in Walkman, televisions, automobiles, washing machines and computers. We have come to rely on them and increasingly have come to expect higher performance at lower cost.

Personal computers clearly illustrate this trend. Anyone who wants to replace a three to five year old computer finds that the trade-in value of his or her computer is surprisingly low. On the bright side, one finds that the complexity and performance of the today's personal computers vastly exceeds that of their old computer and that for about the same purchase price, adjusted for inflation.

While this economic reality reflects the massive growth of the industry, it is hard to even imagine a similar growth in any other industry. For instance, in the automobile industry, no one would even expect a five times faster car with a five times larger capacity at the same price when comparing to what was offered five years ago. Nevertheless, when it comes to personal computers, such expectations are very realistic.

The essential fact which has driven the successful growth of the computer industry is that through industrial skill and technological advances one manages to make smaller and smaller transistors. These devices deliver year after year better performance while consuming less power and because of their smaller size they can also be manufactured at a lower cost per device.

Purpose and Goal of the Text

The purpose of this text is to explore the internal behavior of semiconductor devices, so that we can understand the relation between the device geometry and material parameters on one hand and the resulting electrical characteristics on the other hand.

This text provides the link between the physics of semiconductors and the design of electronic circuits. The material covered in this text is therefore required to successfully design CMOS-based integrated circuits.

The Primary Focus: The MOSFET and CMOS Integrated Circuits

The Metal-Oxide-Silicon Field-Effect-Transistor (MOSFET) is the main subject of this text, since it is already the prevailing device in microprocessors and memory circuits. In addition, the MOSFET is increasingly used in areas as diverse as mainframe computers and power electronics. The MOSFET's advantages over other types of devices are its mature fabrication technology, its successful scaling characteristics and the combination of complementary MOSFETs yielding CMOS circuits.

The fabrication process of silicon devices has evolved over the last 25 years into a mature, reproducible and reliable integrated circuit manufacturing technology. While the focus in this text is on individual devices, one must realize that the manufacturability of millions of such devices on a single substrate is a minimum requirement in today's industry. Silicon has evolved as the material of choice for such devices, for a large part because of its stable oxide, silicon dioxide (SiO_2),

which is used as an insulator, as a surface passivation layer and as a superior gate dielectric.

The scaling of MOSFETs started in the seventies. Since then, the initial 10 micron gate length of the devices was gradually reduced by about a factor two every five years, while in 2000 MOSFETs with a 0.18 micron gate length were manufactured on a large scale. This scaling is expected to continue well into the 21st century, as devices with a gate length smaller than 30 nm have already been demonstrated. While the size reduction is a minimum condition when scaling MOSFETs, successful scaling also requires the reduction of all the other dimensions of the device so that the device indeed delivers superior performance. Devices with record gate lengths are typically not fully scaled, so that several years go by until the large-scale production of such device takes place.

The combination of complementary MOSFETs in logic circuits also called CMOS circuits has the unique advantage that carriers only flow through the devices when the logic circuit changes its logic state. Therefore, there is no associated power dissipation if the logic state must not be changed. The use of CMOS circuits immediately reduces the overall power dissipation by a factor ten, since less than one out of ten gates of a large logic circuit switch at any given time.

Bart Van Zeghbroeck, Boulder, March 2001

Chapter 1: Review of Modern Physics



Examples

Example 1.1

A metal has a workfunction of 4.3 V. What is the minimum photon energy in Joule to emit an electron from this metal through the photo-electric effect? What are the photon frequency in Terahertz and the photon wavelength in micrometer? What is the corresponding photon momentum? What is the velocity of a free electron with the same momentum?

Example 1.2

The spectral density of the sun peaks at a wavelength of 900 nm. If the sun behaves as a black body, what is the temperature of the sun?

Example 1.3

An electron is confined to a 1 micron thin layer of silicon. Assuming that the semiconductor can be adequately described by a one-dimensional quantum well with infinite walls, calculate the lowest possible energy within the material in units of electron volt. If the energy is interpreted as the kinetic energy of the electron, what is the corresponding electron velocity? (The effective mass of electrons in silicon is $0.26 m_0$, where $m_0 = 9.11 \times 10^{-31}$ kg is the free electron rest mass).

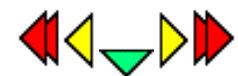
Example 1.4

Consider an infinitely long cylinder with charge density r , dielectric constant ϵ_0 and radius r_0 . What is the electric field in and around the cylinder?

Example 1.5

Calculate the energy relative to the Fermi energy for which the Fermi function equals 5%. Write the answer in units of kT .

Chapter 1: Review of Modern Physics



1.2 Quantum Mechanics

- [1.2.1. Particle-wave duality](#)
- [1.2.2. The photo-electric effect](#)
- [1.2.3. Blackbody radiation](#)
- [1.2.4. The Bohr model](#)
- [1.2.5. Schrödinger's equation](#)
- [1.2.6. Pauli exclusion principle](#)
- [1.2.7. Electronic configuration of the elements](#)

Quantum mechanics emerged in the beginning of the twentieth century as a new discipline because of the need to explain phenomena, which could not be explained using Newtonian mechanics or classical electromagnetic theory. These phenomena include the photoelectric effect, blackbody radiation and the rather complex radiation from an excited hydrogen gas. It is these and other experimental observations which lead to the concepts of quantization of light into photons, the particle-wave duality, the de Broglie wavelength and the fundamental equation describing quantum mechanics, namely the Schrödinger equation. This section also contains a discussion of the energy levels of an infinite one-dimensional quantum well and those of the hydrogen atom.

1.2.1 Particle-wave duality



Quantum mechanics acknowledges the fact that particles exhibit wave properties. For instance, particles can produce interference patterns and can penetrate or "tunnel" through potential barriers. Neither of these effects can be explained using Newtonian mechanics. Photons on the other hand can behave as particles with well-defined energy. These observations blur the classical distinction between waves and particles. Two specific experiments demonstrate the particle-like behavior of light, namely the photoelectric effect and blackbody radiation. Both can only be explained by treating photons as discreet particles with an energy per photon which is proportional to the frequency of the light. The emission spectrum of an excited hydrogen gas demonstrates that electrons confined to an atom can only have discreet energies. Niels Bohr explained the emission spectrum by assuming that the wavelength of an electron wave is inversely proportional to the electron momentum.

The particle and the wave picture are both simplified forms of the wave packet description, a localized wave consisting of a combination of plane waves with different wavelength. As the range of wavelength is compressed to a single value, the wave becomes a plane wave at a single frequency and yields the wave picture. As the range of wavelength is increased, the size of the wave packet is reduced, yielding a localized particle.

1.2.2 The photo-electric effect



The photoelectric effect is by now the "classic" experiment, which demonstrates the quantized nature of light: when applying monochromatic light to a metal in vacuum one finds that electrons are released from the metal. This experiment confirms the notion that electrons are confined to the metal, but can escape when provided sufficient energy, for instance in the form of light. However, the surprising fact is that when illuminating with long wavelengths (typically larger than 400 nm) no electrons are emitted from the metal even if the light intensity is increased. On the other hand, one easily observes electron emission at ultra-violet wavelengths for which the number of electrons emitted does vary with the light intensity. A more detailed analysis reveals that the maximum kinetic energy of the emitted electrons varies linearly with the inverse of the wavelength, for wavelengths shorter than the maximum wavelength.

The experiment is illustrated with Figure 1.2.1:

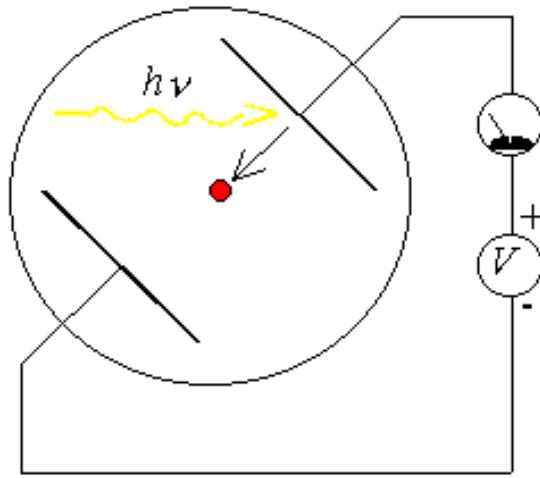


Figure 1.2.1.: Experimental set-up to measure the photoelectric effect.

The experimental apparatus consists of two metal electrodes within a vacuum chamber. Light is incident on one of two electrodes to which an external voltage is applied. The external voltage is adjusted so that the current due to the photo-emitted electrons becomes zero. This voltage corresponds to the maximum kinetic energy, $K.E.$, of the electrons in units of electron volt. That voltage is measured for different wavelengths and is plotted as a function of the inverse of the wavelength as shown in Figure 1.2.2. The resulting graph is a straight line.

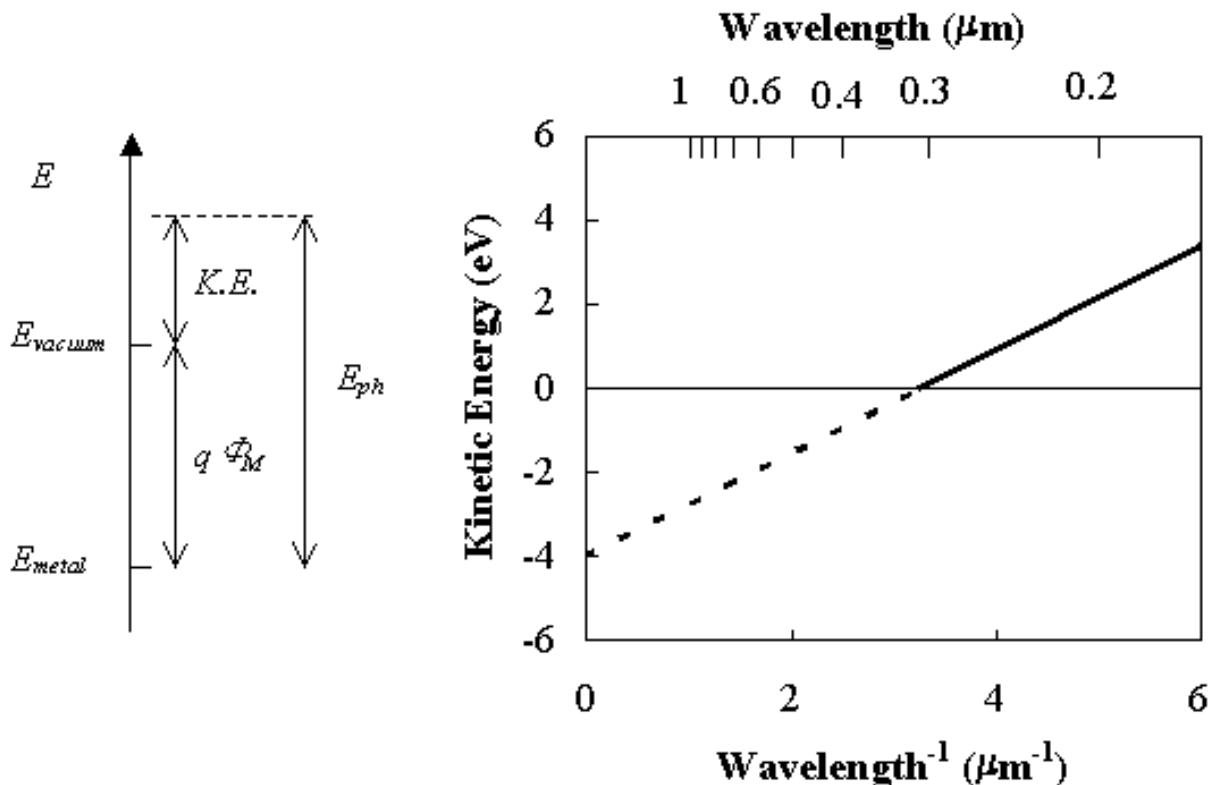


Figure 1.2.2 : Maximum kinetic energy, *K.E.*, of electrons emitted from a metal upon illumination with photon energy, E_{ph} . The energy is plotted versus the inverse of the wavelength of the light.

Albert Einstein explained this experiment by postulating that the energy of light is quantized. He assumed that light consists of individual particles called photons, so that the kinetic energy of the electrons, *K.E.*, equals the energy of the photons, E_{ph} , minus the energy, $q\Phi_M$, required to extract the electrons from the metal. The workfunction, Φ_M , therefore quantifies the potential, which the electrons have to overcome to leave the metal. The slope of the curve was measured to be 1.24 eV/micron, which yielded the following relation for the photon energy, E_{ph} :

$$E_{ph} = h\nu = \frac{hc}{\lambda} \quad (1.2.1)$$

where h is Planck's constant, ν is the frequency of the light, c is the speed of light in vacuum and λ is the wavelength of the light.

While other light-related phenomena such as the interference of two coherent light beams demonstrate the wave characteristics of light, it is the photoelectric effect, which demonstrates the particle-like behavior of light. These experiments lead to the particle-wave duality concept, namely that particles observed in an appropriate environment behave as waves, while waves can also behave as particles. This concept applies to all waves and particles. For instance, coherent electron beams also yield interference patterns similar to those of light beams.

It is the wave-like behavior of particles, which led to the de Broglie wavelength: since particles have wave-like properties, there is an associated wavelength, which is called the de Broglie wavelength and is given by:

$$\lambda = \frac{h}{p} \quad (1.2.2)$$

where λ is the wavelength, h is Planck's constant and p is the particle momentum. This expression enables a correct calculation of the ground energy of an electron in a hydrogen atom using the Bohr model described in Section 1.2.4. One can also show that the same expression applies to photons by combining equation (1.2.1) with $E_{ph} = p c$.

Example 1.1 	<p>A metal has a workfunction of 4.3 V. What is the minimum photon energy in Joule to emit an electron from this metal through the photo-electric effect? What are the photon frequency in Terahertz and the photon wavelength in micrometer? What is the corresponding photon momentum? What is the velocity of a free electron with the same momentum?</p>
Solution	<p>The minimum photon energy, E_{ph}, equals the workfunction, Φ_M, in units of electron volt or 4.3 eV. This also equals:</p> $E_{ph} = q\Phi_M = 1.6 \times 10^{-19} \times 4.3 = 6.89 \times 10^{-19} \text{ Joule}$ <p>The corresponding photon frequency is:</p> $\nu = \frac{6.89 \times 10^{-19}}{6.626 \times 10^{-34}} = 1040 \text{ THz}$ <p>The corresponding wavelength equals:</p> $\lambda = \frac{hc}{E_{ph}} = \frac{6.626 \times 10^{-34} \times 3 \times 10^8}{6.89 \times 10^{-19}} = \frac{1.24 \text{ nm}}{E_{ph} (\text{eV})} = 0.288 \text{ nm}$ <p>The photon momentum, p, is:</p> $p = \frac{h}{\lambda} = \frac{6.626 \times 10^{-34}}{0.288 \times 10^{-9}} = 2.297 \times 10^{-27} \frac{\text{kg m}}{\text{s}}$ <p>And the velocity, v, of a free electron with the same momentum equals:</p> $v = \frac{p}{m_0} = \frac{2.297 \times 10^{-27}}{9.11 \times 10^{-31}} = 2522 \text{ m/s}$ <p>Where m_0 is the free electron mass.</p> $E_{ph} = q\Phi_M = 1.6 \times 10^{-19} \times 4.3 = 6.89 \times 10^{-19} \text{ Joule}$

1.2.3 Blackbody radiation



Another experiment which could not be explained without quantum mechanics is the blackbody radiation experiment: By heating an object to high temperatures one finds that it radiates energy in the form of infra-red, visible and ultra-violet light. The appearance is that of a red glow at temperatures around 800° C which becomes brighter at higher temperatures and eventually looks like white light. The spectrum of the radiation is continuous, which led scientists to initially believe that classical electro-magnetic theory should apply. However, all attempts to describe this phenomenon failed until Max Planck developed the blackbody radiation theory based on the assumption that the energy associated with light is quantized and the energy quantum or photon energy equals:

$$E_{ph} = \hbar \nu = \hbar \omega \quad (1.2.3)$$

Where \hbar is the reduced Planck's constant ($= h/2\pi$), and ω is the radial frequency ($= 2\pi \nu$).

The spectral density, u_ω , or the energy density per unit volume and per unit frequency is given by:

$$u_\omega = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{\exp(\frac{\hbar \omega}{kT}) - 1} \quad (1.2.4)$$

Where k is Boltzmann's constant and T is the temperature. The spectral density is shown versus energy in Figure 1.2.3.

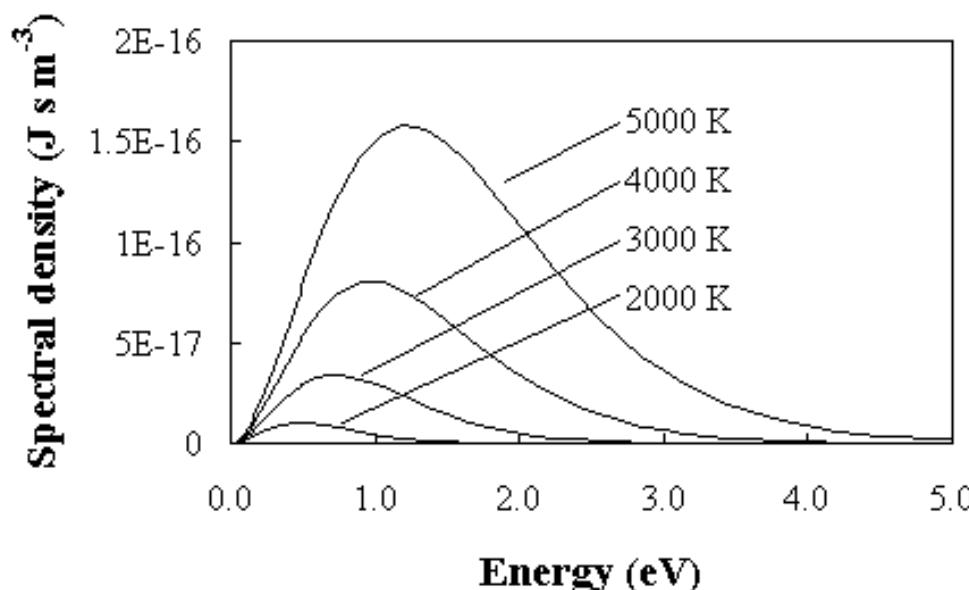


Figure 1.2.3: Spectral density of a blackbody at 2000, 3000, 4000 and 5000 K versus energy. 

The peak value of the blackbody radiation occurs at $2.82 kT$ and increases with the third power of the temperature. Radiation from the sun closely fits that of a black body at 5800 K.

Example 1.2

The spectral density of the sun peaks at a wavelength of 900 nm. If the sun behaves as a black body, what is the temperature of the sun?

Solution

A wavelength of 900 nm corresponds to a photon energy of:

$$E_{ph} = \frac{hc}{\lambda} = \frac{6.626 \times 10^{-34} \times 3 \times 10^8}{900 \times 10^{-9}} = 2.21 \times 10^{-19} \text{ Joule}$$

Since the peak of the spectral density occurs at $2.82 kT$, the corresponding temperature equals:

$$T = \frac{E_{ph}}{2.82 k} = \frac{2.21 \times 10^{-19}}{2.82 \times 1.38 \times 10^{-23}} = 5672 \text{ Kelvin}$$

1.2.4 The Bohr model



The spectrum of electromagnetic radiation from an excited hydrogen gas was yet another experiment, which was difficult to explain since it is discrete rather than continuous. The emitted wavelengths were early on associated with a set of discrete energy levels E_n described by:

$$E_n = -\frac{m_0 q^4}{8 \pi^2 h^2 n^2}, \text{ with } n = 1, 2, \dots \quad (1.2.5)$$

and the emitted photon energies equal the energy difference released when an electron makes a transition from a higher energy E_i to a lower energy E_j .

$$E_{ph} = 13.6 \text{ eV} \left(\frac{1}{j^2} - \frac{1}{i^2} \right), \text{ with } i > j \quad (1.2.6)$$

The maximum photon energy emitted from a hydrogen atom equals 13.6 eV. This energy is also called one Rydberg or one atomic unit. The electron transitions and the resulting photon energies are further illustrated by Figure 1.2.4.

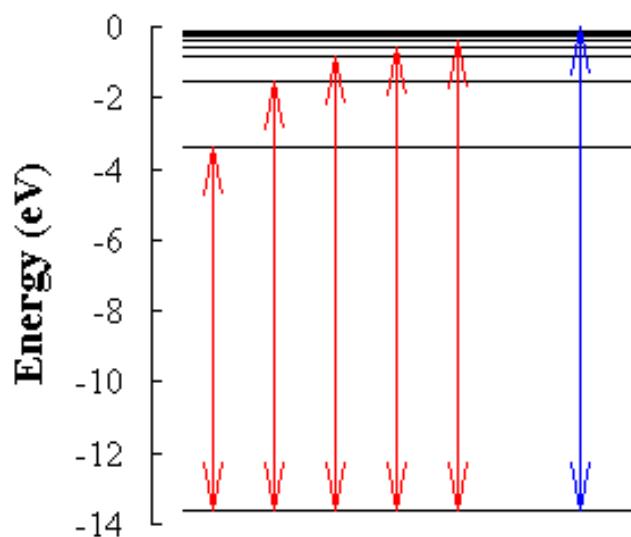


Figure 1.2.4 : Energy levels and possible electronic transitions in a hydrogen atom. Shown are the first six energy levels, as well as six possible transitions involving the lowest energy level ($n = 1$)

However, there was no explanation why the possible energy values were not continuous. No classical theory based on Newtonian mechanics could provide such spectrum. Further more, there was no theory, which could explain these specific values.

Niels Bohr provided a part of the puzzle. He assumed that electrons move along a circular trajectory around the proton like the earth around the sun, as shown in Figure 1.2.5.

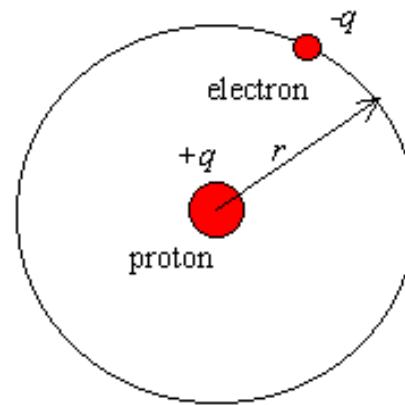


Figure 1.2.5: Trajectory of an electron in a hydrogen atom as used in the Bohr model.

He also assumed that electrons behave within the hydrogen atom as a wave rather than a particle. Therefore, the orbit-like electron trajectories around the proton are limited to those with a length, which equals an integer number of wavelengths so that

$$2\pi r = n\lambda \quad (1.2.7)$$

where r is the radius of the circular electron trajectory and n is a positive integer. The Bohr model also assumes that the momentum of the particle is linked to the de Broglie wavelength (equation (1.2.2))

The model further assumes a circular trajectory and that the centrifugal force equals the electrostatic force, or:

$$m \frac{v^2}{r} = \frac{q^2}{4\pi\epsilon_0 r^2} \quad (1.2.8)$$

Solving for the radius of the trajectory one finds the Bohr radius, a_0 :

$$a_0 = \frac{4\pi\epsilon_0 \hbar^2 n^2}{m q^2} \quad (1.2.9)$$

and the corresponding energy is obtained by adding the kinetic energy and the potential energy of the particle, yielding:

$$E_n = -\frac{m_0 q^4}{8\epsilon_0^2 \hbar^2 n^2}, \text{ with } n = 1, 2, \dots \quad (1.2.10)$$

Where the potential energy is the electrostatic potential of the proton:

$$V(r) = -\frac{q^2}{4\pi\epsilon_0 r} \quad (1.2.11)$$

Note that all the possible energy values are negative. Electrons with positive energy are not bound to the proton and behave as free electrons.

The Bohr model does provide the correct electron energies. However, it leaves many unanswered questions and, more importantly, it does not provide a general method to solve other problems of this type. The wave equation of electrons presented in the next section does provide a way to solve any quantum mechanical problem.



1.2.5 Schrödinger's equation

1.2.5.1. The infinite quantum well

1.2.5.2. The hydrogen atom

A general procedure to solve quantum mechanical problems was proposed by Erwin Schrödinger. Starting from a classical description of the total energy, E , which equals the sum of the kinetic energy, $K.E.$, and potential energy, V , or:

$$E = T + V = \frac{p^2}{2m} + V(x) \quad (1.2.12)$$

He converted this expression into a wave equation by defining a wavefunction, Ψ , and multiplied each term in the equation with that wavefunction:

$$E\Psi = \frac{p^2}{2m}\Psi + V(x)\Psi \quad (1.2.13)$$

$$- \frac{\hbar^2 \partial^2}{\partial x^2}$$

To incorporate the de Broglie wavelength of the particle we now introduce the operator, $-\frac{\hbar^2 \partial^2}{\partial x^2}$, which provides the square of the momentum, p , when applied to a plane wave:

$$- \hbar^2 \frac{\partial^2 \Psi}{\partial x^2} = \hbar^2 k^2 \Psi = p^2 \Psi \quad \text{for } \Psi = e^{i(kx - \omega t)} \quad (1.2.14)$$

Where k is the wavenumber, which equals $2\pi/\lambda$. Without claiming that this is an actual proof we now simply replace the momentum squared, p^2 , in equation (1.2.13) by this operator yielding the time-independent Schrödinger equation.

$$- \frac{\hbar^2}{2m} \frac{d^2 \Psi(x)}{dx^2} + V(x)\Psi(x) = E\Psi(x) \quad (1.2.15)$$

To illustrate the use of Schrödinger's equation, we present two solutions of Schrödinger's equation, that for an infinite quantum well and that for the hydrogen atom.

1.2.5.1. The infinite quantum well

The one-dimensional infinite quantum well represents one of the simplest quantum mechanical structures. We use it here to illustrate some specific properties of quantum mechanical systems. The potential in an infinite well is zero between $x = 0$ and $x = L_x$ and is infinite on either side of the well. The potential and the first five possible energy levels an electron can occupy are shown in Figure 1.2.6:

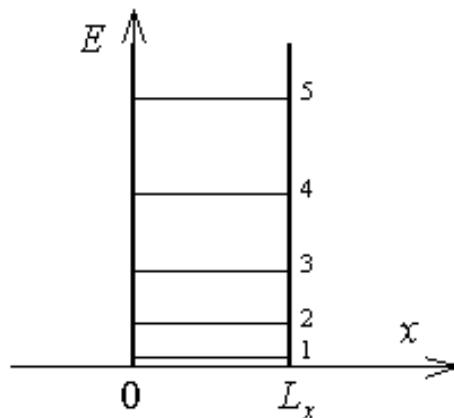


Figure 1.2.6 : Potential energy of an infinite well, with width L_x . Also indicated are the lowest five energy levels in the well.

The energy levels in such a infinite well are given by:

$$E_n = \frac{\hbar^2}{2m^*} \left(\frac{n}{2L_x}\right)^2, \text{ with } n = 1, 2, \dots \quad (1.2.16)$$

where \hbar is Planck's constant and m^* is the effective mass of the particle. n is the quantum number associated with the n^{th} energy level, with energy E_n . Note that the lowest possible energy is not zero although the potential is zero within the well.

Only discreet energy values are obtained as eigenvalues of the Schrödinger equation. The energy difference between adjacent energy levels increases as the energy increases. An electron occupying one of the energy levels can have a positive or negative spin ($s = 1/2$ or $s = -1/2$). Both quantum numbers, n and s , are the only two quantum numbers needed to describe this system.

The wavefunctions corresponding to each energy level are shown in Figure 1.2.7 (a). Each wavefunction has been shifted by the corresponding energy. The probability density function, calculated as $|\Psi|^2$, provides the probability of finding an electron in a certain location in the well. These probability density functions are shown in Figure 1.2.7 (b) for the first five energy levels. For instance, for $n = 2$ the electron is least likely to be in the middle of the well and at the edges of the well. The electron is most likely to be one quarter of the well width away from either edge.

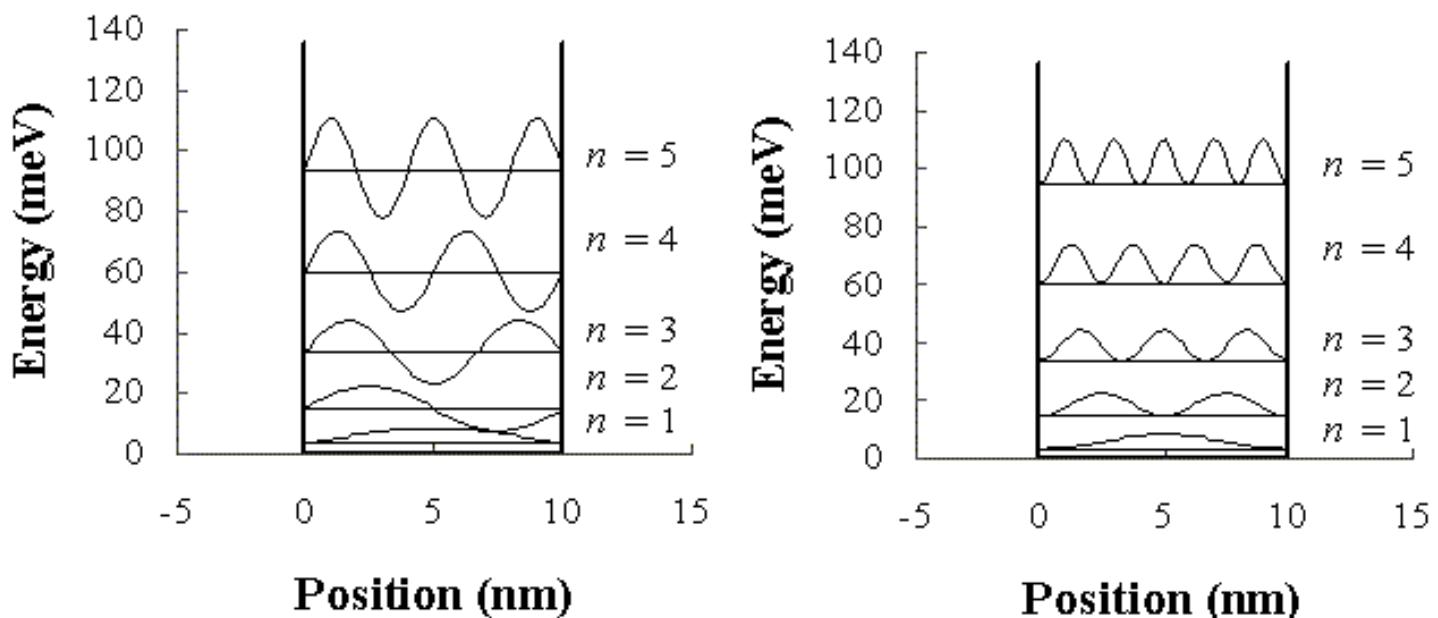


Figure 1.2.7 : Energy levels, wavefunctions (left) and probability density functions (right) in an infinite quantum well. The figure is calculated for a 10 nm wide well containing an electron with mass m_0 . The wavefunctions and the probability density functions are not normalized and shifted by the corresponding electron energy. 

Example 1.3  	<p>An electron is confined to a 1 micron thin layer of silicon. Assuming that the semiconductor can be adequately described by a one-dimensional quantum well with infinite walls, calculate the lowest possible energy within the material in units of electron volt. If the energy is interpreted as the kinetic energy of the electron, what is the corresponding electron velocity? (The effective mass of electrons in silicon is 0.26 m_0, where $m_0 = 9.11 \times 10^{-31}$ kg is the free electron rest mass).</p>
Solution	<p>The lowest energy in the quantum well equals:</p> $E_1 = \frac{\hbar^2}{2m^*} \left(\frac{1}{2L_x}\right)^2 = \frac{(6.626 \times 10^{-34})^2}{2 \times 0.26 \times 9.11 \times 10^{-31}} \left(\frac{1}{2 \times 10^{-6}}\right)^2$ $= 2.32 \times 10^{-25} \text{ Joules} = 1.45 \text{ meV}$ <p>The velocity of an electron with this energy equals:</p> $v = \sqrt{\frac{2E_1}{m^*}} = \sqrt{\frac{2 \times 2.32 \times 10^{-25}}{0.26 \times 9.11 \times 10^{-31}}} = 1.399 \text{ km/s}$

1.2.5.2. The hydrogen atom

The hydrogen atom represents the simplest possible atom since it consists of only one proton and one electron. Nevertheless, the solution to Schrödinger's equation as applied to the potential of the hydrogen atom is rather complex due to the three-dimensional nature of the problem. The potential, $V(r)$ (equation (1.2.11)), is due to the electrostatic force between the positively charged proton and the negatively charged electron. This potential as well as the first three probability density functions ($r^2|\Psi|^2$) of the radially symmetric wavefunctions ($l = 0$) is shown in Figure 1.2.8.

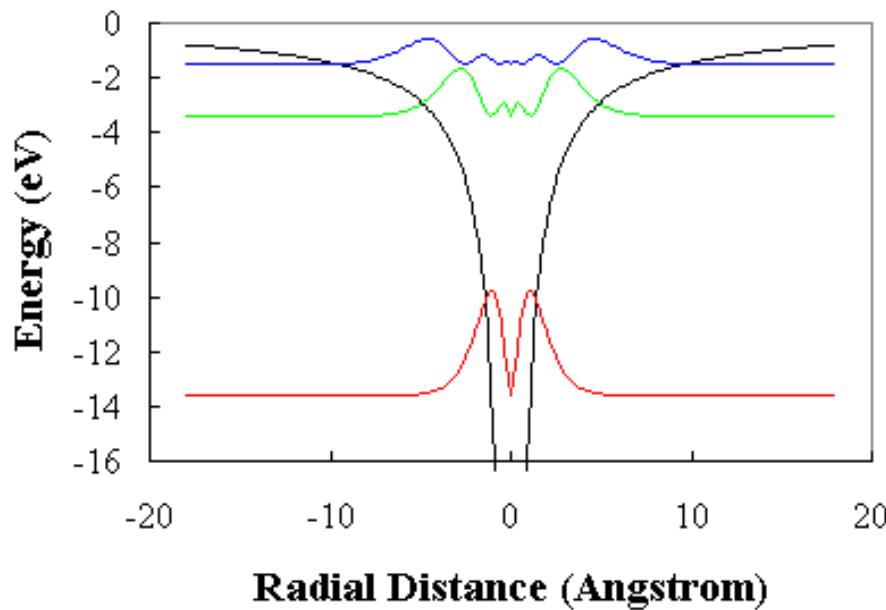


Figure 1.2.8 : Potential energy, $V(x)$, in a hydrogen atom and first three probability densities with $l = 0$. The probability densities are shifted by the corresponding electron energy.

Since the hydrogen atom is a three-dimensional problem, three quantum numbers, labeled n , l , and m , are needed to describe all possible solutions to Schrödinger's equation. The spin of the electron is described by the quantum number s . The energy levels only depend on n , the principal quantum number and are given by equation (1.2.10). The electron wavefunctions however are different for every different set of quantum numbers. While a derivation of the actual wavefunctions is beyond the scope of this text, a list of the possible quantum numbers is needed for further discussion and is therefore provided in Table 1.2.1. For each principal quantum number n , all smaller positive integers are possible values for the angular momentum quantum number l . The quantum number m can take on all integers between l and $-l$, while s can be $\frac{1}{2}$ or $-\frac{1}{2}$. This leads to a maximum of 2 unique sets of quantum numbers for all s orbitals ($l = 0$), 6 for all p orbitals ($l = 1$), 10 for all d orbitals ($l = 2$) and 14 for all f orbitals ($l = 3$).

n	l	m	s
1s	1	0	$\frac{1}{2}, -\frac{1}{2}$
2s	2	0	$\frac{1}{2}, -\frac{1}{2}$
2p	2	1, 0, -1	$\frac{1}{2}, -\frac{1}{2}$
3s	3	0	$\frac{1}{2}, -\frac{1}{2}$
3p	3	1, 0, -1	$\frac{1}{2}, -\frac{1}{2}$
3d	3	2, 1, 0, -1, -2	$\frac{1}{2}, -\frac{1}{2}$
4s	4	0	$\frac{1}{2}, -\frac{1}{2}$
4p	4	1, 0, -1	$\frac{1}{2}, -\frac{1}{2}$
4d	4	2, 1, 0, -1, -2	$\frac{1}{2}, -\frac{1}{2}$
4f	4	3, 2, 1, 0, -1, -2, -3	$\frac{1}{2}, -\frac{1}{2}$

Table 1.2.1: First ten orbitals and corresponding quantum numbers of a hydrogen atom

1.2.6 Pauli exclusion principle

Once the energy levels of an atom are known, one can find the electron configurations of the atom, provided the number of electrons occupying each energy level is known. Electrons are Fermions since they have a half integer spin. They must therefore obey the Pauli exclusion principle. This exclusion principle states that no two Fermions can occupy the same energy level corresponding to a unique set of quantum numbers n, l, m or s . The ground state of an atom is therefore obtained by filling each energy level, starting with the lowest energy, up to the maximum number as allowed by the Pauli exclusion principle.

1.2.7 Electronic configuration of the elements



The electronic configuration of the elements of the periodic table can be constructed using the quantum numbers of the hydrogen atom and the Pauli exclusion principle, starting with the lightest element hydrogen. Hydrogen contains only one proton and one electron. The electron therefore occupies the lowest energy level of the hydrogen atom, characterized by the principal quantum number $n = 1$. The orbital quantum number l equals zero and is referred to as an s orbital (not to be confused with the quantum number for spin, s). The s orbital can accommodate two electrons with opposite spin, but only one is occupied. This leads to the short-hand notation of $1s^1$ for the electronic configuration of hydrogen as listed in Table [1.2.2](#).

Helium is the second element of the periodic table. For this and all other atoms one still uses the same quantum numbers as for the hydrogen atom. This approach is justified since all atom cores can be treated as a single charged particle, which yields a potential very similar to that of a proton. While the electron energies are no longer the same as for the hydrogen atom, the electron wavefunctions are very similar and can be classified in the same way. Since helium contains two electrons it can accommodate two electrons in the $1s$ orbital, hence the notation $1s^2$. Since the s orbitals can only accommodate two electrons, this orbital is now completely filled, so that all other atoms will have more than one filled or partially-filled orbital. The two electrons in the helium atom also fill all available orbitals associated with the first principal quantum number, yielding a filled outer shell. Atoms with a filled outer shell are called noble gases as they are known to be chemically inert.

Lithium contains three electrons and therefore has a completely filled $1s$ orbital and one more electron in the next higher $2s$ orbital. The electronic configuration is therefore $1s^22s^1$ or $[He]2s^1$, where $[He]$ refers to the electronic configuration of helium. Beryllium has four electrons, two in the $1s$ orbital and two in the $2s$ orbital. The next six atoms also have a completely filled $1s$ and $2s$ orbital as well as the remaining number of electrons in the $2p$ orbitals. Neon has six electrons in the $2p$ orbitals, thereby completely filling the outer shell of this noble gas.

The next eight elements follow the same pattern leading to argon, the third noble gas. After that the pattern changes as the underlying $3d$ orbitals of the transition metals (scandium through zinc) are filled before the $4p$ orbitals, leading eventually to the fourth noble gas, krypton. Exceptions are chromium and zinc, which have one more electron in the $3d$ orbital and only one electron in the $4s$ orbital. A similar pattern change occurs for the remaining transition metals, where for the lanthanides and actinides the underlying f orbitals are filled first.

1	Hydrogen	H	$1s^1$								
2	Helium	He	$1s^2$								
3	Lithium	Li	$1s^2$	$2s^1$							
4	Beryllium	Be	$1s^2$	$2s^2$							
5	Boron	B	$1s^2$	$2s^2$	$2p^1$						
6	Carbon	C	$1s^2$	$2s^2$	$2p^2$						
7	Nitrogen	N	$1s^2$	$2s^2$	$2p^3$						
8	Oxygen	O	$1s^2$	$2s^2$	$2p^4$						
9	Fluorine	F	$1s^2$	$2s^2$	$2p^5$						
10	Neon	Ne	$1s^2$	$2s^2$	$2p^6$						
11	Sodium	Na	$1s^2$	$2s^2$	$2p^6$	$3s^1$					
12	Magnesium	Mg	$1s^2$	$2s^2$	$2p^6$	$3s^2$					
13	Aluminum	Al	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^1$				
14	Silicon	Si	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^2$				
15	Phosphorous	P	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^3$				
16	Sulfur	S	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^4$				
17	Chlorine	Cl	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^5$				
18	Argon	Ar	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$				
19	Potassium	K	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$		$4s^1$		
20	Calcium	Ca	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$		$4s^2$		
21	Scandium	Sc	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^1$	$4s^2$		
22	Titanium	Ti	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^2$	$4s^2$		
23	Vanadium	V	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^3$	$4s^2$		
24	Chromium	Cr	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^5$	$4s^1$		
25	Manganese	Mn	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^5$	$4s^2$		
26	Iron	Fe	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^6$	$4s^2$		
27	Cobalt	Co	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^7$	$4s^2$		
28	Nickel	Ni	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^8$	$4s^2$		
29	Copper	Cu	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^{10}$	$4s^1$		
30	Zinc	Zn	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^{10}$	$4s^2$		
31	Gallium	Ga	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^{10}$	$4s^2$	$4p^1$	
32	Germanium	Ge	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^{10}$	$4s^2$	$4p^2$	
33	Arsenic	As	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^{10}$	$4s^2$	$4p^3$	
34	Selenium	Se	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^{10}$	$4s^2$	$4p^4$	
35	Brome	Br	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^{10}$	$4s^2$	$4p^5$	
36	Krypton	Kr	$1s^2$	$2s^2$	$2p^6$	$3s^2$	$3p^6$	$3d^{10}$	$4s^2$	$4p^6$	

Table 1.2.2:

Electronic configuration of the first thirty-six elements of the periodic table.

Chapter 1: Review of Modern Physics



1.4. Statistical Thermodynamics

- [1.4.1.. Thermal equilibrium](#)
- [1.4.2. Laws of thermodynamics](#)
- [1.4.3. The thermodynamic identity](#)
- [1.4.4. The Fermi energy](#)
- [1.4.5. Some useful thermodynamics results](#)

Thermodynamics describes the behavior of systems containing a large number of particles. These systems are characterized by their temperature, volume, number and the type of particles. The state of the system is then further described by its total energy and a variety of other parameters including the entropy. Such a characterization of a system is much simpler than trying to keep track of each particle individually, hence its usefulness. In addition, such a characterization is general in nature so that it can be applied to mechanical, electrical and chemical systems.

The term thermodynamics is somewhat misleading as one deals primarily with systems in thermal equilibrium. These systems have constant temperature, volume and number of particles and their macroscopic parameters do not change over time, so that the dynamics are limited to the microscopic dynamics of the particles within the system.

Statistical thermodynamics is based on the fundamental assumption that all possible configurations of a given system, which satisfy the given boundary conditions such as temperature, volume and number of particles, are equally likely to occur. The overall system will therefore be in the statistically most probable configuration. The entropy of a system is defined as the logarithm of the number of possible configurations. While such definition does not immediately provide insight into the meaning of entropy, it does provide a straightforward analysis since the number of configurations can be calculated for any given system.

Classical thermodynamics provides the same concepts. However, they are obtained through experimental observation. The classical analysis is therefore more tangible compared to the abstract mathematical treatment of the statistical approach.

The study of semiconductor devices requires some specific results, which naturally emerge from statistical thermodynamics. In this section, we review basic thermodynamic principles as well as some specific results. These include the thermal equilibrium concept, the thermodynamic identity, the basic laws of thermodynamics, the thermal energy per particle and the Fermi function.

1.4.1. Thermal equilibrium



A system is in thermal equilibrium if detailed balance is obtained: i.e. every process in the system is exactly balanced by its inverse process so that there is no net effect on the system.

This definition implies that in thermal equilibrium no energy (heat, work or particle energy) is exchanged between the parts within the system or between the system and the environment. Thermal equilibrium is obtained by isolating a system from its environment, removing any internal sources of energy, and waiting for a long enough time until the system does not change any more.

The concept of thermal equilibrium is of interest since various thermodynamic results assume that the system under consideration is in thermal equilibrium. Few systems of interest rigorously satisfy this condition so that we often apply the thermodynamical results to systems that are "close" to thermal equilibrium. Agreement between theories based on this assumption and experiments justify this approach.

1.4.2. Laws of thermodynamics



If two systems are in thermal equilibrium with a third system, they must be in thermal equilibrium with each other.

1. Heat is a form of energy.
2. The second law can be stated either (a) in its classical form or (b) in its statistical form
 - a. Heat can only flow from a higher temperature to a lower temperature.
 - b. The entropy of a closed system tends to remain constant or increases monotonically over time.
- Both forms of the second law could not seem more different. A more rigorous treatment proves the equivalence of both.
3. The entropy of a system approaches a constant as the temperature approaches zero Kelvin.

1.4.3. The thermodynamic identity



The thermodynamic identity states that a change in energy can be caused by adding heat, work or particles. Mathematically this is expressed by:

$$dU = dQ + dW + \mu dN \quad (1.4.1)$$

where U is the total energy, Q is the heat and W is the work. μ is the energy added to a system when adding one particle without adding either heat or work. This energy is also called the electro-chemical potential. N is the number of particles.

1.4.4. The Fermi energy



The Fermi energy, E_F , is the energy associated with a particle, which is in thermal equilibrium with the system of interest. The energy is strictly associated with the particle and does not consist even in part of heat or work. This same quantity is called the electro-chemical potential, μ , in most thermodynamics texts.

1.4.5. Some useful thermodynamics results



Listed below are two results, which will be used while analyzing semiconductor devices. The actual derivation is beyond the scope of this text.

1. The thermal energy of a particle, whose energy depends quadratically on its velocity, equals $kT/2$ per degree of freedom, where k is Boltzmann's constant. This thermal energy is a kinetic energy, which must be added to the potential energy of the particle, and any other kinetic energy. The thermal energy of a non-relativistic electron, which is allowed to move in three dimensions, equals $3/2 kT$.
2. Consider an energy level at energy, E , which is in thermal equilibrium with a large system characterized by a temperature T and Fermi energy E_F . The probability that an electron occupies such energy level is given by:

$$f(E) = \frac{1}{1 + e^{(E - E_F) / kT}} \quad (1.4.2)$$

The function $f(E)$ is called the Fermi function and applies to all particles with half-integer spin. These particles, also called Fermions, obey the Pauli exclusion principle, which states that no two Fermions in a given system can have the exact same set of quantum numbers. Since electrons are Fermions, their probability distribution also equals the Fermi function.

Example 1.5

Calculate the energy relative to the Fermi energy for which the Fermi function equals 5%. Write the answer in units of kT .

Solution

The problem states that:

$$f(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} = 0.05$$

which can be solved yielding:

$$E - E_F = \ln(19)kT = 3kT$$

Chapter 1: Review of Modern Physics



1.3 Electromagnetic Theory

1.3.1. Gauss's law

1.3.2. Poisson's equation

The analysis of most semiconductor devices includes the calculation of the electrostatic potential within the device as a function of the existing charge distribution. Electromagnetic theory and more specifically electrostatic theory are used to obtain the potential. A short description of the necessary tools, namely Gauss's law and Poisson's equation, is provided below.

1.3.1 Gauss's law



Gauss's law is one of Maxwell's equations and provides the relation between the charge density, ρ , and the electric field, \vec{E} . In the absence of time dependent magnetic fields the one-dimensional equation is given by:

$$\frac{d\vec{E}(x)}{dx} = \frac{\rho(x)}{\epsilon} \quad (1.3.1)$$

This equation can be integrated to yield the electric field for a given one-dimensional charge distribution:

$$\vec{E}(x_2) - \vec{E}(x_1) = \int_{x_1}^{x_2} \frac{\rho(x)}{\epsilon} dx \quad (1.3.2)$$

Gauss's law as applied to a three-dimensional charge distribution relates the divergence of the electric field to the charge density:

$$\vec{\nabla} \cdot \vec{E}(x, y, z) = \frac{\rho(x, y, z)}{\epsilon} \quad (1.3.3)$$

This equation can be simplified if the field is constant on a closed surface, A , enclosing a charge Q , yielding:

$$\vec{E} \cdot \vec{A} = \frac{Q}{\epsilon} \quad (1.3.4)$$

Example 1.4



Consider an infinitely long cylinder with charge density r , dielectric constant ϵ_0 and radius r_0 . What is the electric field in and around the cylinder?

Solution	<p>Because of the cylinder symmetry one expects the electric field to be only dependent on the radius, r. Applying Gauss's law one finds:</p> $\vec{\mathcal{E}} \cdot \vec{A} = \mathcal{E} 2\pi r L = \frac{Q}{\epsilon_0} = \frac{\rho \pi r^2 L}{\epsilon_0}$ <p>and</p> $\vec{\mathcal{E}} \cdot \vec{A} = \mathcal{E} 2\pi r L = \frac{Q}{\epsilon_0} = \frac{\rho \pi r_0^2 L}{\epsilon_0}$ <p>where a cylinder with length L was chosen to define the surface A, and edge effects were ignored. The electric field then equals:</p> $\mathcal{E}(r) = \frac{\rho r}{2 \epsilon_0}$ <p>The electric field increases within the cylinder with increasing radius. The electric field decreases outside the cylinder with increasing radius.</p>
----------	---

1.3.2 Poisson's equation



Gauss's law is one of Maxwell's equations and provides the relation between the charge density, ρ , and the electric field, \mathcal{E} . In the absence of time dependent magnetic fields the one-dimensional equation is given by:

$$\frac{d\phi(x)}{dx} = -\mathcal{E}(x) \quad (1.3.5)$$

The electric field vector therefore originates at a point of higher potential and points towards a point of lower potential. The potential can be obtained by integrating the electric field as described by:

$$\phi(x_2) - \phi(x_1) = - \int_{x_1}^{x_2} \mathcal{E}(x) dx \quad (1.3.6)$$

At times, it is convenient to link the charge density to the potential by combining equation (1.3.5) with Gauss's law in the form of equation (1.3.1), yielding:

$$\frac{d^2 \phi(x)}{dx^2} = -\frac{\rho(x)}{\epsilon} \quad (1.3.7)$$

which is referred to as Poisson's equation.

For a three-dimensional field distribution, the gradient of the potential as described by:

$$\vec{\nabla} \phi(x, y, z) = -\vec{\mathcal{E}}(x, y, z) \quad (1.3.8)$$

can be combined with Gauss's law as formulated with equation (1.3.3), yielding a more general form of Poisson's equation:

$$\nabla^2 \phi(x, y, z) = -\frac{\rho(x, y, z)}{\epsilon} \quad (1.3.9)$$

Example 1.5 Calculate the energy relative to the Fermi energy for which the Fermi function equals 5%. Write the answer in units of kT .

Solution The problem states that:

$$f(E) = \frac{1}{1 + \exp\left(\frac{E - E_F}{kT}\right)} = 0.05$$

which can be solved yielding:

$$E - E_F = \ln(19)kT = 3kT$$

Chapter 2: Semiconductor Fundamentals



Examples

Example 2.1

Calculate the maximum fraction of the volume in a simple cubic crystal occupied by the atoms. Assume that the atoms are closely packed and that they can be treated as hard spheres. This fraction is also called the packing density.

Example 2.2

Calculate the energy bandgap of germanium, silicon and gallium arsenide at 300, 400, 500 and 600 K.

Example 2.3

Calculate the number of states per unit energy in a 100 by 100 by 10 nm piece of silicon ($m^* = 1.08 m_0$) 100 meV above the conduction band edge. Write the result in units of eV⁻¹.

Example 2.4

Calculate the effective densities of states in the conduction and valence bands of germanium, silicon and gallium arsenide at 300 K.

Example 2.4b

Calculate the intrinsic carrier density in germanium, silicon and gallium arsenide at 300, 400, 500 and 600 K.

Example 2.5

Calculate the ionization energy for shallow donors and acceptors in germanium and silicon using the hydrogen-like model.

Example 2.6a

A germanium wafer is doped with a shallow donor density of $3n_i/2$. Calculate the electron and hole density.

Example 2.6b

A silicon wafer is doped with a shallow acceptor doping of 10^{16} cm^{-3} . Calculate the electron and hole density.

Example 2.7

A piece of germanium doped with 10^{16} cm^{-3} shallow donors is illuminated with light generating 10^{15} cm^{-3} excess electrons and holes. Calculate the quasi-Fermi energies relative to the intrinsic energy and compare it to the Fermi energy in the absence of illumination.

Example 2.8

Electrons in undoped gallium arsenide have a mobility of $8,800 \text{ cm}^2/\text{V}\cdot\text{s}$. Calculate the average time between collisions. Calculate the distance traveled between two collisions (also called the mean free path). Use an average velocity of 10^7 cm/s .

Example 2.9

A piece of silicon doped with arsenic ($N_d = 10^{17} \text{ cm}^{-3}$) is 100 mm long, 10 mm wide and 1 mm thick. Calculate the resistance of this sample when contacted one each end.

Example 2.10

The hole density in an n-type silicon wafer ($N_d = 10^{17} \text{ cm}^{-3}$) decreases linearly from 10^{14} cm^{-3} to 10^{13} cm^{-3} between $x = 0$ and $x = 1 \mu\text{m}$. Calculate the hole diffusion current density.

Example 2.11

Calculate the electron and hole densities in an n-type silicon wafer ($N_d = 10^{17} \text{ cm}^{-3}$) illuminated uniformly with 10 mW/cm^2 of red light ($E_{\text{ph}} = 1.8 \text{ eV}$). The absorption coefficient of red light in silicon is 10^{-3} cm^{-1} . The minority carrier lifetime is $10 \mu\text{s}$.

Chapter 2: Semiconductor Fundamentals



Equations

$$\vec{r} = k \vec{a}_1 + l \vec{a}_2 + m \vec{a}_3 \quad (2.2.1) \bullet$$

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta} \quad (2.3.1) \bullet$$

$$J_{vb} = \frac{1}{V} \sum_{\substack{\text{filled} \\ \text{states}}} (-q) v_i \quad (2.3.2) \bullet$$

$$J_{vb} = \frac{1}{V} \left(\sum_{\substack{\text{all} \\ \text{states}}} (-q) v_i - \sum_{\substack{\text{empty} \\ \text{states}}} (-q) v_i \right) \quad (2.3.3) \bullet$$

$$J_{vb} = \frac{1}{V} \sum_{\substack{\text{empty} \\ \text{states}}} (+q) v_i \quad (2.3.4) \bullet$$

$$\Psi = A \sin(k_x x) + B \cos(k_x x) \quad (2.4.1) \bullet$$

$$k_x = \frac{n\pi}{L}, n = 1, 2, 3, \dots \quad (2.4.2) \bullet$$

$$N = 2 \times \frac{1}{8} \times \left(\frac{L}{\pi}\right)^3 \times \frac{4}{3} \times \pi \times k^3 \quad (2.4.3) \bullet$$

$$\frac{dN}{dE} = \frac{dN}{dk} \frac{dk}{dE} = \left(\frac{L}{\pi}\right)^3 \pi k^2 \frac{dk}{dE} \quad (2.4.4) \bullet$$

$$E(k) = \frac{\hbar^2 k^2}{2m^*} \quad \bullet(2.4.5)\bullet$$

$$g(E) = \frac{1}{L^3} \frac{dN}{dE} = \frac{8\pi\sqrt{2}}{\hbar^3} m^{*3/2} \sqrt{E}, \text{ for } E \geq 0 \quad \bullet(2.4.6)\bullet$$

$$g_c(E) = \frac{1}{L^3} \frac{dN}{dE} = \frac{8\pi\sqrt{2}}{\hbar^3} m^{*3/2} \sqrt{E - E_c}, \text{ for } E \geq E_c \quad \bullet(2.4.7)\bullet$$

$$f(E) = \frac{1}{1 + e^{(E - E_F)/kT}} \quad \bullet(2.5.1)\bullet$$

$$f_{donor}(E_d) = \frac{1}{1 + \frac{1}{2} e^{(E_d - E_F)/kT}} \quad \bullet(2.5.2)\bullet$$

$$f_{acceptor}(E_a) = \frac{1}{1 + 4e^{(E_a - E_F)/kT}} \quad \bullet(2.5.3)\bullet$$

$$f_{BE}(E) = \frac{1}{e^{(E - E_F)/kT} - 1} \quad \bullet(2.5.4)\bullet$$

$$f_{MB}(E) = \frac{1}{e^{(E - E_F)/kT}} \quad \bullet(2.5.5)\bullet$$

$$n(E) = g_c(E)f(E) \quad \bullet(2.6.1)\bullet$$

$$p(E) = g_v(E)[1 - f(E)] \quad \bullet(2.6.2)\bullet$$

$$g_c(E) = \frac{8\pi\sqrt{2}}{\hbar^3} m_e^{*3/2} \sqrt{E - E_c}, \text{ for } E \geq E_c \quad \bullet(2.6.3)\bullet$$

$$g_v(E) = \frac{8\pi\sqrt{2}}{\hbar^3} m_e^{*3/2} \sqrt{E_v - E}, \text{ for } E \leq E_v \quad \bullet(2.6.4)\bullet$$

$$n = \frac{\text{top of the conduction band}}{\int_{E_c}^{\infty} n(E) dE} = \frac{\text{top of the conduction band}}{\int_{E_c}^{\infty} g_c(E) f(E) dE} \quad \bullet(2.6.5)\bullet$$

$$n_o = \int_{E_c}^{\infty} g_c(E) f(E) dE \quad \bullet(2.6.6)\bullet$$

$$n_o = \int_{E_c}^{\infty} \frac{8\pi\sqrt{2}}{\hbar^3} m_e^{*3/2} \sqrt{E - E_c} \frac{1}{1 + e^{\frac{E - E_F}{kT}}} dE \quad \bullet(2.6.7)\bullet$$

$$p_o = \int_{-\infty}^{E_F} g_v(E) [1 - f(E)] dE \quad \bullet(2.6.8)\bullet$$

$$p_o = \int_{-\infty}^{E_F} \frac{8\pi\sqrt{2}}{\hbar^3} m_h^{*3/2} \sqrt{E_v - E} \frac{1}{1 + e^{\frac{E_F - E}{kT}}} dE \quad \bullet(2.6.9)\bullet$$

$$n_o = \int_{E_c}^{E_F} g_c(E) dE \text{ at } T = 0 \text{ K} \quad \bullet(2.6.10)\bullet$$

$$n_o = \frac{2\sqrt{2}}{3\pi^2} \left(\frac{qm^*}{\hbar^2} \right)^{3/2} (E_F - E_c)^{3/2}, \text{ for } E_F \geq E_c \quad \bullet(2.6.11)\bullet$$

$$n_o \cong \int_{E_c}^{\infty} \frac{8\pi\sqrt{2}}{\hbar^3} m_e^{*3/2} \sqrt{E - E_c} e^{\frac{E_F - E}{kT}} dE = N_e e^{\frac{E_F - E_c}{kT}} \quad \bullet(2.6.12)\bullet$$

$$N_c = 2 \left[\frac{2 \pi m_e^* k T}{h^2} \right]^{3/2} \quad \bullet(2.6.13)\bullet$$

$$p_o \cong \int_{-\infty}^{E_c} \frac{8 \pi \sqrt{2}}{h^3} m_h^{*3/2} \sqrt{E_v - E} e^{-\frac{E-E_F}{kT}} dE = N_v e^{-\frac{E_c-E_F}{kT}} \quad \bullet(2.6.14)\bullet$$

$$N_v = 2 \left[\frac{2 \pi m_h^* k T}{h^2} \right]^{3/2} \quad \bullet(2.6.15)\bullet$$

$$\frac{E_F - E_c}{kT} \cong \ln \frac{n_o}{N_c} + \frac{1}{\sqrt{8}} \frac{n_o}{N_c} - \left(\frac{3}{16} - \frac{\sqrt{3}}{9} \right) \left(\frac{n_o}{N_c} \right)^2 + \dots \quad \bullet(2.6.16)\bullet$$

$$\frac{E_v - E_F}{kT} \cong \ln \frac{p_o}{N_v} + \frac{1}{\sqrt{8}} \frac{p_o}{N_v} - \left(\frac{3}{16} - \frac{\sqrt{3}}{9} \right) \left(\frac{p_o}{N_v} \right)^2 + \dots \quad \bullet(2.6.17)\bullet$$

$$n_i = n_o \Big|_{(E_F = E_i)} = N_c e^{(E_i - E_c)/kT} \quad \bullet(2.6.18)\bullet$$

$$n_i = \sqrt{N_c N_v} e^{-E_g/2kT} \quad \bullet(2.6.19)\bullet$$

$$n_o \cdot p_o = N_c N_v e^{(E_v - E_c)/kT} = n_i^2 \quad \bullet(2.6.20)\bullet$$

$$E_i = \frac{E_c + E_v}{2} + \frac{1}{2} kT \ln \left(\frac{N_v}{N_c} \right) \quad \bullet(2.6.21)\bullet$$

$$E_i = \frac{E_c + E_v}{2} + \frac{3}{4} kT \ln \left(\frac{m_h^*}{m_e} \right) \quad \bullet(2.6.22)\bullet$$

$$n_o = n_i e^{(E_F - E_i)/kT} \quad \bullet(2.6.23)\bullet$$

$$p_o = n_i e^{(E_i - E_F)/kT} \quad \bullet(2.6.24)\bullet$$

$$E_F = E_i + kT \ln \frac{n_o}{n_i} \quad \bullet(2.6.25)\bullet$$

$$E_F = E_i - kT \ln \frac{p_o}{n_i} \quad \bullet(2.6.26)\bullet$$

$$N_d^+ \cong N_d \quad \bullet(2.6.27)\bullet$$

$$N_a^- \cong N_a \quad \bullet(2.6.28)\bullet$$

$$n_o \cong N_d^+ - N_a^- , \text{ if } N_d^+ - N_a^- \gg n_i \quad \bullet(2.6.29)\bullet$$

$$p_o \cong N_a^- - N_d^+ , \text{ if } N_a^- - N_d^+ \gg n_i \quad \bullet(2.6.30)\bullet$$

$$E_c - E_d = 13.6 \frac{m_{cond}^*}{m_0 e^2} \text{ eV} \quad \bullet(2.6.31)\bullet$$

$$\rho = q(p_o - n_o + N_d^+ - N_a^-) = 0 \quad \bullet(2.6.32)\bullet$$

$$n_o = \frac{n_i^2}{n_o} + N_d^+ - N_a^- \quad \bullet(2.6.33)\bullet$$

$$n_o = \frac{N_d^+ - N_a^-}{2} + \sqrt{\left(\frac{N_d^+ - N_a^-}{2}\right)^2 + n_i^2} \quad \bullet(2.6.34)\bullet$$

$$p_o = \frac{N_a^- - N_d^+}{2} + \sqrt{\left(\frac{N_a^- - N_d^+}{2}\right)^2 + n_i^2} \quad \bullet(2.6.35)\bullet$$

$$p_o + N_d^+ = n_o + N_a^- \quad \bullet(2.6.36)\bullet$$

$$n = n_o + \beta n = n_i \exp\left(\frac{F_n - E_i}{kT}\right) \quad \bullet(2.6.37)\bullet$$

$$p = p_o + \beta p = n_i \exp\left(\frac{E_i - F_p}{kT}\right) \quad \bullet(2.6.38)\bullet$$

$$I = \frac{Q}{t_r} = \frac{Q}{L/v} \quad \bullet(2.7.1)\bullet$$

$$\vec{J} = \frac{Q}{AL} \vec{v} = \mathcal{N} \vec{v} = qn\vec{v} \quad \bullet(2.7.2)\bullet$$

$$\vec{F} = m\vec{a} = m \frac{d \langle \vec{v} \rangle}{dt} \quad \bullet(2.7.3)\bullet$$

$$\vec{F} = q\vec{\mathcal{E}} - \frac{m \langle \vec{v} \rangle}{\tau} \quad \bullet(2.7.4)\bullet$$

$$q\vec{\mathcal{E}} = m \frac{d \langle \vec{v} \rangle}{dt} + \frac{m \langle \vec{v} \rangle}{\tau} \quad \bullet(2.7.5)\bullet$$

$$\mu = \frac{\Delta |\vec{v}|}{|\vec{\mathcal{E}}|} = \frac{q \tau}{m} \quad \bullet(2.7.6)\bullet$$

$$\vec{J} = q n \mu_n \vec{\mathcal{E}} \quad \bullet(2.7.7)\bullet$$

$$\mu = \frac{q \tau}{m} \quad \bullet(2.7.8)\bullet$$

$$\mu = \mu_{\min} + \frac{\mu_{\max} - \mu_{\min}}{1 + \left(\frac{N}{N_r}\right)} \quad \bullet(2.7.9)\bullet$$

$$J = q n v_e + q p v_h = q(n \mu_n + p \mu_p) \mathcal{E} \quad \bullet(2.7.10)\bullet$$

$$\frac{\Delta J}{\mathcal{E}} = q(n \mu_n + p \mu_p) \quad \bullet(2.7.11)\bullet$$

$$\sigma = \frac{1}{\mathcal{E}} = \frac{1}{q(n \mu_n + p \mu_p)} \quad \bullet(2.7.12)\bullet$$

$$R_s = \frac{\sigma}{t} \quad \bullet(2.7.13)\bullet$$

$$R = R_s \frac{L}{W} \quad \bullet(2.7.14)\bullet$$

$$v(\mathcal{E}) = \frac{\mu \mathcal{E}}{1 + \frac{\mu \mathcal{E}}{v_{sat}}} \quad \bullet(2.7.15)\bullet$$

$$v_{th} = \frac{l}{\tau} \quad \bullet(2.7.16)\bullet$$

$$\Phi_{n, left \rightarrow right} = \frac{1}{2} v_{th} n(x = -l) \quad \bullet(2.7.17)\bullet$$

$$\Phi_{n, right \rightarrow left} = \frac{1}{2} v_{th} n(x = l) \quad \bullet(2.7.18)\bullet$$

$$\Phi_n = \Phi_{n, left \rightarrow right} - \Phi_{n, right \rightarrow left} = \frac{1}{2} v_{th} [n(x = -l) - n(x = l)] \quad \bullet(2.7.19)\bullet$$

$$\Phi_n = -l v_{th} \frac{n(x = l) - n(x = -l)}{2l} = -l v_{th} \frac{dn}{dx} \quad \bullet(2.7.20)\bullet$$

$$J_n = -q \Phi_n = q l v_{th} \frac{dn}{dx} \quad \bullet(2.7.21)\bullet$$

$$J_n = q D_n \frac{dn}{dx} \quad \bullet(2.7.22)\bullet$$

$$J_p = -q D_p \frac{dp}{dx} \quad \bullet(2.7.23)\bullet$$

$$v_{th} = \frac{l}{\tau} \quad \bullet(2.7.24)\bullet$$

$$\frac{kT}{2} = \frac{m^* v_{th}^2}{2} \quad \bullet(2.7.25)\bullet$$

$$kT = \frac{m^* v_{th}^2}{q} \frac{q \tau}{m^*} = \frac{kT}{q} \mu \quad \bullet(2.7.26)\bullet$$

$$D_n = \mu_n \frac{kT}{q} = \mu_n V_t \quad \bullet(2.7.27)\bullet$$

$$D_p = \mu_p \frac{kT}{q} = \mu_p V_t \quad \bullet(2.7.28)\bullet$$

$$J_n = qn \mu_n \mathcal{E} + q D_n \frac{dn}{dx} \quad \bullet(2.7.29)\bullet$$

$$J_p = qp \mu_p \mathcal{E} - q D_p \frac{dp}{dx} \quad \bullet(2.7.30)\bullet$$

$$I_{total} = A(J_n + J_p) \quad \bullet(2.7.31)\bullet$$

$$U_n = R_n - G_n = \frac{n_p - n_{p0}}{\tau_n} \quad \bullet(2.8.1)\bullet$$

$$U_p = R_p - G_p = \frac{p_n - p_{n0}}{\tau_p} \quad \bullet(2.8.2)\bullet$$

$$U_{b-b} = b(np - n_i^2) \quad \bullet(2.8.3)\bullet$$

$$U_{SHR} = \frac{pn - n_i^2}{p + n + 2n_i \cosh(\frac{E_i - E_t}{kT})} N_t v_{th} \sigma \quad \bullet(2.8.4)\bullet$$

$$U_n = R_n - G_n = \frac{n_p - n_{p0}}{\gamma_n} \quad \bullet(2.8.5)\bullet$$

$$U_p = R_p - G_p = \frac{p_n - p_{n0}}{\gamma_p} \quad \bullet(2.8.6)\bullet$$

$$\gamma_n = \gamma_p = \frac{1}{N_t v_{th} \sigma} \quad \bullet(2.8.7)\bullet$$

$$U_{s,SHR} = \frac{pn - n_i^2}{p + n + 2n_i \cosh(\frac{E_i - E_{st}}{kT})} N_{st} v_{th} \sigma_s \quad \bullet(2.8.8)\bullet$$

$$U_{s,n} = R_{s,n} - G_{s,n} = v_s (n_p - n_{p0}) \quad \bullet(2.8.9)\bullet$$

$$v_s = N_{st} v_{th} \sigma_s \quad \bullet(2.8.10)\bullet$$

$$U_{Auger} = \Gamma_n n (np - n_i^2) + \Gamma_p p (np - n_i^2) \quad \bullet(2.8.11)\bullet$$

$$G_{p,light} = G_{n,light} = \alpha \frac{P_{opt}(x)}{E_{ph} A} \quad \bullet(2.8.12)\bullet$$

$$\frac{dP_{opt}(x)}{dx} = -\alpha \mathcal{P}_{opt}(x) \quad \bullet(2.8.13)\bullet$$

$$\frac{\partial n(x,t)}{\partial t} A dx = \left(\frac{J_n(x)}{-q} - \frac{J_n(x+dx)}{-q} \right) A + (G_n(x,t) - R_n(x,t)) A dx \quad \bullet(2.9.1)\bullet$$

$$J_n(x+dx) = J_n(x) + \frac{d J_n(x)}{dx} dx \quad \bullet(2.9.2)\bullet$$

$$\frac{\partial n(x,t)}{\partial t} = \frac{1}{q} \frac{\partial J_n(x,t)}{\partial x} + G_n(x,t) - R_n(x,t) \quad \bullet(2.9.3)\bullet$$

$$\frac{\partial p(x,t)}{\partial t} = -\frac{1}{q} \frac{\partial J_p(x,t)}{\partial x} + G_p(x,t) - R_p(x,t) \quad \bullet(2.9.4)\bullet$$

$$\begin{aligned} \frac{\partial n(x,t)}{\partial t} = & \\ \mu_n n \frac{\partial \mathcal{E}(x,t)}{\partial x} + \mu_n \mathcal{E} \frac{\partial n(x,t)}{\partial x} + D_n \frac{\partial^2 n(x,t)}{\partial x^2} + G_n(x,t) - R_n(x,t) \end{aligned} \quad \bullet(2.9.5)\bullet$$

$$\begin{aligned} \frac{\partial p(x,t)}{\partial t} = & \\ -\mu_p p \frac{\partial \mathcal{E}(x,t)}{\partial x} - \mu_p \mathcal{E} \frac{\partial p(x,t)}{\partial x} + D_p \frac{\partial^2 p(x,t)}{\partial x^2} + G_p(x,t) - R_p(x,t) \end{aligned} \quad \bullet(2.9.6)\bullet$$

$$\frac{\partial n(x,y,z,t)}{\partial t} = \frac{1}{q} \vec{\nabla} \vec{J}_n(x,y,z,t) + G_n(x,y,z,t) - R_n(x,y,z,t) \quad \bullet(2.9.7)\bullet$$

$$\frac{\partial p(x,y,z,t)}{\partial t} = -\frac{1}{q} \vec{\nabla} \vec{J}_p(x,y,z,t) + G_p(x,y,z,t) - R_p(x,y,z,t) \quad \bullet(2.9.8)\bullet$$

$$\frac{\partial n(x,t)}{\partial t} = D_n \frac{\partial^2 n_p(x,t)}{\partial x^2} - \frac{n_p(x,t) - n_{p0}}{\tau_n} \quad \bullet(2.9.9)\bullet$$

$$\frac{\partial p(x,t)}{\partial t} = D_p \frac{\partial^2 p_n(x,t)}{\partial x^2} - \frac{p_n(x,t) - p_{n0}}{\tau_p} \quad \bullet(2.9.10)\bullet$$

$$0 = D_n \frac{d^2 n_p(x)}{dx^2} - \frac{n_p(x) - n_{p0}}{\tau_n} \quad \bullet(2.9.11)\bullet$$

$$0 = D_p \frac{d^2 p_n(x)}{dx^2} - \frac{p_n(x) - p_{n0}}{\tau_p} \quad \bullet(2.9.12)\bullet$$

$$n_p(x \leq -x_p) = n_{p0} + C e^{-(x+x_p)/L_p} + D e^{(x+x_p)/L_p} \quad \bullet(2.9.13)\bullet$$

$$p_n(x \geq x_n) = p_{n0} + A e^{-(x-x_n)/L_p} + B e^{(x-x_n)/L_p} \quad \bullet(2.9.14)\bullet$$

$$L_n = \sqrt{D_n \tau_n} \quad \bullet(2.9.15)\bullet$$

$$L_p = \sqrt{D_p \tau_p} \quad \bullet(2.9.16)\bullet$$

$$n = n_0 + \delta n \quad \bullet(2.9.17)\bullet$$

$$p = p_0 + \delta p \quad \bullet(2.9.18)\bullet$$

$$0 = \frac{d^2(\delta n_p)}{dx^2} - \frac{\delta n_p}{L_n^2} \quad \bullet(2.9.19)\bullet$$

$$0 = \frac{d^2(\delta p_n)}{dx^2} - \frac{\delta p_n}{L_p^2} \quad \bullet(2.9.20)\bullet$$

$$\rho = q(p - n + N_d^+ - N_a^-) \quad \bullet(2.10.1)\bullet$$

$$\frac{d\varepsilon}{dx} = \frac{\rho}{\varepsilon} \quad \bullet(2.10.2)\bullet$$

$$\frac{d\phi}{dx} = -\varepsilon \quad \bullet(2.10.3)\bullet$$

$$\frac{dE_i}{dx} = q\varepsilon \quad \bullet(2.10.4)\bullet$$

$$n = n_i e^{(F_n - E_i)/kT} \quad \bullet(2.10.5)\bullet$$

$$p = n_i e^{(E_i - F_p)/kT} \quad \bullet(2.10.6)\bullet$$

□

•(2.10.7)•

□

•(2.10.8)•

$$0 = \frac{1}{q} \frac{\partial J_n}{\partial x} - \frac{np - n_i^2}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} \frac{1}{\tau} \quad \bullet(2.10.9)\bullet$$

$$0 = -\frac{1}{q} \frac{\partial J_p}{\partial x} - \frac{np - n_i^2}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} \frac{1}{\tau} \quad \bullet(2.10.10)$$

Chapter 2: Semiconductor Fundamentals



Problems



1. Calculate the packing density of the body centered cubic, the face centered cubic and the diamond lattice, listed in example 2.1.
2. At what temperature does the energy bandgap of silicon equal exactly 1 eV?
3. Prove that the probability of occupying an energy level below the Fermi energy equals the probability that an energy level above the Fermi energy and equally far away from the Fermi energy is not occupied.
4. At what energy (in units of kT) is the Fermi function within 1 % of the Maxwell-Boltzmann distribution function? What is the corresponding probability of occupancy?
5. Calculate the Fermi function at 6.5 eV if $E_F = 6.25$ eV and $T = 300$ K. Repeat at $T = 950$ K assuming that the Fermi energy does not change. At what temperature does the probability that an energy level at $E = 5.95$ eV is empty equal 1 %.
6. Calculate the effective density of states for electrons and holes in germanium, silicon and gallium arsenide at room temperature and at 100 °C. Use the effective masses for density of states calculations.
7. Calculate the intrinsic carrier density in germanium, silicon and gallium arsenide at room temperature (300 K). Repeat at 100 °C. Assume that the energy bandgap is independent of temperature and use the room temperature values.
8. Calculate the position of the intrinsic energy level relative to the midgap energy

$$E_{\text{midgap}} = (E_c + E_v)/2$$

in germanium, silicon and gallium arsenide at 300 K. Repeat at $T = 100$ °C.

9. Calculate the electron and hole density in germanium, silicon and gallium arsenide if the Fermi energy is 0.3 eV above the intrinsic energy level. Repeat if the Fermi energy is 0.3 eV below the conduction band edge. Assume that $T = 300$ K.
10. The equations (2.6.34) and (2.6.35) derived in section 2.6 are only valid for non-degenerate semiconductors (i.e. $E_v + 3kT < E_F < E_c - 3kT$). Where exactly in the derivation was the assumption made that the semiconductor is non-degenerate?
11. A silicon wafer contains 10^{16} cm⁻³ electrons. Calculate the hole density and the position of the intrinsic energy and the Fermi energy at 300 K. Draw the corresponding band diagram to scale, indicating the conduction and valence

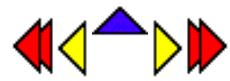
band edge, the intrinsic energy level and the Fermi energy level. Use $n_i = 10^{10} \text{ cm}^{-3}$. 

12. A silicon wafer is doped with 10^{13} cm^{-3} shallow donors and $9 \times 10^{12} \text{ cm}^{-3}$ shallow acceptors. Calculate the electron and hole density at 300 K. Use $n_i = 10^{10} \text{ cm}^{-3}$. 
13. The resistivity of a silicon wafer at room temperature is $5 \Omega\text{cm}$. What is the doping density? Find all possible solutions. 
14. How many phosphorus atoms must be added to decrease the resistivity of n-type silicon at room temperature from $1 \Omega\text{cm}$ to $0.1 \Omega\text{cm}$. Make sure you include the doping dependence of the mobility. State your assumptions. 
15. A piece of n-type silicon ($N_d = 10^{17} \text{ cm}^{-3}$) is uniformly illuminated with green light ($\lambda = 550 \text{ nm}$) so that the power density in the material equals 1 mW/cm^2 . a) Calculate the generation rate of electron-hole pairs using an absorption coefficient of 10^4 cm^{-1} . b) Calculate the excess electron and hole density using the generation rate obtained in (a) and a minority carrier lifetime due to Shockley-Read-Hall recombination of 0.1 ms. c) Calculate the electron and hole quasi-Fermi energies (relative to E_i) based on the excess densities obtained in (b). 
16. A piece of intrinsic silicon is instantaneously heated from 0 K to room temperature (300 K). The minority carrier lifetime due to Shockley-Read-Hall recombination in the material is 1 ms. Calculate the generation rate of electron-hole pairs immediately after reaching room temperature. ($E_t = E_i$). If the generation rate is constant, how long does it take to reach thermal equilibrium? 
17. Calculate the conductivity and resistivity of intrinsic silicon. Use $n_i = 10^{10} \text{ cm}^{-3}$, $\mu_n = 1400 \text{ cm}^2/\text{V}\cdot\text{sec}$ and $\mu_p = 450 \text{ cm}^2/\text{V}\cdot\text{sec}$. 
18. Consider the problem of finding the doping density which results the maximum possible resistivity of silicon at room temperature. ($n_i = 10^{10}$, $\mu_n = 1400 \text{ cm}^2/\text{V}\cdot\text{sec}$ and $\mu_p = 450 \text{ cm}^2/\text{V}\cdot\text{sec}$.)
Should the silicon be doped at all or do you expect the maximum resistivity when dopants are added?
If the silicon should be doped, should it be doped with acceptors or donors (assume that all dopant are shallow).
Calculate the maximum resistivity, the corresponding electron and hole density and the doping density. 
19. The electron density in silicon at room temperature is twice the intrinsic density. Calculate the hole density, the donor density and the Fermi energy relative to the intrinsic energy. Repeat for $n = 5 n_i$ and $n = 10 n_i$. Also repeat for $p = 2 n_i$, $p = 5 n_i$ and $p = 10 n_i$, calculating the electron and acceptor density as well as the Fermi energy relative to the intrinsic energy level. 
20. The expression for the Bohr radius can also be applied to the hydrogen-like atom consisting of an ionized donor and the electron provided by the donor. Modify the expression for the Bohr radius so that it applies to this hydrogen-like atom. Calculate the Bohr radius of an electron orbiting around the ionized donor in silicon. ($\epsilon_r = 11.9$ and $m_e^* = 0.26 m_0$) 
21. Calculate the density of electrons per unit energy (in electron volt) and per unit area (per cubic centimeter) at 1 eV

- above the band minimum. Assume that $m_e^* = 1.08 m_0$. 
22. Calculate the probability that an electron occupies an energy level which is $3kT$ below the Fermi energy. Repeat for an energy level which is $3kT$ above the Fermi energy. 
23. Calculate and plot as a function of energy the product of the probability that an energy level is occupied with the probability that that same energy level is not occupied. Assume that the Fermi energy is zero and that $kT = 1$ eV 
24. The effective mass of electrons in silicon is $0.26 m_0$ and the effective mass of holes is $0.36 m_0$. If the scattering time is the same for both carrier types, what is the ratio of the electron mobility and the hole mobility. 
25. Electrons in silicon carbide have a mobility of $1000 \text{ cm}^2/\text{V}\cdot\text{sec}$. At what value of the electric field do the electrons reach a velocity of $3 \times 10^7 \text{ cm/s}$? Assume that the mobility is constant and independent of the electric field. What voltage is required to obtain this field in a 5 micron thick region? How much time do the electrons need to cross the 5 micron thick region? 
26. A piece of silicon has a resistivity which is specified by the manufacturer to be between 2 and 5 Ohm cm. Assuming that the mobility of electrons is $1400 \text{ cm}^2/\text{V}\cdot\text{sec}$ and that of holes is $450 \text{ cm}^2/\text{V}\cdot\text{sec}$, what is the minimum possible carrier density and what is the corresponding carrier type? Repeat for the maximum possible carrier density. 
27. A silicon wafer has a 2 inch diameter and contains 10^{14} cm^{-3} electrons with a mobility of $1400 \text{ cm}^2/\text{V}\cdot\text{sec}$. How thick should the wafer be so that the resistance between the front and back surface equals 0.1 Ohm. 
28. The electron mobility in germanium is $1000 \text{ cm}^2/\text{V}\cdot\text{sec}$. If this mobility is due to impurity and lattice scattering and the mobility due to lattice scattering only is $1900 \text{ cm}^2/\text{V}\cdot\text{sec}$, what is the mobility due to impurity scattering only? 
29. A piece of n-type silicon is doped with 10^{17} cm^{-3} shallow donors. Calculate the density of electrons per unit energy at $kT/2$ above the conduction band edge. $T = 300$ K. Calculate the electron energy for which the density of electrons per unit energy has a maximum. What is the corresponding probability of occupancy at that maximum? 
30. Phosphorous donor atoms with a concentration of 10^{16} cm^{-3} are added to a piece of silicon. Assume that the phosphorous atoms are distributed homogeneously throughout the silicon. The atomic weight of phosphorous is 31.
 - What is the sample resistivity at 300 K?
 - What proportion by weight does the donor impurity comprise? The density of silicon is 2.33 gram/cm³
 - If 10^{17} atoms cm⁻³ of boron are included in addition to phosphorous, and distributed uniformly, what is the resulting resistivity and type (i.e., p- or n-type material)?
 - Sketch the energy-band diagram under the condition of part c) and show the position of the Fermi energy relative to the valence band edge.
31. Find the equilibrium electron and hole concentrations and the location of the Fermi energy relative to the intrinsic energy in silicon at 27 °C, if the silicon contains the following concentrations of shallow dopants.
 - $1 \times 10^{16} \text{ cm}^{-3}$ boron atoms
 - $3 \times 10^{16} \text{ cm}^{-3}$ arsenic atoms and $2.9 \times 10^{16} \text{ cm}^{-3}$ boron atoms.

32. The electron concentration in a piece of lightly doped, n-type silicon at room temperature varies linearly from 10^{17} cm^{-3} at $x = 0$ to 6×10^{16} cm^{-3} at $x = 2 \mu\text{m}$. Electrons are supplied to keep this concentration constant with time. Calculate the electron current density in the silicon if no electric field is present. Assume $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $T = 300 \text{ K}$.

Chapter 3: Metal-Semiconductor Junctions



Examples

- Example 3.1** Consider a chrome-silicon metal-semiconductor junction with $N_d = 10^{17} \text{ cm}^{-3}$. Calculate the barrier height and the built-in potential. Repeat for a p-type semiconductor with the same doping density.
- Example 3.2** Consider a chrome-silicon metal-semiconductor junction with $N_d = 10^{17} \text{ cm}^{-3}$. Calculate the depletion layer width, the electric field in the silicon at the metal-semiconductor interface, the potential across the semiconductor and the capacitance per unit area for an applied voltage of -5 V.

Example 2.1 Calculate the maximum fraction of the volume in a simple cubic crystal occupied by the atoms. Assume that the atoms are closely packed and that they can be treated as hard spheres. This fraction is also called the packing density.

Solution The atoms in a simple cubic crystal are located at the corners of the units cell, a cube with side a . Adjacent atoms touch each other so that the radius of each atom equals $a/2$. There are eight atoms occupying the corners of the cube, but only one eighth of each is within the unit cell so that the number of atoms equals one per unit cell. The packing density is then obtained from:

$$\frac{\text{Volume of atoms}}{\text{Volume of the unit cell}} = \frac{\frac{4}{3} \mathbf{p} r^3}{a^3} = \frac{\frac{4}{3} \mathbf{p} \left(\frac{a}{2}\right)^3}{a^3} = \frac{\mathbf{p}}{6} = 52\%$$

or about half the volume of the unit cell is occupied by the atoms. The packing density of four cubic crystals is listed in the table below.

	Radius	Atoms/ unit cell	Packing density
Simple cubic	$\frac{a}{2}$	1	$\frac{\mathbf{p}}{6} = 52\%$
Body centered cubic	$\frac{\sqrt{3}a}{4}$	2	$\frac{\mathbf{p}\sqrt{3}}{8} = 68\%$
Face centered cubic	$\frac{\sqrt{2}a}{4}$	4	$\frac{\mathbf{p}\sqrt{2}}{6} = 74\%$
Diamond	$\frac{\sqrt{3}a}{8}$	8	$\frac{\mathbf{p}\sqrt{3}}{16} = 34\%$

Example 2.2 Calculate the energy bandgap of germanium, silicon and gallium arsenide at 300, 400, 500 and 600 K.

Solution The bandgap of silicon at 300 K equals:

$$E_g(300 \text{ K}) = E_g(0 \text{ K}) - \frac{aT^2}{T + b}$$
$$= 1.166 - \frac{0.473 \times 10^{-3} \times (300)^2}{300 + 636} = 1.12 \text{ eV}$$

Similarly one finds the energy bandgap for germanium and gallium arsenide, as well as at different temperatures, yielding:

	Germanium	Silicon	Gallium Arsenide
$T = 300 \text{ K}$	0.66 eV	1.12 eV	1.42 eV
$T = 400 \text{ K}$	0.62 eV	1.09 eV	1.38 eV
$T = 500 \text{ K}$	0.58 eV	1.06 eV	1.33 eV
$T = 600 \text{ K}$	0.54 eV	1.03 eV	1.28 eV

Example 2.3 Calculate the number of states per unit energy in a 100 by 100 by 10 nm piece of silicon ($m^* = 1.08 m_0$) 100 meV above the conduction band edge. Write the result in units of eV⁻¹.

Solution The density of states equals:

$$\begin{aligned}g(E) &= \frac{8\mathbf{p}\sqrt{2}}{h^3} m^{*3/2} \sqrt{E - E_c} \\&= \frac{8\mathbf{p}\sqrt{2} (1.08 \times 9.1 \times 10^{-31})^{3/2}}{(6.626 \times 10^{-34})^3} \sqrt{0.1 \times 1.6 \times 10^{-19}} \\&= 1.51 \times 10^{56} \text{ m}^{-3} \text{J}^{-1}\end{aligned}$$

So that the total number of states per unit energy equals:

$$g(E)V = 1.51 \times 10^{56} \times 10^{-22} \text{ J}^{-1} = 2.41 \times 10^5 \text{ eV}^{-1}$$

Example 2.4 Calculate the effective densities of states in the conduction and valence bands of germanium, silicon and gallium arsenide at 300 K.

Solution The effective density of states in the conduction band of germanium equals:

$$N_c = 2 \left(\frac{2p m_e^* k T}{h^2} \right)^{3/2}$$
$$= 2 \left(\frac{2p 0.55 \times 9.11 \times 10^{-31} \times 1.38 \times 10^{-23} \times 300}{(6.626 \times 10^{-34})^2} \right)^{3/2}$$
$$= 1.02 \times 10^{25} \text{ m}^{-3} = 1.02 \times 10^{19} \text{ cm}^{-3}$$

where the effective mass for density of states was used (see appendix 3 or section 2.3.6). Similarly one finds the effective density of states in the conduction band for other semiconductors and the effective density of states in the valence band:

	Germanium	Silicon	Gallium Arsenide
$N_c (\text{cm}^{-3})$	1.02×10^{19}	2.81×10^{19}	4.35×10^{17}
$N_v (\text{cm}^{-3})$	5.64×10^{18}	1.83×10^{19}	7.57×10^{18}

Note that the effective density of states is temperature dependent and can be obtain from:

$$N_c (T) = N_c (300 \text{ K}) \left(\frac{T}{300} \right)^{3/2}$$

where $N_c (300 \text{ K})$ is the effective density of states at 300 K.

Example
2.4b
Solution

Calculate the intrinsic carrier density in germanium, silicon and gallium arsenide at 300, 400, 500 and 600 K.

The intrinsic carrier density in silicon at 300 K equals:

$$\begin{aligned}n_i(300 \text{ K}) &= \sqrt{N_c N_v} \exp\left(\frac{-E_g}{2kT}\right) \\&= \sqrt{2.81 \times 10^{19} \times 1.83 \times 10^{19}} \exp\left(\frac{-1.12}{2 \times 0.0258}\right) \\&= 8.72 \times 10^9 \text{ cm}^{-3}\end{aligned}$$

Similarly one finds the intrinsic carrier density for germanium and gallium arsenide at different temperatures, yielding:

	Germanium	Silicon	Gallium Arsenide
300 K	2.02×10^{13}	8.72×10^9	2.03×10^6
400 K	1.38×10^{15}	4.52×10^{12}	5.98×10^9
500 K	1.91×10^{16}	2.16×10^{14}	7.98×10^{11}
600 K	1.18×10^{17}	3.07×10^{15}	2.22×10^{13}

Example 2.5 Calculate the ionization energy for shallow donors and acceptors in germanium and silicon using the hydrogen-like model.

Solution Using the effective mass for conductivity calculations (Appendix 3) one finds the ionization energy for shallow donors in germanium to be:

$$E_c - E_d = 13.6 \frac{m_{cond}^*}{m_0 e_r^2} \text{ eV} = 13.6 \frac{0.12}{16^2} \text{ eV} = 6.4 \text{ meV}$$

The calculated ionization energies for donors and acceptors in germanium and silicon are provided below.

	Germanium	Silicon
donors	6.4 meV	13.8 meV
acceptors	11.2 meV	20.5 meV

Note that the actual ionization energies differ from this value and depend on the actual donor atom.

Example 2.6a A germanium wafer is doped with a shallow donor density of $3n_i/2$. Calculate the electron and hole density.

Solution The electron density is obtained from equation (2.6.34)

$$\begin{aligned}n_o &= \frac{N_d^+ - N_a^-}{2} + \sqrt{\left(\frac{N_d^+ - N_a^-}{2}\right)^2 + n_i^2} \\&= n_i \left(\frac{3}{4} + \sqrt{\frac{9}{16} + 1} \right) = 2n_i\end{aligned}$$

and the hole density is obtained using the mass action law:

$$p_o = \frac{n_i^2}{n_o} = \frac{n_i}{2}$$

Example 2.6b A silicon wafer is doped with a shallow acceptor doping of 10^{16} cm^{-3} . Calculate the electron and hole density.

Solution Since the acceptor doping is much larger than the intrinsic density and much smaller than the effective density of states, the hole density equals:

$$p_o \approx N_a^+ = 10^{16} \text{ cm}^{-3}$$

The electron density is then obtained using the mass action law

$$n_o \approx \frac{n_i^2}{N_a^+} = \frac{10^{20}}{10^{16}} = 10^4 \text{ cm}^{-3}$$

The approach described in example 2.6a yields the same result.

Example 2.7 A piece of germanium doped with 10^{16} cm^{-3} shallow donors is illuminated with light generating 10^{15} cm^{-3} excess electrons and holes. Calculate the quasi-Fermi energies relative to the intrinsic energy and compare it to the Fermi energy in the absence of illumination.

Solution The carrier densities when illuminating the semiconductor are:

$$n = n_o + d n = 10^{16} + 10^{15} = 1.1 \times 10^{16} \text{ cm}^{-3}$$

$$p = p_o + d p \approx 10^{15} \text{ cm}^{-3}$$

and the quasi-Fermi energies are:

$$F_n - E_i = kT \ln \frac{n}{n_i} = 0.0259 \times \ln \frac{1.1 \times 10^{16}}{2 \times 10^{13}} = 163 \text{ meV}$$

$$F_p - E_i = -kT \ln \frac{p}{n_i} = 0.0259 \times \ln \frac{1 \times 10^{15}}{2 \times 10^{13}} = -101 \text{ meV}$$

In comparison, the Fermi energy in the absence of light equals

$$E_F - E_i = kT \ln \frac{n_o}{n_i} = 0.0259 \times \ln \frac{10^{16}}{2 \times 10^{13}} = 161 \text{ meV}$$

which is very close to the quasi-Fermi energy of the majority carriers.

Example 2.8 Electrons in undoped gallium arsenide have a mobility of 8,800 cm²/V-s. Calculate the average time between collisions.

Calculate the distance traveled between two collisions (also called the mean free path). Use an average velocity of 10⁷ cm/s.

Solution The collision time, t_c , is obtained from:

$$t_c = \frac{m_n m_e^*}{q} = \frac{0.88 \times 0.067 \times 9.1 \times 10^{-31}}{1.6 \times 10^{-19}} = 0.34 \text{ ps}$$

where the mobility was first converted to MKS units.

The mean free path, l , equals:

$$l = v_{average} t_c = 10^7 \times 0.34 \times 10^{-12} = 34 \text{ nm}$$

Example 2.9 A piece of silicon doped with arsenic ($N_d = 10^{17} \text{ cm}^{-3}$) is 100 μm long, 10 μm wide and 1 μm thick. Calculate the resistance of this sample when contacted one each end.

Solution The resistivity of the silicon equals:

$$r = \frac{1}{qn\mu} = \frac{1}{1.6 \times 10^{-19} \times 10^{17} \times 727} = 0.086 \Omega\text{cm}$$

where the mobility was obtained from Table 2.7.3.

The resistance then equals:

$$R = r \frac{L}{Wt} = 0.086 \times \frac{100 \times 10^{-4}}{10 \times 10^{-4} \times 10^{-4}} = 8.6 \text{ k}\Omega$$

An alternate approach is to first calculate the sheet resistance, R_s :

$$R_s = \frac{r}{t} = \frac{0.086}{10^{-4}} = 860 \Omega/\text{square}$$

From which one then obtains the resistance:

$$R = R_s \frac{L}{W} = 860 \times \frac{100 \times 10^{-4}}{10 \times 10^{-4}} = 8.6 \text{ k}\Omega$$

Example 2.10 The hole density in an n-type silicon wafer ($N_d = 10^{17} \text{ cm}^{-3}$) decreases linearly from 10^{14} cm^{-3} to 10^{13} cm^{-3} between $x = 0$ and $x = 1 \mu\text{m}$. Calculate the hole diffusion current density.

Solution The hole diffusion current density equals:

$$J_p = qD_p \frac{dp}{dx} = 1.6 \times 10^{-19} \times 8.2 \times \frac{9 \times 10^{13}}{10^{-4}} = 1.18 \text{ A/cm}^2$$

where the diffusion constant was calculated using the Einstein relation:

$$D_p = V_t m_p = 0.0259 \times 317 = 8.2 \text{ cm}^2/\text{s}$$

and the hole mobility in the n-type wafer was obtained from Table 2.7.3 as the hole mobility in a p-type material with the same doping density.

Example 2.11 Calculate the electron and hole densities in an n-type silicon wafer ($N_d = 10^{17} \text{ cm}^{-3}$) illuminated uniformly with 10 mW/cm^2 of red light ($E_{ph} = 1.8 \text{ eV}$). The absorption coefficient of red light in silicon is 10^3 cm^{-1} . The minority carrier lifetime is $10 \mu\text{s}$.

Solution The generation rate of electrons and holes equals:

$$G_n = G_p = \alpha \frac{P_{opt}}{E_{ph} A} = 10^{-3} \frac{10^{-2}}{1.8 \times 10^{-19}} = 3.5 \times 10^{13} \text{ cm}^{-3}\text{s}^{-1}$$

where the photon energy was converted into Joules. The excess carrier densities are then obtained from:

$$dn = dp = t_p G_p = 10 \times 10^{-6} \times 3.5 \times 10^{13} = 3.5 \times 10^8 \text{ cm}^{-3}$$

So that the electron and hole densities equal:

$$n = n_o + dn = 10^{17} + 3.5 \times 10^8 = 10^{17} \text{ cm}^{-3}$$

$$p = \frac{n_i^2}{n_o} + dp = \frac{(10^{10})^2}{10^{17}} + 3.5 \times 10^8 = 3.5 \times 10^8 \text{ cm}^{-3}$$

Chapter 2: Semiconductor Fundamentals



2.2. Crystals and crystal structures

2.2.1. Bravais lattices

2.2.2. Common semiconductor crystal structures

2.2.3. Growth of semiconductor crystals

Solid materials are classified by the way the atoms are arranged within the solid. Materials in which atoms are placed randomly are called amorphous. Materials in which atoms are placed in a high ordered structure are called crystalline. Poly-crystalline materials are materials with a high degree of short-range order and no long-range order. These materials consist of small crystalline regions with random orientation called grains, separated by grain boundaries.

Of primary interest in this text are crystalline semiconductors in which atoms are placed in a highly ordered structure. Crystals are categorized by their crystal structure and the underlying lattice. While some crystals have a single atom placed at each lattice point, most crystals have a combination of atoms associated with each lattice point. This combination of atoms is also called the basis.

The classification of lattices, the common semiconductor crystal structures and the growth of single-crystal semiconductors are discussed in the following sections.

2.2.1 Bravais lattices



The Bravais lattices are the distinct lattice types, which when repeated can fill the whole space. The lattice can therefore be generated by three unit vectors, \vec{a}_1, \vec{a}_2 , and \vec{a}_3 and a set of integers k, l and m so that each lattice point, identified by a vector \vec{r} , can be obtained from:

$$\vec{r} = k \vec{a}_1 + l \vec{a}_2 + m \vec{a}_3 \quad (2.2.1)$$

The construction of the lattice points based on a set of unit vectors is illustrated by Figure 2.2.1.

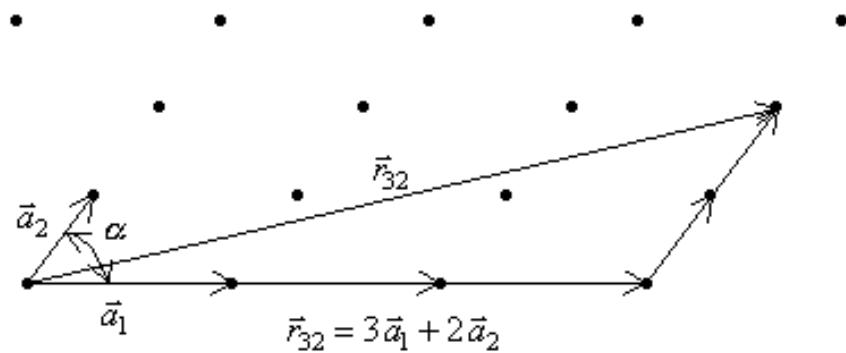
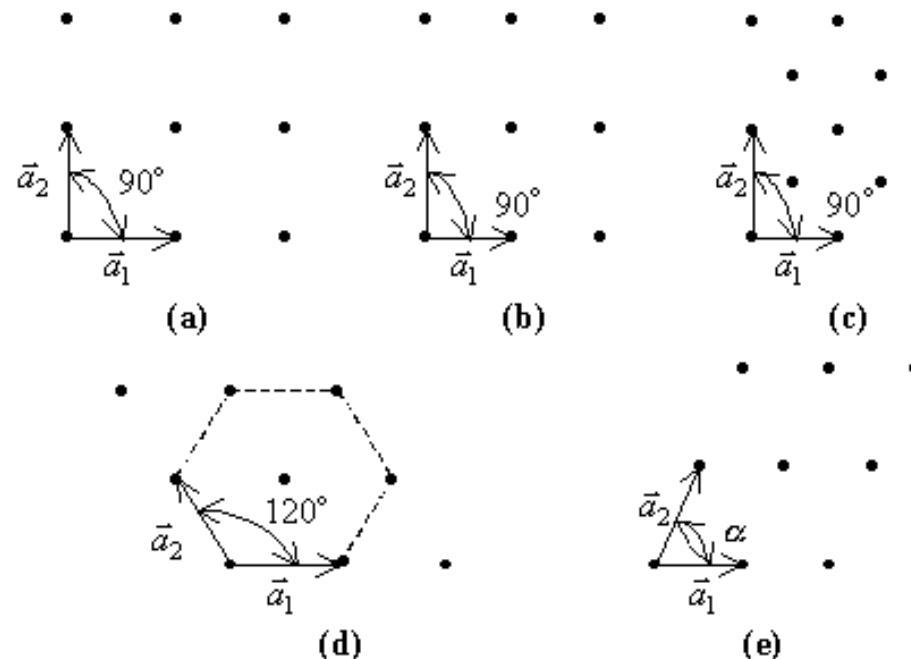


Figure 2.2.1:

The construction of lattice points using unit vectors

In two dimensions, there are five distinct Bravais lattices, while in three dimensions there are fourteen. The lattices in two dimensions are the square lattice, the rectangular lattice, the centered rectangular lattice, the hexagonal lattice and the oblique lattice as shown in Figure 2.2.2. It is customary to organize these lattices in groups which have the same symmetry. An example is the rectangular and the centered rectangular lattice. As can be seen on the figure, all the lattice points of the rectangular lattice can be obtained by a combination of the lattice vectors. The centered rectangular lattice can be constructed in two ways. It can be obtained by starting with the same lattice vectors as those of the rectangular lattice and then adding an additional atom at the center of each rectangle in the lattice. This approach is illustrated by Figure 2.2.2 c). The lattice vectors generate the traditional unit cell and the center atom is obtained by attaching two lattice points to every lattice point of the traditional unit cell. The alternate approach is to define a new set of lattice vectors, one identical to and another starting from the same origin and ending on the center atom. These lattice vectors generate the so-called primitive cell and directly define the centered rectangular lattice.

**Figure 2.2.2.:** The five Bravais lattices of two-dimensional crystals: (a) cubic, (b) rectangular, (c) centered rectangular, (d) hexagonal and (e) oblique

These lattices are listed in Table 2.2.1. a_1 and a_2 are the magnitudes of the unit vectors and α is the angle between them.

Name	Number of Bravais lattices	Conditions
Square	1	$a_1 = a_2, \alpha = 90^\circ$
Rectangular	2	$a_1 \neq a_2, \alpha = 90^\circ$
Hexagonal	1	$a_1 = a_2, \alpha = 120^\circ$
Oblique	1	$a_1 \neq a_2, \alpha \neq 120^\circ, \alpha \neq 90^\circ$

Table 2.2.1.:

Bravais lattices of two-dimensional crystals

The same approach is used for lattices in three dimensions. The fourteen lattices of three-dimensional crystals are classified as shown in Table 2.2.2, where a_1 , a_2 and a_3 are the magnitudes of the unit vectors defining the traditional unit cell and α , β and γ are the angles between these unit vectors.

Name	Number of Bravais lattices	Conditions
Triclinic	1	$a_1 \neq a_2 \neq a_3, \alpha \neq \beta \neq \gamma$
Monoclinic	2	$a_1 \neq a_2 \neq a_3, \alpha = \beta = 90^\circ \neq \gamma$
Orthorhombic	4	$a_1 \neq a_2 \neq a_3, \alpha = \beta = \gamma = 90^\circ$
Tetragonal	2	$a_1 = a_2 \neq a_3, \alpha = \beta = \gamma = 90^\circ$
Cubic	3	$a_1 = a_2 = a_3, \alpha = \beta = \gamma = 90^\circ$
Trigonal	1	$a_1 = a_2 = a_3, \alpha = \beta = \gamma < 120^\circ \neq 90^\circ$
Hexagonal	1	$a_1 = a_2 \neq a_3, \alpha = \beta = 90^\circ, \gamma = 120^\circ$

Table 2.2.2.: Bravais lattices of three-dimensional crystals

The cubic lattices are an important subset of these fourteen Bravais lattices since a large number of semiconductors are cubic. The three cubic Bravais lattices are the simple cubic lattice, the body-centered cubic lattice and the face-centered cubic lattice as shown in Figure 2.2.3. Since all unit vectors identifying the traditional unit cell have the same size, the crystal structure is completely defined by a single number. This number is the lattice constant, a .

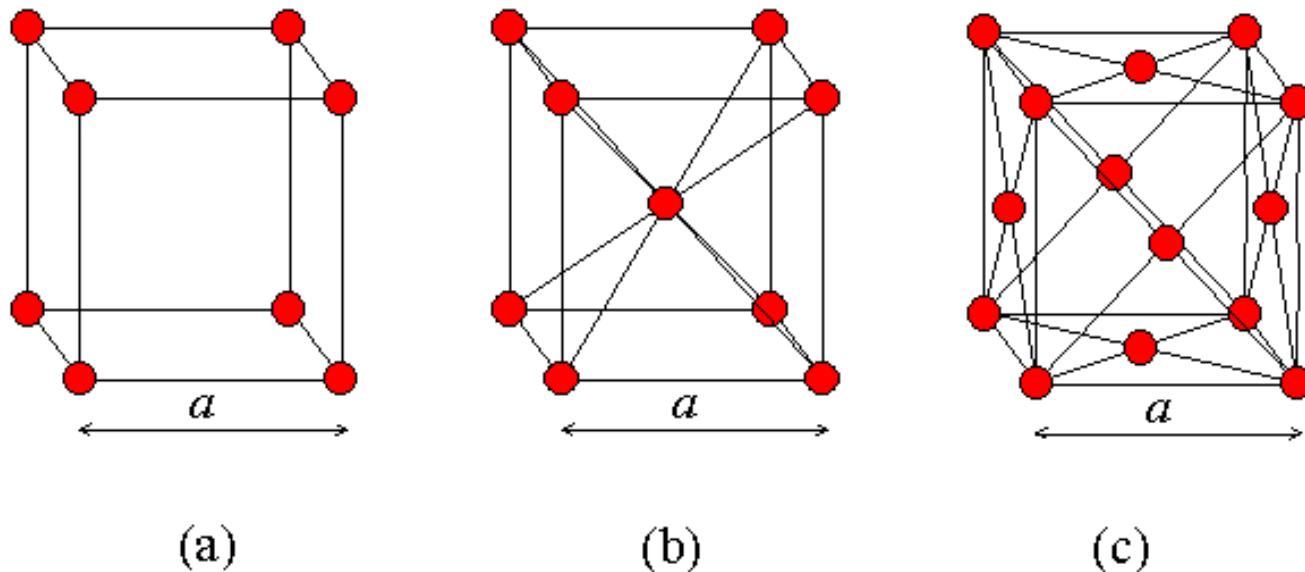
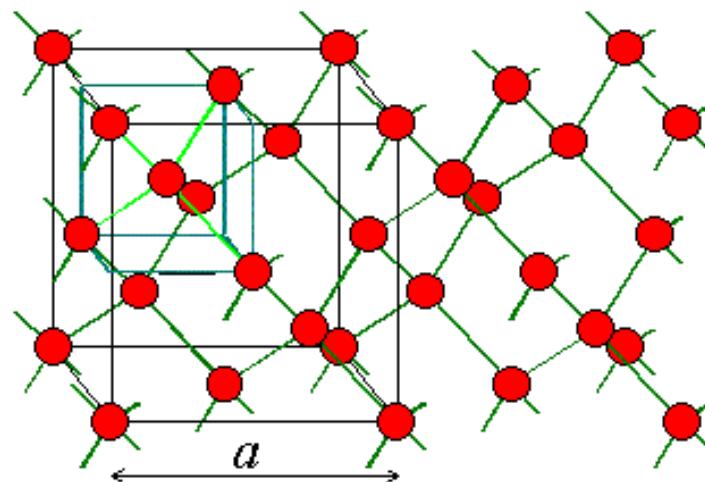


Figure 2.2.3.: The simple cubic (a), the body-centered cubic (b) and the face centered cubic (c) lattice.

2.2.2 Common semiconductor crystal structures



The most common crystal structure among frequently used semiconductors is the diamond lattice, shown in Figure 2.2.4. Each atom in the diamond lattice has a covalent bond with four adjacent atoms, which together form a tetrahedron. This lattice can also be formed from two face-centered-cubic lattices, which are displaced along the body diagonal of the larger cube in Figure 2.2.4 by one quarter of that body diagonal. The diamond lattice therefore is a face-centered-cubic lattice with a basis containing two identical atoms.

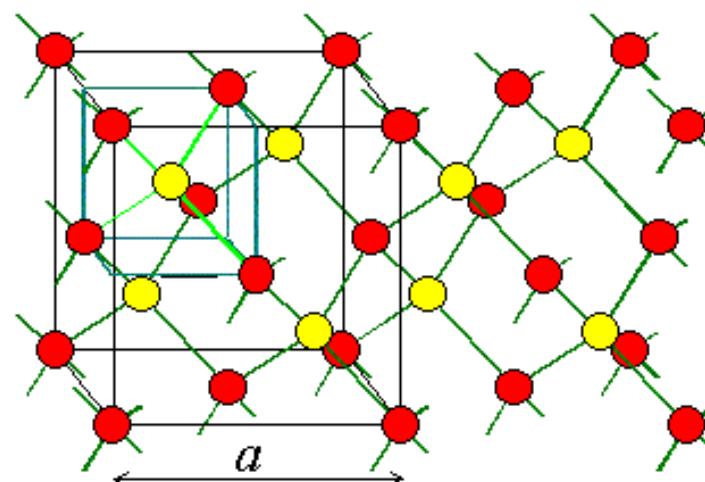
**Figure 2.2.4.:**

The diamond lattice of silicon and germanium

Compound semiconductors such as GaAs and InP have a crystal structure that is similar to that of diamond. However, the lattice contains two different types of atoms. Each atom still has four covalent bonds, but they are bonds with atoms of the other type. This structure is referred to as the zinc-blende lattice, named after zinc-blende (ZnS) as shown in Figure 2.2.5. Both the diamond lattice and the zinc-blende lattice are cubic lattices. A third common crystal structure is the hexagonal structure also referred to as the wurzite crystal structure, which is the hexagonal form of zinc sulfide (ZnS).

Many semiconductor materials can have more than one crystal structure. A large number of compound semiconductors including GaAs, GaN and ZnS can be either cubic or hexagonal. SiC can be cubic or one of several different hexagonal crystal structures.

The cubic crystals are characterized by a single parameter, the lattice constant a , while the hexagonal structures are characterized in the hexagonal plane by a lattice constant a and by the distance between the hexagonal planes, c .

**Figure 2.2.5 :**

The zinc-blende crystal structure of GaAs and InP

Example 2.1

Calculate the maximum fraction of the volume in a simple cubic crystal occupied by the atoms. Assume that the atoms are closely packed and that they can be treated as hard spheres. This fraction is also called the packing density.

Solution

The atoms in a simple cubic crystal are located at the corners of the units cell, a cube with side a . Adjacent atoms touch each other so that the radius of each atom equals $a/2$. There are eight atoms occupying the corners of the cube, but only one eighth of each is within the unit cell so that the number of atoms equals one per unit cell. The packing density is then obtained from:

$$\frac{\text{Volume of atoms}}{\text{Volume of the unit cell}} = \frac{\frac{4}{3} \pi r^3}{a^3} = \frac{\frac{4}{3} \pi (\frac{a}{2})^3}{a^3} = \frac{\pi}{6} = 52\%$$

or about half the volume of the unit cell is occupied by the atoms.

The packing density of four cubic crystals is listed in the table below.

	Radius	Atoms/ unit cell	Packing density
Simple cubic	$\frac{a}{2}$	1	$\frac{\pi}{6} = 52\%$
Body centered cubic	$\frac{\sqrt{3}a}{4}$	2	$\frac{\pi\sqrt{3}}{8} = 68\%$
Face centered cubic	$\frac{\sqrt{2}a}{4}$	4	$\frac{\pi\sqrt{2}}{6} = 74\%$
Diamond	$\frac{\sqrt{3}a}{8}$	8	$\frac{\pi\sqrt{3}}{16} = 34\%$

2.2.3 Growth of semiconductor crystals



Like all crystals, semiconductor crystals can be obtained by cooling the molten semiconductor material. However, this procedure yields poly-crystalline material since crystals start growing in different locations with a different orientation. Instead when growing single-crystalline silicon one starts with a seed crystal and dips one end into the melt. By controlling the temperature difference between the seed crystal and the molten silicon, the seed crystal slowly grows. The result is a large single-crystal silicon boule. Such boules have a cylindrical shape, in part because the seed crystal is rotated during growth and in part because of the cylindrical shape of the crucible containing the melt. The boule is then cut into wafers with a diamond saw and further polished to yield the starting material for silicon device fabrication.

Chapter 2: Semiconductor Fundamentals



2.3 Energy bands

- [2.3.1. Free electron model](#)
- [2.3.2. Periodic potentials](#)
- [2.3.3. Energy bands of semiconductors](#)
- [2.3.4. Metals, insulators and semiconductors](#)
- [2.3.5. Electrons and holes in semiconductors](#)
- [2.3.6. The effective mass concept](#)

Energy bands consisting of a large number of closely spaced energy levels exist in crystalline materials. The bands can be thought of as the collection of the individual energy levels of electrons surrounding each atom. The wavefunctions of the individual electrons, however, overlap with those of electrons confined to neighboring atoms. The Pauli exclusion principle does not allow the electron energy levels to be the same so that one obtains a set of closely spaced energy levels, forming an energy band. The energy band model is crucial to any detailed treatment of semiconductor devices. It provides the framework needed to understand the concept of an energy bandgap and that of conduction in an almost filled band as described by the empty states.

2.3.1 Free electron model



The free electron model of metals has been used to explain the photo-electric effect (see section [1.2.2](#)). This model assumes that electrons are free to move within the metal but are confined to the metal by potential barriers as illustrated by Figure [2.3.1](#). The minimum energy needed to extract an electron from the metal equals $q\Phi_M$, where Φ_M is the workfunction. This model is frequently used when analyzing metals. However, this model does not work well for semiconductors since the effect of the periodic potential due to the atoms in the crystal has been ignored.

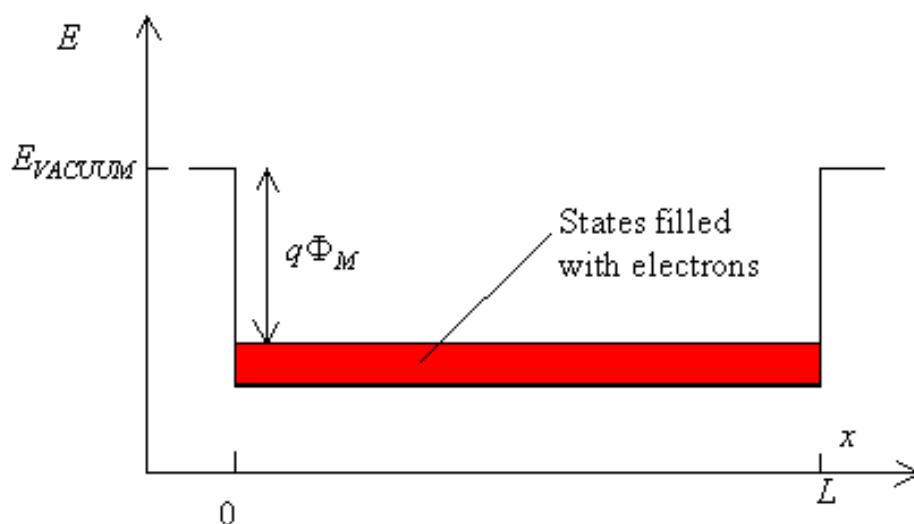


Figure 2.3.1.:

The free electron model of a metal.



2.3.2 Periodic potentials

The analysis of periodic potentials is required to find the energy levels in a semiconductor. This requires the use of periodic wave functions, called Bloch functions which are beyond the scope of this text. The result of this analysis is that the energy levels are grouped in bands, separated by energy band gaps. The behavior of electrons at the top and bottom of such a band is similar to that of a free electron. However, the electrons are affected by the presence of the periodic potential. The combined effect of the periodic potential is included by adjusting the mass of the electron to a different value. This mass will be referred to as the effective mass.

The effect of a periodic arrangement on the electron energy levels is illustrated by Figure 2.3.2. Shown are the energy levels of electrons in a carbon crystal with the atoms arranged in a diamond lattice. These energy levels are plotted as a function of the lattice constant, a .

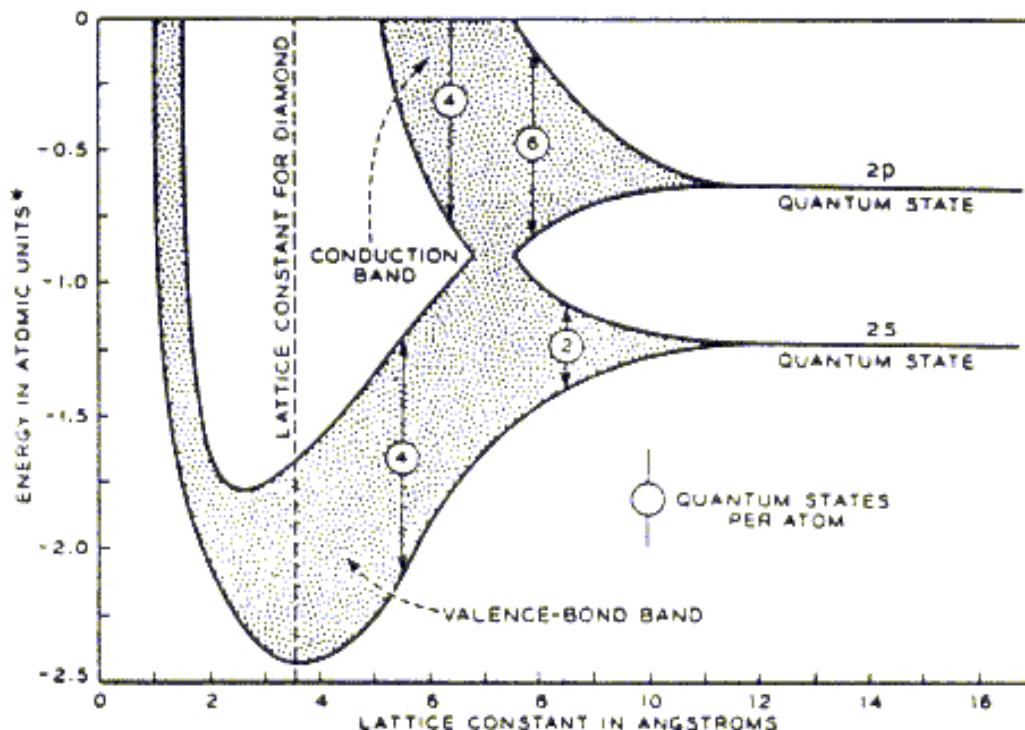


Figure 2.3.2. : Energy bands for diamond versus lattice constant . One atomic unit equals 1 Rydberg = 13.6 eV.

Isolated carbon atoms contain six electrons, which occupy the 1s, 2s and 2p orbital in pairs. The energy of an electron occupying the 2s and 2p orbital is indicated on the figure. The energy of the 1s orbital is not shown. As the lattice constant is reduced, there is an overlap of the electron wavefunctions occupying adjacent atoms. This leads to a splitting of the energy levels consistent with the Pauli exclusion principle. The splitting results in an energy band containing $2N$ states in the 2s band and $6N$ states in the 2p band, where N is the number of atoms in the crystal. A further reduction of the lattice constant causes the 2s and 2p energy bands to merge and split again into two bands containing $4N$ states each. At zero Kelvin, the lower band is completely filled with electrons and labeled as the valence band. The upper band is empty and labeled as the conduction band.



2.3.3 Energy bands of semiconductors

2.3.3.1. Energy band diagrams of common semiconductors

2.3.3.2. Simple energy band diagram of a semiconductor

2.3.3.3. Temperature dependence of the energy bandgap

Complete energy band diagrams of semiconductors are very complex. However, most have features similar to that of the diamond crystal discussed in section 2.3.2. In this section, we first take a closer look at the energy band diagrams of common semiconductors. We then present a simple diagram containing some of the most important feature and discuss the temperature dependence of the energy bandgap.

2.3.3.1. Energy band diagrams of common semiconductors

The energy band diagrams of semiconductors are rather complex. The detailed energy band diagrams of germanium, silicon and gallium arsenide are shown in Figure 2.3.3. The energy is plotted as a function of the wavenumber, k , along the main crystallographic directions in the crystal, since the band diagram depends on the direction in the crystal. The energy band diagrams contain multiple completely-filled and completely-empty bands. In addition, there are multiple partially-filled bands.

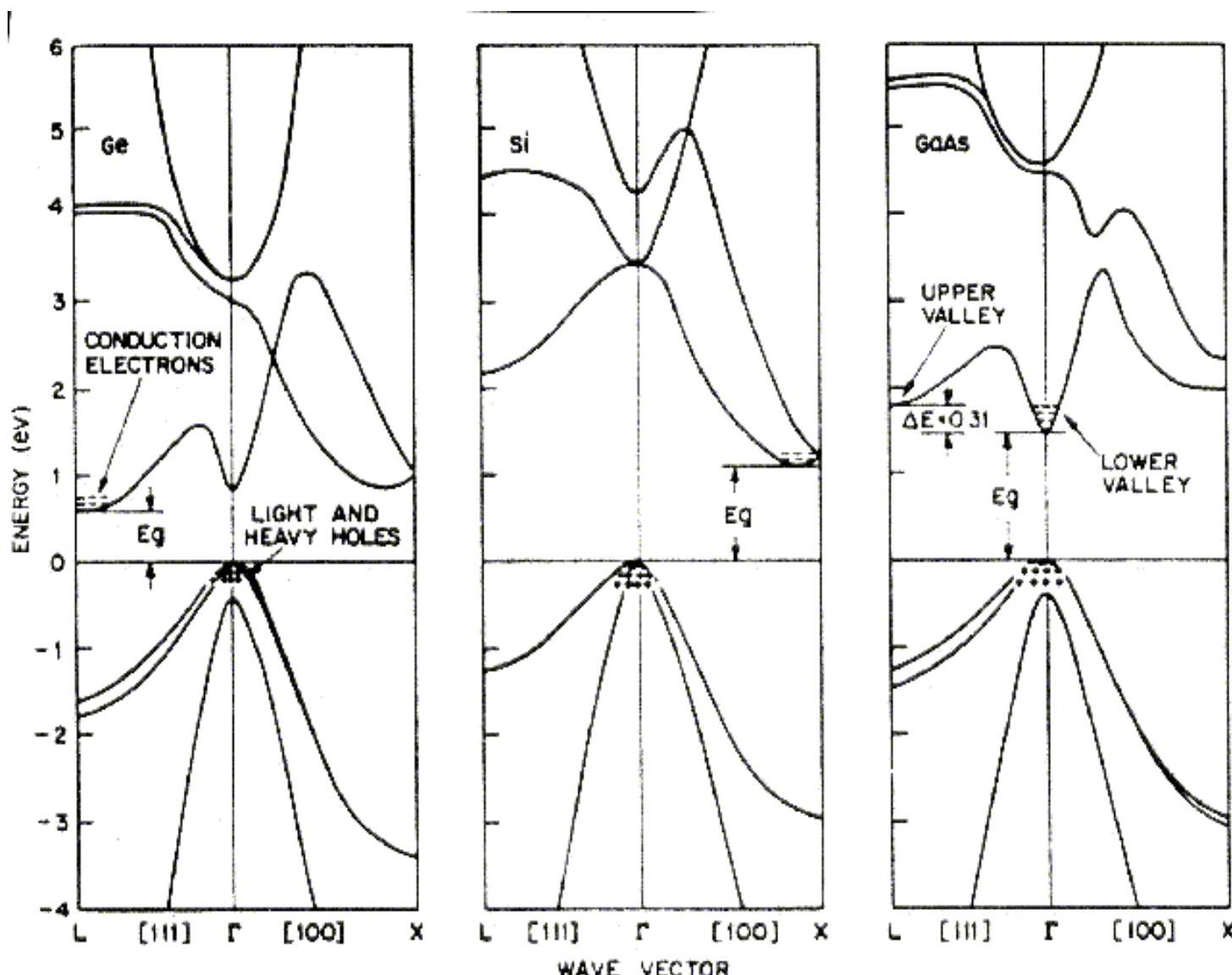


Figure 2.3.3.:

Energy band diagram of (a) germanium, (b) silicon and (c) gallium arsenide

Fortunately, we can simplify the energy band diagram since only the electrons in the highest almost-filled band and the lowest almost-empty band dominate the behavior of the semiconductor. These bands are indicated on the figure by the + and - signs corresponding to the charge of the carriers in those bands.

2.3.3.2. Simple energy band diagram of a semiconductor

The energy band diagrams shown in the previous section are frequently simplified when analyzing semiconductor devices. Since the electronic properties of a semiconductor are dominated by the highest partially empty band and the lowest partially filled band, it is often sufficient to only consider those bands. This leads to a simplified energy band diagram for semiconductors as shown in Figure 2.3.4:

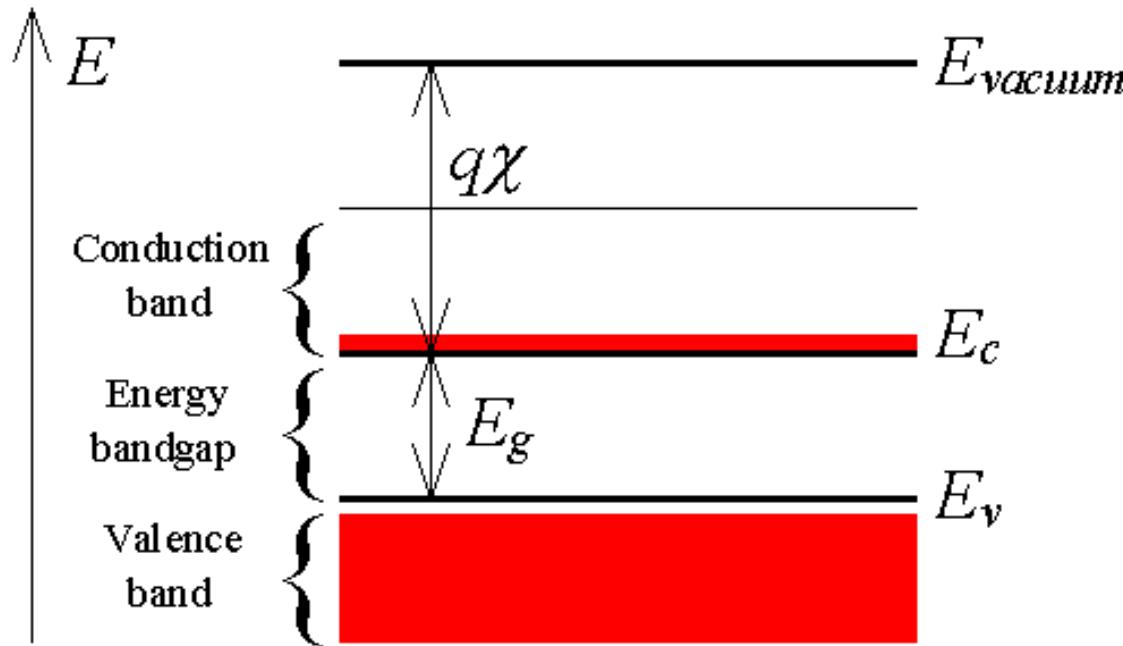


Figure 2.3.4.: A simplified energy band diagram used to describe semiconductors. Shown are the valence and conduction band as indicated by the valence band edge, E_v , and the conduction band edge, E_c . The vacuum level, E_{vacuum} , and the electron affinity, χ , are also indicated on the figure.

The diagram identifies the almost-empty conduction band by a horizontal line. This line indicates the bottom edge of the conduction band and is labeled E_c . Similarly, the top of the valence band is indicated by a horizontal line labeled E_v . The energy bandgap is located between the two lines, which are separated by the bandgap energy E_g . The distance between the conduction band edge, E_c , and the energy of a free electron outside the crystal (called the vacuum level labeled E_{vacuum}) is quantified by the electron affinity, χ multiplied with the electronic charge q .

An important feature of an energy band diagram, which is not included on the simplified diagram, is whether the conduction band minimum and the valence band maximum occur at the same value for the wavenumber. If so, the energy bandgap is called direct. If not, the energy bandgap is called indirect. This distinction is of interest for optoelectronic devices as direct bandgap materials provide more efficient absorption and emission of light. For instance, the smallest bandgap of germanium and silicon is indirect, while gallium arsenide has a direct bandgap as can be seen on Figure 2.3.3.

2.3.3.3. Temperature dependence of the energy bandgap

The energy bandgap of semiconductors tends to decrease as the temperature is increased. This behavior can be better understood if one considers that the interatomic spacing increases when the amplitude of the atomic vibrations increases due to the increased thermal energy. This effect is quantified by the linear expansion coefficient of a material. An increased interatomic spacing decreases the average potential seen by the electrons in the material, which in turn reduces the size of the energy bandgap. A direct modulation of the interatomic distance - such as by applying compressive (tensile) stress - also causes an increase (decrease) of the bandgap.

The temperature dependence of the energy bandgap, E_g , has been experimentally determined yielding the following expression for E_g as a function of the temperature, T :

$$E_g(T) = E_g(0) - \frac{\alpha T^2}{T + \beta} \quad (2.3.1)$$

where $E_g(0)$, α and β are the fitting parameters. These fitting parameters are listed for germanium, silicon and gallium arsenide in Table 2.3.1:

	Germanium	Silicon	GaAs
$E_g(0)$ (eV)	0.7437	1.166	1.519
α (meV/K)	0.477	0.473	0.541
β (K)	235	636	204

Table 2.3.1.: Parameters used to calculate the energy bandgap of germanium, silicon and gallium arsenide (GaAs) as a function of temperature

A plot of the resulting bandgap versus temperature is shown in Figure 2.3.5 for germanium, silicon and gallium arsenide.

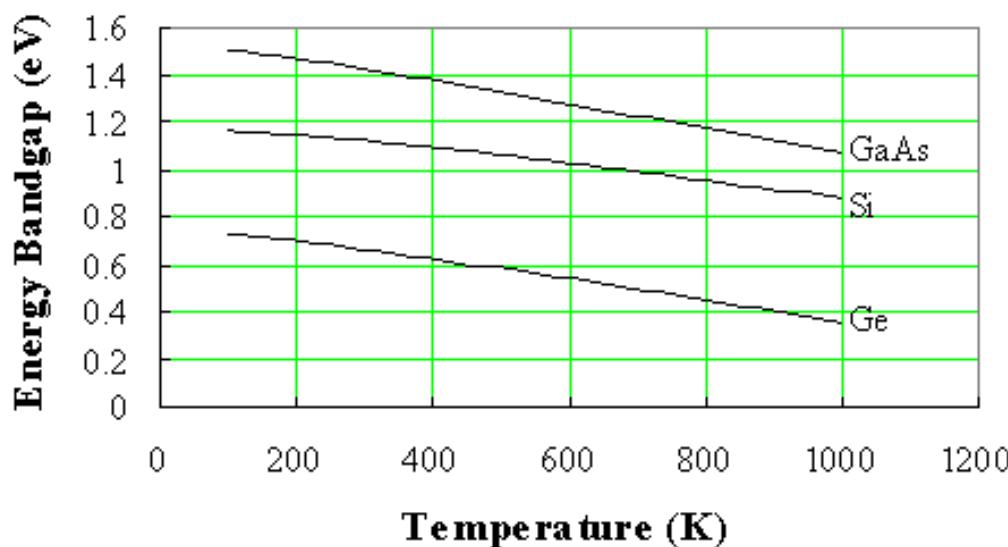


Figure 2.3.5.: Temperature dependence of the energy bandgap of germanium (Ge), silicon (Si) and gallium arsenide (GaAs). 

Example 2.2. Calculate the energy bandgap of germanium, silicon and gallium arsenide at 300, 400, 500 and 600 K.

Solution

The bandgap of silicon at 300 K equals:

$$E_g(300\text{ K}) = E_g(0\text{ K}) - \frac{\alpha T^2}{T + \beta} = 1.166 - \frac{0.473 \times (300)^2}{300 + 636} = 1.12 \text{ eV}$$

Similarly one finds the energy bandgap for germanium and gallium arsenide, as well as at different temperatures, yielding:

	Germanium	Silicon	Gallium Arsenide
$T = 300\text{ K}$	0.66 eV	1.12 eV	1.42 eV
$T = 400\text{ K}$	0.62 eV	1.09 eV	1.38 eV
$T = 500\text{ K}$	0.58 eV	1.06 eV	1.33 eV
$T = 600\text{ K}$	0.54 eV	1.03 eV	1.28 eV

2.3.4 Metals, insulators and semiconductors



Once we know the bandstructure of a given material we still need to find out which energy levels are occupied and whether specific bands are empty, partially filled or completely filled.

Empty bands do not contain electrons. Therefore, they are not expected to contribute to the electrical conductivity of the material. Partially filled bands do contain electrons as well as available energy levels at slightly higher energies. These unoccupied energy levels enable carriers to gain energy when moving in an applied electric field. Electrons in a partially filled band therefore do contribute to the electrical conductivity of the material.

Completely filled bands do contain plenty of electrons but do not contribute to the conductivity of the material. This is because the electrons cannot gain energy since all energy levels are already filled.

In order to find the filled and empty bands we must find out how many electrons can be placed in each band and how many electrons are available. Each band is formed due to the splitting of one or more atomic energy levels. Therefore, the minimum number of states in a band equals twice the number of atoms in the material. The reason for the factor of two is that every energy level can contain two electrons with opposite spin.

To further simplify the analysis, we assume that only the valence electrons (the electrons in the outer shell) are of interest. The core electrons are tightly bound to the atom and are not allowed to freely move in the material.

Four different possible scenarios are shown in Figure [2.3.6](#):

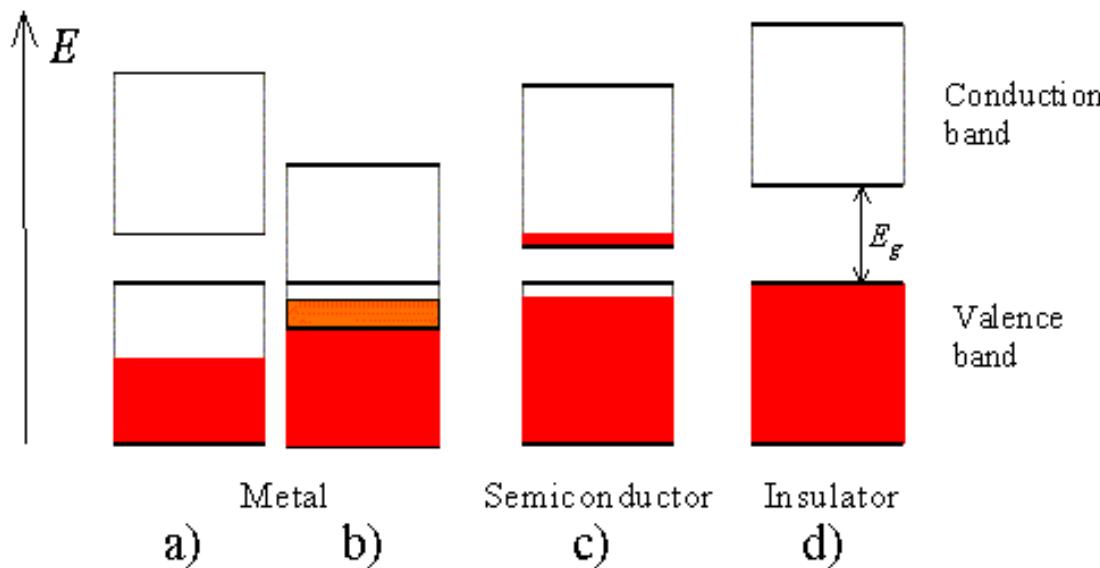


Figure 2.3.6.: Possible energy band diagrams of a crystal. Shown are a) a half filled band, b) two overlapping bands, c) an almost full band separated by a small bandgap from an almost empty band and d) a full band and an empty band separated by a large bandgap.

A half-filled band is shown in Figure 2.3.6 a). This situation occurs in materials consisting of atoms, which contain only one valence electron per atom. Most highly conducting metals including copper, gold and silver satisfy this condition. Materials consisting of atoms that contain two valence electrons can still be highly conducting if the resulting filled band overlaps with an empty band. This scenario is shown in b). No conduction is expected for scenario d) where a completely filled band is separated from the next higher empty band by a larger energy gap. Such materials behave as insulators. Finally, scenario c) depicts the situation in a semiconductor. The completely filled band is now close enough to the next higher empty band that electrons can make it into the next higher band. This yields an almost full band below an almost empty band. We will call the almost full band the valence band since it is occupied by valence electrons. The almost empty band will be called the conduction band, as electrons are free to move in this band and contribute to the conduction of the material.

2.3.5 Electrons and holes in semiconductors



As pointed out in section 2.3.4, semiconductors differ from metals and insulators by the fact that they contain an "almost-empty" conduction band and an "almost-full" valence band. This also means that we will have to deal with the transport of carriers in both bands.

To facilitate the discussion of the transport in the "almost-full" valence band of a semiconductor, we will introduce the concept of holes. It is important for the reader to understand that one could deal with only electrons if one is willing to keep track of all the electrons in the "almost-full" valence band. After all, electrons are the only real particles available in a semiconductor.

The concept of holes is introduced in semiconductors since it is easier to keep track of the missing electrons in an "almost-full" band, rather than keeping track of the actual electrons in that band. We will now first explain the concept of a hole and then point out how the hole concept simplifies the analysis.

Holes are missing electrons. They behave as particles with the same properties as the electrons would have when occupying the same states except that they carry a positive charge. This definition is illustrated further with Figure 2.3.7, which presents the energy band diagram in the presence of an electric field.

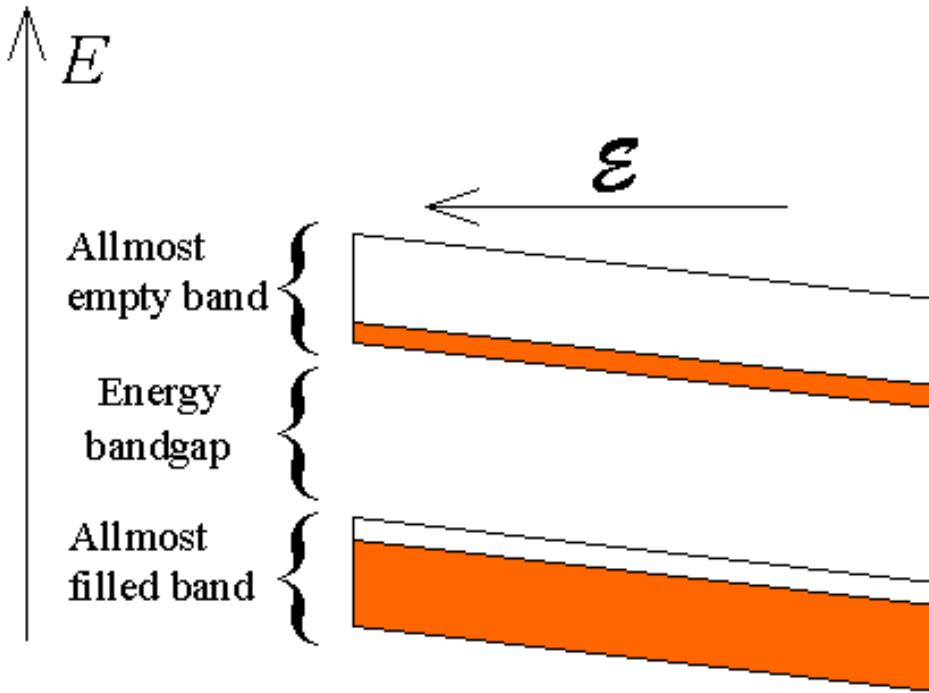


Figure 2.3.7.: Energy band diagram in the presence of a uniform electric field. Shown are the upper almost-empty band and the lower almost-filled band. The tilt of the bands is caused by an externally applied electric field.

A uniform electric field is assumed which causes a constant gradient of the bands.

The electrons in the almost-empty band are negatively charged particles, which therefore move in a direction, which opposes the direction of the field. Electrons therefore move down hill in the upper band. Electrons in the lower band also move in the same direction. The total current density due to the electrons in the valence band can therefore be written as:

$$J_{vb} = \frac{1}{V} \sum_{\text{filled states}} (-q)v_i \quad (2.3.2)$$

where V is the volume of the semiconductor, q is the electronic charge and v is the electron velocity. The sum is taken over all occupied or filled states in the lower band. This equation can be reformulated by first taking the sum over all the states in the lower band and subtracting the current due to the electrons, which are missing in the almost-filled band. This last term therefore represents the sum taken over all the empty states in the lower band, or:

$$J_{vb} = \frac{1}{V} \left(\sum_{\text{all states}} (-q)v_i - \sum_{\text{empty states}} (-q)v_i \right) \quad (2.3.3)$$

The sum over all the states in the lower band has to equal zero since electrons in a completely filled band do not contribute to current, while the remaining term can be written as:

$$J_{vb} = \frac{1}{V} \sum_{\text{empty states}} (+q)v_i \quad (2.3.4)$$

which states that the current is due to positively charged particles associated with the empty states in the almost-filled band. We call these particles holes. Keep in mind that there is no real particle associated with a hole. Instead, the combined behavior of all the electrons, which occupy states in the almost-filled band, is the same as that of positively charge particles associated with the unoccupied states.

The reason the concept of holes simplifies the analysis is that the density of states function of a whole band can be rather complex. However, it can be dramatically simplified if only states close to the band edge need to be considered.



2.3.6 The effective mass concept

Electrons with an energy close to a band minimum behave as free electrons. They accelerate in an applied electric field just like a free electron in vacuum. Their wavefunctions are periodic and extend over the size of the material. The presence of the periodic potential, due to the atoms in the crystal without the valence electrons, changes the properties of the electrons. Therefore, the mass of the electron differs from the free electron mass, m_0 . Because of the anisotropy of the effective mass and the presence of multiple equivalent band minima, we define two types of effective mass, the effective mass for density of states calculations and the effective mass for conductivity calculations. The effective mass values for electrons and holes are listed together with the value of the smallest energy bandgap in Table 2.3.2. Electrons in gallium arsenide have an isotropic effective mass so that the conductivity effective mass equals the density of states effective mass.

		Germanium	Silicon	GaAs
Smallest energy bandgap at 300 K	E_g (eV)	0.66	1.12	1.424
Electron effective mass for density of states calculations	$\frac{^*m_{e,dos}}{m_0}$	0.55	1.08	0.067
Hole effective mass for density of states calculations	$\frac{^*m_{h,dos}}{m_0}$	0.37	0.811	0.45
Electron effective mass for conductivity calculations	$\frac{^*m_{e,cond}}{m_0}$	0.12	0.26	0.067
Hole effective mass for conductivity calculations	$\frac{^*m_{h,cond}}{m_0}$	0.21	0.386	0.34

Table 2.3.2.:

Effective mass of carriers in germanium, silicon and gallium arsenide (GaAs)

Chapter 2: Semiconductor Fundamentals



2.10. The drift-diffusion model

The drift-diffusion model of a semiconductor is frequently used to describe semiconductor devices. It contains all the features described in this chapter.

Starting with Chapter 3, we will apply the drift-diffusion model to a variety of different devices. To facilitate this analysis, we present here a simplified drift-diffusion model, which contains all the essential features. This model results in a set of ten variables and ten equations.

The assumptions of the simplified drift-diffusion model are:

Full ionization: all dopants are assumed to be ionized (shallow dopants)

Non-degenerate: the Fermi energy is assumed to be at least $3 kT$ below/above the conduction/valence band edge.

Steady state: All variables are independent of time

Constant temperature: The temperature is constant throughout the device.

The ten variables are the following:

ρ , the charge density

n , the electron density

p , the hole density

\mathcal{E} , the electric field

ϕ , the potential

E_i , the intrinsic energy

F_n , the electron quasi-Fermi energy

F_p , the hole quasi-Fermi energy

J_n , the electron current density

J_p , the hole current density

The ten equations are:

Charge density equation

$$\rho = q(p - n + N_d^+ - N_a^-) \quad (2.10.1)$$

Electric field and potential equations

$$\frac{d\mathcal{E}}{dx} = \frac{\rho}{\epsilon} \quad (2.10.2)$$

$$\frac{d\phi}{dx} = -\mathcal{E} \quad (2.10.3)$$

$$\frac{dE_i}{dx} = q\mathcal{E} \quad (2.10.4)$$

Carrier density equations

$$n = n_i e^{(F_n - E_i)/kT} \quad (2.10.5)$$

$$p = n_i e^{(E_i - F_p)/kT} \quad (2.10.6)$$

Drift and diffusion current equations

$$J_n = qn\mu_n\mathcal{E} + qD_n \frac{dn}{dx} \quad (2.10.7)$$

$$J_p = qp\mu_p\mathcal{E} - qD_p \frac{dp}{dx} \quad (2.10.8)$$

Continuity equation in steady state with SHR recombination

$$0 = \frac{1}{q} \frac{\partial J_n}{\partial x} - \frac{np - n_i^2}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} \frac{1}{\tau} \quad (2.10.9)$$

$$0 = -\frac{1}{q} \frac{\partial J_p}{\partial x} - \frac{np - n_i^2}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} \frac{1}{\tau} \quad (2.10.10)$$

Chapter 2: Semiconductor Fundamentals



2.9. Continuity equation

[2.9.1. Derivation](#)

[2.9.2. The diffusion equation](#)

[2.9.3. Steady state solution to the diffusion equation](#)

2.9.1. Derivation



The continuity equation describes a basic concept, namely that a change in carrier density over time is due to the difference between the incoming and outgoing flux of carriers plus the generation and minus the recombination. The flow of carriers and recombination and generation rates are illustrated with Figure 2.9.1.

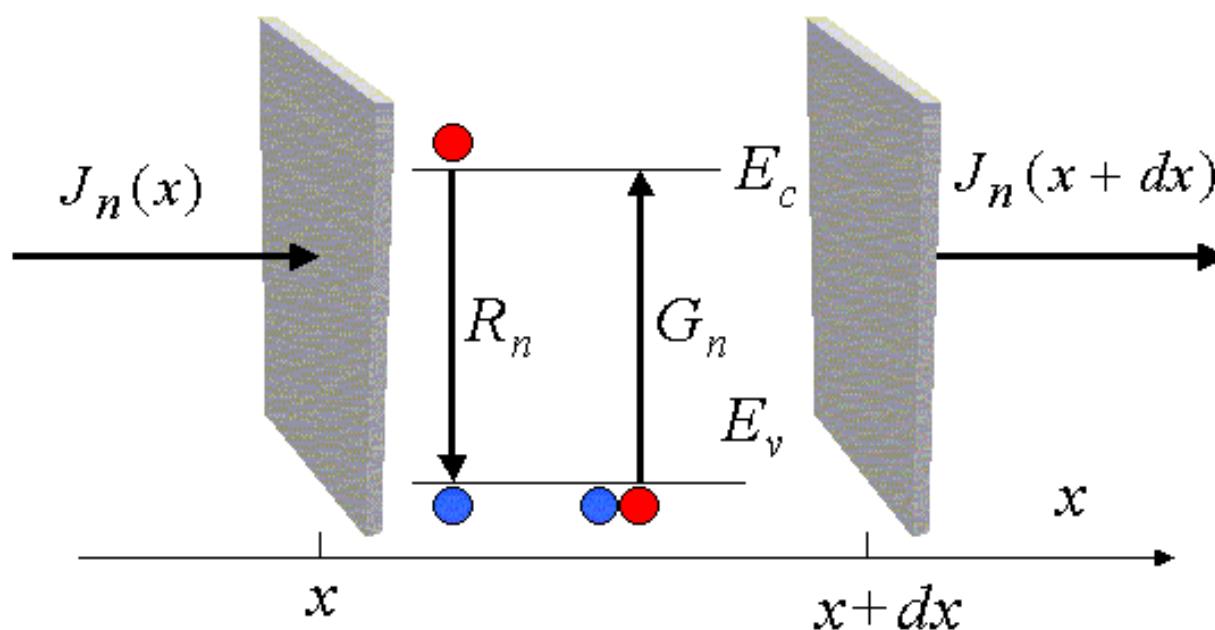


Figure 2.9.1 : Electron currents and possible recombination and generation processes

The rate of change of the carriers between x and $x + dx$ equals the difference between the incoming flux and the outgoing flux plus the generation and minus the recombination:

$$\frac{\partial n(x,t)}{\partial t} A dx = \left(\frac{J_n(x)}{-q} - \frac{J_n(x+dx)}{-q} \right) A + (G_n(x,t) - R_n(x,t)) A dx \quad (2.9.1)$$

where $n(x,t)$ is the carrier density, A is the area, $G_n(x,t)$ is the generation rate and $R_n(x,t)$ is the recombination rate. Using a Taylor series expansion,

$$J_n(x+dx) = J_n(x) + \frac{d J_n(x)}{dx} dx \quad (2.9.2)$$

this equation can be formulated as a function of the derivative of the current:

$$\frac{\partial n(x,t)}{\partial t} = \frac{1}{q} \frac{\partial J_n(x,t)}{\partial x} + G_n(x,t) - R_n(x,t) \quad (2.9.3)$$

and similarly for holes one finds:

$$\frac{\partial p(x,t)}{\partial t} = -\frac{1}{q} \frac{\partial J_p(x,t)}{\partial x} + G_p(x,t) - R_p(x,t) \quad (2.9.4)$$

A solution to these equations can be obtained by substituting the expression for the electron and hole current, (2.7.29) and (2.7.30). This then yields two partial differential equations as a function of the electron density, the hole density and the electric field. The electric field itself is obtained from Gauss's law.

$$\begin{aligned} \frac{\partial n(x,t)}{\partial t} = \\ \mu_n n \frac{\partial \mathbf{E}(x,t)}{\partial x} + \mu_n \mathbf{E} \frac{\partial n(x,t)}{\partial x} + D_n \frac{\partial^2 n(x,t)}{\partial x^2} + G_n(x,t) - R_n(x,t) \end{aligned} \quad (2.9.5)$$

$$\begin{aligned} \frac{\partial p(x,t)}{\partial t} = \\ -\mu_p p \frac{\partial \mathbf{E}(x,t)}{\partial x} - \mu_p \mathbf{E} \frac{\partial p(x,t)}{\partial x} + D_p \frac{\partial^2 p(x,t)}{\partial x^2} + G_p(x,t) - R_p(x,t) \end{aligned} \quad (2.9.6)$$

A generalization in three dimensions yields the following continuity equations for electrons and holes:

$$\frac{\partial n(x,y,z,t)}{\partial t} = \frac{1}{q} \vec{\nabla} \vec{J}_n(x,y,z,t) + G_n(x,y,z,t) - R_n(x,y,z,t) \quad (2.9.7)$$

$$\frac{\partial p(x,y,z,t)}{\partial t} = -\frac{1}{q} \vec{\nabla} \vec{J}_p(x,y,z,t) + G_p(x,y,z,t) - R_p(x,y,z,t) \quad (2.9.8)$$

2.9.2. The diffusion equation



In the quasi-neutral region - a region containing mobile carriers, where the electric field is small - the current is due to diffusion only. In addition, we can use the simple recombination model for the net recombination rate. This leads to the time-dependent diffusion equations for electrons in p-type material and for holes in n-type material:

$$\frac{\partial n(x,t)}{\partial t} = D_n \frac{\partial^2 n_p(x,t)}{\partial x^2} - \frac{n_p(x,t) - n_{p0}}{\tau_n} \quad (2.9.9)$$

$$\frac{\partial p(x,t)}{\partial t} = D_p \frac{\partial^2 p_n(x,t)}{\partial x^2} - \frac{p_n(x,t) - p_{n0}}{\tau_p} \quad (2.9.10)$$

2.9.3. Steady state solution to the diffusion equation



In steady state, the partial derivatives with respect to time are zero, yielding:

$$0 = D_n \frac{d^2 n_p(x)}{dx^2} - \frac{n_p(x) - n_{p0}}{\tau_n} \quad (2.9.11)$$

$$0 = D_p \frac{d^2 p_n(x)}{dx^2} - \frac{p_n(x) - p_{n0}}{\tau_p} \quad (2.9.12)$$

The general solution to these second order differential equations are:

$$n_p(x \leq -x_p) = n_{p0} + C e^{-(x+x_p)/L_p} + D e^{(x+x_p)/L_p} \quad (2.9.13)$$

$$p_n(x \geq x_n) = p_{n0} + A e^{-(x-x_n)/L_p} + B e^{(x-x_n)/L_p} \quad (2.9.14)$$

where L_n and L_p are the diffusion lengths given by:

$$L_n = \sqrt{D_n \tau_n} \quad (2.9.15)$$

$$L_p = \sqrt{D_p \tau_p} \quad (2.9.16)$$

The diffusion constants, D_n and D_p , are obtained using the Einstein relations (2.7.27) and (2.7.28). The diffusion equations can also be written as a function of the excess carrier densities, δn and δp , which are related to the total carrier densities, n and p , and the thermal equilibrium densities, n_0 and p_0 , by:

$$n = n_0 + \delta n \quad (2.9.17)$$

$$p = p_0 + \delta p \quad (2.9.18)$$

yielding:

$$0 = \frac{d^2(\delta n_p)}{dx^2} - \frac{\delta n_p}{L_n^2} \quad (2.9.19)$$

$$0 = \frac{d^2(\delta p_n)}{dx^2} - \frac{\delta p_n}{L_p^2} \quad (2.9.20)$$

The diffusion equation will be used to calculate the diffusion current in p-n junctions and bipolar transistors.

Chapter 2: Semiconductor Fundamentals



2.8. Carrier recombination and generation

2.8.1. Simple recombination-generation model

2.8.2. Band-to-band recombination

2.8.3. Trap assisted recombination

2.8.4. Surface recombination

2.8.5. Auger recombination

2.8.6. Generation due to light

Recombination of electrons and holes is a process by which both carriers annihilate each other: electrons occupy - through one or multiple steps - the empty state associated with a hole. Both carriers eventually disappear in the process. The energy difference between the initial and final state of the electron is released in the process. This leads to one possible classification of the recombination processes. In the case of radiative recombination, this energy is emitted in the form of a photon. In the case of non-radiative recombination, it is passed on to one or more phonons and in Auger recombination it is given off in the form of kinetic energy to another electron. Another classification scheme considers the individual energy levels and particles involved. These different processes are further illustrated with Figure 2.8.1.

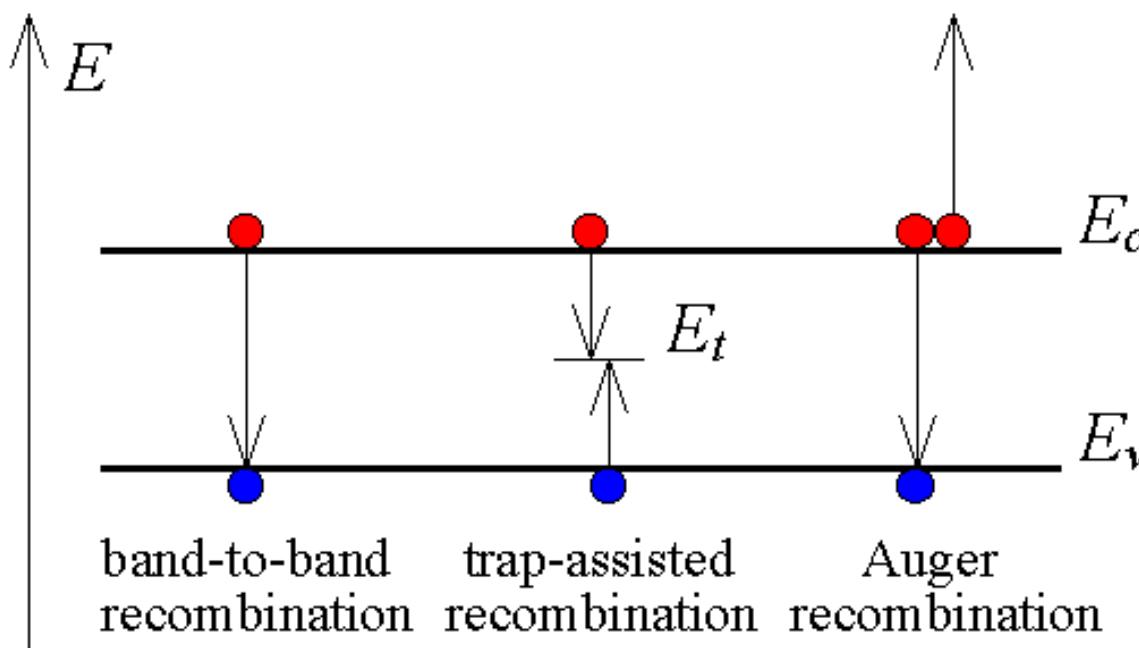


Figure 2.8.1 :

Carrier recombination mechanisms in semiconductors

Band-to-band recombination occurs when an electron falls from its conduction band state into the empty valence band state associated with the hole. This band-to-band transition is typically also a radiative transition in direct bandgap semiconductors.

Trap-assisted recombination occurs when an electron falls into a "trap", an energy level within the bandgap caused by the presence of a foreign atom or a structural defect. Once the trap is filled it cannot accept another electron. The electron occupying the trap, in a second step, falls into an empty valence band state, thereby completing the recombination process. One can envision this process as a two-step transition of an electron from the conduction band to the valence band or as the annihilation of the electron and hole, which meet each other in the trap. We will refer to this process as Shockley-Read-Hall (SRH) recombination.

Auger recombination is a process in which an electron and a hole recombine in a band-to-band transition, but now the resulting energy is given off to another electron or hole. The involvement of a third particle affects the recombination rate so that we need to treat Auger recombination differently from band-to-band recombination.

Each of these recombination mechanisms can be reversed leading to carrier generation rather than recombination. A single expression will be used to describe recombination as well as generation for each of the above mechanisms.

In addition, there are generation mechanisms, which do not have an associated recombination mechanism: generation of carriers by light absorption or a high-energy electron/particle beam. These processes are referred to as ionization processes. Impact ionization, which is the generation mechanism, associated with Auger recombination also belongs to this category. The generation mechanisms are illustrated with Figure 2.8.2.

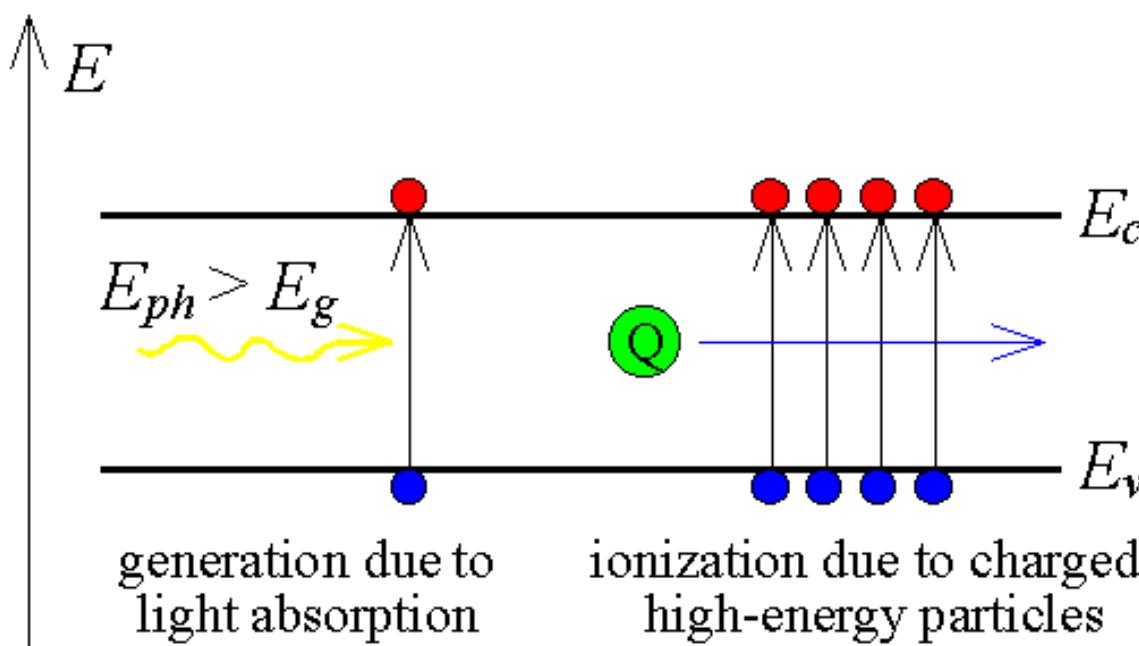


Figure 2.8.2 : Carrier generation due to light absorption and ionization due to high-energy particle beams

Carrier generation due to light absorption occurs if the photon energy is large enough to lift an electron from the valence band into an empty conduction band state, generating one electron-hole pair. The photon energy needs to be larger than the bandgap energy to satisfy this condition. The photon is absorbed in this process and the excess energy, $E_{ph} - E_g$, is added to the electron and the hole in the form of kinetic energy.

Carrier generation or ionization due to a high-energy beam consisting of charged particles is similar except that the available energy can be much larger than the bandgap energy so that multiple electron-hole pairs can be formed. The high-energy particle gradually loses its energy and eventually stops. This generation mechanism is used in semiconductor-based nuclear particle counters. As the number of ionized electron-hole pairs varies with the energy of the particle, one can also use such detector to measure the particle energy.

Finally, there is a generation process called impact ionization, the generation mechanism that is the counterpart of Auger recombination. Impact ionization is caused by an electron/hole with an energy, which is much larger/smaller than the conduction/valence band edge. The detailed mechanism is illustrated with Figure 2.8.3.

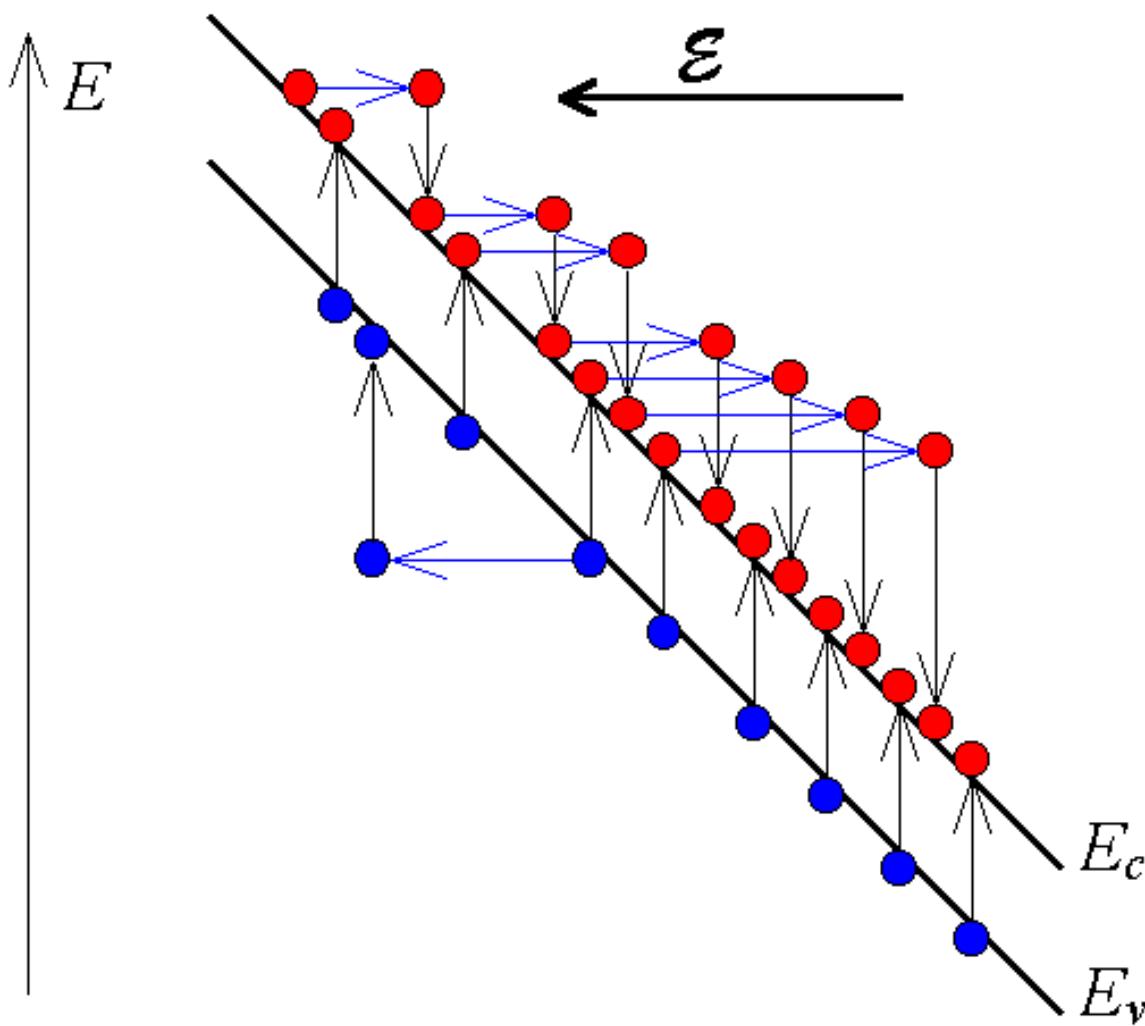


Figure 2.8.3: Impact ionization and avalanche multiplication of electrons and holes in the presence of a large electric field.

The excess energy is given off to generate an electron-hole pair through a band-to-band transition. This generation process causes avalanche multiplication in semiconductor diodes under high reverse bias: As one carrier accelerates in the electric field it gains energy. The kinetic energy is given off to an electron in the valence band, thereby creating an electron-hole pair. The resulting two electrons can create two more electrons which generate four more causing an avalanche multiplication effect. Electrons as well as holes contribute to avalanche multiplication.

2.8.1. Simple recombination-generation model



A simple model for the recombination-generation mechanisms states that the recombination-generation rate is proportional to the excess carrier density. It acknowledges the fact that no recombination takes place if the carrier density equals the thermal equilibrium value. The resulting expression for the recombination of electrons in a p-type semiconductor is given by:

$$U_n = R_n - G_n = \frac{n_p - n_{p0}}{\tau_n} \quad (2.8.1)$$

and similarly for holes in an n-type semiconductor:

$$U_p = R_p - G_p = \frac{p_n - p_{n0}}{\tau_p} \quad (2.8.2)$$

where the parameter τ can be interpreted as the average time after which an excess minority carrier recombines.

We will show for each of the different recombination mechanisms that the recombination rate can be simplified to this form when applied to minority carriers in a "quasi-neutral" semiconductor. The above expressions are therefore only valid under these conditions. The recombination rates of the majority carriers equals that of the minority carriers since in steady state recombination involves an equal number of holes and electrons. Therefore, the recombination rate of the majority carriers depends on the excess-minority-carrier-density as the minority carriers limit the recombination rate.

Recombination in a depletion region and in situations where the hole and electron density are close to each other cannot be described with the simple model and the more elaborate expressions for the individual recombination mechanisms must be used.

2.8.2. Band-to-band recombination



Band-to-band recombination depends on the density of available electrons and holes. Both carrier types need to be available in the recombination process. Therefore, the rate is expected to be proportional to the product of n and p . Also, in thermal equilibrium, the recombination rate must equal the generation rate since there is no net recombination or generation. As the product of n and p equals n_i^2 in thermal equilibrium, the net recombination rate can be expressed as:

$$U_{b-b} = b(np - n_i^2) \quad (2.8.3)$$

where b is the bimolecular recombination constant.

2.8.3. Trap assisted recombination



The net recombination rate for trap-assisted recombination is given by:

$$U_{SHR} = \frac{pn - n_i^2}{p + n + 2n_i \cosh(\frac{E_i - E_t}{kT})} N_t v_{th} \sigma \quad (2.8.4)$$

This expression can be further simplified for $p \gg n$ to:

$$U_n = R_n - G_n = \frac{n_p - n_{p0}}{\tau_n} \quad (2.8.5)$$

and for $n \gg p$ to:

$$U_p = R_p - G_p = \frac{p_n - p_{n0}}{\tau_p} \quad (2.8.6)$$

were

$$\tau_n = \tau_p = \frac{1}{N_t v_{th} \sigma} \quad (2.8.7)$$



2.8.4. Surface recombination

Recombination at semiconductor surfaces and interfaces can have a significant impact on the behavior of devices. This is because surfaces and interfaces typically contain a large number of recombination centers because of the abrupt termination of the semiconductor crystal, which leaves a large number of electrically active dangling bonds. In addition, the surfaces and interfaces are more likely to contain impurities since they are exposed during the device fabrication process. The net recombination rate due to trap-assisted recombination and generation is given by:

$$U_{s,SHR} = \frac{pn - n_i^2}{p + n + 2n_i \cosh(\frac{E_i - E_{st}}{kT})} N_{st} v_{th} \sigma_s \quad (2.8.8)$$

This expression is almost identical to that of Shockley-Hall-Read recombination. The only difference is that the recombination is due to a two-dimensional density of traps, N_{ts} , as the traps only exist at the surface or interface.

This equation can be further simplified for minority carriers in a quasi-neutral region. For instance for electrons in a quasi-neutral p-type region $p \gg n$ and $p \gg n_i$ so that for $E_i = E_{st}$, it can be simplified to:

$$U_{s,n} = R_{s,n} - G_{s,n} = v_s (n_p - n_{p0}) \quad (2.8.9)$$

where the recombination velocity, v_s , is given by:

$$v_s = N_{st} v_{th} \sigma_s \quad (2.8.10)$$

2.8.5. Auger recombination



Auger recombination involves three particles: an electron and a hole, which recombine in a band-to-band transition and give off the resulting energy to another electron or hole. The expression for the net recombination rate is therefore similar to that of band-to-band recombination but includes the density of the electrons or holes, which receive the released energy from the electron-hole annihilation:

$$U_{Auger} = \Gamma_n n (np - n_i^2) + \Gamma_p p (np - n_i^2) \quad (2.8.11)$$

The two terms correspond to the two possible mechanisms.



2.8.6. Generation due to light

Carriers can be generated in semiconductors by illuminating the semiconductor with light. The energy of the incoming photons is used to bring an electron from a lower energy level to a higher energy level. In the case where an electron is removed from the valence band and added to the conduction band, an electron-hole pair is generated. A necessary condition for this to happen is that the energy of the photon, E_{ph} , is larger than the bandgap energy, E_g . As the energy of the photon is given off to the electron, the photon no longer exists.

If each absorbed photon creates one electron-hole pair, the electron and hole generation rates are given by:

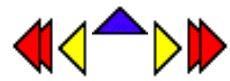
$$G_{p,light} = G_{n,light} = \alpha \frac{P_{opt}(x)}{E_{ph}A} \quad (2.8.12)$$

where α is the absorption coefficient of the material at the energy of the incoming photon. The absorption of light in a semiconductor causes the optical power to decrease with distance. This effect is described mathematically by:

$$\frac{dP_{opt}(x)}{dx} = -\alpha P_{opt}(x) \quad (2.8.13)$$

The calculation of the generation rate of carriers therefore requires first a calculation of the optical power within the structure from which the generation rate can then be obtained using (2.8.12).

Example 2.11	Calculate the electron and hole densities in an n-type silicon wafer ($N_d = 10^{17} \text{ cm}^{-3}$) illuminated uniformly with 10 mW/cm^2 of red light ($E_{ph} = 1.8 \text{ eV}$). The absorption coefficient of red light in silicon is 10^{-3} cm^{-1} . The minority carrier lifetime is $10 \mu\text{s}$.
Solution	<p>The generation rate of electrons and holes equals:</p> $G_n = G_p = \alpha \frac{P_{opt}}{E_{ph}A} = 10^{-3} \frac{10^{-2}}{1.8 \times 1.6 \times 10^{-19}} = 3.5 \times 10^{13} \text{ cm}^{-3} \text{ s}^{-1}$ <p>where the photon energy was converted into Joules. The excess carrier densities are then obtained from:</p> $\delta n = \delta p = \tau_p G_p = 10 \times 10^{-6} \times 3.5 \times 10^{13} = 3.5 \times 10^8 \text{ cm}^{-3} \text{ s}^{-1}$ <p>The excess carrier densities are then obtained from: So that the electron and hole densities equal:</p> $n = n_o + \delta n = 10^{17} + 3.5 \times 10^{13} = 10^{17} \text{ cm}^{-3} \text{ s}^{-1}$



Chapter 3: Metal-Semiconductor Junctions

Equations

$$\phi_B = \Phi_M - \chi, \text{ for an n-type semiconductor} \quad \bullet(3.2.1)\bullet$$

$$\phi_B = \frac{E_g}{q} + \chi - \Phi_M, \text{ for a p-type semiconductor} \quad \bullet(3.2.2)\bullet$$

$$\phi = \Phi_M - \chi - \frac{E_c - E_{F,n}}{q}, \quad \text{n-type} \quad \bullet(3.2.3)\bullet$$

$$\phi = \chi + \frac{E_c - E_{F,p}}{q} - \Phi_M, \quad \text{p-type} \quad \bullet(3.2.4)\bullet$$

$$\phi(x = \infty) - \phi(x = 0) = \phi_i - V_a \quad \bullet(3.2.5)\bullet$$

$$\frac{d^2\phi}{dx^2} = -\frac{\rho}{\epsilon_s} = -\frac{q}{\epsilon_s}(p - n + N_d^+ - N_a^-) \quad \bullet(3.3.1)\bullet$$

$$\frac{d^2\phi}{dx^2} = \frac{2qn_i}{\epsilon_s} \left(\sinh \frac{\phi - \phi_F}{V_t} + \sinh \frac{\phi_F}{V_t} \right) \quad \bullet(3.3.2)\bullet$$

$$\sinh \frac{\phi_F}{V_t} = \frac{N_a^- - N_d^+}{2n_i} \quad \bullet(3.3.3)\bullet$$

$$\begin{aligned} \mathcal{A}(x) &= qN_d & 0 < x < x_d \\ \mathcal{A}(x) &= 0 & x_d < x \end{aligned} \quad \bullet(3.3.4)\bullet$$

$$\begin{aligned} \mathcal{E}(x) &= -\frac{qN_d}{\varepsilon_s} (x_d - x) & 0 < x < x_d \\ \mathcal{E}(x) &= 0 & x_d \leq x \end{aligned} \quad \bullet(3.3.5)\bullet$$

$$\mathcal{E}(x=0) = -\frac{qN_d x_d}{\varepsilon_s} = -\frac{Q_d}{\varepsilon_s} \quad \bullet(3.3.6)\bullet$$

$$\begin{aligned} \mathcal{P}(x) &= 0 & x \leq 0 \\ \mathcal{P}(x) &= \frac{qN_d}{2\varepsilon_s} [x_d^2 - (x_d - x)^2] & 0 < x < x_d \\ \mathcal{P}(x) &= \frac{qN_d x_d^2}{2\varepsilon_s} & x_d \leq x \end{aligned} \quad \bullet(3.3.7)\bullet$$

$$\mathcal{P} - V_a = -\mathcal{P}(x=0) = \frac{qN_d x_d^2}{2\varepsilon_s} \quad \bullet(3.3.8)\bullet$$

$$x_d = \sqrt{\frac{2\varepsilon_s(\mathcal{P} - V_a)}{qN_d}} \quad \bullet(3.3.9)\bullet$$

$$C_j = \left| \frac{dQ_d}{dV_a} \right| = \sqrt{\frac{q\varepsilon_s N_d}{2(\mathcal{P} - V_a)}} = \frac{\varepsilon_s}{x_d} \quad \bullet(3.3.10)\bullet$$

$$\Delta\phi_B = \sqrt{\frac{q\mathcal{E}_{\max}}{4\pi\varepsilon_s}} \quad \bullet(3.3.11)\bullet$$

$$J_n = qvN_c \exp\left(-\frac{\mathcal{P}_B}{V_t}\right) \left(\exp\left(\frac{V_a}{V_t}\right) - 1 \right) \quad \bullet(3.4.1)\bullet$$

$$J_n = qv_R n \Theta \quad \bullet(3.4.2)\bullet$$

$$J_n = \frac{q^2 D_n N_c}{V_t} \sqrt{\frac{2q(\phi - V_a) N_d}{\varepsilon_s}} \exp\left(-\frac{\phi_B}{V_t}\right) \left[\exp\left(\frac{V_a}{V_t}\right) - 1\right] \quad \bullet(3.4.3)\bullet$$

$$S_{\max} = \sqrt{\frac{2q(\phi - V_a) N_d}{\varepsilon_s}} \quad \bullet(3.4.4)\bullet$$

$$J_n = q \mu_n \varepsilon_{\max} N_c \exp\left(-\frac{\phi_B}{V_t}\right) \left[\exp\left(\frac{V_a}{V_t}\right) - 1\right] \quad \bullet(3.4.5)\bullet$$

$$J_{MS} = A^* T^2 e^{-\phi_B/V_t} (e^{V_a/V_t} - 1) \quad \bullet(3.4.6)\bullet$$

$$v_R = \sqrt{\frac{kT}{2\pi m}} \quad \bullet(3.4.7)\bullet$$

$$J_n = q v_R N_c \exp\left(-\frac{\phi_B}{V_t}\right) \left[\exp\left(\frac{V_a}{V_t}\right) - 1\right] \quad \bullet(3.4.8)\bullet$$

$$J_n = q v_R n \Theta \quad \bullet(3.4.9)\bullet$$

$$\Theta = \exp\left(-\frac{4}{3} \frac{\sqrt{2qm^*}}{\hbar} \frac{\phi_B^{3/2}}{\varepsilon}\right) \quad \bullet(3.4.10)\bullet$$

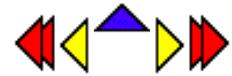


Chapter 3: Metal-Semiconductor Junctions

Bibliography

1. Physics of Semiconductor Devices, Second edition, S. M. Sze, Wiley & Sons, 1981, Chapter 5.
2. Device Electronics for Integrated Circuits, Second edition, R.S. Muller and T. I. Kamins, Wiley & Sons, 1986, Chapter 3.

Chapter 4: p-n Junctions



Equations

$$\phi = V_t \ln \frac{N_d N_a}{n_i^2} \quad (4.2.1) \bullet$$

$$\phi = \phi_i - V_a \quad (4.2.2) \bullet$$

$$\frac{d^2 \phi}{dx^2} = -\frac{\rho}{\epsilon_s} = -\frac{q}{\epsilon_s} (p - n + N_d^+ - N_a^-) \quad (4.3.1) \bullet$$

$$\frac{d^2 \phi}{dx^2} = \frac{2qn_i}{\epsilon_s} (\sinh \frac{\phi - \phi_F}{V_t} + \sinh \frac{\phi_F}{V_t}) \quad (4.3.2) \bullet$$

$$\sinh \frac{\phi_F}{V_t} = \frac{N_a^- - N_d^+}{2n_i} \quad (4.3.3) \bullet$$

$$x_d = x_n + x_p \quad (4.3.4) \bullet$$

$$\rho = q(p - n + N_d^+ - N_a^-) \cong q(N_d^+ - N_a^-), \text{ for } -x_p \leq x \leq x_n \quad (4.3.5) \bullet$$

$$\begin{aligned} \rho(x) &= 0 \\ \rho(x) &= -qN_a \\ \rho(x) &= qN_d \\ \rho(x) &= 0 \end{aligned} \quad (4.3.6) \bullet$$

$$Q_n = qN_d x_n \quad (4.3.7) \bullet$$

$$Q_p = -qN_a x_p \quad \bullet(4.3.8)\bullet$$

$$\frac{d\mathcal{E}(x)}{dx} = \frac{\mathcal{A}(x)}{\varepsilon_s} \cong \frac{q}{\varepsilon_s} (N_d^+(x) - N_a^-(x)), \text{ for } -x_p \leq x \leq x_n \quad \bullet(4.3.9)\bullet$$

$$\mathcal{E}(x) = 0$$

$$\begin{aligned} \mathcal{E}(x) &= -\frac{qN_a(x + x_p)}{\varepsilon_s} \\ \mathcal{E}(x) &= \frac{qN_d(x - x_n)}{\varepsilon_s} \end{aligned} \quad \bullet(4.3.10)\bullet$$

$$\mathcal{E}(x) = 0$$

$$\mathcal{E}(x = 0) = -\frac{qN_a x_p}{\varepsilon_s} = -\frac{qN_d x_n}{\varepsilon_s} \quad \bullet(4.3.11)\bullet$$

$$N_d x_n = N_a x_p \quad \bullet(4.3.12)\bullet$$

$$x_n = x_d \frac{N_a}{N_a + N_d} \quad \bullet(4.3.13)\bullet$$

$$x_p = x_d \frac{N_d}{N_a + N_d} \quad \bullet(4.3.14)\bullet$$

$$\frac{d\mathcal{A}(x)}{dx} = -\mathcal{E}(x) \quad \bullet(4.3.15)\bullet$$

$$\mathcal{A} - V_a = \frac{qN_d x_n^2}{2\varepsilon_s} + \frac{qN_a x_p^2}{2\varepsilon_s} \quad \bullet(4.3.16)\bullet$$

$$x_d = \sqrt{\frac{2 \varepsilon_s}{q} \left(\frac{1}{N_a} + \frac{1}{N_d} \right) (\phi - V_a)} \quad \bullet(4.3.17)\bullet$$

$$x_n = \sqrt{\frac{2 \varepsilon_s}{q} \frac{N_a}{N_d} \frac{1}{N_a + N_d} (\phi - V_a)} \quad \bullet(4.3.18)\bullet$$

$$x_p = \sqrt{\frac{2 \varepsilon_s}{q} \frac{N_d}{N_a} \frac{1}{N_a + N_d} (\phi - V_a)} \quad \bullet(4.3.19)\bullet$$

$$C(V_a) \stackrel{\Delta}{=} \left| \frac{dQ(V_a)}{dV_a} \right| \quad \bullet(4.3.20)\bullet$$

$$C_j = \sqrt{\frac{q \varepsilon_s}{2(\phi - V_a)} \frac{N_a N_d}{N_a + N_d}} \quad \bullet(4.3.21)\bullet$$

$$C_j = \frac{\varepsilon_s}{x_d} \quad \bullet(4.3.22)\bullet$$

$$\frac{1}{C_j^2} = \frac{2}{q \varepsilon_s} \frac{N_a + N_d}{N_a N_d} (\phi - V_a) \quad \bullet(4.3.23)\bullet$$

$$\frac{d(1/C_j^2)}{dV_a} = -\frac{2}{q \varepsilon_s} \frac{N_a + N_d}{N_a N_d} \quad \bullet(4.3.24)\bullet$$

$$N_d = -\frac{2}{q \varepsilon_s} \frac{1}{\frac{d(1/C_j^2)}{dV_a}}, \text{if } N_a \gg N_d \quad \bullet(4.3.25)\bullet$$

□

•(4.3.23)•

$$p_n(x = x_n) = p_{n0} e^{V_a/V_t} \quad \bullet(4.4.1)\bullet$$

$$n_p(x = -x_p) = n_{p0} e^{V_a/V_t} \quad \bullet(4.4.2)\bullet$$

$$p_n(x = w_n) = p_{n0} \quad \bullet(4.4.3)\bullet$$

$$n_p(x = -w_p) = n_{p0} \quad \bullet(4.4.4)\bullet$$

$$p_n(x \geq x_n) = p_{n0} + A^* \cosh \frac{x - x_n}{L_p} + B^* \sinh \frac{x - x_n}{L_p} \quad \bullet(4.4.5)\bullet$$

$$n_p(x \leq -x_p) = n_{p0} + C^* \cosh \frac{x + x_p}{L_n} + D^* \sinh \frac{x + x_p}{L_n} \quad \bullet(4.4.6)\bullet$$

$$p_n(x \geq x_n) = p_{n0} + p_{n0}(e^{V_a/V_t} - 1) \left[\cosh \frac{x - x_n}{L_p} - \coth \frac{w_n}{L_p} \sinh \frac{x - x_n}{L_p} \right] \quad \bullet(4.4.7)\bullet$$

$$n_p(x \leq -x_p) = n_{p0} + n_{p0}(e^{V_a/V_t} - 1) \left[\cosh \frac{x + x_p}{L_n} + \coth \frac{w_p}{L_n} \sinh \frac{x + x_p}{L_n} \right] \quad \bullet(4.4.8)\bullet$$

$$w_n^i = w_n - x_n \quad \bullet(4.4.9)\bullet$$

$$w_p^i = w_p - x_p \quad \bullet(4.4.10)\bullet$$

$$\begin{aligned}
 J_p(x \geq x_n) &= -qD_p \frac{dp}{dx} \\
 &= -\frac{qD_p p_{n0}}{L_p} (e^{V_a/V_t} - 1) [\sinh \frac{x - x_n}{L_p} - \coth \frac{w_n}{L_p} \cosh \frac{x - x_n}{L_p}]
 \end{aligned} \tag{4.4.11}$$

$$\begin{aligned}
 J_n(x \leq -x_p) &= qD_n \frac{dn}{dx} \\
 &= \frac{qD_n n_{p0}}{L_n} (e^{V_a/V_t} - 1) [\sinh \frac{x + x_p}{L_n} + \coth \frac{w_p}{L_n} \cosh \frac{x + x_p}{L_n}]
 \end{aligned} \tag{4.4.12}$$

$$I = A[J_n(x = -x_p) + J_p(x = x_n) + J_r] \cong I_s (e^{V_a/V_t} - 1) \tag{4.4.13}$$

$$I_s = qA \left[\frac{D_n n_{p0}}{L_n} \coth \left(\frac{w_p}{L_n} \right) + \frac{D_p p_{n0}}{L_p} \coth \left(\frac{w_n}{L_p} \right) \right] \tag{4.4.14}$$

$$\coth x = \frac{1}{\tanh x} \cong \frac{1}{x} \text{, for } x \ll 1 \tag{4.4.15}$$

$$p_n(x \geq x_n) = p_{n0} + p_{n0} (e^{V_a/V_t} - 1) \exp \frac{-(x - x_n)}{L_p} \tag{4.4.16}$$

$$n_p(x \leq -x_p) = n_{p0} + n_{p0} (e^{V_a/V_t} - 1) \exp \frac{x + x_p}{L_n} \tag{4.4.17}$$

$$J_p(x \geq x_n) = \frac{qD_p p_{n0}}{L_p} (e^{V_a/V_t} - 1) \exp \frac{-(x - x_n)}{L_p} \tag{4.4.18}$$

$$J_n(x \leq -x_p) = \frac{qD_n n_{p0}}{L_n} (e^{V_a/V_t} - 1) \exp \frac{x + x_p}{L_n} \tag{4.4.19}$$

$$I_s = qA \left[\frac{D_n n_{p0}}{L_n} + \frac{D_p p_{n0}}{L_p} \right] = qA \left[\frac{n_{p0} L_n}{w_n} + \frac{p_{n0} L_p}{w_p} \right] \quad \bullet(4.4.20)\bullet$$

$$I_r \geq qA \frac{p_{n0} x_n}{w_p} (e^{V_a/V_t} - 1) \quad \bullet(4.4.21)\bullet$$

$$I_r \ll I, \text{ for } x_n \ll L_p \quad \bullet(4.4.22)\bullet$$

$$0 = D_n \frac{d^2 n_p}{dx^2}, \text{ and } 0 = D_p \frac{d^2 p_n}{dx^2} \quad \bullet(4.4.23)\bullet$$

$$n_p = A + Bx, \text{ and } p_n = A + Bx \quad \bullet(4.4.24)\bullet$$

$$p_n = p_{n0} + p_{n0} (e^{V_a/V_t} - 1) \left(1 - \frac{x - x_n}{w_n} \right) \quad \bullet(4.4.25)\bullet$$

$$n_p = n_{p0} + n_{p0} (e^{V_a/V_t} - 1) \left(1 + \frac{x - x_p}{w_p} \right) \quad \bullet(4.4.26)\bullet$$

$$I = A [J_n(x = -x_p) + J_p(x = x_n) + J_r] \cong I_s (e^{V_a/V_t} - 1) \quad \bullet(4.4.27)\bullet$$

$$I_s = qA \left[\frac{D_n n_{p0}}{w_p} + \frac{D_p p_{n0}}{w_n} \right] \quad \bullet(4.4.28)\bullet$$

$$J_{b-b} = q \int_{-x_p}^{x_n} U_{b-b} dx \quad \bullet(4.4.29)\bullet$$

$$U_{b-b} = b(np - n_i^2) \quad \bullet(4.4.30)\bullet$$

$$np = n_i e^{(E_F n - E_i)/kT} n_i e^{(E_i - E_F p)/kT} = n_i^2 e^{V_a/V_t} \quad \bullet(4.4.31)\bullet$$

$$J_{b-b} = q \int_{-x_p}^{x_n} n_i^2 (e^{V_a/V_t} - 1) dx = q n_i^2 b w (e^{V_a/V_t} - 1) \quad \bullet(4.4.32)\bullet$$

$$J_{SHR} = q \int_{-x_p}^{x_n} U_{SHR} dx \quad \bullet(4.4.33)\bullet$$

$$J_{SHR} = q \int_{-x_p}^{x_n} \frac{1}{\tau} \frac{n_i^2 (e^{V_a/V_t} - 1)}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} dx \quad \bullet(4.4.34)\bullet$$

$$np = n_i e^{(E_F n - E_i)/kT} n_i e^{(E_i - E_F p)/kT} = n_i^2 e^{V_a/V_t} \quad \bullet(4.4.35)\bullet$$

$$U_{SHR, \max} = \text{MAX} \left(\frac{1}{\tau} \frac{n_i^2 (e^{V_a/V_t} - 1)}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} \right) \cong \frac{n_i}{2\tau} (e^{V_a/2V_t} - 1) \quad \bullet(4.4.36)\bullet$$

$$x' = \frac{\int_{-x_p}^{x_n} \frac{1}{\tau} \frac{n_i^2 (e^{V_a/V_t} - 1)}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} dx}{U_{SHR, \max}} \quad \bullet(4.4.37)\bullet$$

$$J_{SHR} = \frac{q n_i x'}{2\tau} (e^{V_a/2V_t} - 1) \quad \bullet(4.4.38)\bullet$$

$$J = J_s e^{V_a/\eta V_t} \quad \bullet(4.4.39)\bullet$$

$$\gamma = \frac{\log(e)}{V_t \text{ slope}} = \frac{1}{\frac{\text{slope}}{59.6 \text{ mV/decade}}} \quad \bullet(4.4.40)\bullet$$

$$V_a = 2V_t \ln \frac{N_d}{n_i} \quad \bullet(4.4.41)\bullet$$

$$V_a^* = V_a + IR_s \quad \bullet(4.4.42)\bullet$$

$$C = \frac{d\Delta Q}{dV_a} \quad \bullet(4.4.43)\bullet$$

$$\Delta Q_p = \int_{x_n}^{w_n} qA(p_n - p_{n0}) dx \quad \bullet(4.4.44)\bullet$$

$$\Delta Q_p = qA p_{n0} (e^{V_a/V_t} - 1) L_p = I_{s,p} (e^{V_a/V_t} - 1) \tau_p \quad \bullet(4.4.45)\bullet$$

$$I_{s,p} = q \frac{A p_{n0} D_p}{L_p} \quad \bullet(4.4.46)\bullet$$

$$C_{d,p} = \frac{d(I_{s,p} (e^{V_a/V_t} - 1) \tau_p)}{dV_a} = \frac{I_{s,p} e^{V_a/V_t} \tau_p}{V_t} \quad \bullet(4.4.47)\bullet$$

$$C_{d,p} = \frac{I_{s,p} e^{V_a/V_t} t_{r,p}}{V_t} \quad \bullet(4.4.48)\bullet$$

$$t_{r,p} = \frac{w_n^2}{2D_p} \quad \bullet(4.4.49)\bullet$$

$$|\mathcal{E}_{br}| = \frac{4 \times 10^5}{1 - \frac{1}{3} \log(N/10^{16})} \text{ V/cm} \quad \bullet(4.5.1)\bullet$$

$$|V_{br}| = -\mathcal{A} + \frac{|\mathcal{E}_{br}|^2 \varepsilon_s}{2qN} \quad \bullet(4.5.2)\bullet$$

$$w_{br} = \frac{|\mathcal{E}_{br}| \varepsilon_s}{qN} \quad \bullet(4.5.3)\bullet$$

$$dn = \alpha_n n dx \quad \bullet(4.5.4)\bullet$$

$$M = \frac{1}{1 - \int_{x_1}^{x_2} \alpha' dx} \quad \bullet(4.5.5)\bullet$$

$$M = \frac{1}{1 - \left| \frac{V_a}{V_{br}} \right|^n}, \text{ where } 2 < n < 6 \quad \bullet(4.5.6)\bullet$$

$$\Theta = \exp \left(-\frac{4}{3} \frac{\sqrt{2m^*}}{q\hbar} \frac{E_g^{3/2}}{\mathcal{E}} \right) \quad \bullet(4.5.7)\bullet$$

$$J_n = q v_R n \Theta \quad \bullet(4.5.8)\bullet$$

$$I = I_s (e^{V_a/V_t} - 1) - I_{ph} \quad \bullet(4.6.1)\bullet$$

$$I_{ph, \max} = \frac{q}{h\nu} P_{in} \quad \bullet(4.6.2)\bullet$$

$$I_{ph} = (1 - R)(1 - e^{-\alpha d}) \frac{qP_{in}}{h\nu} \quad \bullet(4.6.3)\bullet$$

$$\langle i^2 \rangle = 2qI\Delta f \quad \bullet(4.6.4)\bullet$$

$$\text{Fill Factor} = \frac{I_m V_m}{I_{sc} V_{oc}} \quad \bullet(4.6.5)\bullet$$

$$\text{Roundtrip amplification} = e^{2gL} R_1 R_2 = 1 \quad \bullet(4.6.6)\bullet$$

$$g = \frac{1}{2L} \ln \frac{1}{R_1 R_2} \quad \bullet(4.6.7)\bullet$$

$$P_{out} = \eta \frac{h\nu}{q} (I - I_{th}) \quad \bullet(4.6.8)\bullet$$

Chapter 3: Metal-Semiconductor Junctions

3.2. Structure and principle of operation

$$\mathbf{f}_B = \Phi_M - \mathbf{c} \quad \text{, for an n - type semiconductor} \quad (3.2.1)$$

$$\mathbf{f}_B = \frac{E_g}{q} + \mathbf{c} - \Phi_M \quad \text{, for a p - type semiconductor} \quad (3.2.2)$$

$$\mathbf{f}_i = \Phi_M - \mathbf{c} - \frac{E_c - E_{F,n}}{q} \quad , \quad \text{n - type} \quad (3.2.3)$$

$$\mathbf{f}_i = \frac{E_g}{q} + \mathbf{c} - \frac{E_v - E_{F,p}}{q} - \Phi_M \quad , \quad \text{p - type} \quad (3.2.4)$$

$$\mathbf{f}(x = \infty) - \mathbf{f}(x = 0) = \mathbf{f}_i - V_a \quad (3.2.5)$$

3.3. Electrostatic analysis

$$\frac{d^2\mathbf{f}}{dx^2} = -\frac{\mathbf{r}}{\mathbf{e}_s} = -\frac{q}{\mathbf{e}_s} (p - n + N_d^+ - N_a^-) \quad (3.3.1)$$

$$\frac{d^2\mathbf{f}}{dx^2} = \frac{2qn_i}{\mathbf{e}_s} \left(\sinh \frac{\mathbf{f} - \mathbf{f}_F}{V_t} + \sinh \frac{\mathbf{f}_F}{V_t} \right) \quad (3.3.2)$$

$$\sinh \frac{\mathbf{f}_F}{V_t} = \frac{N_a^- - N_d^+}{2n_i} \quad (3.3.3)$$

$$\begin{aligned} \mathbf{r}(x) &= qN_d & 0 < x < x_d \\ \mathbf{r}(x) &= 0 & x_d < x \end{aligned} \quad (3.3.4)$$

$$\begin{aligned} E(x) &= -\frac{qN_d}{\mathbf{e}_s} (x_d - x) & 0 < x < x_d \\ E(x) &= 0 & x_d \leq x \end{aligned} \quad (3.3.5)$$

$$E(x = 0) = -\frac{qN_d x_d}{\mathbf{e}_s} = -\frac{Q_d}{\mathbf{e}_s} \quad (3.3.6)$$

$$\mathbf{f}(x) = 0 \quad x \leq 0 \quad (3.3.7)$$

$$\mathbf{f}(x) = \frac{qN_d}{2\mathbf{e}_s} [x_d^2 - (x_d - x)^2] \quad 0 < x < x_d$$

$$\mathbf{f}(x) = \frac{qN_d x_d^2}{2\mathbf{e}_s} \quad x_d \leq x$$

$$\mathbf{f}_i - V_a = \mathbf{f}(x = \infty) - \mathbf{f}(x = 0) = \frac{qN_d x_d^2}{2\mathbf{e}_s} \quad (3.3.8)$$

$$x_d = \sqrt{\frac{2\mathbf{e}_s(\mathbf{f}_i - V_a)}{qN_d}} \quad (3.3.9)$$

$$C_j = \left| \frac{dQ_d}{dV_a} \right| = \sqrt{\frac{q\mathbf{e}_s N_d}{2(\mathbf{f}_i - V_a)}} = \frac{\mathbf{e}_s}{x_d} \quad (3.3.10)$$

$$\Delta \mathbf{f}_B = \sqrt{\frac{qE_{\max}}{4p\mathbf{e}_s}} \quad (3.3.11)$$

3.4. Schottky diode current

$$J_n = qvN_c \exp\left(-\frac{\mathbf{f}_B}{V_t}\right) \left(\exp\left(\frac{V_a}{V_t}\right) - 1 \right) \quad (3.4.1)$$

$$J_n = q v_R n \Theta \quad (3.4.2)$$

$$J_n = \frac{qD_n N_c}{V_t} \sqrt{\frac{2q(\mathbf{f}_i - V_a)N_d}{\mathbf{e}_s}} \exp\left(-\frac{\mathbf{f}_B}{V_t}\right) \left[\exp\left(\frac{V_a}{V_t}\right) - 1 \right] \quad (3.4.3)$$

$$E_{\max} = \sqrt{\frac{2q(\mathbf{f}_i - V_a)N_d}{\mathbf{e}_s}} \quad (3.4.4)$$

$$J_n = q \mathbf{m}_n E_{\max} N_c \exp\left(-\frac{\mathbf{f}_B}{V_t}\right) \left[\exp\left(\frac{V_a}{V_t}\right) - 1 \right] \quad (3.4.5)$$

Chapter 3: Metal-Semiconductor Junctions



Problems

1. Consider a gold-GaAs Schottky diode with a capacitance of 1 pF at -1 V. What is the doping density of the GaAs? Also calculate the depletion layer width at zero bias and the field at the surface of the semiconductor at -10 V. The area of the diode is 10^{-5} cm 2 . 
2. Using the work functions listed in table 3.2.1, predict which metal-semiconductor junctions are expected to be ohmic contacts. Use the ideal interface model.
3. Design a platinum-silicon diode with a capacitance of 1 pF and a maximum electric field less than 10^4 V/cm at -10 V bias. Provide a possible doping density and area. Make sure the diode has an area between 10^{-5} and 10^{-7} cm 2 . Is it possible to satisfy all requirements if the doping density equals 10^{17} cm $^{-3}$?
4. A platinum-silicon diode (area = 10^{-4} cm $^{-3}$, $N_d = 10^{17}$ cm $^{-3}$) is part of an LC tuning circuit containing a 100 nH inductance. The applied voltage must be less than 5 V. What is the tuning range of the circuit? The resonant frequency equals

$$\nu = \frac{1}{2\pi \sqrt{LC}}$$

, where L is the inductance and C is the diode capacitance.

Chapter 3: Metal-Semiconductor Junctions



Review Questions

1. What is a flatband diagram?
2. Define the barrier height of a metal-semiconductor junction. Can the barrier height be negative? Explain.
3. Define the built-in potential. Also provide an equation and state the implicit assumption(s).
4. Name three possible reasons why a measured barrier height can differ from the value calculated using equations [\(3.2.1\)](#) or [\(3.2.2\)](#).
5. How does the energy band diagram of a metal-semiconductor junction change under forward and reverse bias? How does the depletion layer width change with bias?
6. What is the full depletion approximation? Why do we need the full depletion approximation?
7. What mechanism(s) cause(s) current in a metal-semiconductor junction?

Chapter 4: p-n Junctions



Problems

1. A silicon p-n junction ($N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$) is biased with $V_a = -3 \text{ V}$. Calculate the built-in potential, the depletion layer width and the maximum electric field of the junction.
2. An abrupt silicon p-n junction consists of a p-type region containing 10^{16} cm^{-3} acceptors and an n-type region containing also 10^{16} cm^{-3} acceptors in addition to 10^{17} cm^{-3} donors.
 - a. Calculate the thermal equilibrium density of electrons and holes in the p-type region as well as both densities in the n-type region.
 - b. Calculate the built-in potential of the p-n junction.
 - c. Calculate the built-in potential of the p-n junction at 100°C .
3. For a p-n junction with a built-in potential of 0.62 V
 - a. What is the potential across the depletion region at an applied voltage, V_a , of 0, 0.5 and -2 Volt?
 - b. If the depletion layer is 1 micrometer at $V_a = 0 \text{ Volt}$, find the maximum electric field in the depletion region.
 - c. Assuming that the net doping density $|N_d - N_a|$ is the same in the n-type and p-type region of the diode, carefully sketch the electric field and the potential as a function of position throughout the depletion region. Add numeric values wherever possible.
4. An abrupt silicon ($n_i = 10^{10} \text{ cm}^{-3}$) p-n junction consists of a p-type region containing 10^{16} cm^{-3} acceptors and an n-type region containing $5 \times 10^{16} \text{ cm}^{-3}$ donors.
 - a. Calculate the built-in potential of this p-n junction.
 - b. Calculate the total width of the depletion region if the applied voltage, V_a equals 0, 0.5 and -2.5 V.
 - c. Calculate maximum electric field in the depletion region at 0, 0.5 and -2.5 V.
 - d. Calculate the potential across the depletion region in the n-type semiconductor at 0, 0.5 and -2.5 V.
5. Consider an abrupt p-n diode in thermal equilibrium with as many donors in the n-type region as acceptors in the p-type region and a maximum electric field of -13 kV/cm and a total depletion layer width of $1 \mu\text{m}$. (assume $\epsilon_s / \epsilon_0 = 12$)
 - a. What is the applied voltage, V_a ?
 - b. What is the built-in potential of the diode?
 - c. What is the donor density in the n-type region and the acceptor density in the p-type region?
 - d. What is the intrinsic carrier density of the semiconductor if the temperature is 300 K ?
6. A silicon ($n_i = 10^{10} \text{ cm}^{-3}$) p-n diode with $N_a = 10^{18} \text{ cm}^{-3}$ has a capacitance of 10^{-8} F/cm^2 at an applied voltage of 0.5 V. Find the donor density.
7. A silicon ($n_i = 10^{10} \text{ cm}^{-3}$) p-n diode has a maximum electric field of -10^6 V/cm and a depletion layer width of $1 \mu\text{m}$. The acceptor density in the p-type region is four times larger than the donor density in the n-type region. Calculate both doping densities.
8. Consider a symmetric silicon p-n diode ($N_a = N_d$)

- a. Calculate the built-in potential if $N_a = 10^{13}$, 10^{15} and 10^{17} cm^{-3} . Also, calculate the doping densities corresponding to a built-in potential of 0.7 V.
 - b. For the same as in part a), calculate the total depletion layer widths, the capacitance per unit area and the maximum electric field in thermal equilibrium.
 - c. For the same as in part a), calculate the total depletion layer widths, the capacitance per unit area and the maximum electric field in thermal equilibrium.
 - d. Repeat part a) and b) with $N_a = 3 N_d$.
9. A one-sided silicon diode has a breakdown voltage of 1000 V for which the maximum electric field at breakdown is 100 kV/cm. What is the maximum possible doping density in the low doped region, the built-in potential, the depletion layer width and the capacitance per unit area? Assume that bulk potential of the highly doped region is $E_g/2$ (= 0.56 V).
10. A silicon p-n junction ($N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$) is biased with $V_a = 0.6$ V. Calculate the ideal diode current assuming that the n-type region is much smaller than the diffusion length with $w_n' = 1 \mu\text{m}$ and assuming a "long" p-type region. Use $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$. The minority carrier lifetime is 10 μs and the diode area is 100 μm by 100 μm .
11. Derive equation [4.4.24](#).
12. Calculate the relative error when using the "short diode" approximation if $L_n = 2 w_p'$ and $L_p = 2 w_n'$.
13. A silicon p-n junction ($N_a = 10^{15} \text{ cm}^{-3}$, $w_p = 1 \mu\text{m}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$, $w_n = 1 \mu\text{m}$) is biased with $V_a = 0.5$ V. Use $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$. The minority carrier lifetime is 10 μs and the diode area is 100 μm by 100 μm .
- a. Calculate the built-in potential of the diode.
 - b. Calculate the depletion layer widths, x_n and x_p , and the widths of the quasi-neutral regions.
 - c. Compare the width of the quasi-neutral regions with the minority-carrier diffusion-lengths and decide whether to use the "long" or "short" diode approximation. Calculate the current through the diode.
 - d. Compare the result of part c) with the current obtained by using the general solution (equation [4.4.24](#))
 - e. Using the approximation chosen in part c) calculate the ratio of the electron current to the hole current traversing the depletion region.
14. An abrupt silicon p-n diode consists of a p-type region containing 10^{18} cm^{-3} acceptors and an n-type region containing 10^{15} cm^{-3} donors.
- a. Calculate the breakdown field in the n-type region.
 - b. Using the breakdown field from part a), calculate the breakdown voltage of the diode.
 - c. What is the depletion layer width at breakdown?
 - d. Discuss edge effects and specify the minimum junction depth needed to avoid these effects.
15. A 1 cm^2 solar cell consists of a p-type region containing 10^{18} cm^{-3} acceptors and an n-type region containing 10^{15} cm^{-3} donors. $w_p' = 0.1 \mu\text{m}$ and $w_n \gg L_p$. Use $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$. The minority carrier lifetime is 10 μs . The diode is illuminated with sun light, yielding a photocurrent density of 30 mA/cm².
- a. Calculate the open circuit voltage and short-circuit current of the solar cell.
 - b. Calculate the maximum power generated by the cell and the corresponding voltage and current.
 - c. Calculate the fill factor of the solar cell.
 - d. Calculate the fill factor for the same cell when it is illuminated by a concentrator so that the photocurrent

density equals 300 A/cm^2 .

Problems

1. Consider a gold-GaAs Schottky diode with a capacitance of 1 pF at -1 V. What is the doping density of the GaAs? Also calculate the depletion layer width at zero bias and the field at the surface of the semiconductor at -10 V bias voltage. The area of the diode is 10^{-5} cm².
2. Using the work functions listed in table 3.2.1, predict which metal-semiconductor junctions are expected to be ohmic contacts. Use the ideal interface model.
3. Design a platinum-silicon diode with a capacitance of 1 pF and a maximum electric field less than 10^4 V/cm at -10 V bias. Provide a possible doping density and area. Make sure the diode has an area between 10^{-5} and 10^{-7} cm². Is it possible to satisfy all requirements if the doping density equals 10^{17} cm⁻³?
4. A platinum-silicon diode (area = 10^{-4} cm⁻², $N_d = 10^{17}$ cm⁻³) is part of an LC tuning circuit containing a 100 nH inductance. The applied voltage must be less than 5 V. What is the tuning range of the circuit? The resonant frequency equals $\nu = \frac{1}{2\pi\sqrt{LC}}$, where L is the inductance and C is the diode capacitance.

Problem 3.1 Consider a gold-GaAs Schottky diode with a capacitance of 1 pF at -1 V. What is the doping density of the GaAs? Also calculate the depletion layer width at zero bias and the field at the surface of the semiconductor at -10 V. The area of the diode is 10^{-5} cm^2 .

Solution The depletion layer width can be calculated from the capacitance yielding:

$$x_d = \frac{\epsilon_s A}{C} = \frac{13.1 \times 8.854 \times 10^{-14} \times 10^{-5}}{10^{-12}} = 0.116 \mu\text{m}$$

From this one can find the doping density:

$$N_d = \frac{2\epsilon_s (F_i - V_a)}{qx_d^2} = \frac{2 \times 13.1 \times 8.854 \times 10^{-14} \times (F_i + 1)}{1.6 \times 10^{-19} \times (0.116 \times 10^{-5})^2}$$

Provided one knows the built-in potential

$$F_i = F_B - V_t \ln \frac{N_c}{N_d} = 4.8 - 4.07 - 0.0259 \ln \frac{4.35 \times 10^{17}}{N_d}$$

Which in turn depends on the doping density.

Starting with $F_i = 0.7$ one finds $N_d = 1.83 \times 10^{17} \text{ cm}^{-3}$ and the corresponding built-in potential $F_i = 0.708$. Further iteration yields the result: $N_d = 1.84 \times 10^{17} \text{ cm}^{-3}$.

The depletion layer width at zero bias equals:

$$x_d = \sqrt{\frac{2\epsilon_s (F_i - V_a)}{qN_d}} = \sqrt{\frac{2 \times 13.1 \times 8.854 \times 10^{-14} \times (0.708 - 0)}{1.6 \times 10^{-19} \times 1.84 \times 10^{17}}} = 0.075 \mu\text{m}$$

And the electric field at the surface for $V_a = -10 \text{ V}$ equals:

$$E(x=0) = \frac{qN_d x_d}{\epsilon_s} = \sqrt{\frac{2(F_i - V_a)qN_d}{\epsilon_s}} = \sqrt{\frac{2 \times 10.7 \times 1.6 \times 10^{-19} \times 1.84 \times 10^{17}}{13.1 \times 8.854 \times 10^{-14}}} = 737 \text{ kV/cm}$$



Chapter 3: Metal-Semiconductor Junctions

3.2. Structure and principle of operation

- [3.2.1. Structure](#)
- [3.2.2. Flatband diagram and built-in potential](#)
- [3.2.3. Thermal equilibrium](#)
- [3.2.4. Forward and reverse bias](#)

3.2.1. Structure



The structure of a metal-semiconductor junction is shown in Figure 3.2.1. It consists of a metal contacting a piece of semiconductor. An ideal Ohmic contact, a contact such that no potential exists between the metal and the semiconductor, is made to the other side of the semiconductor. The sign convention of the applied voltage and current is also shown on Figure 3.2.1.

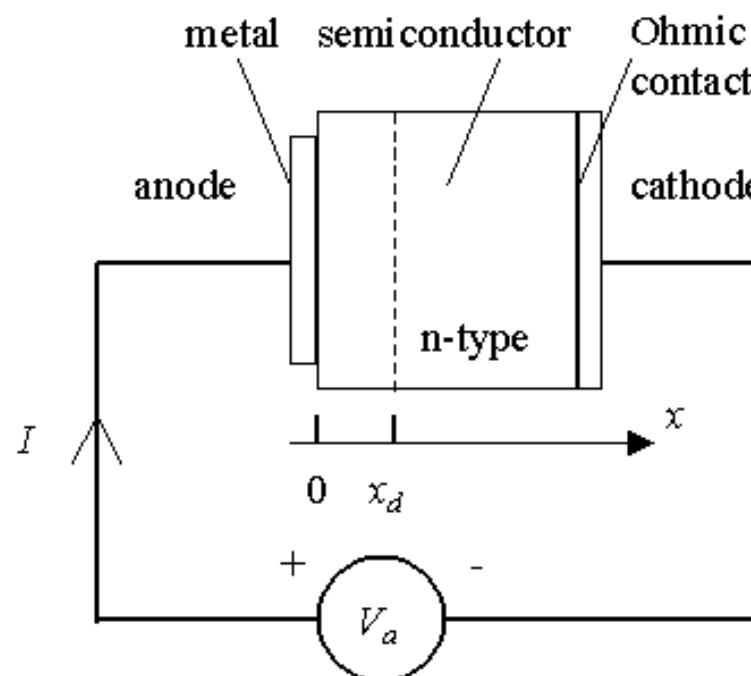


Figure 3.2.1 : Structure and sign convention of a metal-semiconductor junction

3.2.2. Flatband diagram and built-in potential



The barrier between the metal and the semiconductor can be identified on an energy band diagram. To construct such diagram we first consider the energy band diagram of the metal and the semiconductor, and align them using the same vacuum level as shown in Figure 3.2.2 (a). As the metal and semiconductor are brought together, the Fermi energies of the metal and the semiconductor do not change right away. This yields the flatband diagram of Figure 3.2.2 (b).

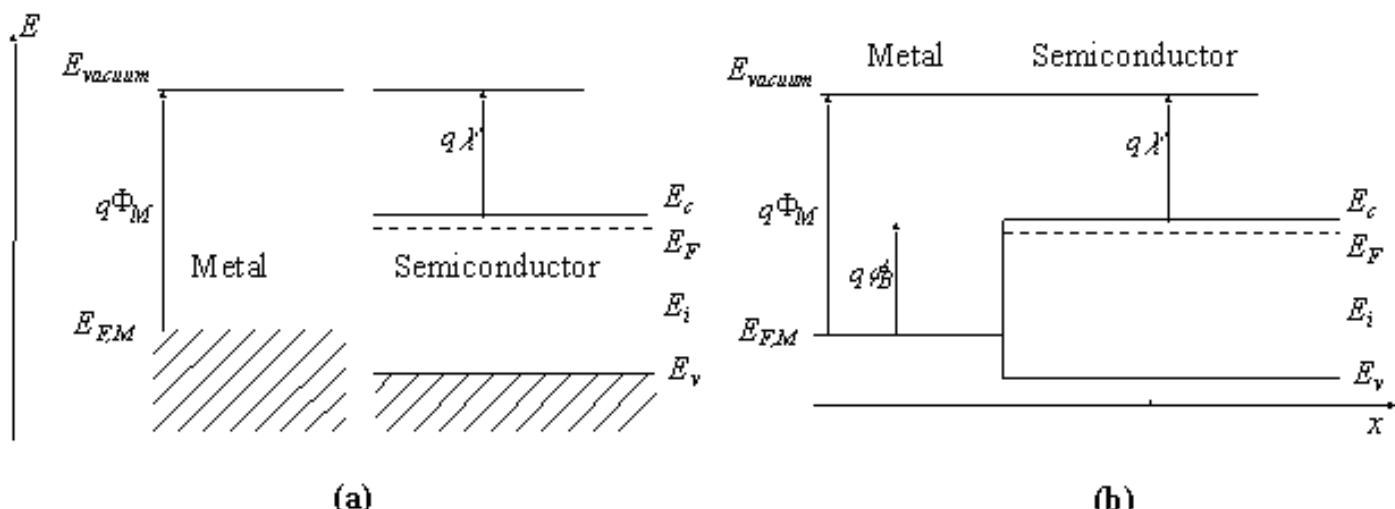


Figure 3.2.2 : Energy band diagram of the metal and the semiconductor before (a) and after (b) contact is made.

The barrier height, ϕ_B , is defined as the potential difference between the Fermi energy of the metal and the band edge where the majority carriers reside. From Figure 3.2.2 (b) one finds that for an n-type semiconductor the barrier height is obtained from:

$$\phi_B = \Phi_M - \chi \text{, for an n - type semiconductor} \quad (3.2.1)$$

Where Φ_M is the work function of the metal and χ is the electron affinity. The work function of selected metals as measured in vacuum can be found in Table 3.2.1. For p-type material, the barrier height is given by the difference between the valence band edge and the Fermi energy in the metal:

$$\phi_B = \frac{E_g}{q} + \chi - \Phi_M \text{, for a p - type semiconductor} \quad (3.2.2)$$

A metal-semiconductor junction will therefore form a barrier for electrons and holes if the Fermi energy of the metal as drawn on the flatband diagram is somewhere between the conduction and valence band edge.

In addition, we define the built-in potential, ϕ_I , as the difference between the Fermi energy of the metal and that of the semiconductor.

$$\phi_I = \Phi_M - \chi - \frac{E_c - E_{F,n}}{q} \text{, n - type} \quad (3.2.3)$$

$$\phi_I = \chi + \frac{E_c - E_{F,p}}{q} - \Phi_M \text{, p - type} \quad (3.2.4)$$

The measured barrier height for selected metal-semiconductor junctions is listed in Table 3.2.1. These experimental barrier heights often differ from the ones calculated using (3.2.1) or (3.2.2). This is due to the detailed behavior of the metal-semiconductor interface. The ideal metal-semiconductor theory assumes that both materials are infinitely pure, that there is no interaction between the two materials nor is there an interfacial layer. Chemical reactions between the metal and the semiconductor alter the barrier height as do interface states at the surface of the semiconductor and interfacial layers. Some general trends however can still be observed. As predicted by (3.2.1), the barrier height on n-type semiconductors increases for metals with a higher work function as can be verified for silicon. Gallium arsenide on the other hand is known to have a large density of surface states so that the barrier height becomes virtually independent of the metal. Furthermore, one finds the barrier heights reported in the literature to vary widely due to different surface cleaning procedures.

	Ag	Al	Au	Cr	Ni	Pt	W
Φ_M (in vacuum)	4.3	4.25	4.8	4.5	4.5	5.3	4.6
n-Ge	0.54	0.48	0.59		0.49		0.48
p-Ge	0.5		0.3				
n-Si	0.78	0.72	0.8	0.61	0.61	0.9	0.67
p-Si	0.54	0.58	0.34	0.5	0.51		0.45
n-GaAs	0.88	0.8	0.9			0.84	0.8
p-GaAs	0.63		0.42				

Table 3.2.1: Workfunction of selected metals and their measured barrier height on germanium, silicon and gallium arsenide.

Example 3.1	Consider a chrome-silicon metal-semiconductor junction with $N_d = 10^{17} \text{ cm}^{-3}$. Calculate the barrier height and the built-in potential. Repeat for a p-type semiconductor with the same doping density.
Solution	<p>The barrier height equals:</p> $\phi_B = \Phi_M - \chi = 4.5 - 4.05 = 0.45 \text{ V}$ <p>Note that this value differs from the one listed in Table 3.2.1 since the work function in vacuum was used. See the discussion in the text for more details.</p> <p>The built-in potential equals:</p> $\phi = \phi_B - V_t \ln \frac{N_c}{N_d} = 0.45 - 0.0259 \ln \frac{2.82 \times 10^{19}}{10^{17}} = 0.30 \text{ V}$ <p>The barrier height for the chrome/p-silicon junction equals:</p> $\phi_B = \chi + \frac{E_g}{q} - \Phi_M = 4.05 + 1.12 - 4.5 = 0.67 \text{ V}$ <p>And the built-in potential equals:</p>

$$\phi_i = \phi_B - V_f \ln \frac{N_v}{N_a} = 0.67 - 0.0259 \ln \frac{1.83 \times 10^{19}}{10^{17}} = 0.53 \text{ V}$$



3.2.3. Thermal equilibrium

The flatband diagram, shown in Figure 3.2.2 (b), is not a thermal equilibrium diagram, since the Fermi energy in the metal differs from that in the semiconductor. Electrons in the n-type semiconductor can lower their energy by traversing the junction. As the electrons leave the semiconductor, a positive charge, due to the ionized donor atoms, stays behind. This charge creates a negative field and lowers the band edges of the semiconductor. Electrons flow into the metal until equilibrium is reached between the diffusion of electrons from the semiconductor into the metal and the drift of electrons caused by the field created by the ionized impurity atoms. This equilibrium is characterized by a constant Fermi energy throughout the structure.

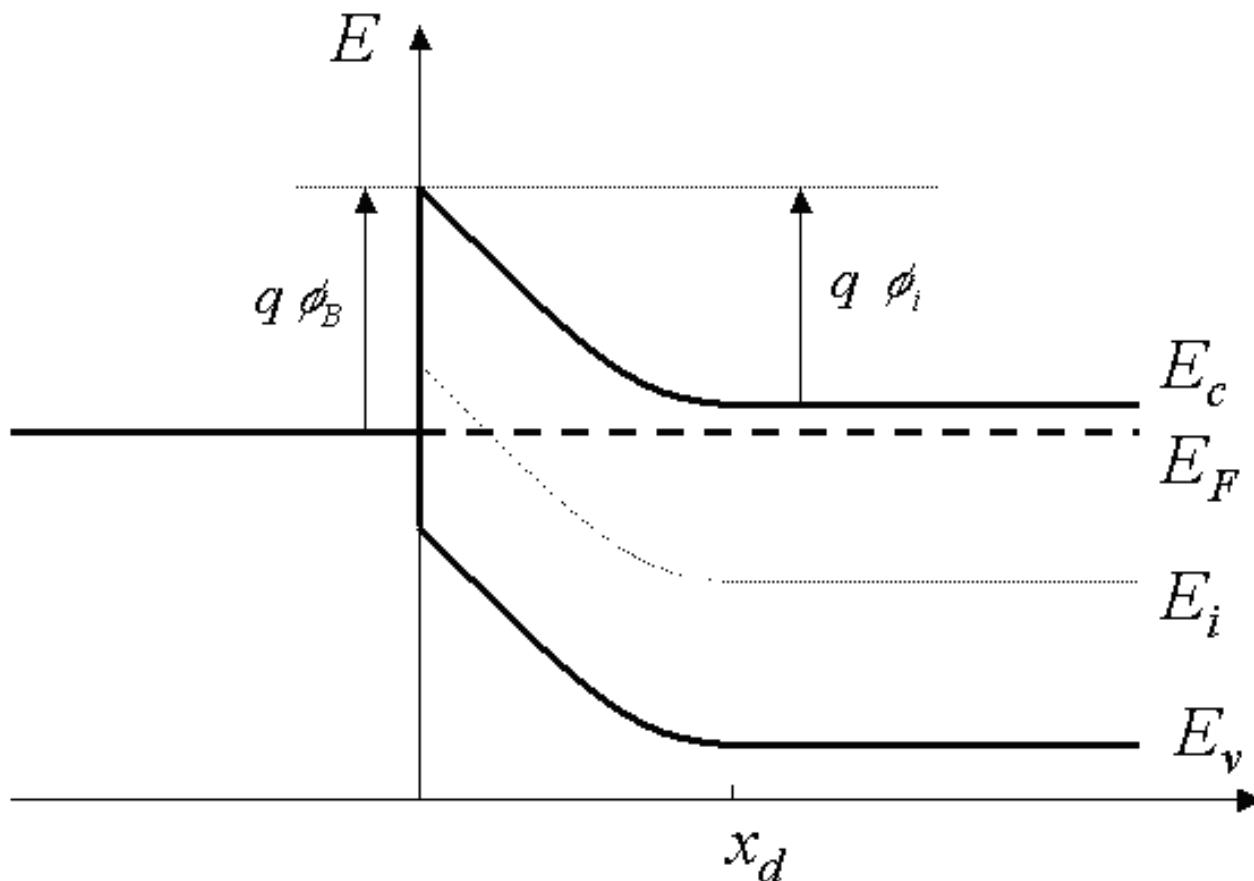


Figure 3.2.3 : Energy band diagram of a metal-semiconductor contact in thermal equilibrium.

It is of interest to note that in thermal equilibrium, i.e. with no external voltage applied, there is a region in the semiconductor close to the junction (), which is depleted of mobile carriers. We call this the depletion region. The potential across the semiconductor equals the built-in potential, ϕ_i .



3.2.4. Forward and reverse bias

Operation of a metal-semiconductor junction under forward and reverse bias is illustrated with Figure 3.2.4. As a positive bias is applied to the metal (Figure 3.2.4 (a)), the Fermi energy of the metal is lowered with respect to the Fermi energy in the semiconductor. This results in a smaller potential drop across the semiconductor. The balance between diffusion and drift is disturbed and more electrons will diffuse towards the metal than the number drifting into the semiconductor. This leads to a positive current through the junction at a voltage comparable to the built-in potential.

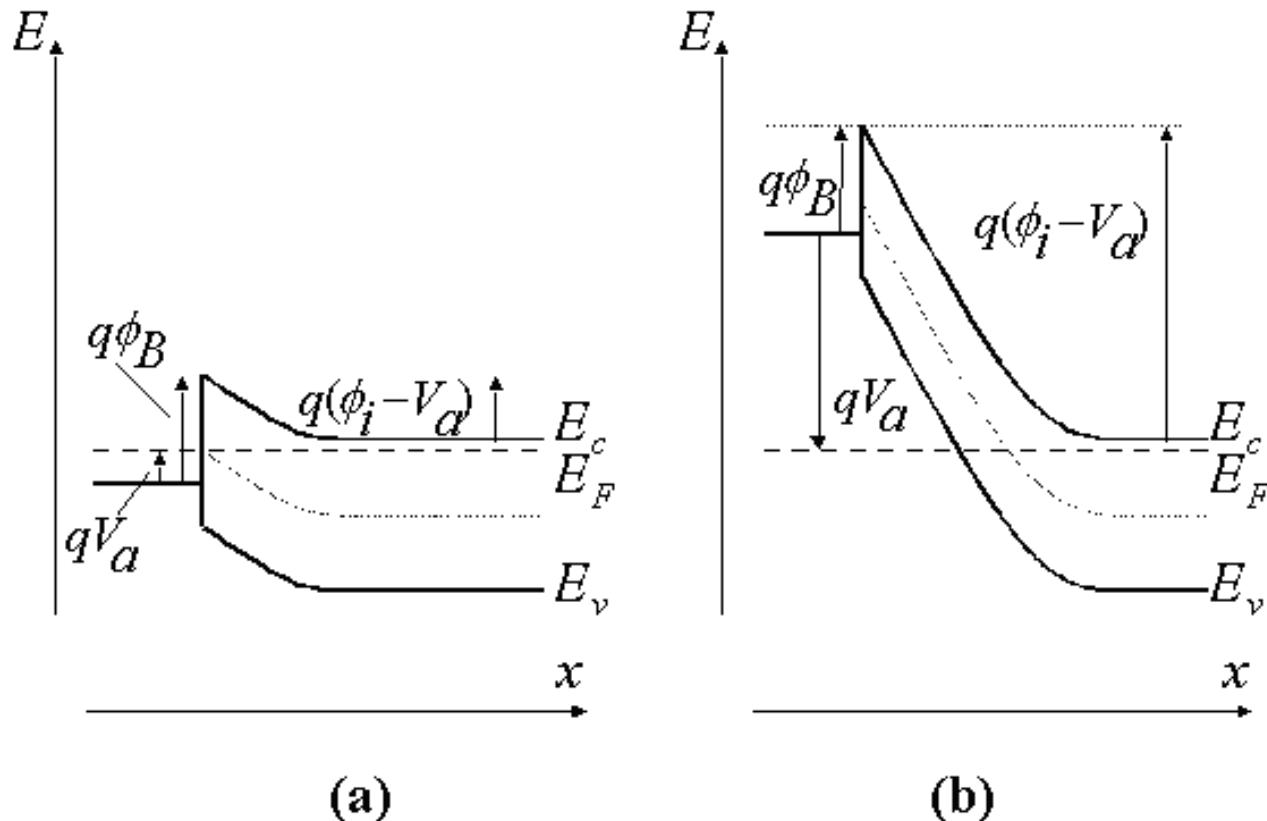


Figure 3.2.4 : Energy band diagram of a metal-semiconductor junction under (a) forward and (b) reverse bias

As a negative voltage is applied (Figure 3.2.4 (b)), the Fermi energy of the metal is raised with respect to the Fermi energy in the semiconductor. The potential across the semiconductor now increases, yielding a larger depletion region and a larger electric field at the interface. The barrier, which restricts the electrons to the metal, is unchanged so that the flow of electrons is limited by that barrier independent of the applied voltage. The metal-semiconductor junction with positive barrier height has therefore a pronounced rectifying behavior. A large current exists under forward bias, while almost no current exists under reverse bias.

The potential across the semiconductor therefore equals the built-in potential, ϕ_i , minus the applied voltage, V_a .

$$\phi(x = \infty) - \phi(x = 0) = \phi_i - V_a \quad (3.2.5)$$

Chapter 4: p-n Junctions



Examples

Example 4.1

An abrupt silicon p-n junction consists of a p-type region containing $2 \times 10^{16} \text{ cm}^{-3}$ acceptors and an n-type region containing also 10^{16} cm^{-3} acceptors in addition to 10^{17} cm^{-3} donors.

Example 4.2

An abrupt silicon ($n_I = 10^{10} \text{ cm}^{-3}$) p-n junction consists of a p-type region containing 10^{16} cm^{-3} acceptors and an n-type region containing $5 \times 10^{16} \text{ cm}^{-3}$ donors.

- Calculate the built-in potential of this p-n junction.
- Calculate the total width of the depletion region if the applied voltage V_a equals 0, 0.5 and -2.5 V.
- Calculate maximum electric field in the depletion region at 0, 0.5 and -2.5 V.
- Calculate the potential across the depletion region in the n-type semiconductor at 0, 0.5 and -2.5 V.

Example 4.3

Consider an abrupt p-n diode with $N_a = 10^{18} \text{ cm}^{-3}$ and $N_d = 10^{16} \text{ cm}^{-3}$. Calculate the junction capacitance at zero bias. The diode area equals 10^{-4} cm^2 . Repeat the problem while treating the diode as a one-sided diode and calculate the relative error.

Example 4.4

An abrupt silicon p-n junction ($N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$) is biased with $V_a = 0.6 \text{ V}$. Calculate the ideal diode current assuming that the n-type region is much smaller than the diffusion length with $w_n' = 1 \mu\text{m}$ and assuming a "long" p-type region. Use $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$. The minority carrier lifetime is $10 \mu\text{s}$ and the diode area is $100 \mu\text{m}$ by $100 \mu\text{m}$.

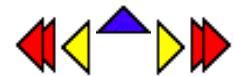
Example 4.5

- Calculate the diffusion capacitance of the diode described in Example 4.4 at zero bias. Use $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$, $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$, $w_p' = 1 \mu\text{m}$ and $w_n' = 1 \text{ mm}$. The minority carrier lifetime equals 0.1 ms.
- For the same diode, find the voltage for which the junction capacitance equals the diffusion capacitance.

Example 4.6

A 1 cm^2 silicon solar cell has a saturation current of 10^{-12} A and is illuminated with sunlight yielding a short-circuit photocurrent of 25 mA. Calculate the solar cell efficiency and fill factor.

Chapter 5: Bipolar Junction Transistors



Examples

Example 5.1

A bipolar transistor with an emitter current of 1 mA has an emitter efficiency of 0.99, a base transport factor of 0.995 and a depletion layer recombination factor of 0.998. Calculate the base current, the collector current, the transport factor and the current gain of the transistor.

Example 5.2

Consider a pnp bipolar transistor with emitter doping of 10^{18} cm^{-3} and base doping of 10^{17} cm^{-3} . The quasi-neutral region width in the emitter is $1 \mu\text{m}$ and $0.2 \mu\text{m}$ in the base. Use $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$. The minority carrier lifetime in the base is 10 ns. Calculate the emitter efficiency, the base transport factor, and the current gain of the transistor biased in the forward active mode. Assume there is no recombination in the depletion region.

Example 5.3

Calculate the saturation voltage of a bipolar transistor biased with a base current of 1 mA and a collector current of 10 mA. Use $\alpha_R = 0.993$ and $\alpha_F = 0.2$.

Example 5.4

Consider a bipolar transistor with a base doping of 10^{17} cm^{-3} and a quasi-neutral base width of $0.2 \mu\text{m}$. Calculate the Early voltage and collector current ideality factor given that the base-emitter capacitance and the base-collector capacitance are 0.2 nF and 0.2 pF . The collector area equals 10^{-4} cm^{-2} .

Example 4.1 An abrupt silicon p-n junction consists of a p-type region containing $2 \times 10^{16} \text{ cm}^{-3}$ acceptors and an n-type region containing also 10^{16} cm^{-3} acceptors in addition to 10^{17} cm^{-3} donors.

- a. Calculate the thermal equilibrium density of electrons and holes in the p-type region as well as both densities in the n-type region.
- b. Calculate the built-in potential of the p-n junction.
- c. Calculate the built-in potential of the p-n junction at 400 K.

Solution

- a. The thermal equilibrium densities are:

In the p-type region:

$$p = N_a = 2 \times 10^{16} \text{ cm}^{-3}$$
$$n = n_i^2/p = 10^{20}/2 \times 10^{16} = 5 \times 10^3 \text{ cm}^{-3}$$

In the n-type region

$$n = N_d - N_a = 9 \times 10^{16} \text{ cm}^{-3}$$
$$p = n_i^2/n = 10^{20}/1 \times 10^{16} = 1.11 \times 10^3 \text{ cm}^{-3}$$

- b. The built-in potential is obtained from

$$F_i = V_t \ln \frac{p_p n_n}{n_i^2} = 0.0259 \ln \frac{2 \times 10^{16} \times 9 \times 10^{16}}{10^{20}} = 0.79 \text{ V}$$

- c. Similarly, the built-in potential at 400 K equals

$$F_i = V_t \ln \frac{p_p n_n}{n_i^2} = 0.0345 \ln \frac{2 \times 10^{16} \times 9 \times 10^{16}}{(4.52 \times 10^{12})^2} = 0.63 \text{ V}$$

where the intrinsic carrier density at 400 K was obtained from Example 2.4 b.

Example 4.2 An abrupt silicon ($n_i = 10^{10} \text{ cm}^{-3}$) p-n junction consists of a p-type region containing 10^{16} cm^{-3} acceptors and an n-type region containing $5 \times 10^{16} \text{ cm}^{-3}$ donors.

- Calculate the built-in potential of this p-n junction.
- Calculate the total width of the depletion region if the applied voltage V_a equals 0, 0.5 and -2.5 V.
- Calculate maximum electric field in the depletion region at 0, 0.5 and -2.5 V.
- Calculate the potential across the depletion region in the n-type semiconductor at 0, 0.5 and -2.5 V.

The built-in potential is calculated from:

$$F_i = V_t \ln \frac{p_n n_p}{n_i^2} = 0.0259 \ln \frac{10^{16} \times 5 \times 10^{16}}{10^{20}} = 0.76 \text{ V}$$

The depletion layer width is obtained from:

$$w = \sqrt{\frac{2e_s}{q} \left(\frac{1}{N_a} + \frac{1}{N_d} \right) (F_i - V_a)}$$

the electric field from

$$E(x=0) = -\frac{2(F_i - V_a)}{w}$$

and the potential across the n-type region equals

$$F_n = \frac{q N_d x_n^2}{2e_s}$$

where

$$x_n = w \frac{N_a}{N_a + N_d}$$

one can also show that

$$F_n = \frac{(F_i - V_a) N_a}{N_a + N_d}$$

This yields the following numeric values:

	$V_a = 0 \text{ V}$	$V_a = 0.5 \text{ V}$	$V_a = -2.5 \text{ V}$
w	0.315 μm	0.143 μm	0.703 μm
E	40 kV/cm	18 kV/cm	89 kV/cm
F_n	0.105 V	0.0216 V	0.522 V

Example 4.3 An abrupt silicon p-n junction ($N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$) is biased with $V_a = 0.6 \text{ V}$. Calculate the ideal diode current assuming that the n-type region is much smaller than the diffusion length with $w_n = 1 \mu\text{m}$ and assuming a "long" p-type region. Use $\mathbf{m}_n = 1000 \text{ cm}^2/\text{V-s}$ and $\mathbf{m}_p = 300 \text{ cm}^2/\text{V-s}$. The minority carrier lifetime is $10 \mu\text{s}$ and the diode area is $100 \mu\text{m}$ by $100 \mu\text{m}$.

The current is calculated from:

$$I = qA \left[\frac{D_n n_{p0}}{L_n} + \frac{D_p p_{n0}}{w_n} \right] (e^{V_a/V_t} - 1)$$

with

$$D_n = \mathbf{m}_n V_t = 1000 \times 0.0258 = 25.8 \text{ cm}^2/\text{V-s}$$

$$D_p = \mathbf{m}_p V_t = 300 \times 0.0258 = 7.75 \text{ cm}^2/\text{V-s}$$

$$n_{p0} = n_i^2 / N_a = 10^{20} / 10^{16} = 10^4 \text{ cm}^{-3}$$

$$p_{n0} = n_i^2 / N_d = 10^{20} / 4 \times 10^{16} = 2.5 \times 10^3 \text{ cm}^{-3}$$

$$L_n = \sqrt{D_n t_n} = \sqrt{25.8 \times 10^{-5}} = 161 \mu\text{m}$$

yielding $I = 40.7 \mu\text{A}$

Note that the hole diffusion current occurs in the "short" n-type region and therefore depends on the quasi-neutral width in that region. The electron diffusion current occurs in the "long" p-type region and therefore depends on the electron diffusion length in that region.

Example 4.4 Consider an abrupt p-n diode with $N_a = 10^{18} \text{ cm}^{-3}$ and $N_d = 10^{16} \text{ cm}^{-3}$. Calculate the junction capacitance at zero bias. The diode area equals 10^{-4} cm^2 . Repeat the problem while treating the diode as a one-sided diode and calculate the relative error.

Solution

The built in potential of the diode equals:

$$f_i = V_t \ln \frac{N_d N_a}{n_i^2} = 0.83 \text{ V}$$

The depletion layer width at zero bias equals:

$$w = \sqrt{\frac{2e_s (f_i - 0)}{qN_d}} = 0.33 \text{ } \mu\text{m}$$

And the junction capacitance at zero bias equals:

$$C_{j0} = \left. \frac{e_s}{w} \right|_{V_a=0} = 3.17 \text{ pF}$$

Repeating the analysis while treating the diode as a one-sided diode, one only has to consider the region with the lower doping density so that

$$w = x_n = \sqrt{\frac{2e_s}{qN_d} (f_i - V_a)} = 0.31 \text{ } \mu\text{m}$$

And the junction capacitance at zero bias equals

$$C_{j0} = \left. \frac{e_s}{w} \right|_{V_a=0} = 3.18 \text{ pF}$$

The relative error equals 0.5 %, which justifies the use of the one-sided approximation.

Example 4.5 a. Calculate the diffusion capacitance of the diode described in Example 4.4 at zero bias. Use $m_i = 1000 \text{ cm}^2/\text{V-s}$, $m_p = 300 \text{ cm}^2/\text{V-s}$, $w_p' = 1 \mu\text{m}$ and $w_n' = 1 \text{ mm}$. The minority carrier lifetime equals 0.1 ms.

b. For the same diode, find the voltage for which the junction capacitance equals the diffusion capacitance.

Solution a. The diffusion capacitance at zero volts equals

$$C_{d,0} = \frac{I_{s,p} t_p}{V_t} + \frac{I_{s,n} t_{r,n}}{V_t} = 1.73 \times 10^{-19} \text{ F}$$

using

$$I_{s,p} = q \frac{A p_{n0} D_p}{L_p}$$

and

$$I_{s,n} = q \frac{A n_{p0} D_n}{w_p}$$

Where the "short" diode expression was used for the capacitance associated with the excess charge due to electrons in the p-type region. The "long" diode expression was used for the capacitance associated with the excess charge due to holes in the n-type region.

The diffusion constants and diffusion lengths equal

$$D_n = m_i \times V_t = 25.8 \text{ cm}^2/\text{s}$$

$$D_p = m_p \times V_t = 7.75 \text{ cm}^2/\text{s}$$

$$L_p = \sqrt{D_p t_p}$$

And the electron transit time in the p-type region equals

$$t_{r,n} = \frac{w_p^2}{2D_n} = 193 \text{ ps}$$

b. The voltage at which the junction capacitance equals the diffusion capacitance is obtained by solving

$$\frac{C_{j0}}{\sqrt{1 - \frac{V_a}{f_i}}} = C_{d,0} e^{V_a/V_t}$$

yielding $V_a = 0.442 \text{ V}$

:

Example 4.6 A 1 cm^2 silicon solar cell has a saturation current of 10^{-12} A and is illuminated with sunlight yielding a short-circuit photocurrent of 25 mA . Calculate the solar cell efficiency and fill factor.

Solution

The maximum power is generated for:

$$\frac{dP}{dV_a} = 0 = I_s (e^{V_m/V_t} - 1) - I_{ph} + \frac{V_m}{V_t} I_s e^{V_m/V_t}$$

where the voltage, V_m , is the voltage corresponding to the maximum power point. This voltage is obtained by solving the following transcendental equation:

$$V_m = V_t \ln \frac{1 + I_{ph}/I_s}{1 + V_m/V_t}$$

Using iteration and a starting value of 0.5 V one obtains the following successive values for V_m :

$$V_m = 0.5, 0.542, 0.540 \text{ V}$$

and the efficiency equals:

$$h = \left| \frac{V_m I_m}{P_{in}} \right| = \frac{0.54 \times 0.024}{0.1} = 13 \%$$

The current, I_m , corresponding to the voltage, V_m , was calculated using equation (4.6.1) and the power of the sun was assumed 100 mW/cm^2 . The fill factor equals:

$$\text{fill factor} = \frac{V_m I_m}{V_{oc} I_{sc}} = \frac{0.54 \times 0.024}{0.62 \times 0.025} = 83 \%$$

where the open circuit voltage is calculated using equation (4.6.1) and $I = 0$. The short circuit current equals the photocurrent.

Example 3.1 Consider a chrome-silicon metal-semiconductor junction with $N_d = 10^{17} \text{ cm}^{-3}$. Calculate the barrier height and the built-in potential. Repeat for a p-type semiconductor with the same doping density.

Solution The barrier height equals:

$$f_B = \Phi_M - c = 4.5 - 4.05 = 0.45 \text{ V}$$

Note that this value differs from the one listed in Table 3.2.1 since the work function in vacuum was used. See the discussion in the text for more details.

The built-in potential equals:

$$f_i = f_B - V_t \ln \frac{N_c}{N_d} = 0.45 - 0.0259 \times \ln \frac{2.82 \times 10^{19}}{10^{17}} = 0.30 \text{ V}$$

The barrier height for the chrome/p-silicon junction equals:

$$f_B = c + \frac{E_g}{q} - \Phi_M = 4.05 + 1.12 - 4.5 = 0.67 \text{ V}$$

And the built-in potential equals:

$$f_i = f_B - V_t \ln \frac{N_v}{N_a} = 0.67 - 0.0259 \times \ln \frac{1.83 \times 10^{19}}{10^{17}} = 0.53 \text{ V}$$

Example 3.2 Consider a chrome-silicon metal-semiconductor junction with $N_d = 10^{17} \text{ cm}^{-3}$. Calculate the depletion layer width, the electric field in the silicon at the metal-semiconductor interface, the potential across the semiconductor and the capacitance per unit area for an applied voltage of -5 V.

Solution The depletion layer width equals:

$$x_d = \sqrt{\frac{2\epsilon_s(f_i - V_a)}{qN_d}}$$

$$= \sqrt{\frac{2 \times 11.9 \times 8.85 \times 10^{-14} \times (0.3 + 5)}{1.6 \times 10^{-19} \times 10^{17}}} = 0.26 \text{ mm}$$

where the built-in potential was already calculated in Example 3.1. The electric field in the semiconductor at the interface is:

$$E(x=0) = \frac{qN_d x_d}{\epsilon_s}$$

$$= \frac{1.6 \times 10^{-19} \times 10^{17} \times 2.6 \times 10^{-5}}{11.9 \times 8.85 \times 10^{-14}} = 4.0 \times 10^5 \text{ V/cm}$$

The potential equals:

$$f(x=x_d) = \frac{qN_d x_d^2}{2\epsilon_s} = f_i - V_a = 5.3 \text{ V}$$

And the capacitance per unit area is obtained from:

$$C_j = \frac{\epsilon_s}{x_d} = \frac{11.9 \times 8.85 \times 10^{-14}}{2.6 \times 10^{-5}} = 40 \text{ nF/cm}^2$$

Chapter 3: Metal-Semiconductor Junctions



3.3. Electrostatic analysis

3.3.1. General discussion - Poisson's equation

3.3.2. Full depletion approximation

3.3.3. Full depletion analysis

3.3.4. Junction capacitance

3.3.5. Schottky barrier lowering



3.3.1. General discussion - Poisson's equation

The electrostatic analysis of a metal-semiconductor junction is of interest since it provides knowledge about the charge and field in the depletion region. It is also required to obtain the capacitance-voltage characteristics of the diode.

The general analysis starts by setting up Poisson's equation:

$$\frac{d^2\phi}{dx^2} = -\frac{\rho}{\epsilon_s} = -\frac{q}{\epsilon_s}(p - n + N_d^+ - N_a^-) \quad (3.3.1)$$

where the charge density, ρ , is written as a function of the electron density, the hole density and the donor and acceptor densities. To solve the equation, we have to express the electron and hole density, n and p , as a function of the potential, ϕ , yielding:

$$\frac{d^2\phi}{dx^2} = \frac{2qn_i}{\epsilon_s} \left(\sinh \frac{\phi - \phi_F}{V_t} + \sinh \frac{\phi_F}{V_t} \right) \quad (3.3.2)$$

with

$$\sinh \frac{\phi_F}{V_t} = \frac{N_a^- - N_d^+}{2n_i} \quad (3.3.3)$$

where the potential is chosen to be zero in the n-type region, where $x \gg x_n$.

This second-order non-linear differential equation (3.3.2) can not be solved analytically. Instead we will make the simplifying assumption that the depletion region is fully depleted and that the adjacent neutral regions contain no charge. This full depletion approximation is the topic of section 3.3.2.



3.3.2. Full depletion approximation

The simple analytic model of the metal-semiconductor junction is based on the full depletion approximation. This approximation is obtained by assuming that the semiconductor is fully depleted over a distance x_d , called the depletion region. While this assumption does not provide an accurate charge distribution, it does provide very reasonable approximate expressions for the electric field and potential throughout the semiconductor. These are derived in section [3.3.3](#).



3.3.3. Full depletion analysis

We now apply the full depletion approximation to an M-S junction containing an n-type semiconductor. We define the depletion region to be between the metal-semiconductor interface ($x = 0$) and the edge of the depletion region ($x = x_d$). The depletion layer width, x_d , is unknown at this point but will later be expressed as a function of the applied voltage.

To find the depletion layer width, we start with the charge density in the semiconductor and calculate the electric field and the potential across the semiconductor as a function of the depletion layer width. We then solve for the depletion layer width by requiring the potential across the semiconductor to equal the difference between the built-in potential and the applied voltage, $\phi_i - V_a$. The different steps of the analysis are illustrated by Figure [3.3.1](#).

As the semiconductor is depleted of mobile carriers within the depletion region, the charge density in that region is due to the ionized donors. Outside the depletion region, the semiconductor is assumed neutral. This yields the following expressions for the charge density, ρ :

$$\begin{aligned} \rho(x) &= qN_d & 0 < x < x_d \\ \rho(x) &= 0 & x_d < x \end{aligned} \quad (3.3.4)$$

where we assumed full ionization so that the ionized donor density equals the donor density, N_d . This charge density is shown in Figure [3.3.1](#) (a). The charge in the semiconductor is exactly balanced by the charge in the metal, Q_M , so that no electric field exists except around the metal-semiconductor interface.

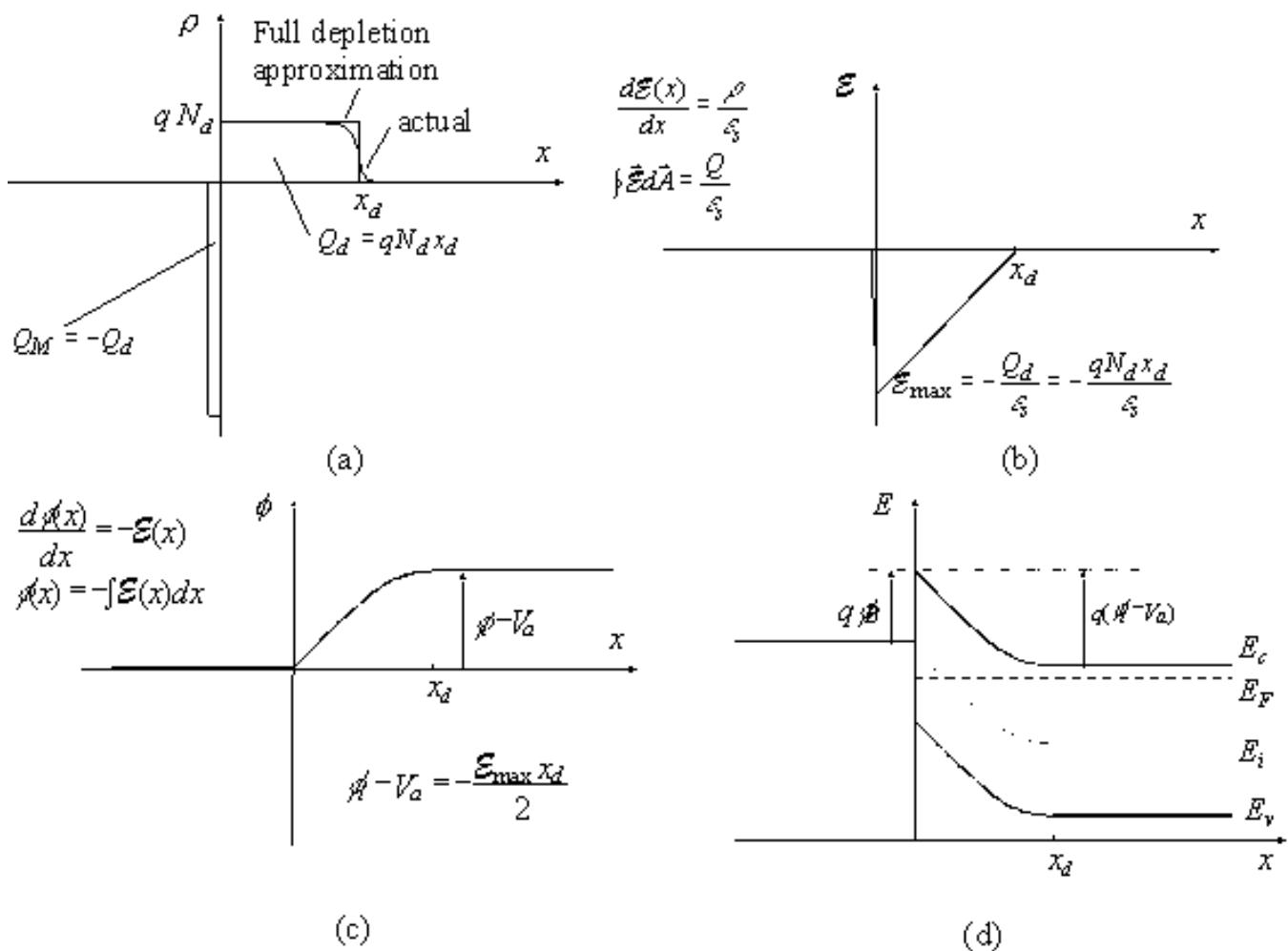


Figure 3.3.1 : (a) Charge density, (b) electric field, (c) potential and (d) energy as obtained with the full depletion analysis.

Using Gauss's law we obtain electric field as a function of position, also shown in Figure 3.3.1 (b):

$$\begin{aligned} \epsilon(x) &= -\frac{qN_d}{\epsilon_s} (x_d - x) & 0 < x < x_d \\ \epsilon(x) &= 0 & x_d \leq x \end{aligned} \quad (3.3.5)$$

where ϵ_s is the dielectric constant of the semiconductor. We also assumed that the electric field is zero outside the depletion region. It is expected to be zero there since a non-zero field would cause the mobile carriers to redistribute until there is no field. The depletion region does not contain mobile carriers so that there can be an electric field. The largest (absolute) value of the electric field is obtained at the interface and is given by:

$$\epsilon(x=0) = -\frac{qN_d x_d}{\epsilon_s} = -\frac{Q_d}{\epsilon_s} \quad (3.3.6)$$

where the electric field was also related to the total charge (per unit area), Q_d , in the depletion layer. Since the electric field is minus the gradient of the potential, one obtains the potential by integrating the expression for the electric field, yielding:

$$\begin{aligned}
 \phi(x) &= 0 & x \leq 0 \\
 \phi(x) &= \frac{qN_d}{2\epsilon_s} [x_d^2 - (x_d - x)^2] & 0 < x < x_d \\
 \phi(x) &= \frac{qN_d x_d^2}{2\epsilon_s} & x_d \leq x
 \end{aligned} \tag{3.3.7}$$

We now assume that the potential across the metal can be neglected. Since the density of free carriers is very high in a metal, the thickness of the charge layer in the metal is very thin. Therefore, the potential across the metal is several orders of magnitude smaller than that across the semiconductor, even though the total amount of charge is the same in both regions.

The total potential difference across the semiconductor equals the built-in potential, ϕ_i , in thermal equilibrium and is further reduced/increased by the applied voltage when a positive/negative voltage is applied to the metal as described by equation (3.2.5). This boundary condition provides the following relation between the semiconductor potential at the surface, the applied voltage and the depletion layer width:

$$\phi_i - V_a = -\phi(x=0) = \frac{qN_d x_d^2}{2\epsilon_s} \tag{3.3.8}$$

Solving this expression for the depletion layer width, x_d , yields:

$$x_d = \sqrt{\frac{2\epsilon_s(\phi_i - V_a)}{qN_d}} \tag{3.3.9}$$

3.3.4. Junction capacitance



In addition, we can obtain the capacitance as a function of the applied voltage by taking the derivative of the charge with respect to the applied voltage yielding:

$$C_j = \left| \frac{dQ_d}{dV_a} \right| = \sqrt{\frac{q\epsilon_s N_d}{2(\phi_i - V_a)}} = \frac{\epsilon_s}{x_d} \tag{3.3.10}$$

The last term in the equation indicates that the expression of a parallel plate capacitor still applies. One can understand this once one realizes that the charge added/removed from the depletion layer as one decreases/increases the applied voltage is added/removed only at the edge of the depletion region. While the parallel plate capacitor expression seems to imply that the capacitance is constant, the metal-semiconductor junction capacitance is not constant since the depletion layer width, x_d , varies with the applied voltage.

Example 3.2

Consider a chrome-silicon metal-semiconductor junction with $N_d = 10^{17} \text{ cm}^{-3}$. Calculate the depletion layer width, the electric field in the silicon at the metal-semiconductor interface, the potential across the semiconductor and the capacitance per unit area for an applied voltage of -5 V.

Solution

The depletion layer width equals:

$$x_d = \sqrt{\frac{2 \epsilon_s (\phi - V_a)}{q N_d}}$$

$$= \sqrt{\frac{2 \times 11.9 \times 8.85 \times 10^{-14} \times (0.3 + 5)}{1.6 \times 10^{-19} \times 10^{17}}} = 0.26 \mu\text{m}$$

where the built-in potential was already calculated in Example 3.1.

The electric field in the semiconductor at the interface is:

$$E(x = 0) = \frac{q N_d x_d}{\epsilon_s}$$

$$= \frac{1.6 \times 10^{-19} \times 10^{17} \times 2.6 \times 10^{-5}}{11.9 \times 8.85 \times 10^{-14}} = 4.0 \times 10^5 \text{ V/cm}$$

The potential equals:

$$\phi(x = x_d) = \frac{q N_d x_d^2}{2 \epsilon_s} = \phi - V_a = 5.3 \text{ V}$$

And the capacitance per unit area is obtained from:

$$C_j = \frac{\epsilon_s}{x_d} = \frac{11.9 \times 8.85 \times 10^{-14}}{2.6 \times 10^{-5}} = 40 \text{ nF/cm}^2$$

3.3.5. Schottky barrier lowering



Image charges build up in the metal electrode of a metal-semiconductor junction as carriers approach the metal-semiconductor interface. The potential associated with these charges reduces the effective barrier height. This barrier reduction tends to be rather small compared to the barrier height itself. Nevertheless this barrier reduction is of interest since it depends on the applied voltage and leads to a voltage dependence of the reverse bias current. Note that this barrier lowering is only experienced by a carrier while approaching the interface and will therefore not be noticeable in a capacitance-voltage measurement.

An energy band diagram of an n-type silicon Schottky barrier including the barrier lowering is shown in Figure 3.3.2:

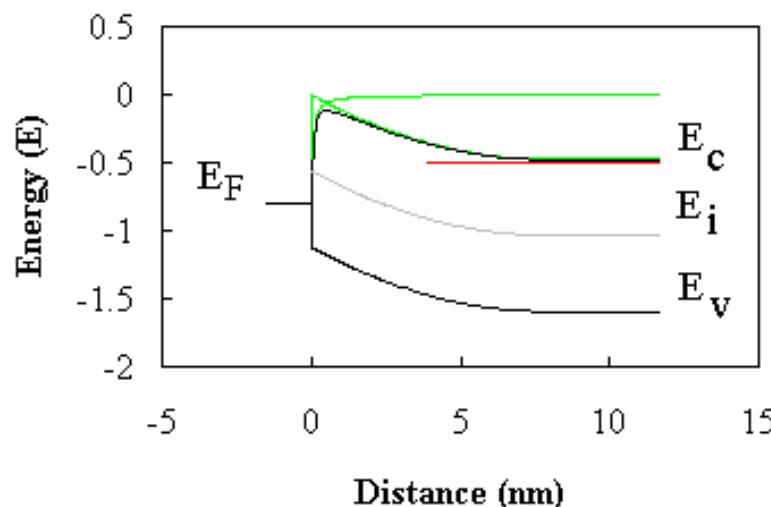


Figure 3.3.2: Energy band diagram of a silicon Schottky barrier with $\phi_B = 0.8$ V and $N_d = 10^{19}$ cm $^{-3}$.

Shown is the energy band diagram obtained using the full-depletion approximation, the potential reduction experienced by electrons, which approach the interface and the resulting conduction band edge. A rounding of the conduction band edge can be observed at the metal-semiconductor interface as well as a reduction of the height of the barrier.

The calculation of the barrier reduction assumes that the charge of an electron close to the metal-semiconductor interface attracts an opposite surface charge, which exactly balances the electron's charge so that the electric field surrounding the electron does not penetrate beyond this surface charge. The time to build-up the surface charge and the time to polarize the semiconductor around the moving electron is assumed to be much shorter than the transit time of the electron. This scenario is based on the assumption that there are no mobile or fixed charges around the electron as it approaches the metal-semiconductor interface. The electron and the induced surface charges are shown in Figure 3.3.3:

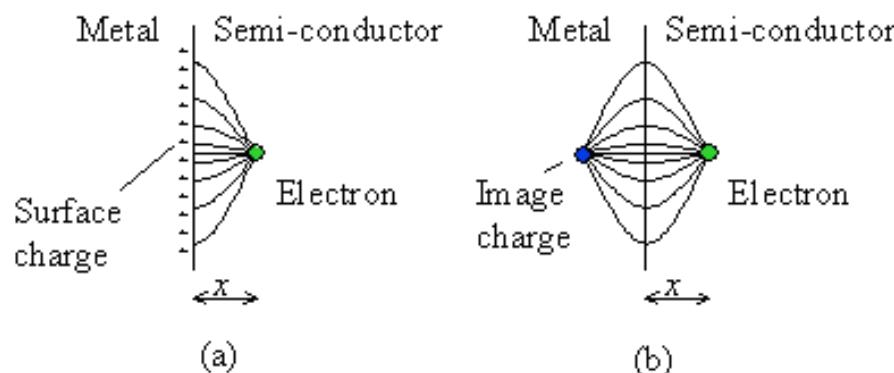
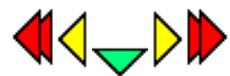


Figure 3.3.3: a) Field lines and surface charges due to an electron in close proximity to a perfect conductor and b) the field lines and image charge of an electron.

It can be shown that the electric field in the semiconductor is identical to that of the carrier itself and another carrier with opposite charge at equal distance but on the opposite side of the interface. This charge is called the image charge. The difference between the actual surface charges and the image charge is that the fields in the metal are distinctly different. The image charge concept is justified on the basis that the electric field lines are perpendicular to the surface of a perfect conductor, so that, in the case of a flat interface, the mirror image of the field lines provides continuous field lines across the interface.

The barrier lowering depends on the square root of the electric field at the interface and is calculated from:

$$\Delta \phi_B = \sqrt{\frac{q\epsilon_{\max}}{4\pi\epsilon_s}} \quad (3.3.11)$$



Chapter 3: Metal-Semiconductor Junctions

3.4. Schottky diode current

3.4.1. Diffusion current

3.4.2. Thermionic emission

3.4.3. Tunneling

The current across a metal-semiconductor junction is mainly due to majority carriers. Three distinctly different mechanisms exist: diffusion of carriers from the semiconductor into the metal, thermionic emission of carriers across the Schottky barrier and quantum-mechanical tunneling through the barrier. The diffusion theory assumes that the driving force is distributed over the length of the depletion layer. The thermionic emission theory on the other hand postulates that only energetic carriers, those, which have an energy equal to or larger than the conduction band energy at the metal-semiconductor interface, contribute to the current flow. Quantum-mechanical tunneling through the barrier takes into account the wave-nature of the electrons, allowing them to penetrate through thin barriers. In a given junction, a combination of all three mechanisms could exist. However, typically one finds that only one limits the current, making it the dominant current mechanism.

The analysis reveals that the diffusion and thermionic emission currents can be written in the following form:

$$J_n = qvN_c \exp\left(-\frac{\phi_B}{V_t}\right) \left(\exp\left(\frac{V_a}{V_t}\right) - 1 \right) \quad (3.4.1)$$

This expression states that the current is the product of the electronic charge, q , a velocity, v , and the density of available carriers in the semiconductor located next to the interface. The velocity equals the mobility multiplied with the field at the interface for the diffusion current and the Richardson velocity (see section 3.4.2) for the thermionic emission current. The minus one term ensures that the current is zero if no voltage is applied as in thermal equilibrium any motion of carriers is balanced by a motion of carriers in the opposite direction.

The tunneling current is of a similar form, namely:

$$J_n = qv_R n \Theta \quad (3.4.2)$$

where v_R is the Richardson velocity and n is the density of carriers in the semiconductor. The tunneling probability term, Θ , is added since the total current depends on the carrier flux arriving at the tunnel barrier multiplied with the probability, Θ , that they tunnel through the barrier.

3.4.1. Diffusion current



This analysis assumes that the depletion layer is large compared to the mean free path, so that the concepts of drift and diffusion are valid. The resulting current density equals:

$$J_n = \frac{q^2 D_n N_c}{V_t} \sqrt{\frac{2q(\phi - V_a) N_d}{\varepsilon_s}} \exp\left(-\frac{\phi_B}{V_t}\right) \left[\exp\left(\frac{V_a}{V_t}\right) - 1 \right] \quad (3.4.3)$$

The current therefore depends exponentially on the applied voltage, V_a , and the barrier height, ϕ_B . The prefactor can more easily be understood if one rewrites it as a function of the electric field at the metal-semiconductor interface, \mathcal{E}_{\max} :

$$\mathcal{E}_{\max} = \sqrt{\frac{2q(\phi - V_a)N_d}{\epsilon_s}} \quad (3.4.4)$$

yielding:

$$J_n = q\mu_n \mathcal{E}_{\max} N_c \exp\left(-\frac{\phi_B}{V_t}\right) \left[\exp\left(\frac{V_a}{V_t}\right) - 1\right] \quad (3.4.5)$$

so that the prefactor equals the drift current at the metal-semiconductor interface, which for zero



3.4.2 Thermionic emission

The thermionic emission theory assumes that electrons, which have an energy larger than the top of the barrier, will cross the barrier provided they move towards the barrier. The actual shape of the barrier is hereby ignored. The current can be expressed as:

$$J_{MS} = A^* T^2 e^{-\phi_B/V_t} (e^{V_a/V_t} - 1) \quad (3.4.6)$$

$$A^* = \frac{4\pi q m^* k^2}{h^3}$$

where A^* is the Richardson constant and ϕ_B is the Schottky barrier height.

The expression for the current due to thermionic emission can also be written as a function of the average velocity with which the electrons at the interface approach the barrier. This velocity is referred to as the Richardson velocity given by:

$$v_R = \sqrt{\frac{kT}{2\pi m}} \quad (3.4.7)$$

So that the current density becomes:

$$J_n = q v_R N_c \exp\left(-\frac{\phi_B}{V_t}\right) \left[\exp\left(\frac{V_a}{V_t}\right) - 1\right] \quad (3.4.8)$$



3.4.3. Tunneling

The tunneling current is obtained from the product of the carrier charge, velocity and density. The velocity equals the Richardson velocity, the velocity with which on average the carriers approach the barrier. The carrier density equals the density of available electrons, n , multiplied with the tunneling probability, Θ , yielding:

$$J_n = q v_R n \Theta \quad (3.4.9)$$

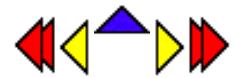
Where the tunneling probability is obtained from:

$$\Theta = \exp\left(-\frac{4}{3} \frac{\sqrt{2qm^*}}{\hbar} \frac{\phi_B^{3/2}}{\mathcal{E}}\right) \quad (3.4.10)$$

and the electric field equals $E = \phi_B/L$.

The tunneling current therefore depends exponentially on the barrier height, ϕ_B , to the 3/2 power.

Chapter 4: p-n Junctions



Bibliography

Chapter 4: p-n Junctions



Review Questions

1. What is a flatband diagram?
2. Discuss the motion of electrons and holes in a p-n junction in thermal equilibrium.
3. Define the built-in potential. Also provide an equation and state the implicit assumption(s).
4. How does the energy band diagram of a p-n junction change under forward and reverse bias?
5. What is the full depletion approximation? Why do we need the full depletion approximation?
6. Derive equation (4.3.17) from (4.3.13), (4.3.14) and (4.3.16).
7. Explain why the capacitance of a p-n junction (4.3.22) equals that of a parallel plate capacitor. How does the capacitance differ from a parallel plate capacitor?
8. How do you extract the doping profile shown in Fig. 4.3.4 from the capacitance shown in Fig. 4.3.3?
9. What mechanism(s) cause(s) current in a p-n junction?
10. How does one calculate the current in a p-n junction?
11. How does one solve the diffusion equation in the quasi-neutral regions?
12. What is the difference between the "long" and "short" diode analysis?
13. When can the recombination/generation current in the depletion region be ignored?
14. Which saturation current is voltage dependent, that for the "long" diode or the one for the "short" diode?
15. Why does one need to include edge effects when calculating the breakdown voltage of a diode?
16. Name two breakdown mechanisms and discuss the temperature dependence of the resulting breakdown voltage.
17. Describe the avalanche breakdown mechanism.
18. Describe tunneling.
19. Illustrate the generation of a photocurrent in a p-n diode by drawing an energy band diagram. Indicate the photo-generated carriers and their direction of motion.

20. Why is the photocurrent negative compared to the forward bias current through the same diode?
21. What limits the quantum efficiency of a photodiode?
22. What is the difference between a solar cell and a photodiode?
23. Why would solar cells be more efficient if the sun were a laser rather than a black body radiator?
24. What limits the power conversion efficiency of a solar cell?
25. Using equation [4.6.1](#) show that the open-circuit voltage increases as the photocurrent increases. Use this result to prove that the power conversion efficiency of a solar increases when using a concentrator which increases the incident power density.
26. Why is silicon not used to fabricate LEDs or laser diodes?
27. Why are planar LEDs so inefficient? How can the efficiency of an LED be improved beyond that of a planar LED?
28. How does the light emitted by an LED differ from that emitted by a laser diode?
29. What is stimulated emission?
30. Why does a laser diode need a waveguide?
31. Explain the lasing condition in words.
32. Describe the power versus current characteristic of a laser diode.

Chapter 5: Bipolar Junction Transistors

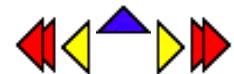


Bibliography



Bibliography

Chapter 5: Bipolar Junction Transistors



Equations

$$w_E^+ = w_E - x_{n,BE} \quad \bullet(5.2.1)\bullet$$

$$w_B^+ = w_B - x_{p,BE} - x_{p,BC} \quad \bullet(5.2.2)\bullet$$

$$w_C^+ = w_C - x_{n,BC} \quad \bullet(5.2.3)\bullet$$

$$x_{n,BE} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BE} - V_{BE})}{q} \frac{N_B}{N_E} \left(\frac{1}{N_B + N_E} \right)} \quad \bullet(5.2.4)\bullet$$

$$x_{p,BE} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BE} - V_{BE})}{q} \frac{N_E}{N_B} \left(\frac{1}{N_B + N_E} \right)} \quad \bullet(5.2.5)\bullet$$

$$x_{p,BC} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BC} - V_{BC})}{q} \frac{N_C}{N_B} \left(\frac{1}{N_B + N_C} \right)} \quad \bullet(5.2.6)\bullet$$

$$x_{n,BC} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BC} - V_{BC})}{q} \frac{N_B}{N_C} \left(\frac{1}{N_B + N_C} \right)} \quad \bullet(5.2.7)\bullet$$

$$\phi_{i,BE} = V_t \ln \frac{N_E N_B}{n_i^2} \quad \bullet(5.2.8)\bullet$$

$$\phi_{i,BC} = V_t \ln \frac{N_C N_B}{n_i^2} \quad \bullet(5.2.9)\bullet$$

$$I_E = I_C + I_B \quad \bullet(5.2.10)\bullet$$

$$I_E = I_{E,n} + I_{E,p} + I_{r,d} \quad \bullet(5.2.11)\bullet$$

$$I_C = I_{E,n} - I_{r,B} \quad \bullet(5.2.12)\bullet$$

$$I_B = I_{E,p} + I_{r,B} + I_{r,d} \quad \bullet(5.2.13)\bullet$$

$$\alpha = \frac{I_C}{I_E} \quad \bullet(5.2.14)\bullet$$

$$\beta = \frac{I_C}{I_B} = \frac{\alpha}{1 - \alpha} \quad \bullet(5.2.15)\bullet$$

$$\alpha = \alpha_T \gamma_E \delta_r \quad \bullet(5.2.16)\bullet$$

$$\gamma_E = \frac{I_{E,n}}{I_{E,n} + I_{E,p}} \quad \bullet(5.2.17)\bullet$$

$$\alpha_T = \frac{I_{E,n} - I_{r,B}}{I_{E,n}} \quad \bullet(5.2.18)\bullet$$

$$\delta_r = \frac{I_E - I_{r,d}}{I_E} \quad \bullet(5.2.19)\bullet$$

□

•(5.2.20)•

$$I_{E,n} = q n_i^2 A_E \left(\frac{D_{n,B}}{N_B w_B} \right) \left(\exp\left(\frac{V_{BE}}{V_t}\right) - 1 \right) \quad •(5.3.1)•$$

$$I_{E,p} = q n_i^2 A_E \left(\frac{D_{p,E}}{N_E w_E} \right) \left(\exp\left(\frac{V_{BE}}{V_t}\right) - 1 \right) \quad •(5.3.2)•$$

$$\Delta Q_{n,B} = q A_E \int_{x_{p,E}}^{w_B - x_{p,C}} (n_p(x) - n_{p0}) dx \quad •(5.3.3)•$$

$$\Delta Q_{n,B} = q A_E \frac{n_i^2}{N_B} \left(\exp\left(\frac{V_{BE}}{V_t}\right) - 1 \right) \frac{w_B}{2} \quad •(5.3.4)•$$

$$I_{E,n} = \frac{\Delta Q_{n,B}}{t_r} \quad •(5.3.5)•$$

$$t_r = \frac{w_B^2}{2 D_{n,B}} \quad •(5.3.6)•$$

$$\frac{\partial n_p(x)}{\partial t} = \frac{1}{q} \frac{\partial J_n(x)}{\partial x} - \frac{n_p(x) - n_{p0}}{\tau_n} \quad •(5.3.7)•$$

$$I_{r,B} = q A_E \int_{x_{p,BE}}^{w_B - x_{p,BC}} \frac{n_p(x) - n_{p0}}{\tau_n} dx \quad •(5.3.8)•$$

$$I_{r,B} = \frac{\Delta Q_{n,B}}{\tau_n} \quad •(5.3.9)•$$

$$\gamma_E = \frac{1}{1 + \frac{D_{p,E} N_B w_B}{D_{n,B} N_E w_E}} \quad \bullet(5.3.10)\bullet$$

$$\beta \equiv \frac{D_{n,B} N_E w_E}{D_{p,E} N_B w_B}, \text{ if } \alpha \equiv \gamma_E \quad \bullet(5.3.11)\bullet$$

$$\alpha_T = 1 - \frac{t_r}{\tau_n} = 1 - \frac{w_B^2}{2 D_{n,B} \tau_n} \quad \bullet(5.3.12)\bullet$$

$$\alpha_T = 1 - \frac{1}{2} \left(\frac{w_B}{L_n} \right)^2 \quad \bullet(5.3.13)\bullet$$

$$I_E = I_F - \alpha_R I_R \quad \bullet(5.3.14)\bullet$$

$$I_B = (1 - \alpha_F) I_F + (1 - \alpha_R) I_R \quad \bullet(5.3.15)\bullet$$

$$I_C = -I_R + \alpha_F I_F \quad \bullet(5.3.16)\bullet$$

$$I_{E,s} \alpha_F = I_{C,s} \alpha_R \quad \bullet(5.3.17)\bullet$$

$$I_F (V_{BE}) \alpha_F = I_R (V_{BC} = V_{BE}) \alpha_R \quad \bullet(5.3.18)\bullet$$

$$V_{CE,sat} = V_{BE} - V_{BC} = V_t \ln \left\{ \frac{I_F}{I_R} \frac{I_{C,s}}{I_{E,s}} \right\} \quad \bullet(5.3.19)\bullet$$

$$V_{CE,sat} = V_t \ln \left\{ \frac{1 + \frac{I_C}{I_B} (1 - \alpha_R)}{\alpha_R \left[1 - \frac{I_C}{I_B} \frac{(1 - \alpha_F)}{\alpha_F} \right]} \right\} \quad \bullet(5.3.20)\bullet$$

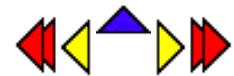
$$\frac{dI_C}{dV_{CE}} \cong - \frac{dI_C}{dV_{BC}} = \frac{I_C}{w_B^'} \frac{dw_B^'}{dV_{BC}} \quad \bullet(5.4.1)\bullet$$

$$\frac{dI_C}{dV_{CE}} \cong \frac{I_C}{|V_A|} \quad \bullet(5.4.2)\bullet$$

$$|V_A| = \frac{\mathcal{Q}_B}{C_{j,BC}} = \frac{qA_C N_B w_B^'}{\varepsilon_s A_C} \frac{1}{x_{p,BC} + x_{n,BC}} \quad \bullet(5.4.3)\bullet$$

$$n = \frac{1}{V_t} \frac{d \ln I_C}{dV_{BE}} \cong 1 + \frac{V_t}{\mathcal{Q}_B} C_{j,BE} \quad \bullet(5.4.4)\bullet$$

Chapter 6: MOS Capacitors



Equations

$$V_{FB} = \Phi_M - \Phi_S \quad \bullet(6.3.1)\bullet$$

$$\Phi_M - \Phi_S = \Phi_M - \mathcal{X} - \frac{E_g}{2q} - V_t \ln\left(\frac{N_a}{n_i}\right) \quad \bullet(6.3.2)\bullet$$

$$\Phi_{poly,S} = V_t \ln\left(\frac{N_{a,poly}}{N_a}\right) \quad \text{p-type polysilicon gate} \quad \bullet(6.3.3)\bullet$$

$$\Phi_{poly,S} = V_t \ln\left(\frac{n_i^2}{N_{d,poly} N_a}\right) \quad \text{n-type polysilicon gate}$$

$$V_{FB} = \Phi_{MS} - \frac{Q_i}{C_{ox}} - \frac{1}{\epsilon_{ox}} \int_0^{t_{ox}} \rho_{ox}(x) x dx \quad \bullet(6.3.4)\bullet$$

$$Q_{inv} = C_{ox} (V_G - V_T) \quad \bullet(6.3.5)\bullet$$

$$Q_{inv} = 0$$

$$Q_d = -q N_a x_d \quad \bullet(6.3.6)\bullet$$

$$\epsilon_s = \frac{q N_a x_d}{\epsilon_s} \quad \bullet(6.3.7)\bullet$$

$$\rho_s = \frac{q N_a x_d^2}{2 \epsilon_s} \quad \bullet(6.3.8)\bullet$$

$$\phi_F = V_t \ln \frac{N_a}{n_i} \quad \bullet(6.3.9)\bullet$$

$$x_d = \sqrt{\frac{2 \varepsilon_s \phi_s}{q N_a}}, \text{ for } 0 \leq \phi_s \leq 2 \phi_F \quad \bullet(6.3.10)\bullet$$

$$Q_{d,T} = -q N_a x_{d,T} \quad \bullet(6.3.11)\bullet$$

$$x_{d,T} = \sqrt{\frac{2 \varepsilon_s (2 \phi_F)}{q N_a}} \quad \bullet(6.3.12)\bullet$$

$$Q_M = -(Q_d + Q_{inv}) \quad \bullet(6.3.13)\bullet$$

$$V_G = V_{FB} + \phi_s + \frac{Q_M}{C_{ox}} = V_{FB} + \phi_s - \frac{Q_d + Q_{inv}}{C_{ox}} \quad \bullet(6.3.14)\bullet$$

$$V_G = V_{FB} + \phi_s + \frac{\sqrt{2 \varepsilon_s q N_a \phi_s}}{C_{ox}}, \text{ for } 0 \leq \phi_s \leq 2 \phi_F \quad \bullet(6.3.15)\bullet$$

$$V_G = V_{FB} + 2 \phi_F + \frac{\sqrt{4 \varepsilon_s q N_a \phi_F}}{C_{ox}} - \frac{Q_{inv}}{C_{ox}} \stackrel{\Delta}{=} V_T - \frac{Q_{inv}}{C_{ox}} \quad \bullet(6.3.16)\bullet$$

$$V_T = V_{FB} + 2 \phi_F + \frac{\sqrt{4 \varepsilon_s q N_a \phi_F}}{C_{ox}} \quad \bullet(6.3.17)\bullet$$

$$C_{LF} = C_{HF} = C_{ox}, \text{ for } V_G \leq V_{FB} \quad \bullet(6.3.18)\bullet$$

$$C_{LF} = C_{HF} = \frac{1}{\frac{1}{C_{ox}} + \frac{x_d}{\varepsilon_s}}, \text{ for } V_{FB} \leq V_G \leq V_T \quad \bullet(6.3.19)\bullet$$

$$x_d = \sqrt{\frac{2\varepsilon_s \phi_s}{qN_d}} \quad \bullet(6.3.20)\bullet$$

$$C_{LF} = C_{ox} \text{ and } C_{HF} = \frac{1}{\frac{1}{C_{ox}} + \frac{x_{d,T}}{\varepsilon_s}}, \text{ for } V_G \geq V_T \quad \bullet(6.3.21)\bullet$$

$$\frac{d^2 \phi}{dx^2} = \frac{q}{\varepsilon_s} (N_a^+ - p) = \frac{qN_a}{\varepsilon_s} (1 - e^{-\frac{\phi}{V_t}}) \cong \frac{qN_a}{\varepsilon_s} \frac{\phi}{V_t} \quad \bullet(6.3.22)\bullet$$

$$\phi = \phi_s e^{-\frac{x}{L_D}}, \text{ with } L_D = \sqrt{\frac{\varepsilon_s V_t}{qN_a}} \quad \bullet(6.3.23)\bullet$$

$$C_{s,FB} = \frac{dQ_s}{d\phi_s} = \frac{d(\varepsilon_s \frac{\phi_s}{L_D})}{d\phi_s} = \frac{\varepsilon_s}{L_D} \quad \bullet(6.3.24)\bullet$$

$$C_{FB} = \frac{1}{\frac{1}{C_{ox}} + \frac{L_D}{\varepsilon_s}} \quad \bullet(6.3.25)\bullet$$

$$time = \frac{|Q_{inv}|}{qG(x_{d,dd} + L_n)} = \frac{C_{ox}(V_G - V_T)}{\frac{qn_i}{2\tau} \left(\sqrt{\frac{2\varepsilon_s \phi_{s,dd}}{qN_a}} + \sqrt{\mu_n V_t \tau} \right)} \quad \bullet(6.3.26)\bullet$$

$$\frac{dV_G}{dt} > \frac{qn_i}{2C_{ox}} \sqrt{\frac{\mu_n V_t}{\tau}} \quad \bullet(6.3.27)\bullet$$

Chapter 5: Bipolar Junction Transistors

5.1. Introduction

5.2. Structure and principle of operation

$$w_E = w_E - x_{n,BE} \quad (5.2.1)$$

$$w_B = w_B - x_{p,BE} - x_{p,BC} \quad (5.2.2)$$

$$w_C = w_C - x_{n,BC} \quad (5.2.3)$$

$$x_{n,BE} = \sqrt{\frac{2e_s(f_{i,BE} - V_{BE})}{q} \frac{N_B}{N_E} \left(\frac{1}{N_B + N_E} \right)} \quad (5.2.4)$$

$$x_{p,BE} = \sqrt{\frac{2e_s(f_{i,BE} - V_{BE})}{q} \frac{N_E}{N_B} \left(\frac{1}{N_B + N_E} \right)} \quad (5.2.5)$$

$$x_{p,BC} = \sqrt{\frac{2e_s(f_{i,BC} - V_{BC})}{q} \frac{N_C}{N_B} \left(\frac{1}{N_B + N_C} \right)} \quad (5.2.6)$$

$$x_{n,BC} = \sqrt{\frac{2e_s(f_{i,BC} - V_{BC})}{q} \frac{N_B}{N_C} \left(\frac{1}{N_B + N_C} \right)} \quad (5.2.7)$$

$$f_{i,BE} = V_t \ln \frac{N_E N_B}{n_i^2} \quad (5.2.8)$$

$$f_{i,BC} = V_t \ln \frac{N_C N_B}{n_i^2} \quad (5.2.9)$$

$$I_E = I_C + I_B \quad (5.2.10)$$

$$I_E = I_{E,n} + I_{E,p} + I_{r,d} \quad (5.2.11)$$

$$I_C = I_{E,n} - I_{r,B} \quad (5.2.12)$$

$$I_B = I_{E,p} + I_{r,B} + I_{r,d} \quad (5.2.13)$$

$$\alpha = \frac{I_C}{I_E} \quad (5.2.14)$$

$$\mathbf{b} = \frac{I_C}{I_B} = \frac{\mathbf{a}}{1 - \mathbf{a}} \quad (5.2.15)$$

$$\mathbf{a} = \mathbf{a}_T \mathbf{g}_E \mathbf{d}_r \quad (5.2.16)$$

$$\mathbf{g}_E = \frac{I_{E,n}}{I_{E,n} + I_{E,p}} \quad (5.2.17)$$

$$\mathbf{a}_T = \frac{I_{E,n} - I_{r,B}}{I_E, n} \quad (5.2.18)$$

$$\mathbf{d}_r = \frac{I_E - I_{r,d}}{I_E} \quad (5.2.19)$$

5.3. Ideal transistor model

$$I_{E,n} = q n_i^2 A_E \left(\frac{D_{n,B}}{N_B w_B} \right) \left(\exp\left(\frac{V_{BE}}{V_t}\right) - 1 \right) \quad (5.3.1)$$

$$I_{E,p} = q n_i^2 A_E \left(\frac{D_{p,E}}{N_E w_E} \right) \left(\exp\left(\frac{V_{BE}}{V_t}\right) - 1 \right) \quad (5.3.2)$$

$$\Delta Q_{n,B} = q A_E \int_{x_{p,E}}^{w_B - x_{p,C}} n_p(x) - n_{p0} dx \quad (5.3.3)$$

$$\Delta Q_{n,B} = q A_E \frac{n_i^2}{N_B} \left(\exp\left(\frac{V_{BE}}{V_t}\right) - 1 \right) \frac{w_B}{2} \quad (5.3.4)$$

$$I_{E,n} = \frac{\Delta Q_{n,B}}{t_r} \quad (5.3.5)$$

$$t_r = \frac{w_B^2}{2 D_{n,B}} \quad (5.3.6)$$

$$\frac{\partial n_p(x)}{\partial t} = \frac{1}{q} \frac{\partial J_n(x)}{\partial x} - \frac{n_p(x) - n_{p0}}{t_n} \quad (5.3.7)$$

$$I_{r,B} = A_E \int_{x_{p,BE}}^{w_B - x_{p,BC}} \frac{\partial J_n(x)}{\partial x} dx = qA_E \int_{x_{p,BE}}^{w_B - x_{p,BC}} \frac{n_p(x) - n_{p0}}{t_n} dx \quad (5.3.8)$$

$$I_{r,B} = \frac{\Delta Q_{n,B}}{t_n} \quad (5.3.9)$$

$$g_E = \frac{1}{1 + \frac{D_{p,E} N_B w_B}{D_{n,B} N_E w_E}} \quad (5.3.10)$$

$$b \equiv \frac{D_{n,B} N_E w_E}{D_{p,E} N_B w_B}, \quad \text{if } \quad a \equiv g_E \quad (5.3.11)$$

$$a_T = 1 - \frac{t_r}{t_n} = 1 - \frac{w_B^2}{2D_{n,B} t_n} \quad (5.3.12)$$

$$a_T = 1 - \frac{1}{2} \left(\frac{w_B}{L_n} \right)^2 \quad (5.3.13)$$

5.3.1. General bias modes of a bipolar transistor

$$I_E = I_F - a_R I_R \quad (5.3.14)$$

$$I_B = (1 - a_F) I_F + (1 - a_R) I_R \quad (5.3.15)$$

$$I_C = -I_R + a_F I_F \quad (5.3.16)$$

$$I_{E,s} a_F = I_{C,s} a_R \quad (5.3.17)$$

$$I_F(V_{BE}) a_F = I_R(V_{BC} = V_{BE}) a_R \quad (5.3.18)$$

$$V_{CE,sat} = V_{BE} - V_{BC} = V_t \ln \left\{ \frac{I_F}{I_R} \frac{I_{C,s}}{I_{E,s}} \right\} \quad (5.3.19)$$

$$V_{CE,sat} = V_t \ln \left\{ \frac{1 + \frac{I_C}{I_B} (1 - a_R)}{a_R \left[1 - \frac{I_C}{I_B} \frac{(1 - a_F)}{a_F} \right]} \right\} \quad (5.3.20)$$

5.4. Non-ideal effects

$$\frac{dI_C}{dV_{CE}} \cong -\frac{dI_C}{dV_{BC}} = \frac{I_C}{w_B} \frac{dw_B}{dV_{BC}} \quad (5.4.1)$$

$$\frac{dI_C}{dV_{CE}} \cong \frac{I_C}{|V_A|} \quad (5.4.2)$$

$$|V_A| = \frac{Q_{p,B}}{C_{j,BC}} = \frac{qA_C N_B w_B}{\frac{e_s A_C}{x_{p,BC} + x_{n,BC}}} \quad (5.4.3)$$

$$n = \frac{1}{V_t \frac{d \ln I_C}{dV_{BE}}} \cong 1 + \frac{V_t}{Q_{p,B}} C_{j,BE} \quad (5.4.4)$$

Chapter 4: p-n Junctions

4.2. Structure and principle of operation

$$\mathbf{f}_i = V_t \ln \frac{N_d N_a}{n_i^2} \quad (4.2.1)$$

$$\mathbf{f} = \mathbf{f}_i - V_a \quad (4.2.2)$$

4.3. Electrostatic analysis of a p-n diode

$$\frac{d^2 \mathbf{f}}{dx^2} = -\frac{\mathbf{r}}{\mathbf{e}_s} = -\frac{q}{\mathbf{e}_s} (p - n + N_d^+ - N_a^-) \quad (4.3.1)$$

$$\frac{d^2 \mathbf{f}}{dx^2} = \frac{2qn_i}{\mathbf{e}_s} \left(\sinh \frac{\mathbf{f} - \mathbf{f}_F}{V_t} + \sinh \frac{\mathbf{f}_F}{V_t} \right) \quad (4.3.2)$$

$$\sinh \frac{\mathbf{f}_F}{V_t} = \frac{N_a^- - N_d^+}{2n_i} \quad (4.3.3)$$

$$x_d = x_n + x_p \quad (4.3.4)$$

$$\mathbf{r} = q(p - n + N_d^+ - N_a^-) \equiv q(N_d^+ - N_a^-), \text{ for } -x_p \leq x \leq x_n \quad (4.3.5)$$

$$\mathbf{r}(x) = 0 \quad , \text{ for } x \leq -x_p \quad (4.3.6)$$

$$\mathbf{r}(x) = -qN_a \quad , \text{ for } -x_p \leq x \leq 0$$

$$\mathbf{r}(x) = qN_d \quad , \text{ for } 0 \leq x \leq x_n$$

$$\mathbf{r}(x) = 0 \quad , \text{ for } x_n \leq x$$

$$Q_n = qN_d x_n \quad (4.3.7)$$

$$Q_p = -qN_a x_p \quad (4.3.8)$$

$$\frac{dE(x)}{dx} = \frac{\mathbf{r}(x)}{\mathbf{e}_s} \equiv \frac{q}{\mathbf{e}_s} (N_d^+(x) - N_a^-(x)), \text{ for } -x_p \leq x \leq x_n \quad (4.3.9)$$

$$E(x) = 0 \quad , \text{ for } x \leq -x_p \quad (4.3.10)$$

$$\begin{aligned} E(x) &= -\frac{qN_a(x+x_p)}{\epsilon_s} \quad , \text{ for } -x_p \leq x \leq 0 \\ E(x) &= \frac{qN_d(x-x_n)}{\epsilon_s} \quad , \text{ for } 0 \leq x \leq x_n \\ E(x) &= 0 \quad , \text{ for } x_n \leq x \end{aligned}$$

$$E(x=0) = -\frac{qN_a x_p}{\epsilon_s} = -\frac{qN_d x_n}{\epsilon_s} \quad (4.3.11)$$

$$N_d x_n = N_a x_p \quad (4.3.12)$$

$$x_n = x_d \frac{N_a}{N_a + N_d} \quad (4.3.13)$$

$$x_p = x_d \frac{N_d}{N_a + N_d} \quad (4.3.14)$$

$$\frac{d\mathbf{f}(x)}{dx} = -E(x) \quad (4.3.15)$$

$$\mathbf{f}_i - V_a = \frac{qN_d x_n^2}{2\epsilon_s} + \frac{qN_a x_p^2}{2\epsilon_s} \quad (4.3.16)$$

$$x_d = \sqrt{\frac{2\epsilon_s}{q} \left(\frac{1}{N_a} + \frac{1}{N_d} \right) (\mathbf{f}_i - V_a)} \quad (4.3.17)$$

$$x_n = \sqrt{\frac{2\epsilon_s}{q} \frac{N_a}{N_d} \frac{1}{N_a + N_d} (\mathbf{f}_i - V_a)} \quad (4.3.18)$$

$$x_p = \sqrt{\frac{2\epsilon_s}{q} \frac{N_d}{N_a} \frac{1}{N_a + N_d} (\mathbf{f}_i - V_a)} \quad (4.3.19)$$

$$C(V_a) \stackrel{\Delta}{=} \left| \frac{dQ(V_a)}{dV_a} \right| \quad (4.3.20)$$

$$C_j = \sqrt{\frac{q\epsilon_s}{2(\mathbf{f}_i - V_a)} \frac{N_a N_d}{N_a + N_d}} \quad (4.3.21)$$

$$C_j = \frac{\mathbf{e}_s}{x_d} \quad (4.3.22)$$

$$\frac{1}{C_j^2} = \frac{2}{q\mathbf{e}_s} \frac{N_a + N_d}{N_a N_d} (\mathbf{f}_i - V_a) \quad (4.3.23)$$

$$\frac{d(1/C_j^2)}{dV_a} = -\frac{2}{q\mathbf{e}_s} \frac{N_a + N_d}{N_a N_d} \quad (4.3.24)$$

$$N_d = -\frac{2}{q\mathbf{e}_s} \frac{1}{\frac{d(1/C_j^2)}{dV_a}}, \text{ if } N_a \gg N_d \quad (4.3.25)$$

4.4. The p-n diode current

$$p_n(x = x_n) = p_{n0} e^{V_a/V_t} \quad (4.4.1)$$

$$n_p(x = -x_p) = n_{p0} e^{V_a/V_t} \quad (4.4.2)$$

$$p_n(x = w_n) = p_{n0} \quad (4.4.3)$$

$$n_p(x = -w_p) = n_{p0} \quad (4.4.4)$$

$$p_n(x \geq x_n) = p_{n0} + A e^{-(x-x_n)/L_p} + B e^{(x-x_n)/L_p} \quad (2.9.13)$$

$$n_p(x \leq -x_p) = n_{p0} + C e^{-(x+x_p)/L_p} + D e^{(x+x_p)/L_p} \quad (2.9.14)$$

$$p_n(x \geq x_n) = p_{n0} + A^* \cosh \frac{x-x_n}{L_p} + B^* \sinh \frac{x-x_n}{L_p} \quad (4.4.5)$$

$$n_p(x \leq -x_p) = n_{p0} + C^* \cosh \frac{x+x_p}{L_n} + D^* \sinh \frac{x+x_p}{L_n} \quad (4.4.6)$$

$$p_n(x \geq x_n) = p_{n0} + p_{n0} (e^{V_a/V_t} - 1) [\cosh \frac{x-x_n}{L_p} - \coth \frac{w_n}{L_p} \sinh \frac{x-x_n}{L_p}] \quad (4.4.7)$$

$$n_p(x \leq -x_p) = n_{p0} + n_{p0} (e^{V_a/V_t} - 1) [\cosh \frac{x+x_p}{L_n} + \coth \frac{w_p}{L_n} \sinh \frac{x+x_p}{L_n}] \quad (4.4.8)$$

$$w_n' = w_n - x_n \quad (4.4.9)$$

$$w_p' = w_p - x_p \quad (4.4.10)$$

$$J_p(x \geq x_n) = -qD_p \frac{dp}{dx} \quad (4.4.11)$$

$$= -\frac{qD_p p_{n0}}{L_p} (e^{V_a/V_t} - 1) [\sinh \frac{x - x_n}{L_p} - \coth \frac{w_p'}{L_p} \cosh \frac{x - x_n}{L_p}]$$

$$J_n(x \leq -x_p) = qD_n \frac{dn}{dx} \quad (4.4.12)$$

$$= \frac{qD_n n_{p0}}{L_n} (e^{V_a/V_t} - 1) [\sinh \frac{x + x_p}{L_n} + \coth \frac{w_p'}{L_n} \cosh \frac{x + x_p}{L_n}]$$

$$I = A[J_n(x = -x_p) + J_p(x = x_n) + J_r] \equiv I_s (e^{V_a/V_t} - 1) \quad (4.4.13)$$

$$I_s = qA \left[\frac{D_n n_{p0}}{L_n} \coth \left(\frac{w_p'}{L_n} \right) + \frac{D_p p_{n0}}{L_p} \coth \left(\frac{w_n'}{L_p} \right) \right] \quad (4.4.14)$$

$$\coth x = \frac{1}{\tanh x} \equiv \frac{1}{x} \text{, for } x \ll 1 \quad (4.4.15)$$

$$p_n(x \geq x_n) = p_{n0} + p_{n0} (e^{V_a/V_t} - 1) \exp \frac{-(x - x_n)}{L_p} \quad (4.4.16)$$

$$n_p(x \leq -x_p) = n_{p0} + n_{p0} (e^{V_a/V_t} - 1) \exp \frac{x + x_p}{L_n} \quad (4.4.17)$$

$$J_p(x \geq x_n) = \frac{qD_p p_{n0}}{L_p} (e^{V_a/V_t} - 1) \exp \frac{-(x - x_n)}{L_p} \quad (4.4.18)$$

$$J_n(x \leq -x_p) = \frac{qD_n n_{p0}}{L_n} (e^{V_a/V_t} - 1) \exp \frac{x + x_p}{L_n} \quad (4.4.19)$$

$$I_s = qA \left[\frac{D_n n_{p0}}{L_n} + \frac{D_p p_{n0}}{L_p} \right] = qA \left[\frac{n_{p0} L_n}{t_n} + \frac{p_{n0} L_p}{t_p} \right] \quad (4.4.20)$$

$$I_r \geq qA \frac{p_{n0} x_n}{t_p} (e^{V_a/V_t} - 1) \quad (4.4.21)$$

$$x_n \ll L_p \quad (4.4.22)$$

$$0 = D_n \frac{d^2 n_p}{dx^2}, \text{ and } 0 = D_p \frac{d^2 p_n}{dx^2} \quad (4.4.23)$$

$$n_p = A + Bx, \text{ and } p_n = C + Dx \quad (4.4.24)$$

$$p_n = p_{n0} + p_{n0} (e^{V_a/V_t} - 1) \left(1 - \frac{x - x_n}{w_n} \right) \quad (4.4.25)$$

$$n_p = n_{p0} + n_{p0} (e^{V_a/V_t} - 1) \left(1 + \frac{x + x_p}{w_p} \right) \quad (4.4.26)$$

$$I = A[J_n(x = -x_p) + J_p(x = x_n) + J_r] \equiv I_s (e^{V_a/V_t} - 1) \quad (4.4.27)$$

$$I_s = qA \left[\frac{D_n n_{p0}}{w_p} + \frac{D_p p_{n0}}{w_n} \right] \quad (4.4.28)$$

$$J_{b-b} = q \int_{-x_p}^{x_n} U_{b-b} dx \quad (4.4.29)$$

$$U_{b-b} = b(np - n_i^2) \quad (4.4.30)$$

$$n_p = n_i e^{(E_{Fn} - E_i)/kT} n_i e^{(E_i - E_{Fp})/kT} = n_i^2 e^{V_a/V_t} \quad (4.4.31)$$

$$J_{b-b} = q \int_{-x_p}^{x_n} n_i^2 (e^{V_a/V_t} - 1) dx = q n_i^2 b x_d (e^{V_a/V_t} - 1) \quad (4.4.32)$$

$$J_{SHR} = q \int_{-x_p}^{x_n} U_{SHR} dx \quad (4.4.33)$$

$$J_{SHR} = q \int_{-x_p}^{x_n} \frac{1}{t} \frac{n_i^2 (e^{V_a/V_t} - 1)}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} dx \quad (4.4.34)$$

$$n_p = n_i e^{(F_n - E_i)/kT} n_i e^{(E_i - F_p)/kT} = n_i^2 e^{V_a/V_t} \quad (4.4.35)$$

$$U_{SHR, \max} = \text{MAX} \left(\frac{1}{t} \frac{n_i^2 (e^{V_a/V_t} - 1)}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} \right) \equiv \frac{n_i}{2t} (e^{V_a/2V_t} - 1) \quad (4.4.36)$$

$$x' = \frac{\int_{-x_p}^{x_n} \frac{1}{t} \frac{n_i^2 (e^{V_a/V_t} - 1)}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} dx}{U_{SHR, \max}} \quad (4.4.37)$$

$$J_{SHR} = \frac{qn_i x'}{2t} (e^{V_a/2V_t} - 1) \quad (4.4.38)$$

$$J = J_s e^{V_a/hV_t} \quad (4.4.39)$$

$$h = \frac{\log(e)}{V_t \text{ slope}} = \frac{1}{\frac{\text{slope}}{59.6 \text{ mV/decade}}} \quad (4.4.40)$$

$$V_a = 2V_t \ln \frac{N_d}{n_i} \quad (4.4.41)$$

$$V_a^* = V_a + IR_s \quad (4.4.42)$$

$$C = \frac{d\Delta Q}{dV_a} \quad (4.4.43)$$

$$\Delta Q_p = \int_{x_n}^{w_n} qA(p_n - p_{n0}) dx \quad (4.4.44)$$

$$\Delta Q_p = qA p_{n0} (e^{V_a/V_t} - 1) L_p = I_{s,p} (e^{V_a/V_t} - 1) t_p \quad (4.4.45)$$

$$I_{s,p} = q \frac{A p_{n0} D_p}{L_p} \quad (4.4.46)$$

$$C_{d,p} = \frac{d(I_{s,p} (e^{V_a/V_t} - 1) t_p)}{dV_a} = \frac{I_{s,p} e^{V_a/V_t} t_p}{V_t} \quad (4.4.47)$$

$$C_{d,p} = \frac{I_{s,p} e^{V_a/V_t} t_{r,p}}{V_t} \quad (4.4.48)$$

$$t_{r,p} = \frac{w_p^2}{2D_p} \quad (4.4.49)$$

4.5. Reverse bias breakdown

$$|E_{br}| = \frac{4 \times 10^5}{1 - \frac{1}{3} \log(N/10^{16})} \text{ V/cm} \quad (4.5.1)$$

$$|V_{br}| = -F_i + \frac{|E_{br}|^2 \epsilon_s}{2qN} \quad (4.5.2)$$

$$x_{d,br} = \frac{|E_{br}| \mathbf{e}_s}{qN} \quad (4.5.3)$$

$$dn = \mathbf{a}_n n dx \quad (4.5.4)$$

$$M = \frac{1}{1 - \int_{x_1}^{x_2} \mathbf{a} dx} \quad (4.5.5)$$

$$M = \frac{1}{1 - \left| \frac{V_a}{V_{br}} \right|^n}, \text{ where } 2 < n < 6 \quad (4.5.6)$$

$$\Theta = \exp \left(-\frac{4}{3} \frac{\sqrt{2m^*}}{q\hbar} \frac{E_g^{3/2}}{E} \right) \quad (4.5.7)$$

$$J_n = q v_R n \Theta \quad (4.5.8)$$

4.6. Optoelectronic devices

$$I = I_s (e^{V_a/V_t} - 1) - I_{ph} \quad (4.6.1)$$

$$I_{ph,\max} = \frac{q}{h\mathbf{n}} P_{in} \quad (4.6.2)$$

$$I_{ph} = (1 - R)(1 - e^{-\mathbf{a} d}) \frac{qP_{in}}{h\mathbf{n}} \quad (4.6.3)$$

$$\langle i^2 \rangle = 2q I \Delta f \quad (4.6.4)$$

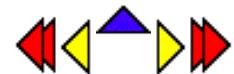
$$\text{Fill Factor} = \frac{I_m V_m}{I_{sc} V_{oc}} \quad (4.6.5)$$

$$\text{Roundtrip amplification} = e^{2gL} R_1 R_2 = 1 \quad (4.6.6)$$

$$g = \frac{1}{2L} \ln \frac{1}{R_1 R_2} \quad (4.6.7)$$

$$P_{out} = h \frac{h\mathbf{n}}{q} (I - I_{th}) \quad (4.6.8)$$

Chapter 5: Bipolar Junction Transistors



Review Questions

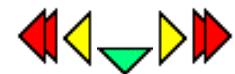
1. Describe the motion of electrons and holes in a pnp bipolar transistor biased in the forward active mode with $V_{BC} = 0$.
2. What is the definition of the emitter efficiency? Explain in words and provide the corresponding equation.
3. What is the definition of the base transport factor? Explain in words and provide the corresponding equation.
4. Derive the relation between the current gain and the transport factor.
5. How does recombination in the quasi-neutral base region affect the emitter, base and collector current?
6. How does recombination in the base-emitter depletion region affect the emitter, base and collector current?
7. Explain the four different bias modes of a bipolar transistor.
8. Explain why a transistor can have a current gain larger than one in the common emitter mode. Provide the necessary and sufficient conditions needed to obtain a current gain larger than one.
9. What is the Early effect and how does it affect the transistor characteristics?

Review Questions

1. What is a flatband diagram?
2. Discuss the motion of electrons and holes in a p-n junction in thermal equilibrium.
3. Define the built-in potential. Also provide an equation and state the implicit assumption(s).
4. How does the energy band diagram of a p-n junction change under forward and reverse bias?
5. What is the full depletion approximation? Why do we need the full depletion approximation?
6. Derive equation (4.3.17) from (4.3.13), (4.3.14) and (4.3.16).
7. Explain why the capacitance of a p-n junction (4.3.22) equals that of a parallel plate capacitor. How does the capacitance differ from a parallel plate capacitor?
8. How do you extract the doping profile shown in Fig. 4.3.4 from the capacitance shown in Fig. 4.3.3?
9. What mechanism(s) cause(s) current in a p-n junction?
10. How does one calculate the current in a p-n junction?
11. How does one solve the diffusion equation in the quasi-neutral regions?
12. What is the difference between the "long" and "short" diode analysis?
13. When can the recombination/generation current in the depletion region be ignored?
14. Which saturation current is voltage dependent, that for the "long" diode or the one for the "short" diode?
15. Why does one need to include edge effects when calculating the breakdown voltage of a diode?
16. Name two breakdown mechanisms and discuss the temperature dependence of the resulting breakdown voltage.
17. Describe the avalanche breakdown mechanism.
18. Describe tunneling.
19. Illustrate the generation of a photocurrent in a p-n diode by drawing an energy band diagram. Indicate the photo-generated carriers and their direction of motion.
20. Why is the photocurrent negative compared to the forward bias current through the same diode?
21. What limits the quantum efficiency of a photodiode?
22. What is the difference between a solar cell and a photodiode?
23. Why would solar cells be more efficient if the sun were a laser rather than a black body radiator?

24. What limits the power conversion efficiency of a solar cell?
25. Using equation 4.6.1 show that the open-circuit voltage increases as the photocurrent increases. Use this result to prove that the power conversion efficiency of a solar increases when using a concentrator which increases the incident power density.
26. Why is silicon not used to fabricated LEDs or laser diodes?
27. Why are planar LEDs so inefficient? How can the efficiency of an LED be improved beyond that of a planar LED?
28. How does the light emitted by an LED differ from that emitted by a laser diode?
29. What is stimulated emission?
30. Why does a laser diode need a waveguide?
31. Explain the lasing condition in words.
32. Describe the power versus current characteristic of a laser diode.

Chapter 4: p-n Junctions



4.3. Electrostatic analysis of a p-n diode

4.3.1. General discussion - Poisson's equation

4.3.2. The full-depletion approximation

4.3.3. Full depletion analysis

4.3.4. Junction capacitance

The electrostatic analysis of a p-n diode is of interest since it provides knowledge about the charge density and the electric field in the depletion region. It is also required to obtain the capacitance-voltage characteristics of the diode. The analysis is very similar to that of a metal-semiconductor junction ([section 3.3](#)). A key difference is that a p-n diode contains two depletion regions of opposite type.

4.3.1. General discussion - Poisson's equation



The general analysis starts by setting up Poisson's equation:

$$\frac{d^2 \phi}{dx^2} = -\frac{\rho}{\epsilon_s} = -\frac{q}{\epsilon_s} (p - n + N_d^+ - N_a^-) \quad (4.3.1)$$

where the charge density, ρ , is written as a function of the electron density, the hole density and the donor and acceptor densities. To solve the equation, we have to express the electron and hole density, n and p , as a function of the potential, ϕ , yielding:

$$\frac{d^2 \phi}{dx^2} = \frac{2qn_i}{\epsilon_s} \left(\sinh \frac{\phi - \phi_F}{V_t} + \sinh \frac{\phi_F}{V_t} \right) \quad (4.3.2)$$

with

$$\sinh \frac{\phi_F}{V_t} = \frac{N_a^- - N_d^+}{2n_i} \quad (4.3.3)$$

where the potential is chosen to be zero in the n -type region, far away from the p-n interface.

This second-order non-linear differential equation [\(4.3.2\)](#) can not be solved analytically. Instead we will make the simplifying assumption that the depletion region is fully depleted and that the adjacent neutral regions contain no charge. This full depletion approximation is the topic of the next section.

4.3.2. The full-depletion approximation



The full-depletion approximation assumes that the depletion region around the metallurgical junction has well-defined edges. It also assumes that the transition between the depleted and the quasi-neutral region is abrupt. We define the quasi-neutral region as the region adjacent to the depletion region where the electric field is small and the free carrier density is close to the net doping density.

The full-depletion approximation is justified by the fact that the carrier densities change exponentially with the position of the Fermi energy relative to the band edges. For example, as the distance between the Fermi energy and the conduction band edge is increased by 59 meV, the electron concentration at room temperature decreases to one tenth of its original value. The charge in the depletion layer is then quickly dominated by the remaining ionized impurities, yielding a constant charge density for uniformly doped regions.

We will therefore start the electrostatic analysis using an abrupt charge density profile, while introducing two unknowns, namely the depletion layer width in the *p*-type region, x_p , and the depletion region width in the *n*-type region, x_n . The sum of the two depletion layer widths in each region is the total depletion layer width x_d , or:

$$x_d = x_n + x_p \quad (4.3.4)$$

From the charge density, we then calculate the electric field and the potential across the depletion region. A first relationship between the two unknowns is obtained by setting the positive charge in the depletion layer equal to the negative charge. This is required since the electric field in both quasi-neutral regions must be zero. A second relationship between the two unknowns is obtained by relating the potential across the depletion layer width to the applied voltage. The combination of both relations yields a solution for x_p and x_n , from which all other parameters can be obtained.

4.3.3. Full depletion analysis



Once the full-depletion approximation is made it is easy to find the charge density profile: It equals the sum of the charges due to the holes, electrons, ionized acceptors and ionized holes:

$$\rho = q(p - n + N_d^+ - N_a^-) \cong q(N_d^+ - N_a^-), \text{ for } -x_p \leq x \leq x_n \quad (4.3.5)$$

where it is assumed that no free carriers are present within the depletion region. For an abrupt p-n diode with doping densities, N_a and N_d , the charge density is then given by:

$$\begin{aligned} \rho(x) &= 0 \\ \rho(x) &= -qN_a \\ \rho(x) &= qN_d \\ \rho(x) &= 0 \end{aligned} \quad (4.3.6)$$

This charge density, ρ , is shown in Figure 4.3.1 (a).

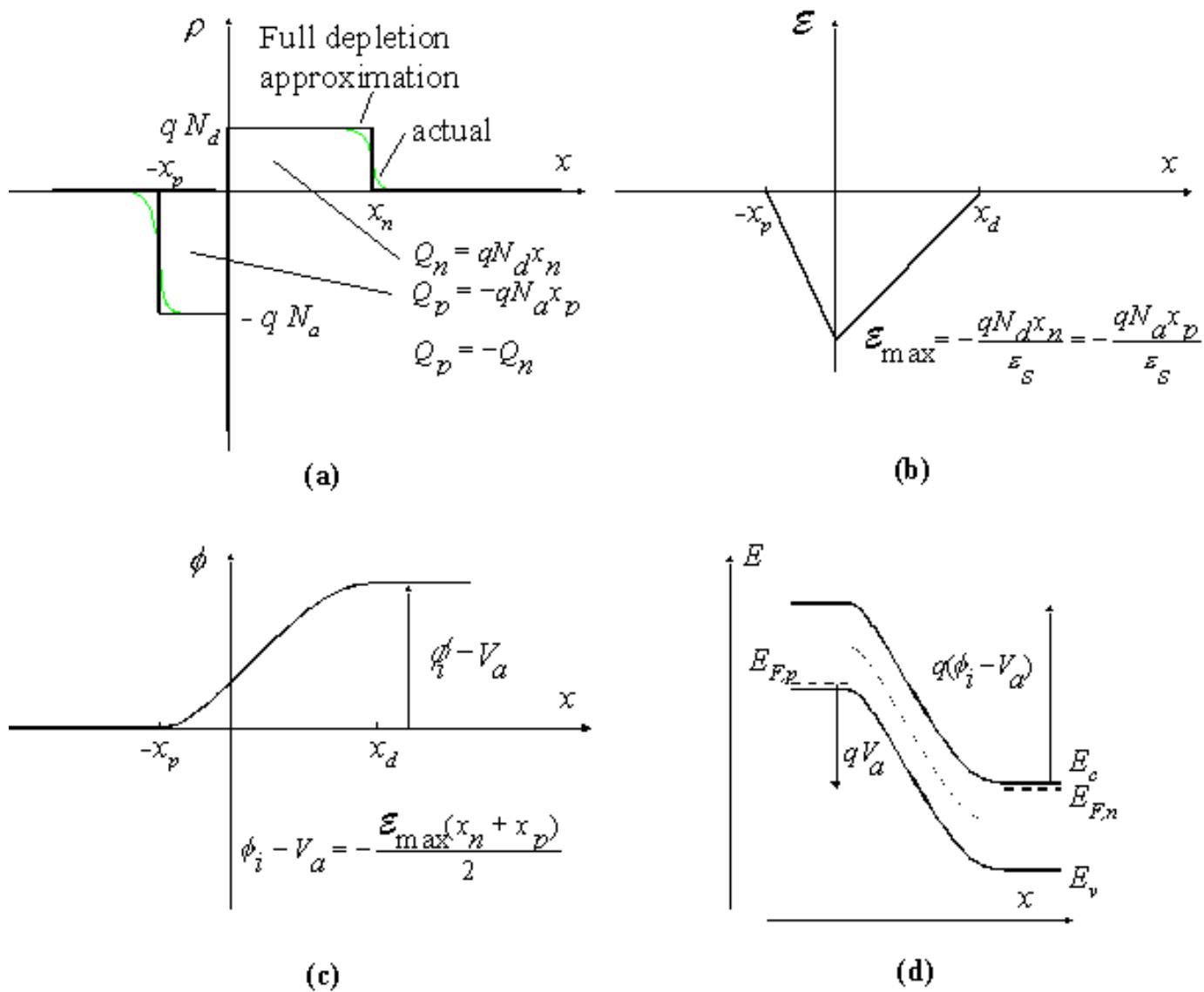


Figure 4.3.1: (a) Charge density in a p-n junction, (b) Electric field, (c) Potential and (d) Energy band diagram

As can be seen from Figure 4.3.1 (a), the charge density is constant in each region, as dictated by the full-depletion approximation. The total charge per unit area in each region is also indicated on the figure. The charge in the n-type region, Q_n , and the charge in the p-type region, Q_p , are given by:

$$Q_n = qN_d x_n \quad (4.3.7)$$

$$Q_p = -qN_a x_p \quad (4.3.8)$$

The electric field is obtained from the charge density using Gauss's law, which states that the field gradient equals the charge density divided by the dielectric constant or:

$$\frac{d\epsilon(x)}{dx} = \frac{\rho(x)}{\epsilon_s} \cong \frac{q}{\epsilon_s} (N_d^+(x) - N_a^-(x)), \text{ for } -x_p \leq x \leq x_n \quad (4.3.9)$$

The electric field is obtained by integrating equation (4.3.9). The boundary conditions consistent with the full depletion approximation are that the electric field is zero at both edges of the depletion region, namely at $x = -x_p$ and $x = x_n$. The electric field has to be zero outside the depletion region since any field would cause the free carriers to move thereby eliminating the electric field. Integration of the charge density in an abrupt p-n diode as shown in Figure 4.3.1 (a) is given by:

$$\mathcal{E}(x) = 0$$

$$\begin{aligned}\mathcal{E}(x) &= -\frac{qN_a(x + x_p)}{\epsilon_s} \\ \mathcal{E}(x) &= \frac{qN_d(x - x_n)}{\epsilon_s}\end{aligned}\tag{4.3.10}$$

$$\mathcal{E}(x) = 0$$

The electric field varies linearly in the depletion region and reaches a maximum value at $x = 0$ as can be seen on Figure 4.3.1(b). This maximum field can be calculated on either side of the depletion region, yielding:

$$\mathcal{E}(x = 0) = -\frac{qN_a x_p}{\epsilon_s} = -\frac{qN_d x_n}{\epsilon_s}\tag{4.3.11}$$

This provides the first relationship between the two unknowns, x_p and x_n , namely:

$$N_d x_n = N_a x_p\tag{4.3.12}$$

This equation expresses the fact that the total positive charge in the n -type depletion region, Q_n , exactly balances the total negative charge in the p -type depletion region, Q_p . We can then combine equation (4.3.4) with expression (4.3.12) for the total depletion-layer width, x_d , yielding:

$$x_n = x_d \frac{N_a}{N_a + N_d}\tag{4.3.13}$$

and

$$x_p = x_d \frac{N_d}{N_a + N_d}\tag{4.3.14}$$

The potential in the semiconductor is obtained from the electric field using:

$$\frac{d\phi(x)}{dx} = -\mathcal{E}(x)\tag{4.3.15}$$

We therefore integrate the electric field yielding a piece-wise parabolic potential versus position as shown in Figure 4.3.1(c)

The total potential across the semiconductor must equal the difference between the built-in potential and the applied voltage, which provides a second relation between x_p and x_n , namely:

$$\phi - V_a = \frac{qN_d x_n^2}{2\epsilon_s} + \frac{qN_a x_p^2}{2\epsilon_s}\tag{4.3.16}$$

The depletion layer width is obtained by substituting the expressions for x_p and x_n , (4.3.13) and (4.3.14), into the expression for the potential across the depletion region, yielding:

$$x_d = \sqrt{\frac{2\epsilon_s}{q} \left(\frac{1}{N_a} + \frac{1}{N_d} \right) (\phi - V_a)}\tag{4.3.17}$$

from which the solutions for the individual depletion layer widths, x_p and x_n are obtained:

$$x_n = \sqrt{\frac{2\epsilon_s}{q} \frac{N_a}{N_d} \frac{1}{N_a + N_d} (\phi - V_a)} \quad (4.3.18)$$

$$x_p = \sqrt{\frac{2\epsilon_s}{q} \frac{N_d}{N_a} \frac{1}{N_a + N_d} (\phi - V_a)} \quad (4.3.19)$$

Example 4.2	<p>An abrupt silicon ($n_i = 10^{10} \text{ cm}^{-3}$) p-n junction consists of a p-type region containing 10^{16} cm^{-3} acceptors and an n-type region containing $5 \times 10^{16} \text{ cm}^{-3}$ donors.</p> <ol style="list-style-type: none"> Calculate the built-in potential of this p-n junction. Calculate the total width of the depletion region if the applied voltage V_a equals 0, 0.5 and -2.5 V. Calculate maximum electric field in the depletion region at 0, 0.5 and -2.5 V. Calculate the potential across the depletion region in the n-type semiconductor at 0, 0.5 and -2.5 V.
Solution	<p>The built-in potential is calculated from:</p> $\phi = V_t \ln \frac{p_n n_p}{n_i^2} = 0.0259 \ln \frac{10^{16} \times 5 \times 10^{16}}{10^{20}} = 0.76 \text{ V}$ <p>The depletion layer width is obtained from:</p> $x_d = \sqrt{\frac{2\epsilon_s}{q} \left(\frac{1}{N_a} + \frac{1}{N_d} \right) (\phi - V_a)}$ <p>the electric field from</p> $\mathcal{E}(x = 0) = -\frac{2(\phi - V_a)}{x_d}$ <p>and the potential across the n-type region equals</p> $\phi_n = \frac{q N_d x_n^2}{2\epsilon_s}$ <p>where</p> $x_n = x_d \frac{N_a}{N_a + N_d}$ <p>one can also show that:</p>

$$\mathcal{A} = \frac{(\mathcal{A} - V_a) N_a}{N_a + N_d}$$

This yields the following numeric values:

	$V_a = 0 \text{ V}$	$V_a = 0.5 \text{ V}$	$V_a = -2.5 \text{ V}$
x_d	$0.315 \mu\text{m}$	$0.143 \mu\text{m}$	$0.703 \mu\text{m}$
ϵ	40 kV/cm	18 kV/cm	89 kV/cm
ϕ_n	0.105 V	0.0216 V	0.522 V



4.3.4. Junction capacitance

Any variation of the charge within a p-n diode with an applied voltage variation yields a capacitance, which must be added to the circuit model of a p-n diode. This capacitance related to the depletion layer charge in a p-n diode is called the junction capacitance.

The capacitance versus applied voltage is by definition the change in charge for a change in applied voltage, or:

$$C(V_a) \stackrel{\Delta}{=} \left| \frac{dQ(V_a)}{dV_a} \right| \quad (4.3.20)$$

The absolute value sign is added in the definition so that either the positive or the negative charge can be used in the calculation, as they are equal in magnitude. Using equation (4.3.7) and (4.3.18) one obtains:

$$C_j = \sqrt{\frac{q \epsilon_s}{2(\mathcal{A} - V_a)} \frac{N_a N_d}{N_a + N_d}} \quad (4.3.21)$$

A comparison with equation (4.3.17), which provides the depletion layer width, x_d , as a function of voltage, reveals that the expression for the junction capacitance, C_j , seems to be identical to that of a parallel plate capacitor, namely:

$$C_j = \frac{\epsilon_s}{x_d} \quad (4.3.22)$$

The difference, however, is that the depletion layer width and hence the capacitance is voltage dependent. The parallel plate expression still applies since charge is only added at the edge of the depletion regions. The distance between the added negative and positive charge equals the depletion layer width, x_d . A capacitance versus voltage measurement can be used to obtain the built-in voltage and the doping density of a one-sided p-n diode. When plotting the inverse of the capacitance squared, one expects a linear dependence as expressed by:

$$\frac{1}{C_j^2} = \frac{2}{q \epsilon_s} \frac{N_a + N_d}{N_a N_d} (\mathcal{A} - V_a) \quad (4.3.23)$$

The capacitance-voltage characteristic and the corresponding $1/C^2$ curve are shown in Figure 4.3.2.

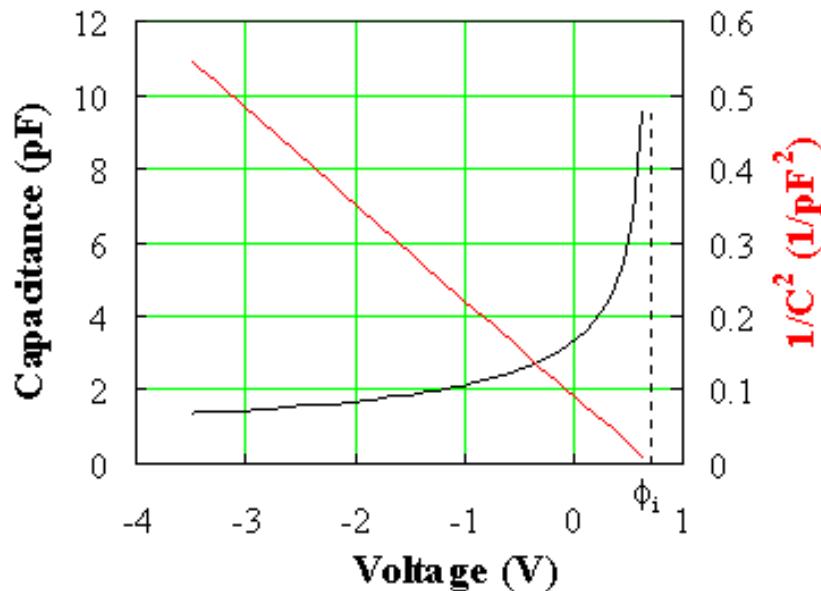


Figure 4.3.2 : Capacitance and $1/C^2$ versus voltage of a p-n diode with $N_a = 10^{16} \text{ cm}^{-3}$, $N_d = 10^{17} \text{ cm}^{-3}$ and an area of 10^{-4} cm^2 .

The built-in voltage is obtained at the intersection of the $1/C^2$ curve and the horizontal axis, while the doping density is obtained from the slope of the curve.

$$\frac{d(1/C_j^2)}{dV_a} = -\frac{2}{q\epsilon_s} \frac{N_a + N_d}{N_a N_d} \quad (4.3.24)$$

Example 4.3	Consider an abrupt p-n diode with $N_a = 10^{18} \text{ cm}^{-3}$ and $N_d = 10^{16} \text{ cm}^{-3}$. Calculate the junction capacitance at zero bias. The diode area equals 10^{-4} cm^2 . Repeat the problem while treating the diode as a one-sided diode and calculate the relative error.
Solution	<p>The built in potential of the diode equals:</p> $\phi_i = V_t \ln \frac{N_d N_a}{n_i^2} = 0.83 \text{ V}$ <p>The depletion layer width at zero bias equals:</p> $x_d = \sqrt{\frac{2 \epsilon_s (\phi_i - 0)}{q N_d}} = 0.33 \text{ } \mu\text{m}$ <p>And the junction capacitance at zero bias equals:</p> $C_{j0} = \frac{\epsilon_s}{x_d} \bigg _{V_a=0} = 3.17 \text{ pF}$ <p>Repeating the analysis while treating the diode as a one-sided diode, one only has to consider the region</p>

with the lower doping density so that

$$x_d \approx x_n = \sqrt{\frac{2 \epsilon_s}{q N_d} (A - V_a)} = 0.31 \mu\text{m}$$

And the junction capacitance at zero bias equals

$$C_{j0} = \left. \frac{\epsilon_s}{x_d} \right|_{V_a=0} = 3.18 \text{ pF}$$

The relative error equals 0.5 %, which justifies the use of the one-sided approximation.

A capacitance-voltage measurement also provides the doping density profile of one-sided p-n diodes. For a p⁺-n diode, one obtains the doping density from:

$$N_d = -\frac{2}{q \epsilon_s} \frac{1}{d(1/C_j^2)} \frac{dV_a}{dV_a} \text{, if } N_a \gg N_d \quad (4.3.25)$$

while the depth equals the depletion layer width which is obtained from $x_d = \epsilon_s A / C_j$. Both the doping density and the corresponding depth can be obtained at each voltage, yielding a doping density profile. Note that the capacitance in equations (4.3.21), (4.3.22), (4.3.23), and (4.3.25) is a capacitance per unit area.

As an example, we consider the measured capacitance-voltage data obtained on a 6H-SiC p-n diode. The diode consists of a highly doped p-type region on a lightly doped n-type region on top of a highly doped n-type substrate. The measured capacitance as well as $1/C^2$ is plotted as a function of the applied voltage. The dotted line forms a reasonable fit at voltages close to zero from which one can conclude that the doping density is almost constant close to the p-n interface. At large negative voltages the capacitance becomes almost constant which corresponds to a high doping density according to equation (4.3.25).

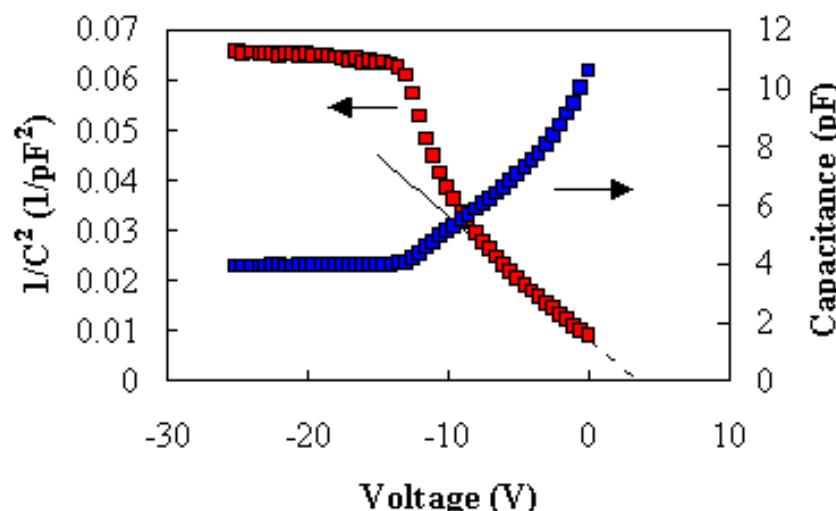
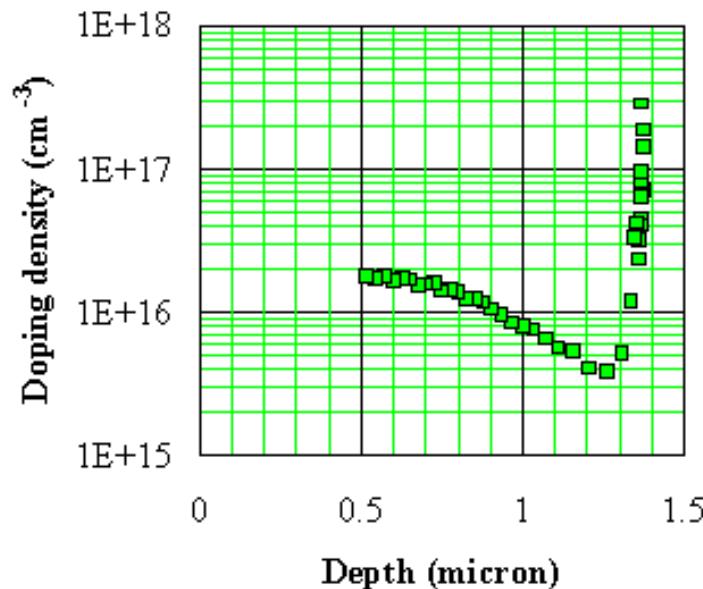
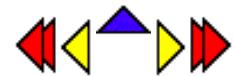


Figure 4.3.3 : Capacitance and $1/C^2$ versus voltage of a 6H-SiC p-n diode.

The doping profile calculated from the data presented in Figure 4.3.3 is shown in Figure 4.3.4. The figure confirms the presence of the highly doped substrate and yields the thickness of the n-type layer. No information is obtained at the interface ($x = 0$) as is typical for doping profiles obtained from C-V measurements. This is because the capacitance measurement is limited to small forward bias voltages since the forward bias current and the diffusion capacitance affect the accuracy of the capacitance measurement.

**Figure 4.3.4 :** Doping profile corresponding to the measured data, shown in Figure 4.3.3.

Chapter 5: Bipolar Junction Transistors



Problems

1. A silicon npn bipolar transistor with $N_E = 10^{18} \text{ cm}^{-3}$, $N_B = 10^{17} \text{ cm}^{-3}$ and $N_C = 10^{16} \text{ cm}^{-3}$, $w_E = 1 \mu\text{m}$, $w_B = 0.5 \mu\text{m}$, and $w_C = 4 \mu\text{m}$ is biased with $V_{BE} = 0.6 \text{ V}$ and $V_{CB} = 0 \text{ V}$. Use $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$, $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\tau_n = \tau_p = 100 \text{ ns}$. The emitter area equals 10^{-4} cm^2 .
 - a. Calculate the width of the quasi-neutral regions in the emitter, base and collector.
 - b. Calculate the minority-carrier diffusion lengths in the emitter, base and collector. Calculate the ratio of the minority-carrier diffusion length and the quasi-neutral region width in each region.
 - c. Calculate the excess-minority-carrier charge density per unit area in the emitter, base and collector.
 - d. Calculate the emitter current while ignoring the recombination in the depletion region.
 - e. Calculate the base transit time and the current due to recombination of electrons in the base.
 - f. Calculate the emitter efficiency and the base transport factor.
 - g. Calculate the transport factor and the current gain assuming there is no recombination in the depletion regions.
 - h. Calculate the collector capacitance, the majority-carrier charge density in the base and the Early voltage.
2. A silicon npn bipolar transistor has an emitter doping, $N_E = 2 \times 10^{18} \text{ cm}^{-3}$, an emitter width $w_E = 1 \mu\text{m}$, and a base doping of $2 \times 10^{17} \text{ cm}^{-3}$. A current gain of 100 and an early voltage of 100 V is desired. Using $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$, $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\tau_n = \tau_p = 100 \text{ ns}$, find the corresponding base width and collector doping. The emitter area equals 10^{-4} cm^2 .

Chapter 6: MOS Capacitors



Problems

1. Consider an aluminum-SiO₂-silicon MOS capacitor ($\Phi_M = 4.1$ V, $\epsilon_{ox}/\epsilon_0 = 3.9$, $\chi = 4.05$ V and $N_a = 10^{17}$ cm⁻³) MOS capacitor with $t_{ox} = 5$ nm.
 - a. Calculate the flatband voltage and threshold voltage.
 - b. Repeat for an n-type silicon substrate with $N_d = 10^{16}$ cm⁻³.
 - c. Repeat with a surface charge of 10⁻⁷ C/cm²
 - d. Repeat with a charge density in the oxide of 10⁻¹ C/cm³
- A high-frequency capacitance voltage measurement of a silicon MOS structure was fitted by the following expression:

$$C(V_G) = 6 \text{ pF} + 12 \text{ pF}/(1 + \exp(V_G))$$

- a. Calculate the oxide capacitance per unit area and the oxide thickness. The area of the capacitor is 100 x 100 micron and the relative dielectric constant equals 3.9.

From the minimum capacitance, calculate the maximum depletion layer width and the substrate doping density.

- Calculate the bulk potential.
- Calculate the flatband capacitance and the flatband voltage.
- Calculate the threshold voltage.
- An MOS capacitor with an oxide thickness of 20 nm has an oxide capacitance, which is three times larger than the minimum high-frequency capacitance in inversion. Find the substrate doping density.
- A CMOS gate requires n-type and p-type MOS capacitors with a threshold voltage of 2 and -2 Volt respectively. If the gate oxide is 50 nm what are the required substrate doping densities? Assume the gate electrode is aluminum. Repeat for a p+ poly-silicon gate.
- Consider a p-MOS capacitor (with an n-type substrate) and with an aluminum gate. Find the doping density for which the threshold voltage is 3 times larger than the flat band voltage. $t_{ox} = 25$ nm. Repeat for a capacitor with 10¹¹ cm⁻² electronic charges at the oxide-semiconductor interface.
- A silicon p-MOS capacitor ($N_d = 4 \times 10^{16}$ cm⁻³, $t_{ox} = 40$ nm) is biased halfway between the flatband and threshold voltage. Calculate the applied voltage and the corresponding capacitance

Problems

1. A silicon npn bipolar transistor with $N_E = 10^{18} \text{ cm}^{-3}$, $N_B = 10^{17} \text{ cm}^{-3}$ and $N_C = 10^{16} \text{ cm}^{-3}$, $w_E = 1 \mu\text{m}$, $w_B = 0.5 \mu\text{m}$, and $w_C = 4 \mu\text{m}$ is biased with $V_{BE} = 0.6 \text{ V}$ and $V_{CB} = 0 \text{ V}$. Use $\mathbf{m}_n = 1000 \text{ cm}^2/\text{V-s}$, $\mathbf{m}_p = 300 \text{ cm}^2/\text{V-s}$ and $t_n = t_p = 100 \text{ ns}$. The emitter area equals 10^{-4} cm^2 .
 - a) Calculate the width of the quasi-neutral regions in the emitter, base and collector.
 - b) Calculate the minority-carrier diffusion lengths in the emitter, base and collector. Calculate the ratio of the minority-carrier diffusion length and the quasi-neutral region width in each region.
 - c) Calculate the excess-minority-carrier charge density per unit area in the emitter, base and collector.
 - d) Calculate the emitter current while ignoring the recombination in the depletion region.
 - e) Calculate the base transit time and the current due to recombination of electrons in the base.
 - f) Calculate the emitter efficiency and the base transport factor.
 - g) Calculate the transport factor and the current gain assuming there is no recombination in the depletion regions.
 - h) Calculate the collector capacitance, the majority-carrier charge density in the base and the Early voltage.
2. A silicon npn bipolar transistor has an emitter doping, $N_E = 2 \times 10^{18} \text{ cm}^{-3}$, an emitter width $w_E = 1 \mu\text{m}$, and a base doping of $2 \times 10^{17} \text{ cm}^{-3}$. A current gain of 100 and an early voltage of 100 V is desired. Using $\mathbf{m}_n = 1000 \text{ cm}^2/\text{V-s}$, $\mathbf{m}_p = 300 \text{ cm}^2/\text{V-s}$ and $t_n = t_p = 100 \text{ ns}$, find the corresponding base width and base doping. The emitter area equals 10^{-4} cm^2 .

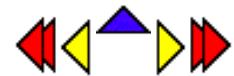
Problems

1. A silicon p-n junction ($N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$) is biased with $V_a = -3 \text{ V}$. Calculate the built-in potential, the depletion layer width and the maximum electric field of the junction.
2. An abrupt silicon p-n junction consists of a p-type region containing 10^{16} cm^{-3} acceptors and an n-type region containing also 10^{16} cm^{-3} acceptors in addition to 10^{17} cm^{-3} donors.
 - a) Calculate the thermal equilibrium density of electrons and holes in the p-type region as well as both densities in the n-type region.
 - b) Calculate the built-in potential of the p-n junction.
 - c) Calculate the built-in potential of the p-n junction at 100°C .
3. For a p-n junction with a built-in potential of 0.62 V
 - a) What is the potential across the depletion region at an applied voltage, V_a , of 0, 0.5 and -2 Volt?
 - b) If the depletion layer is 1 micrometer at $V_a = 0$ Volt, find the maximum electric field in the depletion region.
 - c) Assuming that the net doping density $|N_d - N_a|$ is the same in the n-type and p-type region of the diode, carefully sketch the electric field and the potential as a function of position throughout the depletion region. Add numeric values wherever possible.
4. An abrupt silicon ($n_i = 10^{10} \text{ cm}^{-3}$) p-n junction consists of a p-type region containing 10^{16} cm^{-3} acceptors and an n-type region containing $5 \times 10^{16} \text{ cm}^{-3}$ donors.
 - a) Calculate the built-in potential of this p-n junction.
 - b) Calculate the total width of the depletion region if the applied voltage V_a equals 0, 0.5 and -2.5 V.
 - c) Calculate maximum electric field in the depletion region at 0, 0.5 and -2.5 V.
 - d) Calculate the potential across the depletion region in the n-type semiconductor at 0, 0.5 and -2.5 V.
5. Consider an abrupt p-n diode in thermal equilibrium with as many donors in the n-type region as acceptors in the p-type region and a maximum electric field of -13 kV/cm and a total depletion layer width of $1 \mu\text{m}$. (assume $\epsilon_s/\epsilon_0 = 12$)
 - a) What is the applied voltage, V_a ?
 - b) What is the built-in potential of the diode?
 - c) What is the donor density in the n-type region and the acceptor density in the p-type region?
 - d) What is the intrinsic carrier density of the semiconductor if the temperature is 300 K ?

6. A silicon ($n_i = 10^{10} \text{ cm}^{-3}$) p-n diode with $N_a = 10^{18} \text{ cm}^{-3}$ has a capacitance of 10^{-8} F/cm^2 at an applied voltage of 0.5 V. Find the donor density.
7. A silicon ($n_i = 10^{10} \text{ cm}^{-3}$) p-n diode has a maximum electric field of -10^6 V/cm and a depletion layer width of $1 \mu\text{m}$. The acceptor density in the p-type region is four times larger than the donor density in the n-type region. Calculate both doping densities.
8. Consider a symmetric silicon p-n diode ($N_a = N_d$)
 - a) Calculate the built-in potential if $N_a = 10^{13}, 10^{15}$ and 10^{17} cm^{-3} . Also, calculate the doping densities corresponding to a built-in potential of 0.7 V.
 - b) For the same as in part a), calculate the total depletion layer widths, the capacitance per unit area and the maximum electric field in thermal equilibrium.
 - c) Repeat part a) and b) with $N_a = 3 N_d$.
9. A one-sided silicon diode has a breakdown voltage of 1000 V for which the maximum electric field at breakdown is 100 kV/cm . What is the maximum possible doping density in the low doped region, the built-in potential, the depletion layer width and the capacitance per unit area? Assume that bulk potential of the highly doped region is $E_g/2$ ($= 0.56 \text{ V}$).
10. A silicon p-n junction ($N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$) is biased with $V_a = 0.6 \text{ V}$. Calculate the ideal diode current assuming that the n-type region is much smaller than the diffusion length with $w_n' = 1 \mu\text{m}$ and assuming a "long" p-type region. Use $m_i = 1000 \text{ cm}^2/\text{V-s}$ and $m_p = 300 \text{ cm}^2/\text{V-s}$. The minority carrier lifetime is $10 \mu\text{s}$ and the diode area is $100 \mu\text{m}$ by $100 \mu\text{m}$.
11. Derive equation 4.4.14.
12. Calculate the relative error when using the "short diode" approximation if $L_n = 2 w_p'$ and $L_p = 2 w_n'$.
13. A silicon p-n junction ($N_a = 10^{15} \text{ cm}^{-3}$, $w_p = 1 \mu\text{m}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$, $w_n = 1 \mu\text{m}$) is biased with $V_a = 0.5 \text{ V}$. Use $m_i = 1000 \text{ cm}^2/\text{V-s}$ and $m_p = 300 \text{ cm}^2/\text{V-s}$. The minority carrier lifetime is $10 \mu\text{s}$ and the diode area is $100 \mu\text{m}$ by $100 \mu\text{m}$.
 - a) Calculate the built-in potential of the diode.
 - b) Calculate the depletion layer widths, x_n and x_p , and the widths of the quasi-neutral regions.
 - c) Compare the width of the quasi-neutral regions with the minority-carrier diffusion-lengths and decide whether to use the "long" or "short" diode approximation. Calculate the current through the diode.
 - d) Compare the result of part c) with the current obtained by using the general solution (equation 4.4.14)

- e) Using the approximation chosen in part c) calculate the ratio of the electron current to the hole current traversing the depletion region.
14. An abrupt silicon p-n diode consists of a p-type region containing 10^{18} cm^{-3} acceptors and an n-type region containing 10^{15} cm^{-3} donors.
- Calculate the breakdown field in the n-type region.
 - Using the breakdown field from part a), calculate the breakdown voltage of the diode.
 - What is the depletion layer width at breakdown?
 - Discuss edge effects and specify the minimum junction depth needed to avoid these effects.
15. A 1 cm^2 solar cell consists of a p-type region containing 10^{18} cm^{-3} acceptors and an n-type region containing 10^{15} cm^{-3} donors. $w_p = 0.1 \mu\text{m}$ and $w_n \gg L_p$. Use $m_i = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $m_b = 300 \text{ cm}^2/\text{V}\cdot\text{s}$. The minority carrier lifetime is $10 \mu\text{s}$. The diode is illuminated with sun light, yielding a photocurrent density of 30 mA/cm^2 .
- Calculate the open circuit voltage and short-circuit current of the solar cell.
 - Calculate the maximum power generated by the cell and the corresponding voltage and current.
 - Calculate the fill factor of the solar cell.
 - Calculate the fill factor for the same cell when a concentrator illuminates it so that the photocurrent density equals 300 A/cm^2 .

Chapter 6: MOS Capacitors



Examples

Example 6.1



Calculate the flatband voltage of a silicon nMOS capacitor with a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$ and an aluminum gate ($\Phi_M = 4.1 \text{ V}$). Assume there is no fixed charge in the oxide or at the oxide-silicon interface.

Example 6.2



Calculate the threshold voltage of a silicon nMOS capacitor with a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$, a 20 nm thick oxide ($\epsilon_{ox} = 3.9 \epsilon_0$) and an aluminum gate ($\Phi_M = 4.1 \text{ V}$). Assume there is no fixed charge in the oxide or at the oxide-silicon interface.

Example 6.3



Calculate the oxide capacitance, the flatband capacitance and the high frequency capacitance in inversion of a silicon nMOS capacitor with a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$, a 20 nm thick oxide ($\epsilon_{ox} = 3.9 \epsilon_0$) and an aluminum gate ($\Phi_M = 4.1 \text{ V}$).

Chapter 7: MOS Field-Effect-Transistors



Examples

Example 7.1

Calculate the drain current of a silicon nMOSFET with $V_T = 1$ V, $W = 10$ μm , $L = 1$ μm and $t_{\text{ox}} = 20$ nm. The device is biased with $V_{\text{GS}} = 3$ V and $V_{\text{DS}} = 5$ V. Use the quadratic model, a surface mobility of $300 \text{ cm}^2/\text{V}\cdot\text{s}$ and set $V_{\text{BS}} = 0$ V.

Also calculate the transconductance at $V_{\text{GS}} = 3$ V and $V_{\text{DS}} = 5$ V and compare it to the output conductance at $V_{\text{GS}} = 3$ V and $V_{\text{DS}} = 0$ V.

Example 7.2

Repeat example 7.1 using the variable depletion layer model. Use $V_{\text{FB}} = -0.807$ V and $N_a = 10^{17} \text{ cm}^{-3}$.

Example 7.3

Calculate the threshold voltage of a silicon nMOSFET when applying a substrate voltage, $V_{\text{BS}} = 0, -2.5, -5, -7.5$ and -10 V. The capacitor has a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$, a 20 nm thick oxide ($\epsilon_{\text{ox}} = 3.9 \epsilon_0$) and an aluminum gate ($\Phi_M = 4.1$ V). Assume there is no fixed charge in the oxide or at the oxide-silicon interface.

Example 6.1 Calculate the flatband voltage of a silicon nMOS capacitor with a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$ and an aluminum gate ($\Phi_M = 4.1 \text{ V}$). Assume there is no fixed charge in the oxide or at the oxide-silicon interface.

Solution The flatband voltage equals the work function difference since there is no charge in the oxide or at the oxide-semiconductor interface.

$$V_{FB} = \Phi_{MS} = \Phi_M - \mathbf{c} - \frac{E_g}{2q} - V_t \ln \frac{N_a}{n_i}$$
$$= 4.1 - 4.05 - 0.56 - 0.026 \times \ln \frac{10^{17}}{10^{10}} = -0.93 \text{ V}$$

The flatband voltages for nMOS and pMOS capacitors with an aluminum or a poly-silicon gate are listed in the table below.

	Aluminum	p^+ poly	n^+ poly
nMOS	-0.93 V	0.14 V	-0.98 V
pMOS	-0.09 V	0.98 V	-0.14 V

Example 6.2 Calculate the threshold voltage of a silicon nMOS capacitor with a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$, a 20 nm thick oxide ($\epsilon_{ox} = 3.9 \epsilon_0$) and an aluminum gate ($\Phi_M = 4.1 \text{ V}$). Assume there is no fixed charge in the oxide or at the oxide-silicon interface.

Solution

The threshold voltage equals:

$$V_T = V_{FB} + 2f_F + \frac{\sqrt{4\epsilon_s q N_a f_F}}{C_{ox}}$$

$$= -0.93 + 2 \times 0.42$$

$$+ \frac{\sqrt{4 \times 11.9 \times 8.85 \times 10^{-14} \times 1.6 \times 10^{-19} \times 10^{17} \times 0.42}}{3.9 \times 8.85 \times 10^{-14} / 20 \times 10^{-7}}$$

$$= -0.09 \text{ V}$$

Where the flatband voltage was already calculated in example 6.1. The threshold voltage voltages for nMOS and pMOS capacitors with an aluminum or a poly-silicon gate are listed in the table below.

	Aluminum	p^+ poly	n^+ poly
nMOS	-0.09 V	0.98 V	-0.14 V
pMOS	-0.93 V	0.14 V	-0.98 V

Example 6.3 Calculate the oxide capacitance, the flatband capacitance and the high frequency capacitance in inversion of a silicon nMOS capacitor with a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$, a 20 nm thick oxide ($\epsilon_{ox} = 3.9 \epsilon_0$) and an aluminum gate ($\Phi_M = 4.1 \text{ V}$).

Solution The oxide capacitance equals:

$$C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = \frac{3.9 \times 8.85 \times 10^{-14}}{2 \times 10^{-6}} = 173 \text{ nF/cm}^2$$

The flatband capacitance equals:

$$C_{FB} = \frac{1}{\frac{1}{C_{ox}} + \frac{L_D}{\epsilon_s}} = \frac{1}{\frac{1}{173 \times 10^{-9}} + \frac{1.3 \times 10^{-6}}{11.9 \times 8.85 \times 10^{-14}}} = 142 \text{ nF/cm}^2$$

where the Debye length is obtained from:

$$L_D = \sqrt{\frac{\epsilon_s V_t}{q N_a}} = \sqrt{\frac{11.9 \times 8.85 \times 10^{-14} \times 0.0259}{1.6 \times 10^{-19} \times 10^{17}}} = 13 \text{ nm}$$

The high frequency capacitance in inversion equals:

$$C_{HF,inv} = \frac{1}{\frac{1}{C_{ox}} + \frac{x_{d,T}}{\epsilon_s}} = \frac{1}{\frac{1}{173 \times 10^{-9}} + \frac{1.05 \times 10^{-5}}{11.9 \times 8.85 \times 10^{-14}}} = 63 \text{ nF/cm}^2$$

and the depletion layer width at threshold equals:

$$x_{d,T} = \sqrt{\frac{2\epsilon_s (2f_F)}{q N_a}} = \sqrt{\frac{2 \times 11.9 \times 8.85 \times 10^{-14} \times 2 \times 0.419}{1.6 \times 10^{-19} \times 10^{17}}} = 105 \text{ nm}$$

The bulk potential, f_F , was already calculated in example 6.1

Example 5.1 A bipolar transistor with an emitter current of 1 mA has an emitter efficiency of 0.99, a base transport factor of 0.995 and a depletion layer recombination factor of 0.998. Calculate the base current, the collector current, the transport factor and the current gain of the transistor.

Solution The transport factor and current gain are:

$$a = g_E a_T d_r = 0.99 \times 0.995 \times 0.998 = 0.983$$

and

$$b = \frac{a}{1-a} = 58.1$$

The collector current then equals

$$I_C = a I_E = 0.983 \text{ mA}$$

And the base current is obtained from:

$$I_B = I_E - I_C = 17 \mu\text{A}$$

Example 5.2 Consider a pnp bipolar transistor with emitter doping of 10^{18} cm^{-3} and base doping of 10^{17} cm^{-3} . The quasi-neutral region width in the emitter is $1 \mu\text{m}$ and $0.2 \mu\text{m}$ in the base. Use $m_i = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $m_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$. The minority carrier lifetime in the base is 10 ns.

Calculate the emitter efficiency, the base transport factor, and the current gain of the transistor biased in the forward active mode. Assume there is no recombination in the depletion region.

Solution

The emitter efficiency is obtained from:

$$g_E = \frac{1}{1 + \frac{D_{p,E} N_B w_B}{D_{n,B} N_E w_E}} = 0.994$$

The base transport factor equals:

$$a_T = 1 - \frac{w_B^2}{2D_{n,B} t_n} = 0.9992$$

The current gain then becomes:

$$b = \frac{a}{1-a} = 147.5$$

where the transport factor, a , was calculated as the product of the emitter efficiency and the base transport factor:

$$a = g_E a_T = 0.994 \times 0.9992 = 0.993$$

Example 5.3 Calculate the saturation voltage of a bipolar transistor biased with a base current of 1 mA and a collector current of 10 mA. Use $\alpha_R = 0.993$ and $\alpha_F = 0.2$.

Solution The saturation voltage equals:

$$V_{CE,sat} = V_t \ln \left\{ \frac{1 + \frac{I_C}{I_B} (1 - \alpha_R)}{\alpha_R \left[1 - \frac{I_C}{I_B} \frac{(1 - \alpha_F)}{\alpha_F} \right]} \right\} = 0.1 \text{ V}$$

Example 5.4 Consider a bipolar transistor with a base doping of 10^{17} cm^{-3} and a quasi-neutral base width of $0.2 \mu\text{m}$. Calculate the Early voltage and collector current ideality factor given that the base-emitter capacitance and the base-collector capacitance are 0.2 nF and 0.2 pF . The collector area equals 10^{-4} cm^{-2} .

Solution

The Early voltage equals:

$$|V_A| = \frac{Q_B}{C_{j,BC}} = \frac{qA_C N_B w_B}{C_{j,BC}} = 160 \text{ V}$$

The saturation voltage equals:

$$n \equiv 1 + \frac{V_t}{Q_B} C_{j,BE} = 1.16$$

Chapter 4: p-n Junctions



4.2. Structure and principle of operation

4.2.1. Structure

4.2.2. Thermal equilibrium

4.2.3. The built-in potential

4.2.4. Forward and reverse bias

A p-n junction consists of two semiconductor regions with opposite doping type as shown in Figure 4.2.1. The region on the left is *p*-type with an acceptor density N_a , while the region on the right is *n*-type with a donor density N_d . The dopants are assumed to be shallow, so that the electron (hole) density in the *n*-type (*p*-type) region is approximately equal to the donor (acceptor) density.

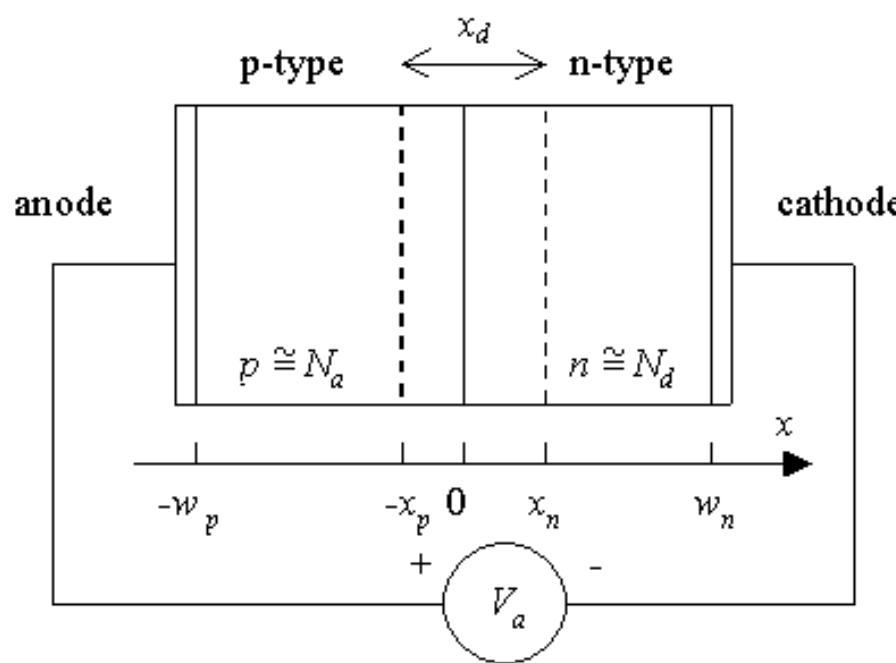


Figure 4.2.1 :

Cross-section of a p-n junction

We will assume, unless stated otherwise, that the doped regions are uniformly doped and that the transition between the two regions is abrupt. We will refer to this structure as being an abrupt p-n junction.

Frequently we will deal with p-n junctions in which one side is distinctly higher-doped than the other. We will find that in such a case only the low-doped region needs to be considered, since it primarily determines the device characteristics. We will refer to such a structure as a one-sided abrupt p-n junction.

The junction is biased with a voltage V_a as shown in Figure 4.2.1. We will call the junction forward-biased if a positive voltage is applied to the *p*-doped region and reversed-biased if a negative voltage is applied to the *p*-doped region. The contact to the *p*-type region is also called the anode, while the contact to the *n*-type region is called the cathode, in reference to the anions or positive carriers and cations or negative carriers in each of these regions.



4.2.1. Flatband diagram

The principle of operation will be explained using a gedanken experiment, an experiment, which is in principle possible but not necessarily executable in practice. We imagine that one can bring both semiconductor regions together, aligning both the conduction and valence band energies of each region. This yields the so-called flatband diagram shown in Figure 4.2.2.

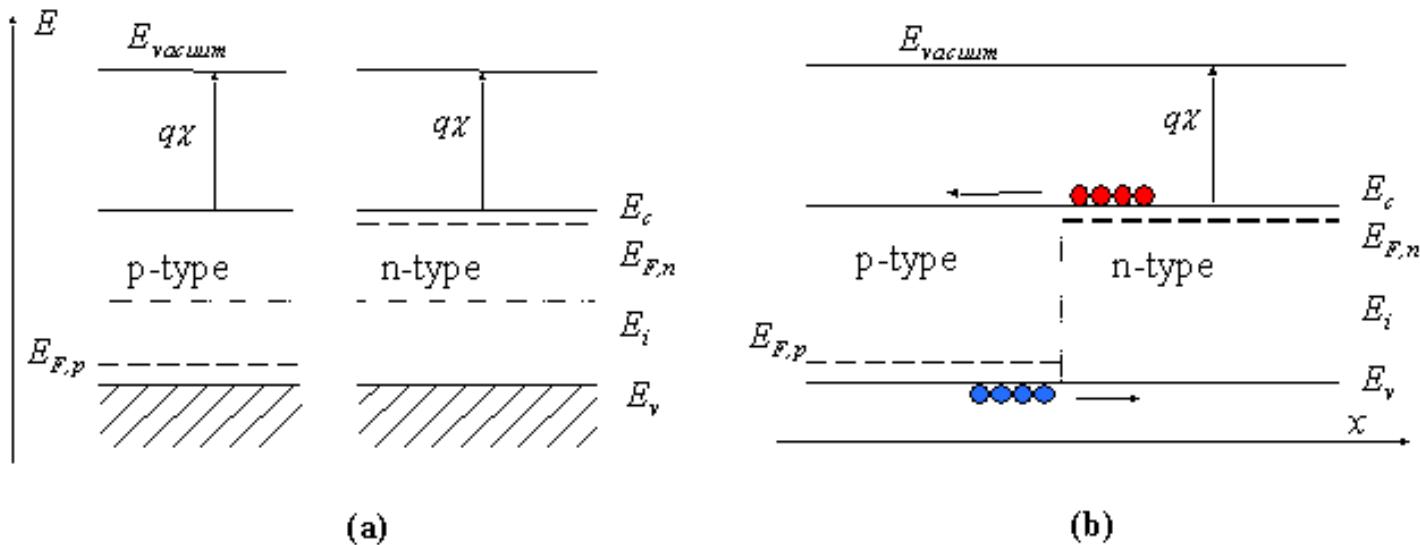


Figure 4.2.2 : Energy band diagram of a p-n junction (a) before and (b) after merging the n-type and p-type regions

Note that this does not automatically align the Fermi energies, E_{Fn} and E_{Fp} . Also, note that this flatband diagram is not an equilibrium diagram since both electrons and holes can lower their energy by crossing the junction. A motion of electrons and holes is therefore expected before thermal equilibrium is obtained. The diagram shown in Figure 4.2.2 (b) is called a flatband diagram. This name refers to the horizontal band edges. It also implies that there is no field in the semiconductor and no charge.



4.2.2. Thermal equilibrium

To reach thermal equilibrium, electrons/holes close to the metallurgical junction diffuse across the junction into the p-type/n-type region where hardly any electrons/holes are present. This process leaves the ionized donors (acceptors) behind, creating a region around the junction, which is depleted of mobile carriers. We call this region the depletion region, extending from $x = -x_p$ to $x = x_n$. The charge due to the ionized donors and acceptors causes an electric field, which in turn causes a drift of carriers in the opposite direction. The diffusion of carriers continues until the drift current balances the diffusion current, thereby reaching thermal equilibrium as indicated by a constant Fermi energy. This situation is shown in Figure 4.2.3:

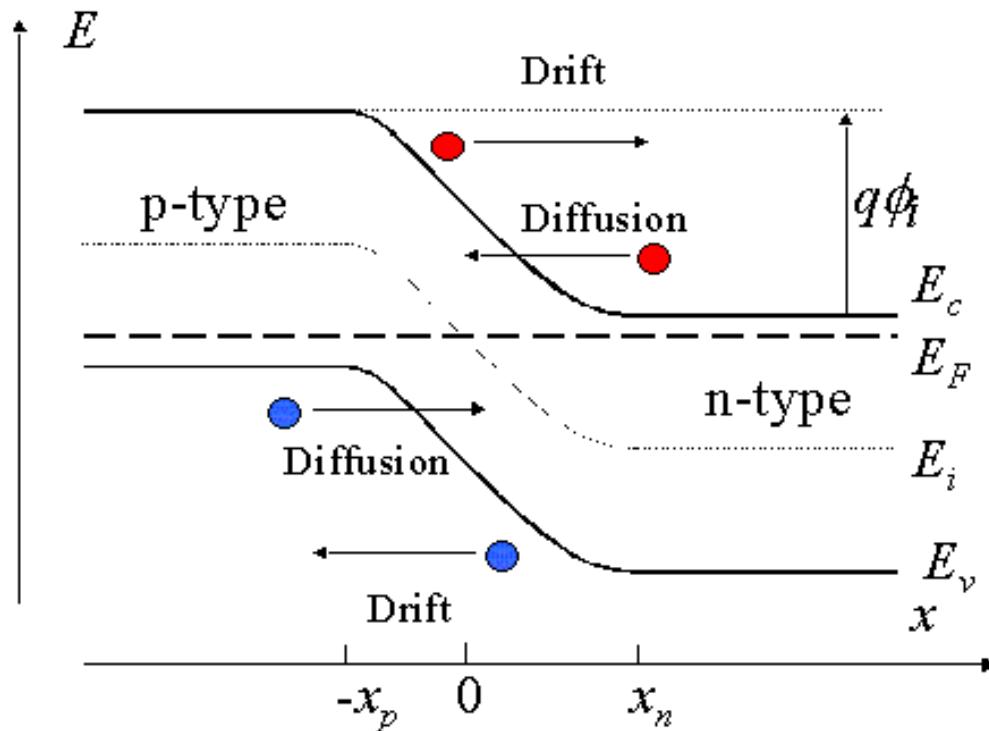


Figure 4.2.3 : Energy band diagram of a p-n junction in thermal equilibrium

While in thermal equilibrium no external voltage is applied between the n-type and p-type material, there is an internal potential, ϕ_i , which is caused by the workfunction difference between the n-type and p-type semiconductors. This potential equals the built-in potential, which will be further discussed in the next section.

4.2.3. The built-in potential



The built-in potential in a semiconductor equals the potential across the depletion region in thermal equilibrium. Since thermal equilibrium implies that the Fermi energy is constant throughout the p-n diode, the built-in potential equals the difference in the Fermi energies, E_{Fn} and E_{Fp} , divided by the electronic charge. It also equals the sum of the bulk potentials of each region, ϕ_n and ϕ_p , since the bulk potential quantifies the distance between the Fermi energy and the intrinsic energy. This yields the following expression for the built-in potential.

$$\phi_i = V_t \ln \frac{N_d N_a}{n_i^2} \quad (4.2.1)$$

Example 4.1	<p>An abrupt silicon p-n junction consists of a p-type region containing $2 \times 10^{16} \text{ cm}^{-3}$ acceptors and an n-type region containing also 10^{16} cm^{-3} acceptors in addition to 10^{17} cm^{-3} donors.</p> <ol style="list-style-type: none"> Calculate the thermal equilibrium density of electrons and holes in the p-type region as well as both densities in the n-type region. Calculate the built-in potential of the p-n junction. Calculate the built-in potential of the p-n junction at 400 K.
-------------	---

Solution

a. The thermal equilibrium densities are:

In the *p*-type region:

$$p = N_a = 2 \times 10^{16} \text{ cm}^{-3}$$

$$n = n_i^2/p = 10^{20}/2 \times 10^{16} = 5 \times 10^3 \text{ cm}^{-3}$$

In the *n*-type region

$$n = N_d - N_a = 9 \times 10^{16} \text{ cm}^{-3}$$

$$p = n_i^2/n = 10^{20}/(1 \times 10^{16}) = 1.11 \times 10^3 \text{ cm}^{-3}$$

b. The built-in potential is obtained from:

$$\phi_B = V_t \ln \frac{p_n n_p}{n_i^2} = 0.0259 \ln \frac{2 \times 10^{16} \times 9 \times 10^{16}}{10^{20}} = 0.79 \text{ V}$$

c. Similarly, the built-in potential at 400 K equals:

$$\phi_B = V_t \ln \frac{p_n n_p}{n_i^2} = 0.0345 \ln \frac{2 \times 10^{16} \times 9 \times 10^{16}}{(4.52 \times 10^{12})^2} = 0.63 \text{ V}$$

where the intrinsic carrier density at 400 K was obtained from example 2.4.b

4.2.4. Forward and reverse bias



We now consider a p-n diode with an applied bias voltage, V_a . A forward bias corresponds to applying a positive voltage to the anode (p-type region) relative to the cathode (n-type region). A reverse bias corresponds to a negative voltage applied to the cathode. Both bias modes are illustrated with Figure 4.2.4. The applied voltage is proportional to the difference between the Fermi energy in the n-type and p-type quasi-neutral regions.

As a negative voltage is applied, the potential across the semiconductor increases and so does the depletion layer width. As a positive voltage is applied, the potential across the semiconductor decreases and with it the depletion layer width. The total potential across the semiconductor equals the built-in potential minus the applied voltage, or:

$$\phi = \phi_B - V_a \quad (4.2.1)$$

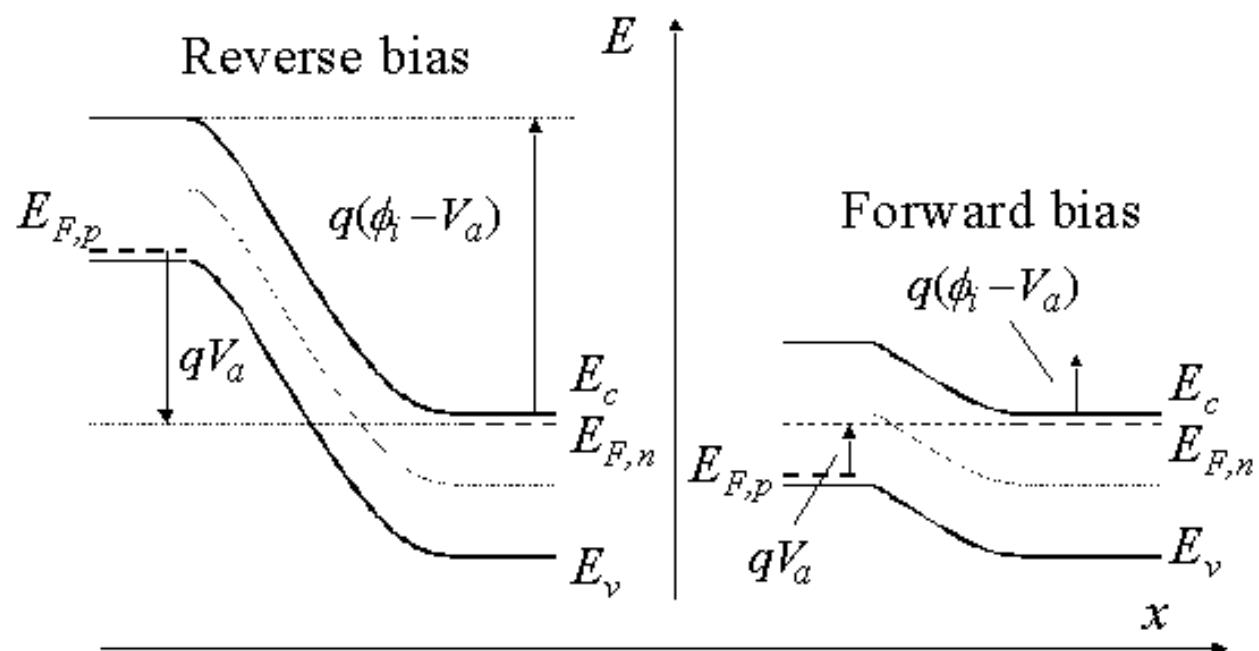
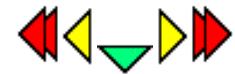


Figure 4.2.4: Energy band diagram of a p-n junction under reverse and forward bias

Chapter 4: p-n Junctions



4.4. The p-n diode current

- [4.4.1. General discussion](#)
- [4.4.2. The ideal diode current](#)
- [4.4.3. Recombination-Generation current](#)
- [4.4.4. I-V characteristics of real p-n diodes](#)
- [4.4.5. The diffusion capacitance](#)

4.4.1. General discussion



The current in a p-n diode is due to carrier recombination or generation somewhere within the p-n diode structure. Under forward bias, the diode current is due to recombination. This recombination can occur within the quasi-neutral semiconductor, within the depletion region or at the metal-semiconductor Ohmic contacts. Under reverse bias, the current is due to generation. Carrier generation due to light will further increase the current under forward as well as reverse bias.

In this section, we first derive the ideal diode current. We will also distinguish between the "long" diode and "short" diode case. The "long" diode expression applies to p-n diodes in which recombination/generation occurs in the quasi-neutral region only. This is the case if the quasi-neutral region is much larger than the carrier diffusion length. The "short" diode expression applies to p-n diodes in which recombination/generation occurs at the contacts only. In a short diode, the quasi-neutral region is much smaller than the diffusion length. In addition to the "long" and "short" diode expressions, we also present the general result, which deals with recombination/generation in a finite quasi-neutral region.

Next, we derive expressions for the recombination/generation in the depletion region. Here we have to distinguish between the different recombination mechanisms - band-to-band recombination and Shockley-Hall-Read recombination - as they lead to different current-voltage characteristics.

4.4.2. The ideal diode current



4.4.2.1. General discussion and overview

When calculating the current in a p-n diode one needs to know the carrier density and the electric field throughout the p-n diode which can then be used to obtain the drift and diffusion current. Unfortunately, this requires the knowledge of the quasi-Fermi energies, which is only known if the currents are known. The straightforward approach is to simply solve the drift-diffusion equation listed in section [2.10](#) simultaneously. This approach however does not yield an analytic solution.

To avoid this problem we will assume that the electron and hole quasi-Fermi energies in the depletion region equal those in the adjacent n-type and p-type quasi-neutral regions. We will derive an expression for "long" and "short" diodes as well as a general expression, which is to be used if the quasi-neutral region is comparable in size to the diffusion length.

4.4.2.2. Assumptions and boundary conditions

The electric field and potential are obtained by using the full depletion approximation. Assuming that the quasi-Fermi energies are constant throughout the depletion region, one obtains the minority carrier densities at the edges of the depletion region, yielding:

$$p_n(x = x_n) = p_{n0} e^{V_a/V_t} \quad (4.4.1)$$

and

$$n_p(x = -x_p) = n_{p0} e^{V_a/V_t} \quad (4.4.2)$$

These equations can be verified to yield the thermal-equilibrium carrier density for zero applied voltage. In addition, an increase of the applied voltage will increase the separation between the two quasi-Fermi energies by the applied voltage multiplied with the electronic charge and the carrier density depends exponentially on this quantity.

The carrier density at the metal contacts is assumed to equal the thermal-equilibrium carrier density. This assumption implies that excess carriers immediately recombine when reaching either of the two metal-semiconductor contacts. As recombination is typically higher at a semiconductor surface and is further enhanced by the presence of the metal, this is found to be a reasonable assumption. This results in the following set of boundary conditions:

$$p_n(x = w_n) = p_{n0} \quad (4.4.3)$$

and

$$n_p(x = -w_p) = n_{p0} \quad (4.4.4)$$

4.4.2.3. General current expression

The general expression for the ideal diode current is obtained by applying the boundary conditions to the general solution of the diffusion equation for each of the quasi-neutral regions, as described by equation (2.9.13) and (2.9.14):

$$p_n(x \geq x_n) = p_{n0} + A e^{-(x-x_n)/L_p} + B e^{(x-x_n)/L_p} \quad (2.9.13)$$

$$n_p(x \leq -x_p) = n_{p0} + C e^{-(x+x_p)/L_p} + D e^{(x+x_p)/L_p} \quad (2.9.14)$$

The boundary conditions at the edge of the depletion regions are described by (4.4.1), (4.4.2), (4.4.3) and (4.4.4).

Before applying the boundary conditions, it is convenient to rewrite the general solution in terms of hyperbolic functions:

$$p_n(x \geq x_n) = p_{n0} + A^* \cosh \frac{x - x_n}{L_p} + B^* \sinh \frac{x - x_n}{L_p} \quad (4.4.5)$$

$$n_p(x \leq -x_p) = n_{p0} + C^* \cosh \frac{x + x_p}{L_n} + D^* \sinh \frac{x + x_p}{L_n} \quad (4.4.6)$$

where A^* , B^* , C^* and D^* are constants whose value remains to be determined. Applying the boundary conditions then yields:

$$p_n(x \geq x_n) = p_{n0} + p_{n0} (e^{V_a/V_t} - 1) \left[\cosh \frac{x - x_n}{L_p} - \coth \frac{w_n}{L_p} \sinh \frac{x - x_n}{L_p} \right] \quad (4.4.7)$$

$$n_p(x \leq -x_p) = n_{p0} + n_{p0} (e^{V_a/V_t} - 1) \left[\cosh \frac{x + x_p}{L_n} + \coth \frac{w_p}{L_n} \sinh \frac{x + x_p}{L_n} \right] \quad (4.4.8)$$

Where the quasi-neutral region widths, w_n' and w_p' , are defined as:

$$w_n' = w_n - x_n \quad (4.4.9)$$

and

$$w_p' = w_p - x_p \quad (4.4.10)$$

The current density in each region is obtained by calculating the diffusion current density using equations (2.7.22) and (2.7.23):

$$\begin{aligned} J_p(x \geq x_n) &= -qD_p \frac{dp}{dx} \\ &= -\frac{qD_p p_{n0}}{L_p} (e^{V_a/V_t} - 1) \left[\sinh \frac{x - x_n}{L_p} - \coth \frac{w_n'}{L_p} \cosh \frac{x - x_n}{L_p} \right] \end{aligned} \quad (4.4.11)$$

$$\begin{aligned} J_n(x \leq -x_p) &= qD_n \frac{dn}{dx} \\ &= \frac{qD_n n_{p0}}{L_n} (e^{V_a/V_t} - 1) \left[\sinh \frac{x + x_p}{L_n} + \coth \frac{w_p'}{L_n} \cosh \frac{x + x_p}{L_n} \right] \end{aligned} \quad (4.4.12)$$

The total current must be constant throughout the structure since a steady state case is assumed. No charge can accumulate or disappear somewhere in the structure so that the charge flow must be constant throughout the diode. The total current then equals the sum of the maximum electron current in the p-type region, the maximum hole current in the n-type regions and the current due to recombination within the depletion region. The maximum currents in the quasi-neutral regions occur at either side of the depletion region and can therefore be calculated from equations (4.4.11) and (4.4.12). Since we do not know the current due to recombination in the depletion region we will simply assume that it can be ignored. Later we will more closely examine this assumption. The total current is then given by:

$$I = A[J_n(x = -x_p) + J_p(x = x_n) + J_r] \cong I_s (e^{V_a/V_t} - 1) \quad (4.4.13)$$

where I_s can be written in the following form:

$$I_s = qA \left[\frac{D_n n_{p0}}{L_n} \coth \left(\frac{w_p'}{L_n} \right) + \frac{D_p p_{n0}}{L_p} \coth \left(\frac{w_n'}{L_p} \right) \right] \quad (4.4.14)$$

4.4.2.4. The p-n diode with a "long" quasi-neutral region

A diode with a "long" quasi-neutral region has a quasi-neutral region which is much larger than the minority-carrier diffusion length in that region, or $w_n' > L_p$ and $w_p' > L_n$. The general solution can be simplified under those conditions using:

$$\coth x = \frac{1}{\tanh x} \cong \frac{1}{x}, \text{ for } x \ll 1 \quad (4.4.15)$$

Yielding the following carrier densities, current densities and saturation currents:

$$p_n(x \geq x_n) = p_{n0} + p_{n0} (e^{V_a/V_t} - 1) \exp \frac{-(x - x_n)}{L_p} \quad (4.4.16)$$

$$n_p(x \leq -x_p) = n_{p0} + n_{p0}(e^{V_a/V_t} - 1) \exp \frac{x + x_p}{L_n} \quad (4.4.17)$$

$$J_p(x \geq x_n) = \frac{qD_p p_{n0}}{L_p} (e^{V_a/V_t} - 1) \exp \frac{-(x - x_n)}{L_p} \quad (4.4.18)$$

$$J_n(x \leq -x_p) = \frac{qD_n n_{p0}}{L_n} (e^{V_a/V_t} - 1) \exp \frac{x + x_p}{L_n} \quad (4.4.19)$$

$$I_s = qA \left[\frac{D_n n_{p0}}{L_n} + \frac{D_p p_{n0}}{L_p} \right] = qA \left[\frac{n_{p0} L_n}{\tau_n} + \frac{p_{n0} L_p}{\tau_p} \right] \quad (4.4.20)$$

We now come back to our assumption that the current due to recombination in the depletion region can be simply ignored. Given that there is recombination in the quasi-neutral region, it would be unreasonable to suggest that the recombination rate would simply drop to zero in the depletion region. Instead, we assume that the recombination rate is constant in the depletion region. To further simplify the analysis we will consider a p⁺-n junction so that we only need to consider the recombination in the n-type region. The current due to recombination in the depletion region is then given by:

$$I_r \geq qA \frac{p_{n0} x_n}{\tau_p} (e^{V_a/V_t} - 1) \quad (4.4.21)$$

so that I_r can be ignored if:

$$I_r \ll I, \text{ for } x_n \ll L_p \quad (4.4.22)$$

A necessary, but not sufficient requirement is therefore that the depletion region width is much smaller than the diffusion length for the ideal diode assumption to be valid. Silicon and germanium p-n diodes usually satisfy this requirement, while gallium arsenide p-n diodes rarely do because of the short carrier lifetime and diffusion length.

As an example we now consider a silicon p-n diode with $N_a = 1.5 \times 10^{14} \text{ cm}^{-3}$ and $N_d = 10^{14} \text{ cm}^{-3}$. The minority carrier lifetime was chosen to be very short, namely 400 ps, so that most features of interest can be easily observed. We start by examining the electron and hole density throughout the p-n diode, shown in Figure 4.4.1:

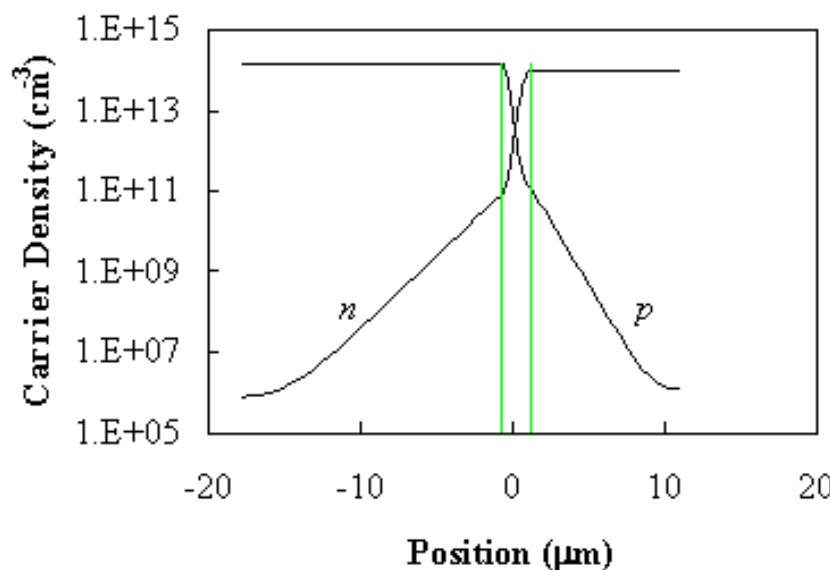


Figure 4.4.1 : Electron and hole density throughout a forward biased p-n diode.

The majority carrier densities in the quasi-neutral region simply equal the doping density. The minority carrier densities in the quasi-neutral regions are obtained from the equations (4.4.16) and (4.4.17). The electron and hole densities in the depletion region are calculating using the assumption that the electron/hole quasi-Fermi energy in the depletion region equals the electron/hole quasi-Fermi energy in the quasi-neutral n-type/p-type region. The corresponding band diagram is shown in Figure 4.4.2:

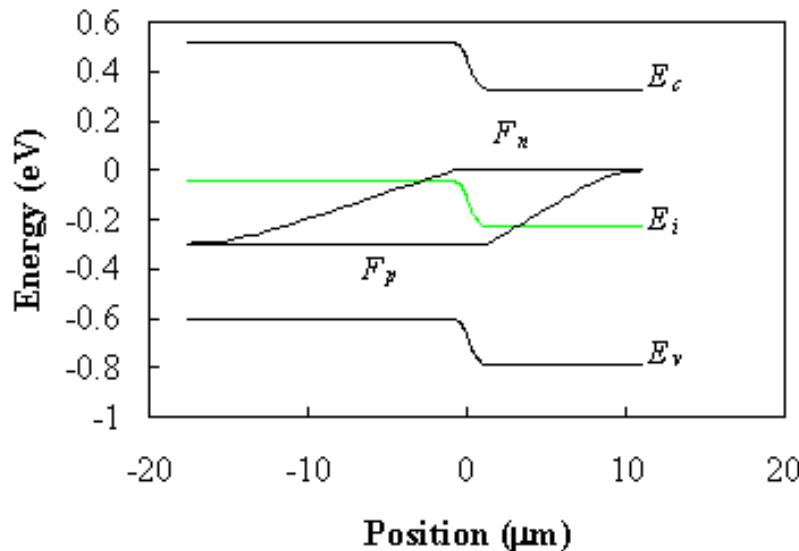


Figure 4.4.2 : Energy band diagram of a p-n diode. Shown are the conduction band edge, E_c , and the valence band edge, E_v , the intrinsic energy, E_i , the electron quasi-Fermi energy, F_n , and the hole quasi-Fermi energy, F_p .

The quasi-Fermi energies were obtained by combining (4.4.16) and (4.4.17) with (2.6.37) and (2.6.38). Note that the quasi-Fermi energies vary linearly within the quasi-neutral regions.

Next, we discuss the current density. Shown in Figure 4.4.3 is the electron and hole current density as calculated using (4.4.18) and (4.4.19). The current due to recombination in the depletion region was assumed to be constant.

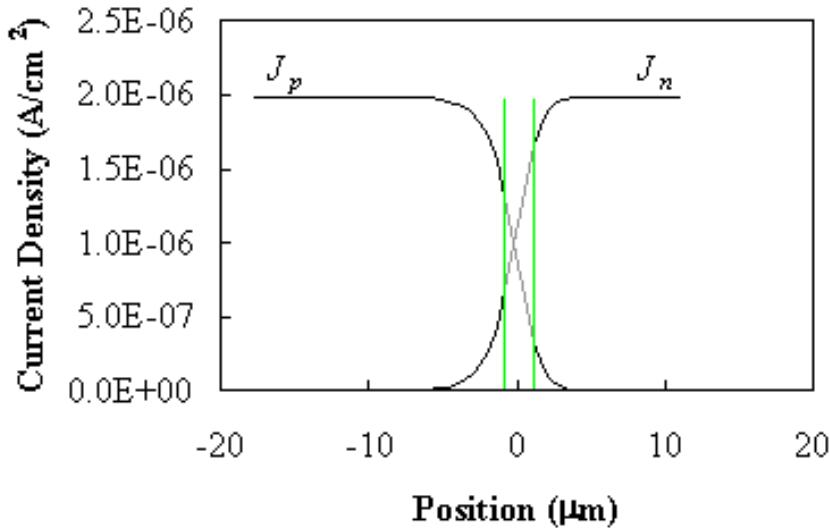


Figure 4.4.3 : Electron and hole current density versus position. The vertical lines indicate the edges of the depletion region.

4.4.2.5. The p-n diode with a "short" quasi-neutral region

A "short" diode is a diode with quasi-neutral regions, which are much shorter than the minority-carrier diffusion lengths. As the quasi-neutral region is much smaller than the diffusion length one finds that recombination in the quasi-neutral region is negligible so that the diffusion equations are reduced to:

$$0 = D_n \frac{d^2 n_p}{dx^2}, \text{ and } 0 = D_p \frac{d^2 p_n}{dx^2} \quad (4.4.23)$$

The resulting carrier density varies linearly throughout the quasi-neutral region and in general is given by:

$$n_p = A + Bx, \text{ and } p_n = A + Bx \quad (4.4.24)$$

where A, B, C and D are constants obtained by satisfying the boundary conditions. Applying the same boundary conditions at the edge of the depletion region as above (equations (4.4.3) and (4.4.4)) and requiring thermal equilibrium at the contacts yields:

$$p_n = p_{n0} + p_{n0} (e^{V_a/V_t} - 1) \left(1 - \frac{x - x_n}{w_n} \right) \quad (4.4.25)$$

$$n_p = n_{p0} + n_{p0} (e^{V_a/V_t} - 1) \left(1 + \frac{x + x_p}{w_p} \right) \quad (4.4.26)$$

for the hole and electron density in the n-type quasi-neutral region.

The current in a "short" diode is again obtained by adding the maximum diffusion currents in each of the quasi-neutral regions and ignoring the current due to recombination in the depletion region, yielding:

$$I = A[J_n(x = -x_p) + J_p(x = x_n) + J_r] \cong I_s (e^{V_a/V_t} - 1) \quad (4.4.27)$$

where the saturation current, I_s is given by:

$$I_s = qA \left[\frac{D_n n_p 0}{w_p} + \frac{D_p p_n 0}{w_n} \right] \quad (4.4.28)$$

A comparison of the "short" diode expression with the "long" diode expression reveals that they are the same except for the use of either the diffusion length or the quasi-neutral region width in the denominator, whichever is smaller.

Given that we now have two approximate expressions, it is of interest to know when to use one or the other. To this end, we now consider a one-sided n⁺-p diode. The p-type semiconductor has a width w_p and we normalize the excess electron density relative to its value at $x = 0$. The Ohmic contact to the p-type region is ideal so that the excess density is zero at $x = w_p'$. The normalized excess carrier density is shown in Figure 4.4.4 for different values of the diffusion length.

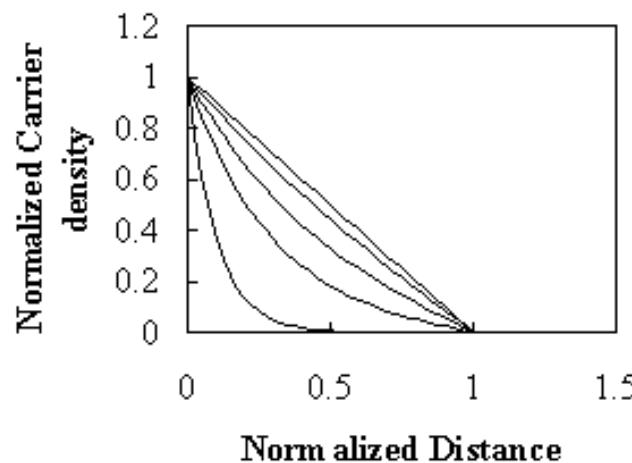


Figure 4.4.4 : Excess electron density versus position as obtained by solving the diffusion equation with $\delta n(x = 0) = 1$ and $\delta n(x/w_p' = 1) = 0$. The ratio of the diffusion length to the width of the quasi-neutral region is varied from 0.1 (Bottom curve), 0.3, 0.5, 1 and 10 (top curve)

The figure illustrates how the excess electron density changes as the diffusion length is varied relative to the width of the quasi-neutral region. For the case where the diffusion length is much smaller than the width ($L_n \ll w_p'$), the electron density decays exponentially and the "long" diode expression can be used. If the diffusion length is much longer than the width ($L_n \gg w_p'$), the electron density reduces linearly with position and the "short" diode expression can be used. If the diffusion length is comparable to the width of the quasi-neutral region width one must use the general expression. A numeric analysis reveals that the error is less than 10 % when using the short diode expression with $L_n > w_p'/2$ and when using the long diode expression with $L_n > 2 w_p'$.

Example 4.4	<p>An abrupt silicon p-n junction ($N_a = 10^{16} \text{ cm}^{-3}$ and $N_d = 4 \times 10^{16} \text{ cm}^{-3}$) is biased with $V_a = 0.6 \text{ V}$. Calculate the ideal diode current assuming that the n-type region is much smaller than the diffusion length with $w_n' = 1 \mu\text{m}$ and assuming a "long" p-type region. Use $\mu_n = 1000 \text{ cm}^2/\text{V-s}$ and $\mu_p = 300 \text{ cm}^2/\text{V-s}$. The minority carrier lifetime is 10 μs and the diode area is 100 μm by 100 μm.</p>
-------------	--

Solution

The current is calculated from:

$$I = qA \left[\frac{D_n n_{p0}}{L_n} + \frac{D_p p_{n0}}{w_n} \right] (e^{V_a/V_t} - 1)$$

with

- $D_n = \mu_n V_t = 1000 \times 0.0258 = 25.8 \text{ cm}^2/\text{V-s}$
- $D_p = \mu_p V_t = 300 \times 0.0258 = 7.75 \text{ cm}^2/\text{V-s}$
- $n_{p0} = n_i^2/N_a = 10^{20}/10^{16} = 10^4 \text{ cm}^{-3}$
- $p_{n0} = n_i^2/N_d = 10^{20}/4 \times 10^{16} = 2.5 \times 10^3 \text{ cm}^{-3}$

$$L_n = \sqrt{D_n \tau_n} = \sqrt{25.8 \times 10^{-5}} = 161 \mu\text{m}$$

yielding $I = 40.7 \text{ mA}$

Note that the hole diffusion current occurs in the "short" n-type region and therefore depends on the quasi-neutral width in that region. The electron diffusion current occurs in the "long" p-type region and therefore depends on the electron diffusion length in that region.

4.4.3. Recombination-Generation current



4.4.3.1. Band-to-band Recombination-Generation current

The recombination/generation current due to band-to-band recombination/generation is obtained by integrating the net recombination rate, U_{b-b} , over the depletion region:

$$J_{b-b} = q \int_{-x_p}^{x_n} U_{b-b} dx \quad (4.4.29)$$

where the net recombination rate is given by:

$$U_{b-b} = b(np - n_i^2) \quad (4.4.30)$$

The carrier densities can be related to the constant quasi-Fermi energies and the product is independent of position:

$$np = n_i e^{(E_{Fn} - E_i)/kT} n_i e^{(E_i - E_{Fp})/kT} = n_i^2 e^{V_a/V_t} \quad (4.4.31)$$

This allows the integral to be solved analytically yielding:

$$J_{b-b} = q \int_{-x_p}^{x_n} n_i^2 (e^{V_a/V_t} - 1) dx = q n_i^2 b w (e^{V_a/V_t} - 1) \quad (4.4.32)$$

The current due to band-to-band recombination has therefore the same voltage dependence as the ideal diode current and simply adds an additional term to the expression for the saturation current.

4.4.3.2. Shockley-Hall-Read Recombination-Generation current

The current due to trap-assisted recombination in the depletion region is also obtained by integrating the trap-assisted recombination rate over the depletion region width:

$$J_{SHR} = q \int_{-x_p}^{x_n} U_{SHR} dx \quad (4.4.33)$$

Substituting the expression for the recombination rate yields:

$$J_{SHR} = q \int_{-x_p}^{x_n} \frac{1}{\tau} \frac{n_i^2 (e^{V_a/V_t} - 1)}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} dx \quad (4.4.34)$$

where the product of the electron and hole densities was obtained by assuming that the quasi-Fermi energies are constant throughout the depletion region, which leads to:

$$n_p = n_i e^{(E_n - E_i)/kT} n_i e^{(E_i - E_p)/kT} = n_i^2 e^{V_a/V_t} \quad (4.4.35)$$

The maximum recombination rate is obtained when the electron and hole densities are equal and therefore equals the square root of the product yielding:

$$U_{SHR, \max} = \text{MAX} \left(\frac{1}{\tau} \frac{n_i^2 (e^{V_a/V_t} - 1)}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} \right) \cong \frac{n_i}{2\tau} (e^{V_a/2V_t} - 1) \quad (4.4.36)$$

From which an effective width can be defined which, when multiplied with the maximum recombination rate, equals the integral of the recombination rate over the depletion region. This effective width, x' , is then defined by:

$$x' = \frac{\int_{-x_p}^{x_n} \frac{1}{\tau} \frac{n_i^2 (e^{V_a/V_t} - 1)}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} dx}{U_{SHR, \max}} \quad (4.4.37)$$

and the associated current due to trap-assisted recombination in the depletion region is given by:

$$J_{SHR} = \frac{q n_i x'}{2\tau} (e^{V_a/2V_t} - 1) \quad (4.4.38)$$

This does not provide an actual solution since the effective width, x' , still must be determined by performing a numeric integration. Nevertheless, the above expression provides a way to obtain an upper estimate by substituting the depletion layer width, x_d , as it is always larger than the effective width.



4.4.4. I-V characteristics of real p-n diodes

The forward biased I - V characteristics of real p-n diodes are further affected by high injection and the series resistance of the diode. To illustrate these effects while summarizing the current mechanisms discussed previously we consider the I - V characteristics of a silicon p^+ - n diode with $N_d = 4 \times 10^{14} \text{ cm}^{-3}$, $\tau_p = 10 \text{ ms}$, and $\mu_p = 450 \text{ cm}^2/\text{V}\cdot\text{s}$. The I-V characteristics are plotted on a semi-logarithmic scale and four different regions can be distinguished as indicated on Figure 4.4.5. First, there is the ideal diode region where the current increases by one order of magnitude as the voltage is increased by 60 mV. This region is referred to as having an ideality factor, n , of one. This ideality factor is obtained by fitting a section of the curve to the following expression for the current:

$$J = J_s e^{V_a / nV_t} \quad (4.4.39)$$

The ideality factor can also be obtained from the slope of the curve on a semi-logarithmic scale using:

$$\gamma = \frac{\log(e)}{V_t \text{ slope}} = \frac{1}{\text{slope}} \quad (4.4.40)$$

where the slope is in units of V/decade. To the left of the ideal diode region there is the region where the current is dominated by the trap-assisted recombination in the depletion region described in section 4.4.3.2. This part of the curve has an ideality factor of two. To the right of the ideal diode region, the current becomes limited by high injection effects and by the series resistance.

High injection occurs in a forward biased p-n diode when the injected minority carrier density exceeds the doping density. High injection will therefore occur first in the lowest doped region of the diode since that region has the highest minority carrier density.

Using equations (4.4.1) and (4.4.2), one finds that high injection occurs in a p^+ - n diode for the following applied voltage:

$$V_a = 2V_t \ln \frac{N_d}{n_i} \quad (4.4.41)$$

or at $V_a = 0.55 \text{ V}$ for the diode of Figure 4.4.5 as can be verified on the figure as the voltage where the ideality factor changes from one to two. For higher forward bias voltages, the current does no longer increase exponentially with voltage. Instead, it increases linearly due to the series resistance of the diode. This series resistance can be due to the contact resistance between the metal and the semiconductor, due to the resistivity of the semiconductor or due to the series resistance of the connecting wires. This series resistance increases the external voltage, V_a^* , relative to the internal voltage, V_a , considered so far.

$$V_a^* = V_a + IR_s \quad (4.4.42)$$

Where I is the diode current and R_s is the value of the series resistance.

These four regions can be observed in most p-n diodes although the high-injection region rarely occurs, as the series resistance tends to limit the current first.

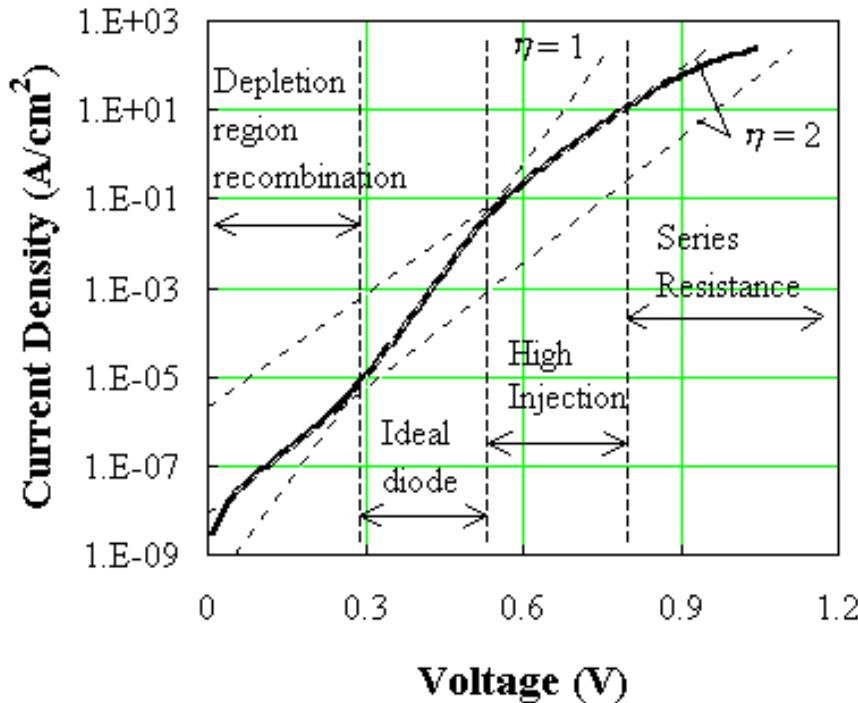


Figure 4.4.5: Current-Voltage characteristics of a silicon diode under forward bias.

4.4.5. The diffusion capacitance



As a p-n diode is forward biased, the minority carrier distribution in the quasi-neutral region increases dramatically. In addition, to preserve quasi-neutrality, the majority carrier density increases by the same amount. This effect leads to an additional capacitance called the diffusion capacitance.

The diffusion capacitance is calculated from the change in charge with voltage:

$$C = \frac{d\Delta Q}{dV_a} \quad (4.4.43)$$

Where the charge, ΔQ , is due to the excess carriers. Unlike a parallel plate capacitor, the positive and negative charge is no longer separated in space. Instead, the electrons and holes are separated by the energy bandgap. Nevertheless, these voltage dependent charges yield a capacitance just as the one associated with a parallel plate capacitor. The excess minority-carrier charge is obtained by integrating the charge density over the quasi-neutral region:

$$\Delta Q_p = \int_{x_n}^{w_n} qA(p_n - p_{n0}) dx \quad (4.4.44)$$

We now distinguish between the two limiting cases as discussed when calculating the ideal diode current, namely the "long" diode and a "short" diode. The carrier distribution, $p_n(x)$, in a "long" diode is illustrated by Figure 4.4.6 (a).

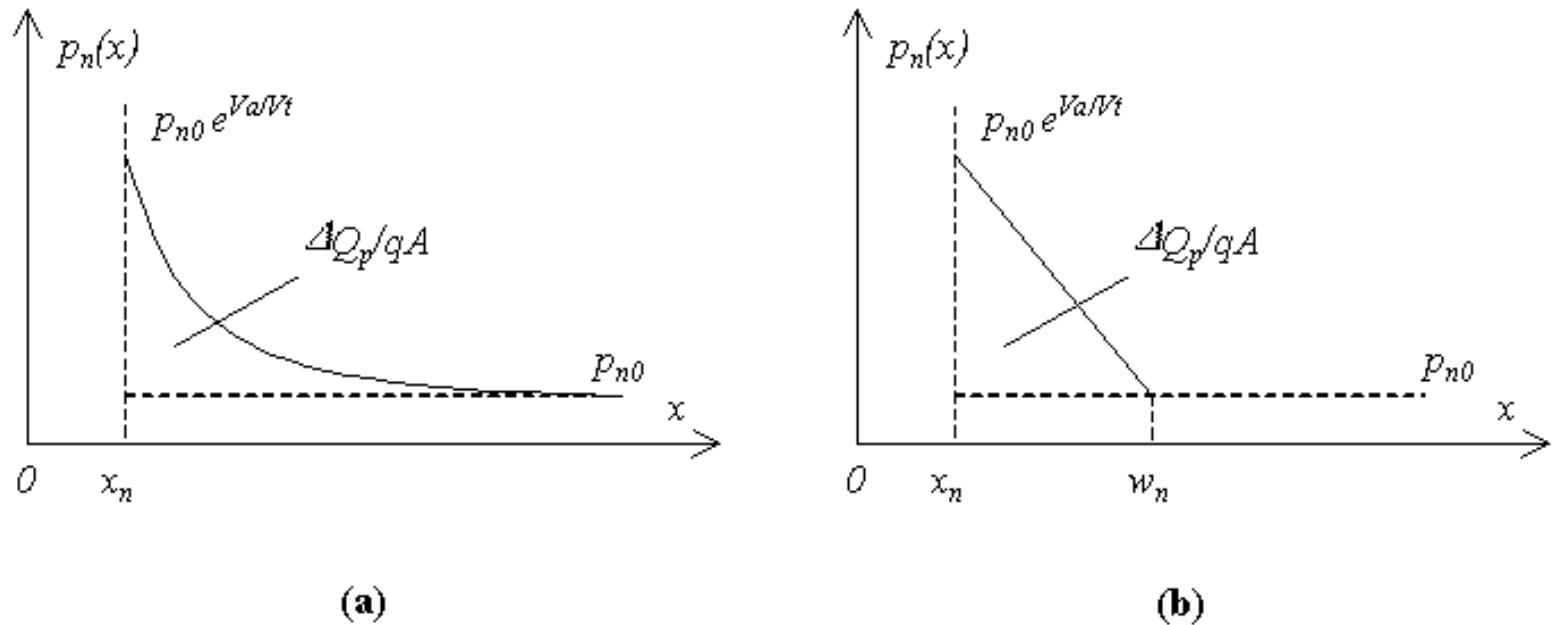


Figure 4.4.6: Minority carrier distribution in (a) a "long" diode, and (b) a "short" diode. The excess minority-carrier charge, ΔQ_p , in the quasi-neutral region, is proportional to the area defined by the solid and dotted lines.

Using equation (4.4.18), the excess charge, ΔQ_p , becomes:

$$\Delta Q_p = q A p_{n0} (e^{V_a/V_t} - 1) L_p = I_{s,p} (e^{V_a/V_t} - 1) \tau_p \quad (4.4.45)$$

where $I_{s,p}$ is the saturation current for holes, given by:

$$I_{s,p} = q \frac{A p_{n0} D_p}{L_p} \quad (4.4.46)$$

Equation (4.4.45) directly links the excess charge to the diffusion current. Since all injected minority carriers recombine in the quasi-neutral region, the current equals the excess charge divided by the average time needed to recombine with the majority carriers, i.e. the carrier lifetime, τ_p . This relation is the corner stone of the charge control model of p-n diodes and bipolar junction transistors.

The diffusion capacitance then equals:

$$C_{d,p} = \frac{d(I_{s,p} (e^{V_a/V_t} - 1) \tau_p)}{dV_a} = \frac{I_{s,p} e^{V_a/V_t} \tau_p}{V_t} \quad (4.4.47)$$

Similarly, for a "short" diode, as illustrated by Figure 4.4.6 (b), one obtains:

$$C_{d,p} = \frac{I_{s,p} e^{V_a/V_t} t_{r,p}}{V_t} \quad (4.4.48)$$

Where $t_{r,p}$ is the hole transit time given by:

$$t_{r,p} = \frac{w_n^2}{2 D_p} \quad (4.4.49)$$

Again, the excess charge can be related to the current. However, in the case of a "short" diode all minority carriers flow through the quasi-neutral region and do not recombine with the majority carriers. The current therefore equals the excess charge divided by the average time needed to traverse the quasi-neutral region, i.e. the transit time, $t_{r,p}$.

The total diffusion capacitance is obtained by adding the diffusion capacitance of the n-type quasi-neutral region to that of the p-type quasi-neutral region.

The total capacitance of the junction equals the sum of the junction capacitance, discussed in section [4.3.4](#), and the diffusion capacitance. For reverse biased voltages and small forward bias voltages, one finds that the junction capacitance is dominant. As the forward bias voltage is further increased the diffusion capacitance increases exponentially and eventually becomes larger than the junction capacitance.

Example 4.5	<ol style="list-style-type: none"> Calculate the diffusion capacitance of the diode described in Example 4.4 at zero bias. Use $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$, $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$, $w_p' = 1 \mu\text{m}$ and $w_n' = 1 \text{ mm}$. The minority carrier lifetime equals 0.1 ms. For the same diode, find the voltage for which the junction capacitance equals the diffusion capacitance.
-------------	--

Solution	<ol style="list-style-type: none"> The diffusion capacitance at zero volts equals $C_{d,0} = \frac{I_{s,p} \tau_p}{V_t} + \frac{I_{s,n} \tau_r n}{V_t} = 1.73 \times 10^{-19} \text{ F}$ using $I_{s,p} = q \frac{A p n_0 D_p}{L_p}$ and $I_{s,n} = q \frac{A n p_0 D_n}{w_p}$ where the "short" diode expression was used for the capacitance associated with the excess charge due to electrons in the p-type region. The "long" diode expression was used for the capacitance associated with the excess charge due to holes in the n-type region. The diffusion constants and diffusion lengths equal $D_n = \mu_n \times V_t = 25.8 \text{ cm}^2/\text{s}$ $D_p = \mu_p \times V_t = 7.75 \text{ cm}^2/\text{s}$ $L_p = \sqrt{D_p \tau_p}$ And the electron transit time in the p-type region equals
----------	---

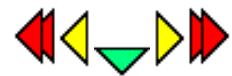
$$t_{r,n} = \frac{w_p^2}{2D_n} = 193 \text{ ps}$$

- b. The voltage at which the junction capacitance equals the diffusion capacitance is obtained by solving

$$\frac{C_{j0}}{\sqrt{1 - \frac{V_a}{V_t}}} = C_{d,0} e^{V_a/V_t}$$

yielding $V_a = 0.442 \text{ V}$

Chapter 4: p-n Junctions



4.5. Reverse bias breakdown

[4.5.1. General breakdown characteristics](#)

[4.5.2. Edge effects](#)

[4.5.3. Avalanche breakdown](#)

[4.5.4. Zener breakdown](#)

4.5.1. General breakdown characteristics



The maximum reverse bias voltage that can be applied to a p-n diode is limited by breakdown. Breakdown is characterized by the rapid increase of the current under reverse bias. The corresponding applied voltage is referred to as the breakdown voltage.

The breakdown voltage is a key parameter of power devices. The breakdown of logic devices is equally important as one typically reduces the device dimensions without reducing the applied voltages, thereby increasing the internal electric field.

Two mechanisms can cause breakdown, namely avalanche multiplication and quantum mechanical tunneling of carriers through the bandgap. Neither of the two breakdown mechanisms is destructive. However heating caused by the large breakdown current and high breakdown voltage causes the diode to be destroyed unless sufficient heat sinking is provided.

Breakdown in silicon at room temperature can be predicted using the following empirical expression for the electric field at breakdown.

$$|\mathcal{E}_{br}| = \frac{4 \times 10^5}{1 - \frac{1}{3} \log(N/10^{16})} \text{ V/cm} \quad (4.5.1)$$

Assuming a one-sided abrupt p-n diode, the corresponding breakdown voltage can then be calculated, yielding:

$$|V_{br}| = -A + \frac{|\mathcal{E}_{br}|^2 \epsilon_s}{2qN} \quad (4.5.2)$$

The resulting breakdown voltage is inversely proportional to the square of the doping density if one ignores the weak doping dependence of the electric field at breakdown. The corresponding depletion layer width equals:

$$w_{br} = \frac{|\mathcal{E}_{br}| \epsilon_s}{qN} \quad (4.5.3)$$

4.5.2. Edge effects



Few p-n diodes are truly planar and typically have higher electric fields at the edges. Since the diodes will break down in the regions where the breakdown field is reached first, one has to take into account the radius of curvature of the metallurgical junction at the edges. Most doping processes including diffusion and ion implantation yield a radius of curvature on the order of the junction depth, x_j . The p-n diode interface can then be approximated as having a cylindrical shape along a straight edge and a spherical at a corner of a rectangular pattern. Both structures can be solved analytically as a function of the doping density, N , and the radius of curvature, x_j .

The resulting breakdown voltages and depletion layer widths are plotted below as a function of the doping density of an abrupt one-sided junction.

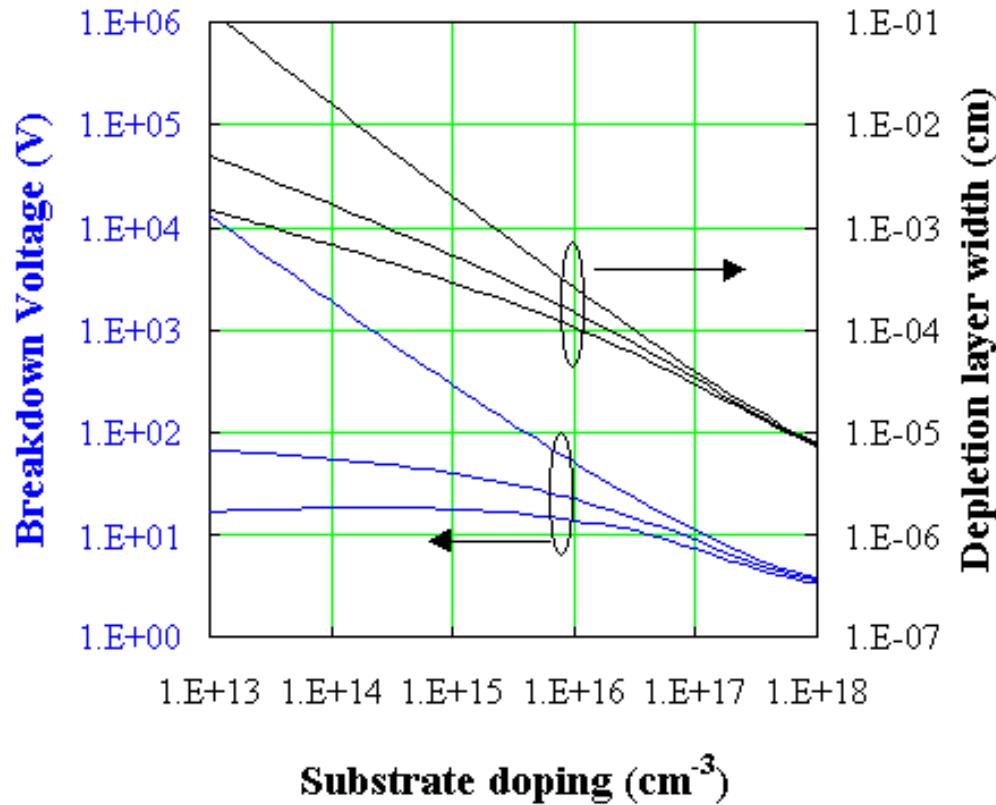


Figure 4.5.1 : Breakdown voltage and depletion layer width at breakdown versus doping density of an abrupt one-sided p-n diode. Shown are the voltage and width for a planar (top curves), cylindrical (middle curves) and spherical (bottom curves) junction.

4.5.3. Avalanche breakdown



Avalanche breakdown is caused by impact ionization of electron-hole pairs. This process was described previously in section 2.8. When applying a high electric field, carriers gain kinetic energy and generate additional electron-hole pairs through impact ionization. The ionization rate is quantified by the ionization constants of electrons and holes, α_n and α_p . These ionization constants are defined as the change of the carrier density with position divided by the carrier density or:

$$dn = \alpha_n n dx \quad (4.5.4)$$

The ionization causes a generation of additional electrons and holes. Assuming that the ionization coefficients of electrons and holes are the same, the multiplication factor M , can be calculated from:

$$M = \frac{1}{1 - \int_{x_1}^{x_2} \alpha dx} \quad (4.5.5)$$

The integral is taken between x_1 and x_2 , the region within the depletion layer where the electric field is assumed constant and large enough to cause impact ionization. Outside this range, the electric field is assumed to be too low to cause impact ionization. The equation for the multiplication factor reaches infinity if the integral equals one. This condition can be interpreted as follows: For each electron coming to the high field at point x_1 one additional electron-hole pair is generated arriving at point x_2 . This hole drifts in the opposite direction and generates an additional electron-hole pair at the starting point x_1 . One initial electron therefore yields an infinite number of electrons arriving at x_2 , hence an infinite multiplication factor.

The multiplication factor is commonly expressed as a function of the applied voltage and the breakdown voltage using the following empirical relation:

$$M = \frac{1}{1 - \left| \frac{V_a}{V_{br}} \right|^n}, \text{ where } 2 < n < 6 \quad (4.5.6)$$

4.5.4. Zener breakdown



Quantum mechanical tunneling of carriers through the bandgap is the dominant breakdown mechanism for highly doped p-n junctions. The analysis is identical to that of [tunneling in a metal-semiconductor junction](#) where the barrier height is replaced by the energy bandgap of the material.

The tunneling probability equals:

$$\Theta = \exp \left(-\frac{4}{3} \frac{\sqrt{2m^*}}{q\hbar} \frac{E_g^{3/2}}{\epsilon} \right) \quad (4.5.7)$$

where the electric field equals $\epsilon = E_g/(qL)$.

The tunneling current is obtained from the product of the carrier charge, velocity and carrier density. The velocity equals the Richardson velocity, the velocity with which on average the carriers approach the barrier while the carrier density equals the density of available electrons multiplied with the tunneling probability, yielding:

$$J_n = q \nu_R n \Theta \quad (4.5.8)$$

The tunneling current therefore depends exponentially on the bandgap energy to the 3/2 power.

Chapter 4: p-n Junctions



4.6. Optoelectronic devices

4.6.1. Photodiodes

4.6.2. Solar cells

4.6.3. LEDs

4.6.4. Laser diodes

P-n junctions are an integral part of several optoelectronic devices. These include photodiodes, solar cells light emitting diodes (LEDs) and semiconductor lasers. In this section we discuss the principle of operation of these devices and derive an expression for key parameters.

4.6.1. Photodiodes



Photodiodes and crystalline solar cells are essentially the same as the p-n diodes, which have been described in this chapter. However, the diode is exposed to light, which yields a photocurrent in addition to the diode current so that the total diode current is given by:

$$I = I_s (e^{V_a/V_t} - 1) + I_{ph} \quad (4.6.1)$$

where the additional photocurrent, I_{ph} , is due to photogeneration of electrons and holes shown in Figure 4.6.1. These electrons and holes are pulled into the region where they are majority carriers by the electric field in the depletion region.

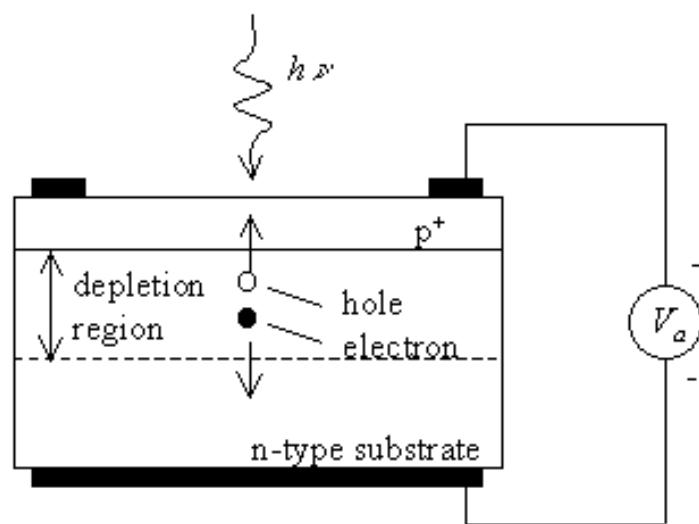


Figure 4.6.1:

Motion of photo-generated carriers in a p-n photodiode.

The photo-generated carriers cause a photocurrent, which opposes the diode current under forward bias. Therefore, the diode can be used as a photodetector - using a reverse or zero bias voltage - as the measured photocurrent is proportional to the incident light intensity. The diode can also be used as a solar cell - using a forward bias - to generate electrical power.

The primary characteristics of a photodiode are the responsivity, the dark current and the bandwidth. The responsivity is the photocurrent divided by the incident optical power. The maximum photocurrent in a photodiode equals

$$I_{ph, \max} = \frac{q}{h\nu} P_{in} \quad (4.6.2)$$

Where P_{in} is the incident optical power. This maximum photocurrent occurs when each incoming photon creates one electron-hole pair, which contributes to the photocurrent. The photocurrent in the presence of a reflection R at the surface of the photodiode and an absorption over a thickness d in a material with an absorption coefficient α is given by:

$$I_{ph} = (1 - R)(1 - e^{-\alpha d}) \frac{qP_{in}}{h\nu} \quad (4.6.3)$$

The photocurrent is further reduced if photo-generated electron-hole pairs recombine within the photodiode instead of being swept into the regions where they are majority carriers.

The dark current is the current through the diode in the absence of light. This current is due to the ideal diode current, the generation/recombination of carriers in the depletion region and any surface leakage, which occurs in the diode. The dark current obviously limits the minimum power detected by the photodiode, since a photocurrent much smaller than the dark current would be hard to measure.

However, the true limitation is the shot noise generated by the current through the diode. The shot noise as quantified by the average of the square of the noise current is given by:

$$\langle i^2 \rangle = 2qI\Delta f \quad (4.6.4)$$

Where I is the diode current and Δf is the bandwidth of the detector. The bandwidth of the diode is affected by the transit time of the photo-generated carriers through the diode and by the capacitance of the diode. The carrier transit time yields the intrinsic bandwidth of the diode while the capacitance together with the impedance of the amplifier or the transmission line connected to the diode yields a the parasitic RC delay.

4.6.2. Solar cells



Solar cells are typically illuminated with sunlight and are intended to convert the solar energy into electrical energy. The solar energy is in the form of electromagnetic radiation, more specifically "black-body" radiation as described in section [1.2.3](#). The sun's spectrum is consistent with that of a black body at a temperature of 5800 K. The radiation spectrum has a peak at 0.8 eV. A significant part of the spectrum is in the visible range of the spectrum (400 - 700 nm). The power density is approximately 100 mW/cm².

Only part of the solar spectrum actually makes it to the earth's surface. Scattering and absorption in the earth's atmosphere, and the incident angle affect the incident power density. Therefore, the available power density depends on the time of the day, the season and the latitude of a specific location.

Of the solar light, which does reach a solar cell, only photons with energy larger than the energy bandgap of the semiconductor generate electron-hole pairs. In addition, one finds that the voltage across the solar cell at the point where it delivers its maximum power is less than the bandgap energy in electron volt. The overall power-conversion efficiency of single-crystalline solar cells ranges from 10 to 30 % yielding 10 to 30 mW/cm².

The calculation of the maximum power of a solar cell is illustrated by Figure [4.6.2](#) and Figure [4.6.3](#). The sign convention of the current and voltage is shown as well. It considers a current coming out of the cell to be positive as it leads to electrical power generation. The power generated depends on the solar cell itself and the load connected to it. As an example, a resistive load is shown in the diagram below.

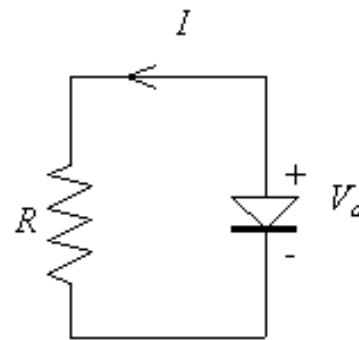


Figure 4.6.2 : Circuit diagram and sign convention of a p-n diode solar cell connected to a resistive load.

The current and the power as function of the forward bias voltage across the diode are shown in Figure 4.6.3 for a photocurrent of 1 mA:

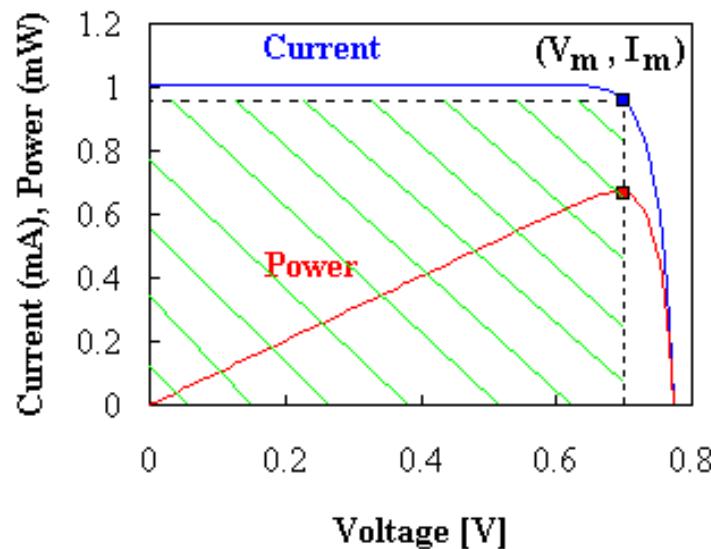


Figure 4.6.3 : Current-Voltage (I - V) and Power-Voltage (P - V) characteristics of a p-n diode solar cell with $I_{ph} = 1$ mA and $I_s = 10^{-10}$ A. The crosshatched area indicates the power generated by the solar cell. The markers indicate the voltage and current, V_m and I_m , for which the maximum power, P_m is generated. 

We identify the open-circuit voltage, V_{oc} , as the voltage across the illuminated cell at zero current. The short-circuit current, I_{sc} , is the current through the illuminated cell if the voltage across the cell is zero. The short-circuit current is close to the photocurrent while the open-circuit voltage is close to the turn-on voltage of the diode as measured on a current scale similar to that of the photocurrent.

The power equals the product of the diode voltage and current and at first increases linearly with the diode voltage but then rapidly goes to zero around the turn-on voltage of the diode. The maximum power is obtained at a voltage labeled as V_m with I_m being the current at that voltage.

The fill factor of the solar cell is defined as the ratio of the maximum power of the cell to the product of the open-circuit voltage, V_{oc} , and the short-circuit current, I_{sc} , or:

$$\text{Fill Factor} = \frac{I_m V_m}{I_{sc} V_{oc}} \quad (4.6.5)$$

Example 4.6	A 1 cm ² silicon solar cell has a saturation current of 10 ⁻¹² A and is illuminated with sunlight yielding a short-circuit photocurrent of 25 mA. Calculate the solar cell efficiency and fill factor.
Solution	<p>The maximum power is generated for:</p> $\frac{dP}{dV_a} = 0 = I_s (e^{V_m/V_t} - 1) - I_{ph} + \frac{V_m}{V_t} I_s e^{V_m/V_t}$ <p>where the voltage, V_m, is the voltage corresponding to the maximum power point. This voltage is obtained by solving the following transcendental equation:</p> $V_m = V_t \ln \frac{1 + I_{ph}/I_s}{1 + V_m/V_t}$ <p>Using iteration and a starting value of 0.5 V one obtains the following successive values for V_m:</p> $V_m = 0.5, 0.542, 0.540 \text{ V}$ <p>and the efficiency equals:</p> $\eta = \left \frac{V_m I_m}{P_{in}} \right = \frac{0.54 \times 0.024}{0.1} = 13\%$ <p>The current, I_m, corresponding to the voltage, V_m, was calculated using equation (4.6.1) and the power of the sun was assumed 100 mW/cm². The fill factor equals:</p> $\text{fill factor} = \frac{V_m I_m}{V_{oc} I_{sc}} = \frac{0.54 \times 0.024}{0.62 \times 0.025} = 83\%$ <p>where the open circuit voltage is calculated using equation (4.6.1) and $I = 0$. The short circuit current equals the photocurrent.</p>

4.6.3. LEDs



Light emitting diodes are p-n diodes in which the recombination of electrons and holes yields a photon. This radiative recombination process occurs primarily in direct bandgap semiconductors where the lowest conduction band minimum and the highest valence band maximum occur at $k = 0$, where k is the wavenumber. Examples of direct bandgap semiconductors are GaAs, InP, GaP, GaN while most group IV semiconductors including Si, Ge and SiC are indirect bandgap semiconductors.

The radiative recombination process is in competition with non-radiative recombination processes such as trap-assisted recombination. Radiative recombination dominates at high minority-carrier densities. Using a quantum well, a thin region with a lower bandgap, positioned at the metallurgical junction, one can obtain high carrier densities at low current densities. These quantum well LEDs have high internal quantum efficiency as almost every electron injected in the quantum well recombines with a hole yielding a photon.

The external quantum efficiency of planar LEDs is much lower than unity due to total internal reflection. As the photons are generated in the semiconductor, which has a high refractive index, only photons traveling normal to the semiconductor-air interface can exit the semiconductor. For GaAs with a refractive index of 3.5, the angle for total internal reflection equals 17° so that only a few percent of the generated photons can escape the semiconductor. This effect can be avoided by having a spherical semiconductor shape, which ensures that most photons are normal to the interface. The external quantum efficiency can thereby be increased to values larger than 50%.

4.6.4. Laser diodes



Laser diodes are very similar to LEDs since they also consist of a p-n diode with an active region where electrons and holes recombine resulting in light emission. However, a laser diode also contains an optical cavity where stimulated emission takes place. The laser cavity consists of a waveguide terminated on each end by a mirror. As an example, the structure of an edge-emitting laser diode is shown in Figure 4.6.4. Photons, which are emitted into the waveguide, can travel back and forth in this waveguide provided they are reflected at the mirrors.

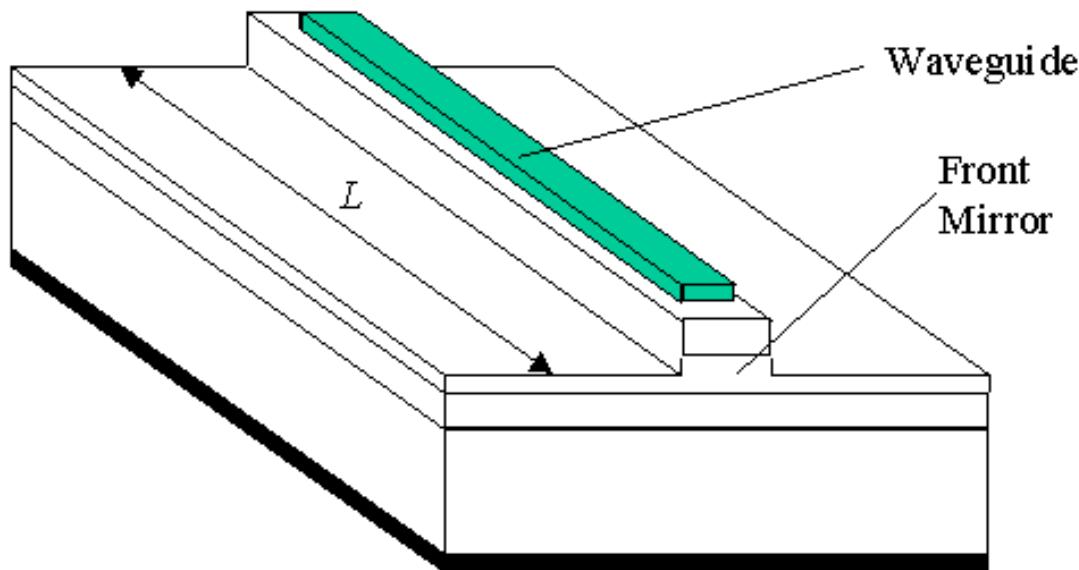


Figure 4.6.4 : Structure of an edge-emitting laser diode.

The light in the waveguide is amplified by stimulated emission. Stimulated emission is a process where a photon triggers the radiative recombination of an electron and hole thereby creating an additional photon with the same energy and phase as the incident photon. This process is illustrated with Figure 4.6.5. This "cloning" of photons results in a coherent beam.

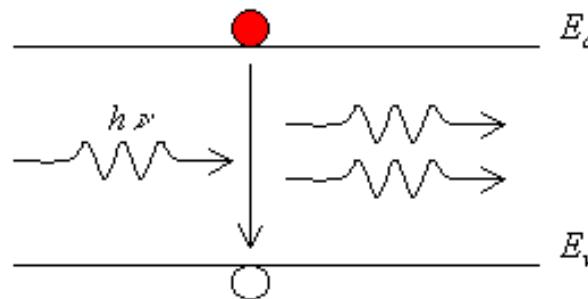


Figure 4.6.5 :

Stimulated emission of a photon.

The stimulated emission process yields an increase in photons as they travel along the waveguide. Combined with the waveguide losses, stimulated emission yields a net gain per unit length, g . The number of photons can therefore be maintained if the roundtrip amplification in a cavity of length, L , including the partial reflection at the mirrors with reflectivity R_1 and R_2 equals unity.

This yields the following lasing condition:

$$\text{Roundtrip amplification} = e^{2gL} R_1 R_2 = 1 \quad (4.6.6)$$

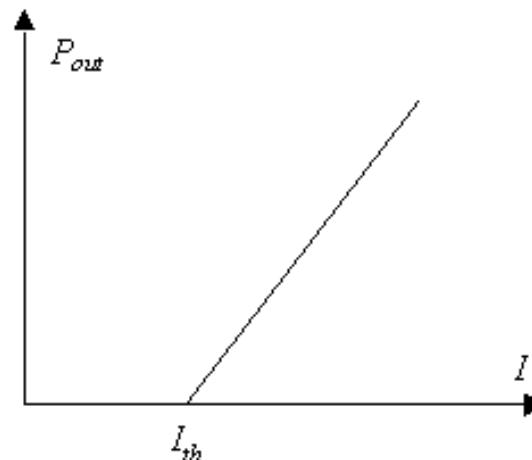
If the roundtrip amplification is less than one then the number of photons steadily decreases. If the roundtrip amplification is larger than one, the number of photons increases as the photons travel back and forth in the cavity and no steady state value would be obtained. The gain required for lasing therefore equals:

$$g = \frac{1}{2L} \ln \frac{1}{R_1 R_2} \quad (4.6.7)$$

Initially, the gain is negative if no current is applied to the laser diode as absorption dominates in the waveguide. As the laser current is increased, the absorption first decreases and the gain increases. The current for which the gain satisfies the lasing condition is the threshold current of the laser, I_{th} . Below the threshold current very little light is emitted by the laser structure. For an applied current larger than the threshold current, the output power, P_{out} , increases linearly with the applied current, as each additional incoming electron-hole pair is converted into an additional photon. The output power therefore equals:

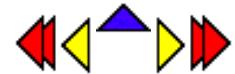
$$P_{out} = \eta \frac{h\nu}{q} (I - I_{th}) \quad (4.6.8)$$

where $h\nu$ is the energy per photon. The factor, η , indicates that only a fraction of the generated photons contribute to the output power of the laser as photons are partially lost through the other mirror and throughout the waveguide.

**Figure 4.6.6 :**

Output power from a laser diode versus the applied current.

Chapter 1: Review of Modern Physics



Equations

$$E_{ph} = h\nu = \frac{hc}{\lambda} \quad \bullet(1.2.1)\bullet$$

$$\lambda = \frac{h}{p} \quad \bullet(1.2.2)\bullet$$

$$E_{ph} = h\nu = \hbar\omega \quad \bullet(1.2.3)\bullet$$

$$u_{\omega} = \frac{\hbar}{\pi^2 c^3} \frac{\omega^3}{\exp(\frac{\hbar\omega}{kT}) - 1} \quad \bullet(1.2.4)\bullet$$

$$E_n = -\frac{m_0 q^4}{8\epsilon_0 h^2 n^2}, \text{ with } n = 1, 2, \dots \quad \bullet(1.2.5)\bullet$$

$$E_{ph} = 13.6 \text{ eV} \left(\frac{1}{j^2} - \frac{1}{i^2} \right), \text{ with } i > j \quad \bullet(1.2.6)\bullet$$

$$2\pi r = n\lambda \quad \bullet(1.2.7)\bullet$$

$$m \frac{v^2}{r} = \frac{q^2}{4\pi\epsilon_0 r^2} \quad \bullet(1.2.8)\bullet$$

$$a_0 = \frac{\epsilon_0 h^2 n^2}{8\pi m_0 q} \quad \bullet(1.2.9)\bullet$$

$$E_n = -\frac{m_0 q^4}{8 \epsilon_0^2 \hbar^2 n^2}, \text{ with } n = 1, 2, \dots \quad \bullet(1.2.10)\bullet$$

$$V(r) = -\frac{q^2}{4 \pi \epsilon_0 r} \quad \bullet(1.2.11)\bullet$$

$$E = T + V = \frac{p^2}{2m} + V(x) \quad \bullet(1.2.12)\bullet$$

$$E\Psi = \frac{p^2}{2m}\Psi + V(x)\Psi \quad \bullet(1.2.13)\bullet$$

$$-\hbar^2 \frac{\partial^2 \Psi}{\partial x^2} = \hbar^2 k^2 \Psi = p^2 \Psi \text{ for } \Psi = e^{i(kx - \omega t)} \quad \bullet(1.2.14)\bullet$$

$$-\frac{\hbar^2}{2m} \frac{d^2 \Psi(x)}{dx^2} + V(x)\Psi(x) = E\Psi(x) \quad \bullet(1.2.15)\bullet$$

$$E_n = \frac{\hbar^2}{2m} \left(\frac{n}{2L_x}\right)^2, \text{ with } n = 1, 2, \dots \quad \bullet(1.2.16)\bullet$$

$$\frac{d\mathcal{E}(x)}{dx} = \frac{\mathcal{A}(x)}{\varepsilon} \quad \bullet(1.3.1)\bullet$$

$$\mathcal{E}(x_2) - \mathcal{E}(x_1) = \int_{x_1}^{x_2} \frac{\mathcal{A}(x)}{\varepsilon} dx \quad \bullet(1.3.2)\bullet$$

$$\vec{\nabla} \mathcal{E}(x, y, z) = \frac{\mathcal{A}(x, y, z)}{\varepsilon} \quad \bullet(1.3.3)\bullet$$

$$\vec{\mathcal{E}} \cdot \vec{A} = \frac{Q}{\varepsilon} \quad \bullet(1.3.4)\bullet$$

$$\frac{d \phi(x)}{dx} = -\mathcal{E}(x) \quad \bullet(1.3.5)\bullet$$

$$\phi(x_2) - \phi(x_1) = - \int_{x_1}^{x_2} \mathcal{E}(x) dx \quad \bullet(1.3.6)\bullet$$

$$\frac{d^2 \phi(x)}{dx^2} = - \frac{\mathcal{A}(x)}{\varepsilon} \quad \bullet(1.3.7)\bullet$$

$$\vec{\nabla} \phi(x, y, z) = -\mathcal{E}(x, y, z) \quad \bullet(1.3.8)\bullet$$

$$\nabla^2 \phi(x, y, z) = - \frac{\mathcal{A}(x, y, z)}{\varepsilon} \quad \bullet(1.3.9)\bullet$$

$$dU = dQ + dW + \mu dN \quad \bullet(1.4.1)\bullet$$

$$f(E) = \frac{1}{1 + e^{(E - E_F) / kT}} \quad \bullet(1.4.2)$$

Chapter 2: Semiconductor Fundamentals



Bibliography



1. Physical Properties of crystals, J.F. Nye, Oxford University Press, 1979.
2. Crystal physics, H.J. Juretschke, W.A. Benjamin, Inc., 1974
3. Introduction to Solid State Physics, C. Kittel, Wiley & sons,
4. Electrons and holes in semiconductors, W. Shockley, D. Van Nostrand Company, 1959
5. Electrons in Solids Third edition, R.H. Bube, Academic Press, 1992.
6. Solid State Electronic Devices, Fourth edition, B. G. Streetman, Prentice Hall, 1995
7. Solid State Theory, W. A. Harrison, Dover publications, 1974
8. Thermal Physics, second edition, C. Kittel and H. Kroemer, Freeman, 1980

Example 1.1 A metal has a workfunction of 4.3 V. What is the minimum photon energy in Joule to emit an electron from this metal through the photo-electric effect? What are the photon frequency in Terahertz and the photon wavelength in micrometer? What is the corresponding photon momentum? What is the velocity of a free electron with the same momentum?

Solution

The minimum photon energy, E_{ph} , equals the workfunction, Φ_M , in units of electron volt or 4.3 eV. This also equals

$$E_{ph} = q\Phi_M = 1.6 \times 10^{-19} \times 4.3 = 6.89 \times 10^{-19} \text{ Joule}$$

The corresponding photon frequency is:

$$n = \frac{E_{ph}}{h} = \frac{6.89 \times 10^{-19}}{6.626 \times 10^{-34}} = 1040 \text{ THz}$$

The corresponding wavelength equals:

$$l = \frac{hc}{E_{ph}} = \frac{6.626 \times 10^{-34} \times 3 \times 10^8}{6.89 \times 10^{-19}} = \frac{1.24 \text{ nm}}{E_{ph} (\text{eV})} = 0.288 \text{ nm}$$

The photon momentum, p , is:

$$p = \frac{h}{l} = \frac{6.626 \times 10^{-34}}{0.288 \times 10^{-6}} = 2.297 \times 10^{-27} \frac{\text{kg m}}{\text{s}}$$

And the velocity, v , of a free electron with the same momentum equals

$$v = \frac{p}{m_0} = \frac{2.297 \times 10^{-27}}{9.11 \times 10^{-31}} = 2522 \text{ m/s}$$

Where m_0 is the free electron mass.

Example 1.2 The spectral density of the sun peaks at a wavelength of 900 nm. If the sun behaves as a black body, what is the temperature of the sun?

Solution A wavelength of 900 nm corresponds to a photon energy of:

$$E_{ph} = \frac{hc}{\lambda} = \frac{6.626 \times 10^{-34} \times 3 \times 10^8}{900 \times 10^{-9}} = 2.21 \times 10^{-19} \text{ Joule}$$

Since the peak of the spectral density occurs at $2.82 kT$, the corresponding temperature equals:

$$T = \frac{E_{ph}}{2.82 k} = \frac{2.21 \times 10^{-19}}{2.82 \times 1.38 \times 10^{-23}} = 5672 \text{ Kelvin}$$

Example 1.3 An electron is confined to a 1 micron thin layer of silicon. Assuming that the semiconductor can be adequately described by a one-dimensional quantum well with infinite walls, calculate the lowest possible energy within the material in units of electron volt. If the energy is interpreted as the kinetic energy of the electron, what is the corresponding electron velocity? (The effective mass of electrons in silicon is 0.26 m_0 , where $m_0 = 9.11 \times 10^{-31}$ kg is the free electron rest mass).

Solution The lowest energy in the quantum well equals:

$$E_1 = \frac{h^2}{2m^*} \left(\frac{1}{2L_x} \right)^2 = \frac{(6.626 \times 10^{-34})^2}{2 \times 0.26 \times 9.11 \times 10^{-31}} \left(\frac{1}{2 \times 10^{-6}} \right)^2$$
$$= 2.32 \times 10^{-25} \text{ Joules} = 1.45 \mu\text{eV}$$

The velocity of an electron with this energy equals:

$$v = \sqrt{\frac{2E_1}{m^*}} = \sqrt{\frac{2 \times 2.32 \times 10^{-25}}{0.26 \times 9.11 \times 10^{-31}}} = 1.399 \text{ km/s}$$

Chapter 2: Semiconductor Fundamentals



2.7. Carrier Transport

- [2.7.1. Carrier drift](#)
- [2.7.2. Carrier Mobility](#)
- [2.7.3. Velocity saturation](#)
- [2.7.4. Carrier diffusion](#)

A motion of free carriers in a semiconductor leads to a current. This motion can be caused by an electric field due to an externally applied voltage, since the carriers are charged particles. We will refer to this as carrier drift. In addition, carriers also move from regions where the carrier density is high to regions where the carrier density is low. This carrier transport mechanism is due to the thermal energy and the associated random motion of the carriers. We will refer to this transport mechanism as carrier diffusion. The total current in a semiconductor equals the sum of the drift and the diffusion current.

As one applies an electric field to a semiconductor, the electrostatic force causes the carriers to first accelerate and then reach a constant average velocity, v , due to collisions with impurities and lattice vibrations. The ratio of the velocity to the applied field is called the mobility. The velocity saturates at high electric fields reaching the saturation velocity. Additional scattering occurs when carriers flow at the surface of a semiconductor, resulting in a lower mobility due to surface or interface scattering mechanisms.

Diffusion of carriers is obtained by creating a carrier density gradient. Such gradient can be obtained by varying the doping density in a semiconductor or by applying a thermal gradient.

Both carrier transport mechanisms are related since the same particles and scattering mechanisms are involved. This leads to a relationship between the mobility and the diffusion constant called the Einstein relation.

2.7.1. Carrier drift



- [2.7.1.1 Impurity scattering](#)
- [2.7.1.2 Lattice scattering](#)
- [2.7.1.3 Surface scattering](#)

The motion of a carrier drifting in a semiconductor due to an applied electric field is illustrated in Figure 2.7.1. The field causes the carrier to move with a velocity, v .

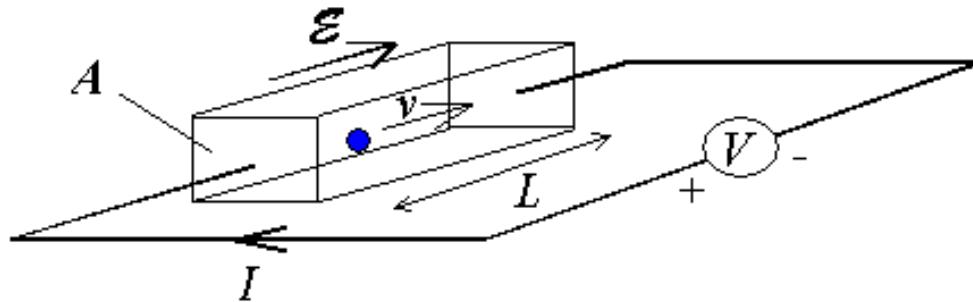


Figure 2.7.1 : Drift of a carrier due to an applied electric field.

Assuming that all the carriers in the semiconductor move with the same velocity, the current can be expressed as the total charge in the semiconductor divided by the time needed to travel from one electrode to the other, or:

$$I = \frac{Q}{t_r} = \frac{Q}{L/v} \quad (2.7.1)$$

where τ_r is the transit time of a particle, traveling with velocity, v , over the distance L . The current density can then be rewritten as a function of either the charge density, ρ , or the density of carriers, n in the semiconductor:

$$\vec{J} = \frac{Q}{AL} \vec{v} = \rho \vec{v} = qn\vec{v} \quad (2.7.2)$$

Carriers however do not follow a straight path along the electric field lines, but instead bounce around in the semiconductor and constantly change direction and velocity due to scattering. This behavior occurs even when no electric field is applied and is due to the thermal energy of the electrons. Electrons in a non-degenerate and non-relativistic electron gas have a thermal energy, which equals $kT/2$ per particle per degree of freedom. A typical thermal velocity at room temperature is around 10^7 cm/s, which exceeds the typical drift velocity in semiconductors. The carrier motion in the semiconductor in the absence and in the presence of an electric field can therefore be visualized as in Figure 2.7.2.

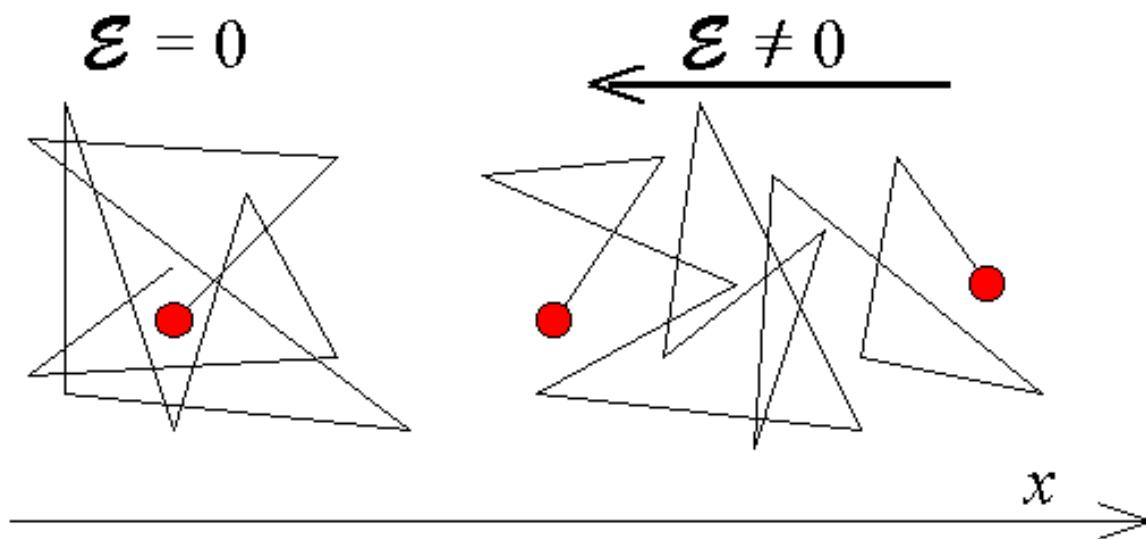


Figure 2.7.2 : Random motion of carriers in a semiconductor with and without an applied electric field.

In the absence of an applied electric field, the carrier exhibits random motion and the carriers move quickly through the semiconductor and frequently change direction. When an electric field is applied, the random motion still occurs but in addition, there is on average a net motion along the direction of the field.

We now analyze the carrier motion considering only the average velocity, $\langle \vec{v} \rangle$ of the carriers. Applying Newton's law, we state that the acceleration of the carriers is proportional to the applied force:

$$\vec{F} = m\vec{a} = m \frac{d \langle \vec{v} \rangle}{dt} \quad (2.7.3)$$

The force consists of the difference between the electrostatic force and the scattering force due to the loss of momentum at the time of scattering. This scattering force equals the momentum divided by the average time between scattering events, so that:

$$\vec{F} = q\vec{E} - \frac{m \langle \vec{v} \rangle}{\tau} \quad (2.7.4)$$

Combining both relations yields an expression for the average particle velocity:

$$q\vec{E} = m \frac{d \langle \vec{v} \rangle}{dt} + \frac{m \langle \vec{v} \rangle}{\tau} \quad (2.7.5)$$

We now consider only the steady state situation in which the particle has already accelerated and has reached a constant average velocity. Under such conditions, the velocity is proportional to the applied electric field and we define the mobility as the velocity to field ratio:

$$\mu = \frac{\Delta |\vec{v}|}{|\vec{E}|} = \frac{q \tau}{m} \quad (2.7.6)$$

The mobility of a particle in a semiconductor is therefore expected to be large if its mass is small and the time between scattering events is large.

The drift current, described by (2.7.2), can then be rewritten as a function of the mobility, yielding:

$$\vec{J} = qn\mu_n \vec{E} \quad (2.7.7)$$

Throughout this derivation, we simply considered the mass, m , of the particle. However in order to incorporate the effect of the periodic potential of the atoms in the semiconductor we must use the effective mass, m^* , rather than the free particle mass:

$$\mu = \frac{q \tau}{m^*} \quad (2.7.8)$$

Example 2.8	Electrons in undoped gallium arsenide have a mobility of 8,800 cm ² /V-s. Calculate the average time between collisions. Calculate the distance traveled between two collisions (also called the mean free path). Use an average velocity of 10 ⁷ cm/s.
-------------	---

Solution	<p>The collision time, τ_c, is obtained from:</p> $\tau_c = \frac{\mu_n m_e^*}{q} = \frac{0.88 \times 0.067 \times 9.1 \times 10^{-31}}{1.6 \times 10^{-19}} = 0.34 \text{ ps}$ <p>where the mobility was first converted in MKS units.</p> <p>The mean free path, l, equals:</p> $l = v_{average} \tau_c = 10^7 \times 0.34 \times 10^{-12} = 34 \text{ nm}$
----------	--

2.7.1.1 Impurity scattering

By impurities, we mean foreign atoms in the solid, which are efficient scattering centers especially when they have a net charge. Ionized donors and acceptors in a semiconductor are a common example of such impurities. The amount of scattering due to electrostatic forces between the carrier and the ionized impurity depends on the interaction time and the number of impurities. Larger impurity concentrations result in a lower mobility. The dependence on the interaction time helps to explain the temperature dependence. The interaction time is directly linked to the relative velocity of the carrier and the impurity, which is related to the thermal velocity of the carriers. This thermal velocity increases with the ambient temperature so that the interaction time increases. Thereby, the amount of scattering decreases, resulting in a mobility increase with temperature. To first order, the mobility due to impurity scattering is proportional to $T^{3/2}/N_I$, where N_I is the density of charged impurities.

2.7.1.2 Lattice scattering

Scattering by lattice waves includes the absorption or emission of either acoustical or optical phonons. Since the density of phonons in a solid increases with temperature, the scattering time due to this mechanism will decrease with temperature as will the mobility. Theoretical calculations reveal that the mobility in non-polar semiconductors, such as silicon and germanium, is dominated by acoustic phonon interaction. The resulting mobility is expected to be proportional to $T^{-3/2}$, while the mobility due to optical phonon scattering only is expected to be proportional to $T^{-1/2}$. Experimental values of the temperature dependence of the mobility in germanium, silicon and gallium arsenide are provided in Table 2.7.1.

	Germanium	Silicon	Gallium Arsenide
Electron mobility	$\propto T^{-1.7}$	$\propto T^{-2.4}$	$\propto T^{-1.0}$
Hole mobility	$\propto T^{-2.3}$	$\propto T^{-2.2}$	$\propto T^{-2.1}$

Table 2.7.1 : Temperature dependence of the mobility in germanium, silicon and gallium arsenide due to phonon scattering

2.7.1.3 Surface scattering

The surface and interface mobility of carriers is affected by the nature of the adjacent layer or surface. Even if the carrier does not transfer into the adjacent region, its wavefunction does extend over 1 to 10 nanometer, so that there is a non-zero probability for the particle to be in the adjacent region. The net mobility is then a combination of the mobility in both layers. For carriers in the inversion layer of a MOSFET, one finds that the mobility can be up to three times lower than the bulk value. This is due to the distinctly lower mobility of electrons in the amorphous silicon. The presence of charged surface states further reduces the mobility just as ionized impurities would.

2.7.2. Carrier Mobility



[2.7.2.1 Doping dependence](#)

[2.7.2.2 Conductivity and Resistivity](#)

2.7.2.1 Doping dependence

The mobility of electrons and holes in silicon at room temperature is shown in Figure 2.7.3.

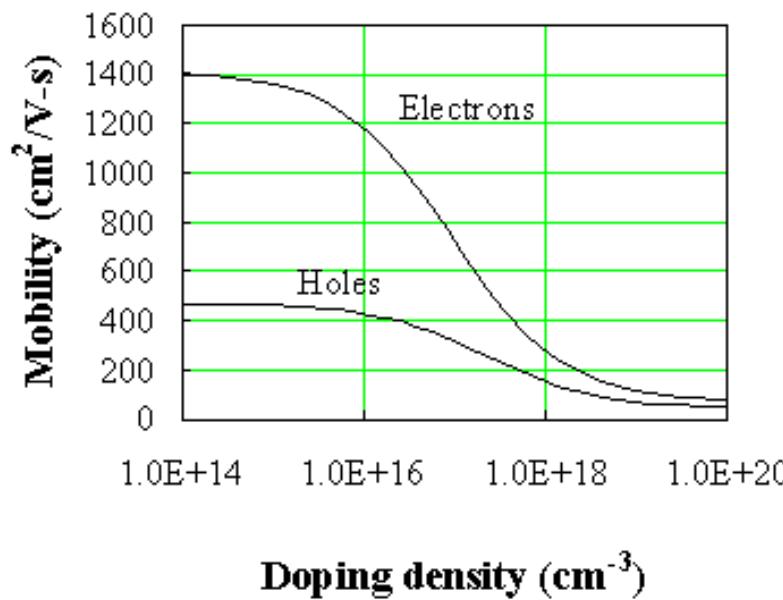


Figure 2.7.3 : Electron and hole mobility versus doping density for silicon

The electron mobility and hole mobility have a similar doping dependence: For low doping concentrations, the mobility is almost constant and is primarily limited by phonon scattering. At higher doping concentrations, the mobility decreases due to ionized impurity scattering with the ionized doping atoms. The actual mobility also depends on the type of dopant. Figure 2.7.3 is for phosphorous and boron doped silicon.

Note that the mobility is linked to the total number of ionized impurities or the sum of the donor and acceptor densities. The free carrier density, as described in section 2.6.4.1 is to first order related to the difference between the donor and acceptor concentration.

The minority carrier mobility also depends on the total impurity density. The minority-carrier mobility can be approximated by the majority-carrier mobility in a material with the same number of impurities. The mobility at a particular doping density is obtained from the following empiric expression:

$$\mu = \mu_{\min} + \frac{\mu_{\max} - \mu_{\min}}{1 + \left(\frac{N}{N_r}\right)^{\alpha}} \quad (2.7.9)$$

	Arsenic	Phosphorous	Boron
μ_{\min} ($\text{cm}^2/\text{V}\cdot\text{s}$)	52.2	68.5	44.9
μ_{\max} ($\text{cm}^2/\text{V}\cdot\text{s}$)	1417	1414	470.5
N_r (cm^{-3})	9.68×10^{16}	9.20×10^{16}	2.23×10^{17}
α	0.68	0.711	0.719

Table 2.7.2 : Parameters for calculation of the mobility as a function of the doping density

The resulting mobilities in units of $\text{cm}^2/\text{V}\cdot\text{s}$ are listed for different doping densities in Table 2.7.3.

<i>N</i>	Arsenic	Phosphorous	Boron
10^{15} cm^{-3}	1359	1362	462
10^{16} cm^{-3}	1177	1184	429
10^{17} cm^{-3}	727	721	317
10^{18} cm^{-3}	284	277	153
10^{19} cm^{-3}	108	115	71

Table 2.7.3 :

Mobility in silicon for different doping densities

2.7.2.2 Conductivity and Resistivity

The conductivity of a material is defined as the current density divided by the applied electric field. Since the current density equals the product of the charge of the mobile carriers, their density and velocity, it can be expressed as a function of the electric field using the mobility. To include the contribution of electrons as well as holes to the conductivity, we add the current density due to holes to that of the electrons, or:

$$J = q n v_e + q p v_h = q(n \mu_n + p \mu_p) E \quad (2.7.10)$$

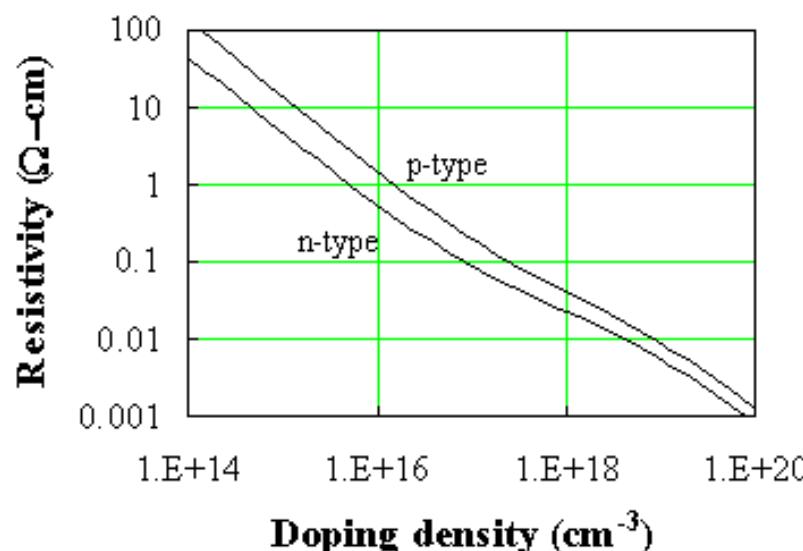
The conductivity due to electrons and holes is then obtained from:

$$\sigma = \frac{\Delta J}{E} = q(n \mu_n + p \mu_p) \quad (2.7.11)$$

The resistivity is defined as the inverse of the conductivity, namely:

$$\rho = \frac{1}{\sigma} = \frac{1}{q(\mu_n n + \mu_p p)} \quad (2.7.12)$$

The resulting resistivity as calculated with equation (2.7.12) is shown in Figure 2.7.4.

**Figure 2.7.4 :**Resistivity of n-type and p-type silicon versus doping density 

The sheet resistance concept is used to characterize both wafers and thin doped layers, since it is typically easier to measure the sheet resistance rather than the resistivity of the material. The sheet resistance of a uniformly-doped layer with resistivity, ρ , and thickness, t , is given by their ratio:

$$R_s = \frac{\rho}{t} \quad (2.7.13)$$

While the unit of the sheet resistance is Ohms, one refers to it as Ohms per square. This nomenclature comes in handy when the resistance of a rectangular piece of material with length, L , and width W must be obtained. It equals the product of the sheet resistance and the number of squares or:

$$R = R_s \frac{L}{W} \quad (2.7.14)$$

where the number of squares equals the length divided by the width. Figure 2.7.5 provides, as an example, the sheet resistance of a 14 mil thick silicon wafer which is n-type or p-type.

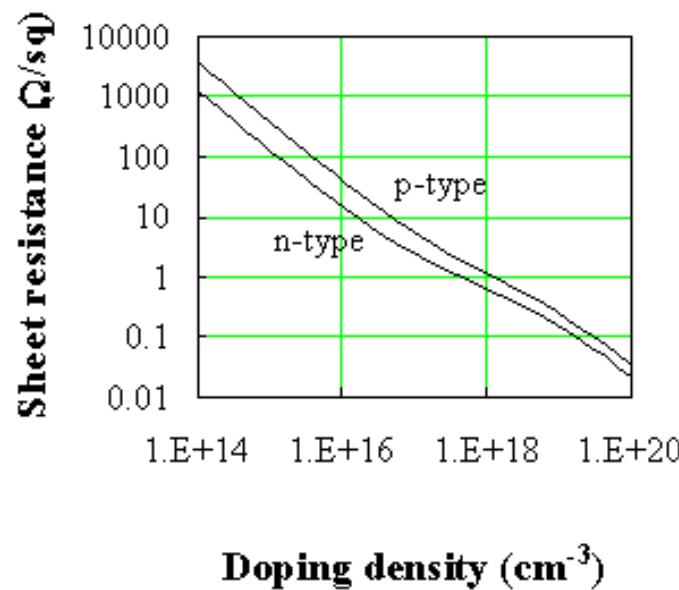


Figure 2.7.5 : Sheet resistance of a 14 mil thick n-type and p-type silicon wafer versus doping density. 

Example 2.9	A piece of silicon doped with arsenic ($N_d = 10^{17} \text{ cm}^{-3}$) is 100 μm long, 10 μm wide and 1 μm thick. Calculate the resistance of this sample when contacted one each end.
Solution	<p>The resistivity of the silicon equals:</p> $\rho = \frac{1}{qn\mu_n} = \frac{1}{1.6 \times 10^{-19} \times 10^{17} \times 727} = 0.086 \Omega\text{cm}$ <p>where the mobility was obtained from Table 2.7.3.</p> <p>The resistance then equals:</p> $R = \rho \frac{L}{Wt} = 0.086 \times \frac{100 \times 10^{-4}}{10 \times 10^{-4} \times 10^{-4}} = 8.6 \text{ k}\Omega$

An alternate approach is to first calculate the sheet resistance, R_s :

$$R_s = \frac{\rho}{t} = \frac{0.086}{10^{-4}} = 860 \Omega/\text{square}$$

From which one then obtains the resistance:

$$R = R_s \frac{L}{W} = 860 \times \frac{100 \times 10^{-4}}{10 \times 10^{-4}} = 8.6 \text{ k}\Omega$$



2.7.3. Velocity saturation



The linear relationship between the average carrier velocity and the applied field breaks down when high fields are applied. As the electric field is increased, the average carrier velocity and the average carrier energy increases as well. When the carrier energy increases beyond the optical phonon energy, the probability of emitting an optical phonon increases abruptly. This mechanism causes the carrier velocity to saturate with increasing electric field. For carriers in silicon and other materials, which do not contain accessible higher bands, the velocity versus field relation can be described by:

$$v(\mathcal{E}) = \frac{\mu \mathcal{E}}{1 + \frac{\mu \mathcal{E}}{v_{sat}}} \quad (2.7.15)$$

The maximum obtainable velocity, v_{sat} , is referred to as the saturation velocity.



2.7.4. Carrier diffusion

[2.7.4.1 Diffusion current](#)

[2.7.4.2 Total current](#)

Carrier diffusion is due to the thermal energy, kT , which causes the carriers to move at random even when no field is applied. This random motion does not yield a net motion of carriers nor does it yield a net current in material with a uniform carrier density as any carrier which leaves a specific location is on average replace by another one. However if a carrier gradient is present, the diffusion process will attempt to make the carrier density uniform: carriers diffuse from regions where the density is high to regions where the density is low. The diffusion process is not unlike the motion of sand on a vibrating table; hills as well as valleys are smoothed out over time.

In this section we will first derive the expression for the current due to diffusion and then combine it with the drift current to obtain the total drift-diffusion current.

2.7.4.1 Diffusion current

The derivation is based on the basic notion that carriers at non-zero temperature (Kelvin) have an additional thermal energy, which equals $kT/2$ per degree of freedom. It is the thermal energy, which drives the diffusion process. At $T = 0$ K there is no diffusion.

While one should recognize that the random nature of the thermal energy would normally require a statistical treatment of the carriers, we instead will use average values to describe the process. Such approach is justified on the basis that a more elaborate statistical approach yields the same results. To further simplify the derivation, we will derive the diffusion current for a one-dimensional semiconductor in which carriers can only move along one direction.

We now introduce the average values of the variables of interest, namely the thermal velocity, v_{th} , the collision time, τ_c , and the mean free path, l . The thermal velocity is the average velocity of the carriers going in the positive or negative direction. The collision time is the time during which carriers will move with the same velocity before a collision occurs with an atom or with another carrier. The mean free path is the average length a carrier will travel between collisions. These three averages are related by:

$$v_{th} = \frac{l}{\tau} \quad (2.7.16)$$

Consider now the situation illustrated with Figure 2.7.6.

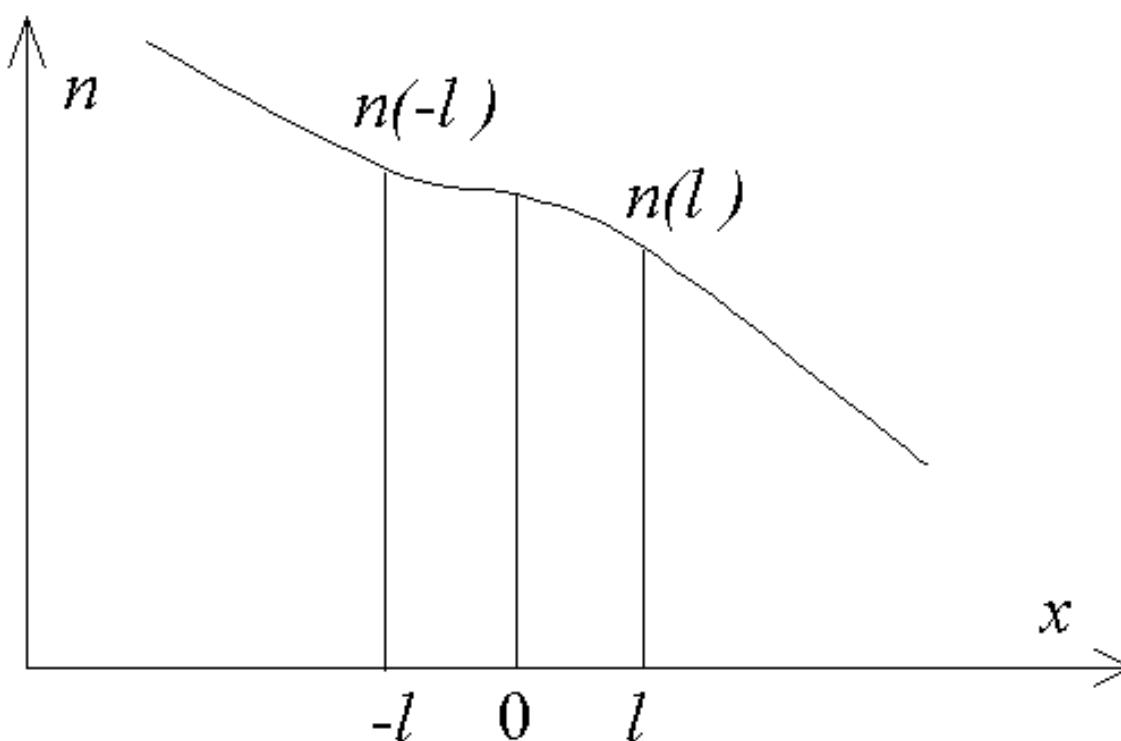


Figure 2.7.6 : Carrier density profile used to derive the diffusion current expression

Shown is a variable carrier density, $n(x)$. Of interest are the carrier densities which are one mean free path away from $x = 0$, since the carriers which will arrive at $x = 0$ originate either at $x = -l$ or $x = l$. The flux at $x = 0$ due to carriers which originate at $x = -l$ and move from left to right equals:

$$\Phi_{n, \text{left} \rightarrow \text{right}} = \frac{1}{2} v_{th} n(x = -l) \quad (2.7.17)$$

where the factor 1/2 is due to the fact that only half of the carriers move to the left while the other half moves to the right. The flux at $x = 0$ due to carriers, which originate at $x = +l$ and move from right to left, equals:

$$\Phi_{n, \text{right} \rightarrow \text{left}} = \frac{1}{2} v_{th} n(x = l) \quad (2.7.18)$$

The total flux of carriers moving from left to right at $x = 0$ therefore equals:

$$\Phi_n = \Phi_{n, left \rightarrow right} - \Phi_{n, right \rightarrow left} = \frac{1}{2} v_{th} [n(x = -l) - n(x = l)] \quad (2.7.19)$$

Where the flux due to carriers moving from right to left is subtracted from the flux due to carriers moving from left to right. Given that the mean free path is small we can write the difference in densities divided by the distance between $x = -l$ and $x = l$ as the derivative of the carrier density:

$$\Phi_n = -l v_{th} \frac{n(x = l) - n(x = -l)}{2l} = -l v_{th} \frac{dn}{dx} \quad (2.7.20)$$

The electron diffusion current equals this flux times the charge of an electron, or:

$$J_n = -q \Phi_n = q l v_{th} \frac{dn}{dx} \quad (2.7.21)$$

Typically, we will replace the product of the thermal velocity and the mean free path by a single parameter, namely the diffusion constant, D_n .

$$J_n = q D_n \frac{dn}{dx} \quad (2.7.22)$$

Repeating the same derivation for holes yields:

$$J_p = -q D_p \frac{dp}{dx} \quad (2.7.23)$$

We now further explore the relation between the diffusion constant and the mobility. At first, it seems that there should be no relation between the two since the driving force is distinctly different: diffusion is caused by thermal energy while an externally applied field causes drift. However one essential parameter in the analysis, namely the collision time, τ_c , should be independent of what causes the carrier motion.

We now combine the relation between the velocity, mean free path and collision time,

$$v_{th} = \frac{l}{\tau} \quad (2.7.24)$$

with the result from thermodynamics, stating that electrons carry a thermal energy which equals $kT/2$ for each degree of freedom. Applied to a one-dimensional situation, this leads to:

$$\frac{kT}{2} = \frac{m^* v_{th}^2}{2} \quad (2.7.25)$$

We now use these relations to rewrite the product of the thermal velocity and the mean free path as a function of the carrier mobility:

$$l v_{th} = \frac{m^* v_{th}^2}{q} \frac{q \tau}{m^*} = \frac{kT}{q} \mu \quad (2.7.26)$$

Using the definition of the diffusion constant we then obtain the following expressions which are often referred to as the Einstein relations:

$$D_n = \mu_n \frac{kT}{q} = \mu_n V_t \quad (2.7.27)$$

$$D_p = \mu_p \frac{kT}{q} = \mu_p V_t \quad (2.7.28)$$

Example 2.10	The hole density in an n-type silicon wafer ($N_d = 10^{17} \text{ cm}^{-3}$) decreases linearly from 10^{14} cm^{-3} to 10^{13} cm^{-3} between $x = 0$ and $x = 1 \mu\text{m}$. Calculate the hole diffusion current density.
--------------	--

Solution

The hole diffusion current density equals:

$$J_p = qD_p \frac{dp}{dx} = 1.6 \times 10^{-19} \times 8.2 \times \frac{9 \times 10^{13}}{10^{-4}} = 1.18 \text{ A/cm}^2$$

where the diffusion constant was calculated using the Einstein relation:

$$D_p = V_t \mu_p = 0.0259 \times 317 = 8.2 \text{ cm}^2/\text{s}$$

and the hole mobility in the n-type wafer was obtained from Table 2.7.3 as the hole mobility in a p-type material with the same doping density.

2.7.4.2 Total current

The total electron current is obtained by adding the current due to diffusion to the drift current, yielding:

$$J_n = qn \mu_n \mathcal{E} + qD_n \frac{dn}{dx} \quad (2.7.29)$$

and similarly for holes:

$$J_p = qp \mu_p \mathcal{E} - qD_p \frac{dp}{dx} \quad (2.7.30)$$

The total current is the sum of the electron and hole current densities multiplied with the area, A , perpendicular to the direction of the carrier flow:

$$I_{total} = A(J_n + J_p) \quad (2.7.31)$$

Chapter 2: Semiconductor Fundamentals



2.6. Carrier densities

- [2.6.1. General discussion](#)
- [2.6.2. Calculation of the Fermi integral](#)
- [2.6.3. Intrinsic semiconductors](#)
- [2.6.4. Doped semiconductors](#)
- [2.6.5. Non-equilibrium carrier densities](#)

Now that we have discussed the density of states and the distribution functions, we have all the necessary tools to calculate the carrier density in a semiconductor.

2.6.1 General discussion



The density of electrons in a semiconductor is related to the density of available states and the probability that each of these states is occupied. The density of occupied states per unit volume and energy, $n(E)$, is simply the product of the density of states in the conduction band, $g_c(E)$ and the Fermi-Dirac probability function, $f(E)$, (also called the Fermi function):

$$n(E) = g_c(E)f(E) \quad (2.6.1)$$

Since holes correspond to empty states in the valence band, the probability of having a hole equals the probability that a particular state is not filled, so that the hole density per unit energy, $p(E)$, equals:

$$p(E) = g_v(E)[1 - f(E)] \quad (2.6.2)$$

Where $g_v(E)$ is the density of states in the valence band. The density of carriers is then obtained by integrating the density of carriers per unit energy over all possible energies within a band. A general expression is derived as well as an approximate analytic solution, which is valid for non-degenerate semiconductors. In addition, we also present the Joyce-Dixon approximation, an approximate solution useful when describing degenerate semiconductors.

The density of states in a semiconductor was obtained by solving the Schrödinger equation for the particles in the semiconductor. Rather than using the actual and very complex potential in the semiconductor, we use the simple particle-in-a box model, where one assumes that the particle is free to move within the material.

For an electron which behaves as a free particle with effective mass, m^* , the density of states was derived in section [2.4](#), yielding:

$$g_c(E) = \frac{8\pi\sqrt{2}}{h^3} m_e^{*3/2} \sqrt{E - E_c}, \text{ for } E \geq E_c \quad (2.6.3)$$

where E_c is the bottom of the conduction band below which the density of states is zero. The density of states for holes in the valence band is given by:

$$g_v(E) = \frac{8\pi\sqrt{2}}{h^3} m_e^{*3/2} \sqrt{E_v - E}, \text{ for } E \leq E_v \quad (2.6.4)$$



2.6.2. Calculation of the Fermi integral

2.6.2.1 Carrier density at zero Kelvin

2.6.2.2 Non-degenerate semiconductors

2.6.2.3 Degenerate semiconductors

The carrier density in a semiconductor, is obtained by integrating the product of the density of states and the probability density function over all possible states. For electrons in the conduction band the integral is taken from the bottom of the conduction band, labeled, E_c , to the top of the conduction band:

$$n = \frac{\int_{E_c}^{\text{top of the conduction band}} n(E) dE}{\int_{E_c}^{\text{top of the conduction band}} g_c(E) f(E) dE} \quad (2.6.5)$$

Where $g_c(E)$ is the density of states in the conduction band and $f(E)$ is the Fermi function.

This general expression is illustrated with Figure 2.6.1 for a parabolic density of states function with $E_c = 0$. The figure shows the density of states function, $g_c(E)$, the Fermi function, $f(E)$, as well as the product of both, which is the density of electrons per unit volume and per unit energy, $n(E)$. The integral corresponds to the crosshatched area.

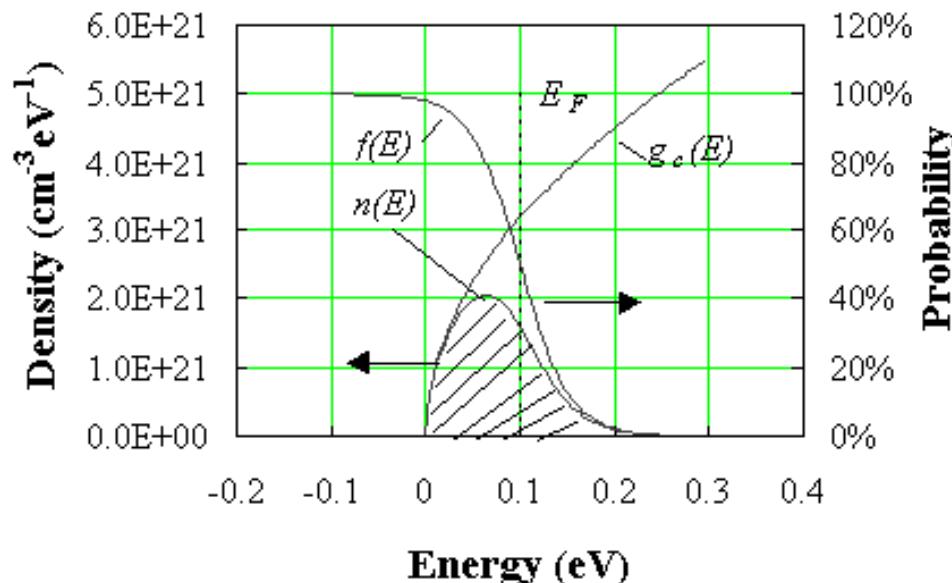


Figure 2.6.1 : The carrier density integral. Shown are the density of states, $g_c(E)$, the density per unit energy, $n(E)$, and the probability of occupancy, $f(E)$. The carrier density, n_o , equals the crosshatched area.

The actual location of the top of the conduction band does not need to be known as the Fermi function goes to zero at higher energies. The upper limit can therefore be replaced by infinity. We also relabeled the carrier density as no to indicate that the carrier density is the carrier density in thermal equilibrium.

$$n_o = \int_{E_c}^{\infty} g_c(E) f(E) dE \quad (2.6.6)$$

Using equations (2.6.3) and (2.5.1) this integral becomes:

$$n_o = \int_{E_c}^{\infty} \frac{8\pi\sqrt{2}}{h^3} m_e^{*3/2} \sqrt{E - E_c} \frac{1}{1 + e^{\frac{E - E_F}{kT}}} dE \quad (2.6.7)$$

While this integral can not be solved analytically at non-zero temperatures, we can obtain either a numeric solution or an approximate analytical solution. Similarly for holes one obtains:

$$p_o = \int_{-\infty}^{E_v} g_v(E) [1 - f(E)] dE \quad (2.6.8)$$

and

$$p_o = \int_{-\infty}^{E_v} \frac{8\pi\sqrt{2}}{h^3} m_h^{*3/2} \sqrt{E_v - E} \frac{1}{1 + e^{\frac{E_F - E}{kT}}} dE \quad (2.6.9)$$

The calculation of the electron and hole density in a semiconductor is further illustrated by Figure 2.6.2.

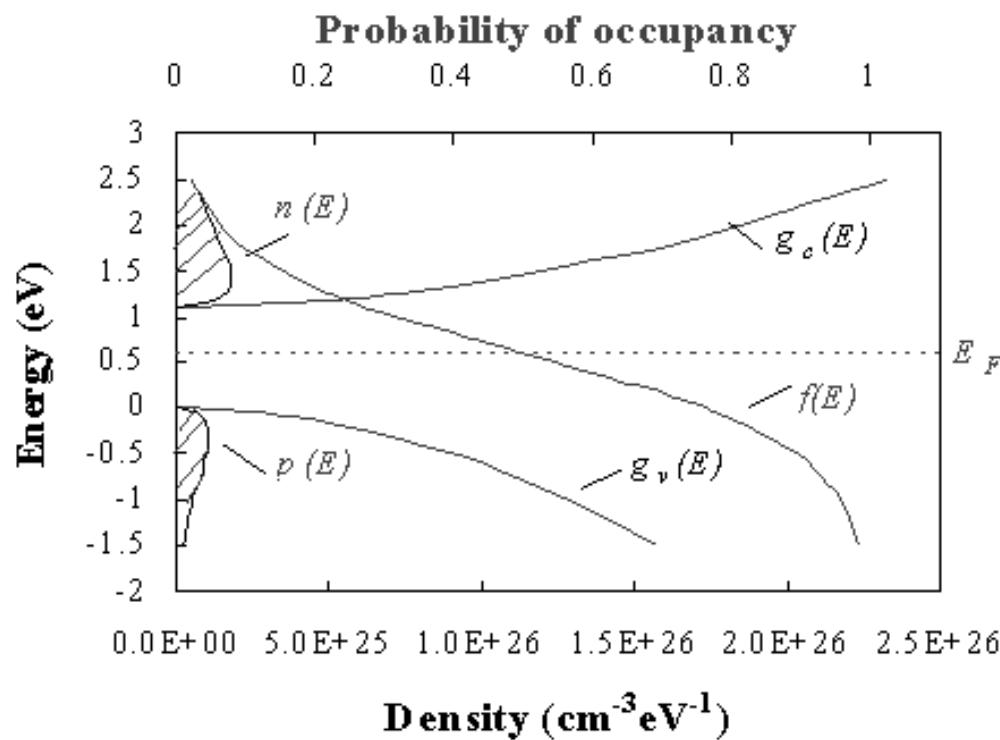


Figure 2.6.2 : The density of states and carrier densities in the conduction and valence band. Shown are the electron and hole density per unit energy, $n(E)$ and $p(E)$, the density of states in the conduction and valence band, $g_c(E)$ and $g_v(E)$ and the probability of occupancy, $f(E)$. The crosshatched area indicates the electron and hole densities. 

Indicated on the figure are the density of states in the conduction and valence band, the Fermi distribution function and the electron and hole densities per unit energy. The crosshatched areas indicate the thermal-equilibrium carrier densities. From the figure, one can easily see that the electron density will increase as the Fermi energy is increased. The hole density decreases with increasing Fermi energy. As the Fermi energy is decreased, the electron density decreases and the hole density increases.

2.6.2.1 Carrier density at zero Kelvin

Equation (2.6.7) can be solved analytically at $T = 0$ K, since the Fermi function at $T = 0$ K equals one for all energies below the Fermi energy and 0 for all energies larger than the Fermi energy. Equation (2.6.7) can therefore be simplified to:

$$n_o = \int_{E_c}^{E_F} g_c(E) dE \text{ at } T = 0 \text{ K} \quad (2.6.10)$$

and integration yields:

$$n_o = \frac{2}{3} \frac{\sqrt{2}}{\pi^2} \left(\frac{qm^*}{\hbar^2} \right)^{3/2} (E_F - E_c)^{3/2}, \text{ for } E_F \geq E_c \quad (2.6.11)$$

This expression can be used to approximate the carrier density in heavily degenerate semiconductors provided $kT \ll (E_F - E_c) > 0$

2.6.2.2 Non-degenerate semiconductors

Non-degenerate semiconductors are defined as semiconductors for which the Fermi energy is at least $3kT$ away from either band edge. The reason we restrict ourselves to non-degenerate semiconductors is that this definition allows the Fermi function to be replaced by a simple exponential function, i.e. the Maxwell-Boltzmann distribution function. The carrier density integral can then be solved analytically yielding:

$$n_o \cong \int_{E_c}^{\infty} \frac{8\pi\sqrt{2}}{\hbar^3} m_e^{*3/2} \sqrt{E - E_c} e^{\frac{E_F - E}{kT}} dE = N_c e^{\frac{E_F - E_c}{kT}} \quad (2.6.12)$$

with

$$N_c = 2 \left[\frac{2\pi m_e^{*kT}}{\hbar^2} \right]^{3/2} \quad (2.6.13)$$

where N_c is the effective density of states in the conduction band. Similarly for holes, one can approximate the hole density integral as:

$$p_o \cong \int_{-\infty}^{E_v} \frac{8\pi\sqrt{2}}{\hbar^3} m_h^{*3/2} \sqrt{E_v - E} e^{\frac{E - E_F}{kT}} dE = N_v e^{\frac{E_v - E_F}{kT}} \quad (2.6.14)$$

with

$$N_v = 2 \left[\frac{2\pi m_h^{*kT}}{\hbar^2} \right]^{3/2} \quad (2.6.15)$$

where N_v is the effective density of states in the valence band.

Example 2.4

Calculate the effective densities of states in the conduction and valence bands of germanium, silicon and gallium arsenide at 300 K.

Solution	<p>The effective density of states in the conduction band of germanium equals: where the effective mass for density of states was used (Appendix 3). Similarly one finds the effective densities for silicon and gallium arsenide and those of the valence band:</p> $N_c = 2 \left(\frac{2 \pi m_e^* k T}{h^2} \right)^{3/2}$ $= 2 \left(\frac{2 \pi 0.55 \times 9.11 \times 10^{-31} \times 1.38 \times 10^{-23} \times 300}{(6.626 \times 10^{-34})^2} \right)^{3/2}$ $= 1.02 \times 10^{25} \text{ m}^{-3} = 1.02 \times 10^{19} \text{ cm}^{-3}$ <p>Calculate the effective densities of states in the conduction and valence bands of germanium, silicon and gallium arsenide at 300 K.</p> <table border="1" data-bbox="489 696 1338 908"> <thead> <tr> <th></th><th>Germanium</th><th>Silicon</th><th>Gallium Arsenide</th></tr> </thead> <tbody> <tr> <td>$N_c \text{ (cm}^{-3}\text{)}$</td><td>$1.02 \times 10^{19}$</td><td>$2.81 \times 10^{19}$</td><td>$4.35 \times 10^{17}$</td></tr> <tr> <td>$N_v \text{ (cm}^{-3}\text{)}$</td><td>$5.64 \times 10^{18}$</td><td>$1.83 \times 10^{19}$</td><td>$7.57 \times 10^{18}$</td></tr> </tbody> </table> <p>Note that the effective density of states is temperature dependent and can be obtain from:</p> $N_c(T) = N_c(300 \text{ K}) \left(\frac{T}{300} \right)^{3/2}$ <p>where $N_c(300 \text{ K})$ is the effective density of states at 300 K.</p>		Germanium	Silicon	Gallium Arsenide	$N_c \text{ (cm}^{-3}\text{)}$	1.02×10^{19}	2.81×10^{19}	4.35×10^{17}	$N_v \text{ (cm}^{-3}\text{)}$	5.64×10^{18}	1.83×10^{19}	7.57×10^{18}
	Germanium	Silicon	Gallium Arsenide										
$N_c \text{ (cm}^{-3}\text{)}$	1.02×10^{19}	2.81×10^{19}	4.35×10^{17}										
$N_v \text{ (cm}^{-3}\text{)}$	5.64×10^{18}	1.83×10^{19}	7.57×10^{18}										

2.6.2.3 Degenerate semiconductors

A useful approximate expression applicable to degenerate semiconductors was obtained by Joyce and Dixon and is given by:

$$\frac{E_F - E_c}{kT} \cong \ln \frac{n_o}{N_c} + \frac{1}{\sqrt{8}} \frac{n_o}{N_c} - \left(\frac{3}{16} - \frac{\sqrt{3}}{9} \right) \left(\frac{n_o}{N_c} \right)^2 + \dots \quad (2.6.16)$$

for electrons and by:

$$\frac{E_v - E_F}{kT} \cong \ln \frac{p_o}{N_v} + \frac{1}{\sqrt{8}} \frac{p_o}{N_v} - \left(\frac{3}{16} - \frac{\sqrt{3}}{9} \right) \left(\frac{p_o}{N_v} \right)^2 + \dots \quad (2.6.17)$$

for holes.

2.6.3. Intrinsic semiconductors



2.6.3.1 Intrinsic carrier density2.6.3.2 Mass action law2.6.3.3 Intrinsic Fermi energy2.6.3.4 Intrinsic material as reference

Intrinsic semiconductors are semiconductors, which do not contain impurities. They do contain electrons as well as holes. The electron density equals the hole density since the thermal activation of an electron from the valence band to the conduction band yields a free electron in the conduction band as well as a free hole in the valence band. We will identify the intrinsic hole and electron density using the symbol n_i , and refer to it as the intrinsic carrier density.

2.6.3.1 Intrinsic carrier density

Intrinsic semiconductors are usually non-degenerate, so that the expressions for the electron ([2.6.12](#)) and hole ([2.6.14](#)) densities in non-degenerate semiconductors apply. Labeling the Fermi energy of intrinsic material as E_i , we can then write two relations between the intrinsic carrier density and the intrinsic Fermi energy, namely:

$$n_i = n_o \Big|_{(E_F = E_i)} = N_c e^{(E_i - E_c)/kT} \quad (2.6.18)$$

It is possible to eliminate the intrinsic Fermi energy from both equations, simply by multiplying both equations and taking the square root. This provides an expression for the intrinsic carrier density as a function of the effective density of states in the conduction and valence band, and the bandgap energy $E_g = E_c - E_v$.

$$n_i = \sqrt{N_c N_v} e^{-E_g/2kT} \quad (2.6.19)$$

The temperature dependence of the intrinsic carrier density is dominated by the exponential dependence on the energy bandgap. In addition, one has to consider the temperature dependence of the effective densities of states and that of the energy bandgap. A plot of the intrinsic carrier density versus temperature is shown in Figure [2.6.3](#). The temperature dependence of the effective masses was ignored.

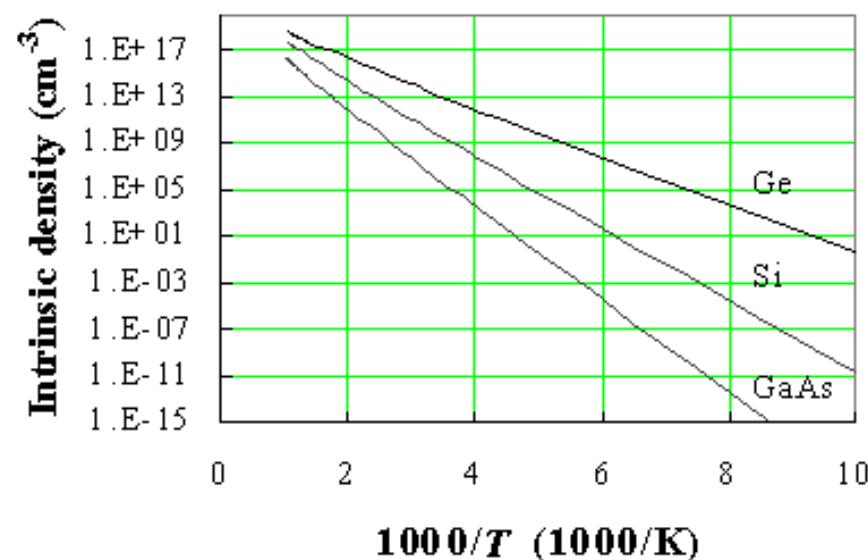


Figure 2.6.3 : Intrinsic carrier density versus temperature in gallium arsenide (GaAs), silicon (Si) and germanium (Ge).



Example 2.4b	Calculate the intrinsic carrier density in germanium, silicon and gallium arsenide at 300, 400, 500 and 600 K.																				
Solution	<p>The intrinsic carrier density in silicon at 300 K equals:</p> $n_i(300 \text{ K}) = \sqrt{N_c N_v} \exp\left(\frac{E_g}{2kT}\right)$ $= \sqrt{2.81 \times 10^{19} \times 1.83 \times 10^{19}} \exp\left(\frac{1.12}{2 \times 0.0258}\right)$ $= 8.72 \times 10^9 \text{ cm}^{-3}$ <p>Similarly one finds the intrinsic carrier density for germanium and gallium arsenide at different temperatures, yielding:</p> <table style="width: 100%; text-align: center;"> <thead> <tr> <th></th> <th>Germanium</th> <th>Silicon</th> <th>Gallium Arsenide</th> </tr> </thead> <tbody> <tr> <td>300 K</td> <td>2.02×10^{13}</td> <td>8.72×10^9</td> <td>2.03×10^6</td> </tr> <tr> <td>400 K</td> <td>1.38×10^{15}</td> <td>4.52×10^{12}</td> <td>5.98×10^9</td> </tr> <tr> <td>500 K</td> <td>1.91×10^{16}</td> <td>2.16×10^{14}</td> <td>7.98×10^{11}</td> </tr> <tr> <td>600 K</td> <td>1.18×10^{17}</td> <td>3.07×10^{15}</td> <td>2.22×10^{13}</td> </tr> </tbody> </table>		Germanium	Silicon	Gallium Arsenide	300 K	2.02×10^{13}	8.72×10^9	2.03×10^6	400 K	1.38×10^{15}	4.52×10^{12}	5.98×10^9	500 K	1.91×10^{16}	2.16×10^{14}	7.98×10^{11}	600 K	1.18×10^{17}	3.07×10^{15}	2.22×10^{13}
	Germanium	Silicon	Gallium Arsenide																		
300 K	2.02×10^{13}	8.72×10^9	2.03×10^6																		
400 K	1.38×10^{15}	4.52×10^{12}	5.98×10^9																		
500 K	1.91×10^{16}	2.16×10^{14}	7.98×10^{11}																		
600 K	1.18×10^{17}	3.07×10^{15}	2.22×10^{13}																		

Note that the values at 300 K as calculated in example 2.4 are not identical to those listed in Appendix 3. This is due to an accumulation of assumptions in the derivation. The numbers in Appendix 3 are obtained from careful measurements and should therefore be used instead of those calculated in example 2.4.

2.6.3.2 Mass action law

Using the same approach as in section 2.6.3.1, one can prove that the product of the electron and hole density equals the square of the intrinsic carrier density for any non-degenerate semiconductor. By multiplying the expressions for the electron and hole densities in a non-degenerate semiconductor, as in equations (2.6.12) and (2.6.14), one obtains:

$$n_o \cdot p_o = N_c N_v e^{(E_v - E_c)/kT} = n_i^2 \quad (2.6.20)$$

This property is referred to as the mass action law. It is a powerful relation, which enables to quickly find the hole density if the electron density is known or vice versa. This relation is only valid for non-degenerate semiconductors in thermal equilibrium

2.6.3.3 Intrinsic Fermi energy

The above equations for the intrinsic electron and hole density can be solved for the intrinsic Fermi energy, yielding:

$$E_i = \frac{E_c + E_v}{2} + \frac{1}{2} kT \ln\left(\frac{N_v}{N_c}\right) \quad (2.6.21)$$

The intrinsic Fermi energy is typically close to the midgap energy, half way between the conduction and valence band edge. The intrinsic Fermi energy can also be expressed as a function of the effective masses of the electrons and holes in the semiconductor. For this we use equations (2.6.13) and (2.6.15) for the effective density of states in the conduction and valence band, yielding:

$$E_i = \frac{E_c + E_v}{2} + \frac{3}{4} kT \ln\left(\frac{m_h^*}{m_e^*}\right) \quad (2.6.22)$$

2.6.3.4 Intrinsic material as reference

Dividing the expressions for the carrier densities (2.6.12) and (2.6.14), by the one for the intrinsic density (2.6.18) allows to write the carrier densities as a function of the intrinsic density, n_i , and the intrinsic Fermi energy, E_i , or:

$$n_o = n_i e^{(E_F - E_i)/kT} \quad (2.6.23)$$

and

$$p_o = n_i e^{(E_i - E_F)/kT} \quad (2.6.24)$$

We will use primarily these two equations to find the electron and hole density in a semiconductor in thermal equilibrium. The same relations can also be rewritten to obtain the Fermi energy from either carrier density, namely:

$$E_F = E_i + kT \ln \frac{n_o}{n_i} \quad (2.6.25)$$

and

$$E_F = E_i - kT \ln \frac{p_o}{n_i} \quad (2.6.26)$$

2.6.4. Doped semiconductors



[2.6.4.1 Dopants and impurities](#)

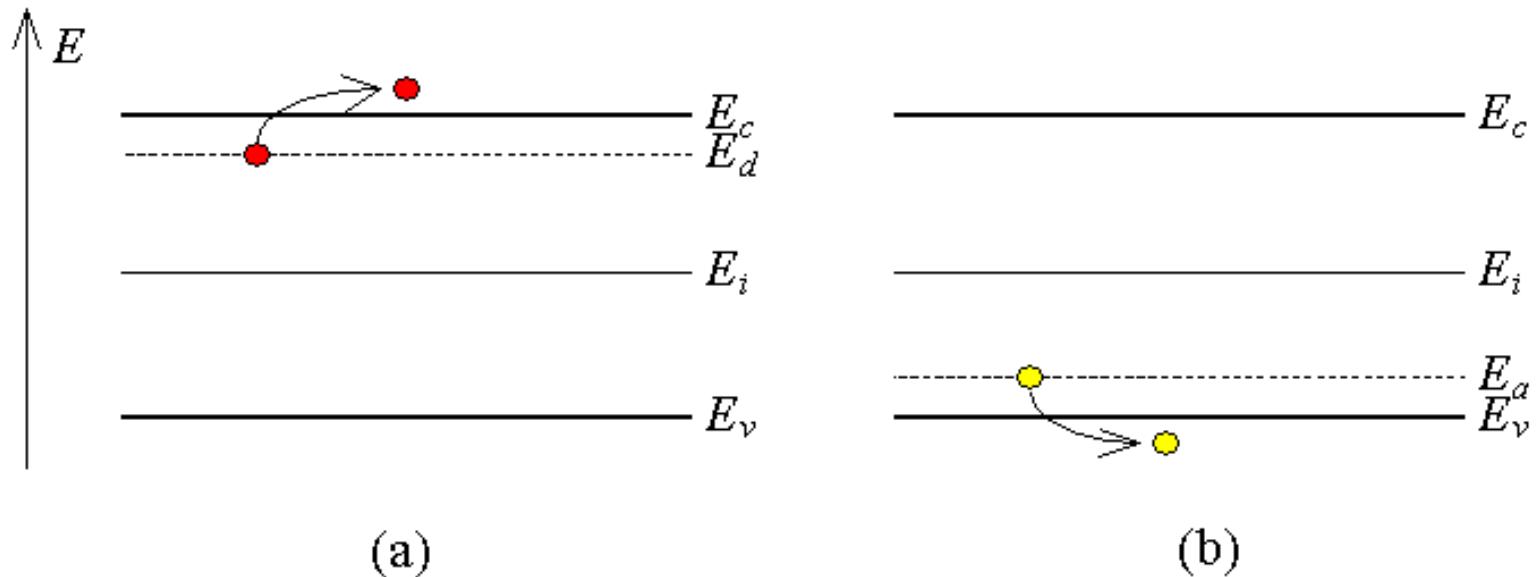
[2.6.4.2 Ionization energy model](#)

[2.6.4.3 Analysis of non-degenerately doped semiconductors](#)

[2.6.4.4 General analysis](#)

Doped semiconductors are semiconductors, which contain impurities, foreign atoms incorporated into the crystal structure of the semiconductor. Either these impurities can be unintentional, due to lack of control during the growth of the semiconductor, or they can be added on purpose to provide free carriers in the semiconductor.

The generation of free carriers requires not only that impurities are present, but also that the impurities give off electrons to the conduction band in which case they are called donors. If they give off holes to the valence band, they are called acceptors (since they effectively accept an electron from the filled valence band). The ionization of shallow donors and acceptors are illustrated by Figure 2.6.4. Indicated are the donor and acceptor energies, E_d and E_a . The donor energy level is filled prior to ionization. Ionization causes the donor to be emptied, yielding an electron in the conduction band and a positively charged donor ion. The acceptor energy is empty prior to ionization. Ionization of the acceptor corresponds to the empty acceptor level being filled by an electron from the filled valence band. This is equivalent to a hole given off by the acceptor atom to the valence band.

**Figure 2.6.4 :**

Ionization of a) a shallow donor and b) a shallow acceptor

A semiconductor doped with impurities, which are ionized (meaning that the impurity atoms either have donated or accepted an electron), will therefore contain free carriers. Shallow impurities are impurities, which require little energy - typically around the thermal energy, kT , or less - to ionize. Deep impurities require energies much larger than the thermal energy to ionize so that only a fraction of the impurities present in the semiconductor contribute to free carriers. Deep impurities, which are more than five times the thermal energy away from either band edge, are very unlikely to ionize. Such impurities can be effective recombination centers, in which electrons and holes fall and annihilate each other. Such deep impurities are also called traps.

Ionized donors provide free electrons in a semiconductor, which is then called n-type, while ionized acceptors provide free holes in a semiconductor, which we refer to as being a p-type semiconductor.

2.6.4.1 Dopants and impurities

The ionization of the impurities is dependent on the thermal energy and the position of the impurity level within the energy band gap as described by the impurity distribution functions discussed in section 2.5.3.

Shallow impurities readily ionize so that the free carrier density equals the impurity concentration. For shallow donors this implies that the electron density equals the donor concentration, or:

$$N_d^+ \cong N_d \quad (2.6.27)$$

While for shallow acceptors the hole density equals the acceptor concentration, or:

$$N_a^- \cong N_a \quad (2.6.28)$$

If a semiconductor contains both shallow donors and shallow acceptors it is called compensated since equal amounts of donor and acceptor atoms compensate each other, yielding no free carriers. The presence of shallow donors and shallow acceptors in a semiconductor cause the electrons given off by the donor atoms to fall into the acceptor state, which ionizes the acceptor atoms without yielding a free electron or hole. The resulting carrier density in compensated material, which contains both shallow donors and shallow acceptors, is approximately equal to the difference between the donor and acceptor concentration if the donor concentration is larger, yielding n-type material, or:

$$n_o \cong N_d^+ - N_a^-, \text{ if } N_d^+ - N_a^- \gg n_i \quad (2.6.29)$$

If the acceptor concentration is larger than the donor concentration, the hole density of the resulting p-type material equals the difference between the acceptor and donor concentration, or:

$$p_o \cong N_a^- - N_d^+, \text{ if } N_a^- - N_d^+ \gg n_i \quad (2.6.30)$$

2.6.4.2 Ionization energy model

The energy required to remove an electron from a donor atom can be approximated using a hydrogen-like model. After all, the donor atom consists of a positively charged ion and an electron just like the proton and electron of the hydrogen atom. The difference however is that the average distance, r , between the electron and the donor ion is much larger since the electron occupies one of the outer orbitals. This is illustrated by Figure 2.6.5.

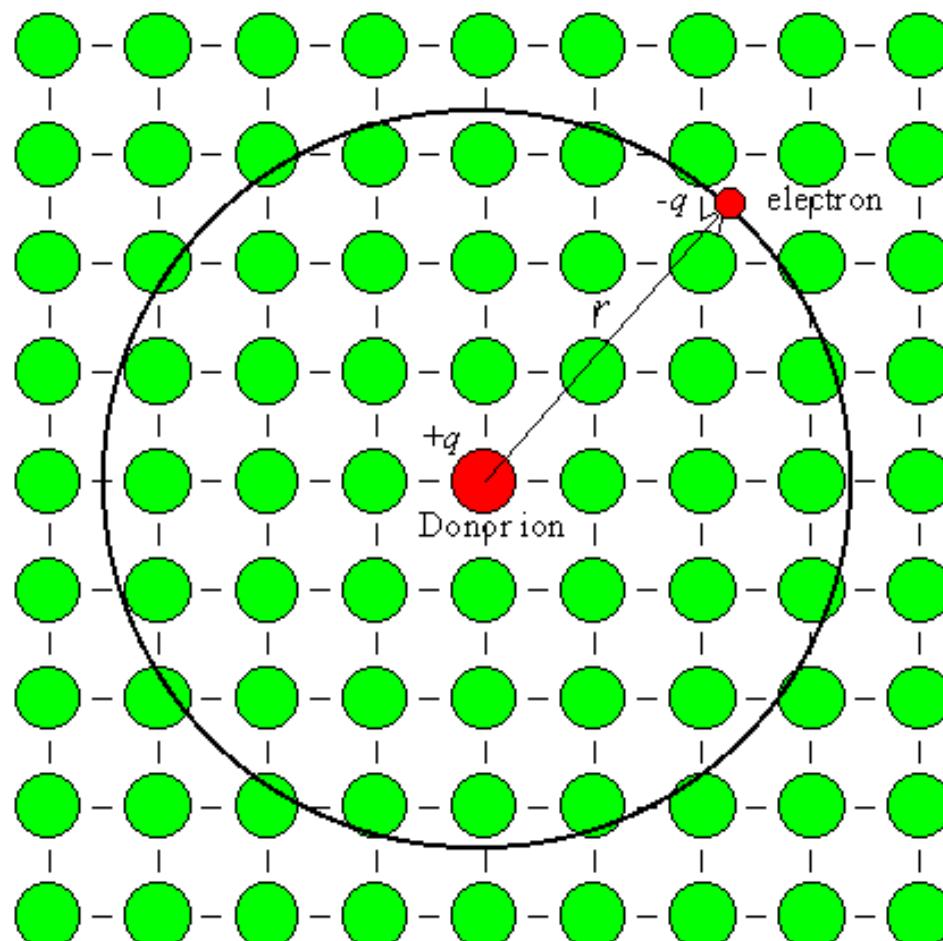


Figure 2.6.5: Trajectory of an electron bound to a donor ion within a semiconductor crystal.

For shallow donors, this distance, r , is much larger than the inter-atomic spacing of the semiconductor crystal. The ionization energy, E_d , can be estimated by modifying equation (1.2.10), which describes the electron energy in a Hydrogen atom, yielding:

$$E_c - E_d = 13.6 \frac{m_{\text{cond}}^*}{m_0 \epsilon_r^2} \text{ eV} \quad (2.6.31)$$

where m_{cond}^* is the effective mass for conductivity calculations and ϵ_r is the relative dielectric constant of the semiconductor. The ionization energy is calculated as the difference between the energy of a free electron and that of an electron occupying the lowest energy level, E_1 .

Example 2.5	Calculate the ionization energy for shallow donors and acceptors in germanium and silicon using the hydrogen-like model.									
Solution	<p>Using the effective mass for conductivity calculations (Appendix 3) one finds the ionization energy for shallow donors in germanium to be:</p> $E_c - E_d = 13.6 \frac{m_{\text{cond}}^*}{m_0 \epsilon_r^2} \text{ eV} = 13.6 \frac{0.12}{16^2} \text{ eV} = 6.4 \text{ meV}$ <p>The calculated ionization energies for donors and acceptors in germanium and silicon are provided below.</p> <table style="margin-left: auto; margin-right: auto;"> <tr> <td></td> <td style="text-align: center;">Germanium</td> <td style="text-align: center;">Silicon</td> </tr> <tr> <td>donors</td> <td style="text-align: center;">6.4 meV</td> <td style="text-align: center;">13.8 meV</td> </tr> <tr> <td>acceptors</td> <td style="text-align: center;">11.2 meV</td> <td style="text-align: center;">20.5 meV</td> </tr> </table> <p>Note that the actual ionization energies differ from these values and depend on the actual donor atom.</p>		Germanium	Silicon	donors	6.4 meV	13.8 meV	acceptors	11.2 meV	20.5 meV
	Germanium	Silicon								
donors	6.4 meV	13.8 meV								
acceptors	11.2 meV	20.5 meV								

2.6.4.3 Analysis of non-degenerately doped semiconductors

The calculation of the electron density starts by assuming that the semiconductor is neutral, so that there is a zero charge density in the material. This is a reasonable assumption since a net charge density would result in an electric field. This electric field would move any mobile charge so that it eliminates any charge imbalance.

The charge density in a semiconductor depends on the free electron and hole density and on the ionized impurity densities. Ionized donors, which have given off an electron, are positively charged. Ionized acceptors, which have accepted an electron, are negatively charged. The total charge density is therefore given by:

$$\rho = q(p_o - n_o + N_d^+ - N_a^-) = 0 \quad (2.6.32)$$

The hole concentration in thermal equilibrium can be written as a function of the electron density by using the mass action law (2.6.20). This yields the following relation between the electron density and the ionized impurity densities:

$$n_o = \frac{n_i^2}{n_o} + N_d^+ - N_a^- \quad (2.6.33)$$

Note that the use of the mass action law restricts the validity of this derivation to non-degenerate semiconductors as defined in section [2.6.2.2](#). Solving this quadratic equation yields a solution for the electron density, namely:

$$n_o = \frac{N_d^+ - N_a^-}{2} + \sqrt{\left(\frac{N_d^+ - N_a^-}{2}\right)^2 + n_i^2} \quad (2.6.34)$$

The same derivation can be repeated for holes, yielding:

$$p_o = \frac{N_a^- - N_d^+}{2} + \sqrt{\left(\frac{N_a^- - N_d^+}{2}\right)^2 + n_i^2} \quad (2.6.35)$$

The above expressions provide the free carrier densities for compensated semiconductors assuming that all donors and acceptors are ionized.

From the carrier densities, one then obtains the Fermi energies using equations [\(2.6.25\)](#) and [\(2.6.26\)](#) which are repeated below:

$$E_F = E_i + kT \ln \frac{n_o}{n_i} \quad (2.6.25)$$

or

$$E_F = E_i - kT \ln \frac{p_o}{n_i} \quad (2.6.26)$$

The Fermi energies in n-type and p-type silicon as a function of doping density is shown in Figure [2.6.6](#) for different temperatures:

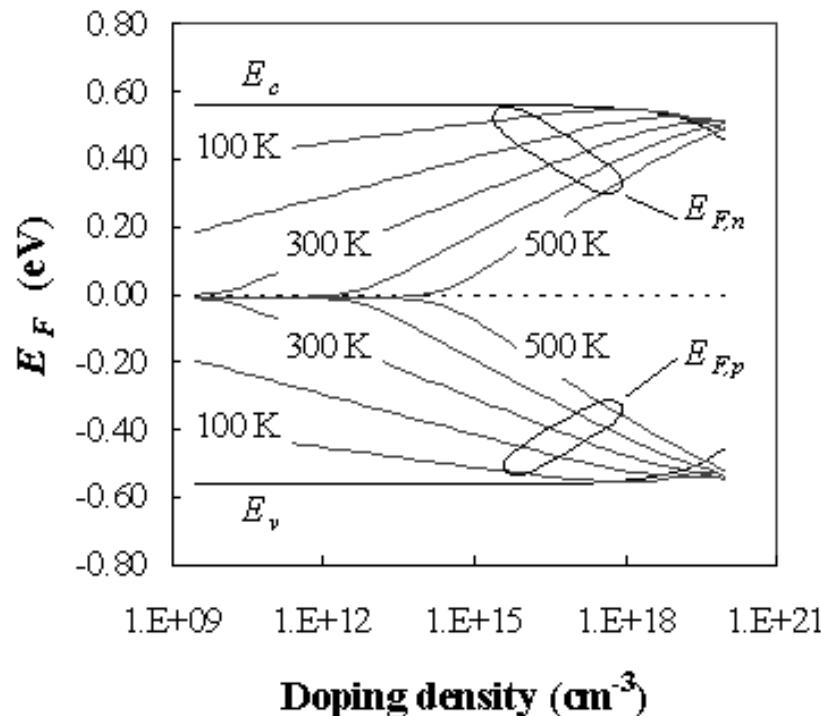


Figure 2.6.6 : Fermi energy of n-type and p-type silicon, $E_{F,n}$ and $E_{F,p}$, as a function of doping density at 100, 200, 300, 400 and 500 K. Shown are the conduction and valence band edges, E_c and E_v . The midgap energy is set to zero. 

Figure 2.6.6 illustrates how the Fermi energies vary with doping density. The Fermi energy varies linearly, when plotting the density on a logarithmic scale, up to a doping density of 10^{18} cm^{-3} . This simple dependence requires that the semiconductor is neither intrinsic nor degenerate and that all the dopants are ionized. For compensated material, containing only shallow dopants, one uses the net doping density, $|N_d - N_a|$.

Example 2.6a	A germanium wafer is doped with a shallow donor density of $3n_i/2$. Calculate the electron and hole density.
Solution	The electron density is obtained from equation (2.6.34) and the hole density is obtained using the mass action law:
Example 2.6b	A silicon wafer is doped with a shallow acceptor doping of 10^{16} cm^{-3} . Calculate the electron and hole density.
Solution	Since the acceptor doping is much larger than the intrinsic density and much smaller than the effective density of states, the hole density equals: The electron density is then obtained using the mass action law. The approach described in example 2.6a yields the same result.

2.6.4.4 General analysis

A more general analysis takes also into account the fact that the ionization of the impurities is not 100%, but instead is given by the impurity distribution functions provided in section 2.5.3.

The analysis again assumes that there is no net charge in the semiconductor (charge neutrality). This also means that the total density of positively charged particles (holes and ionized donors) must equals the total density of negatively charged particles (electrons and ionized acceptors) yielding:

$$p_o + N_d^+ = n_o + N_a^- \quad (2.6.36)$$

The electron and hole densities are then written as a function of the Fermi energy. For non-degenerate semiconductors one uses equations (2.6.12) and (2.6.14), while the ionized impurity densities equal the impurity density multiplied with the probability of occupancy for the acceptors and one minus the probability of occupancy for the donors. The Joyce-Dixon approximation, described in section 2.6.2.3 is used to calculate the degenerate carrier densities.

A graphical solution to equation (2.6.36) above can be obtained by plotting both sides of the equation as a function of the Fermi energy as illustrated in Figure 2.6.7.

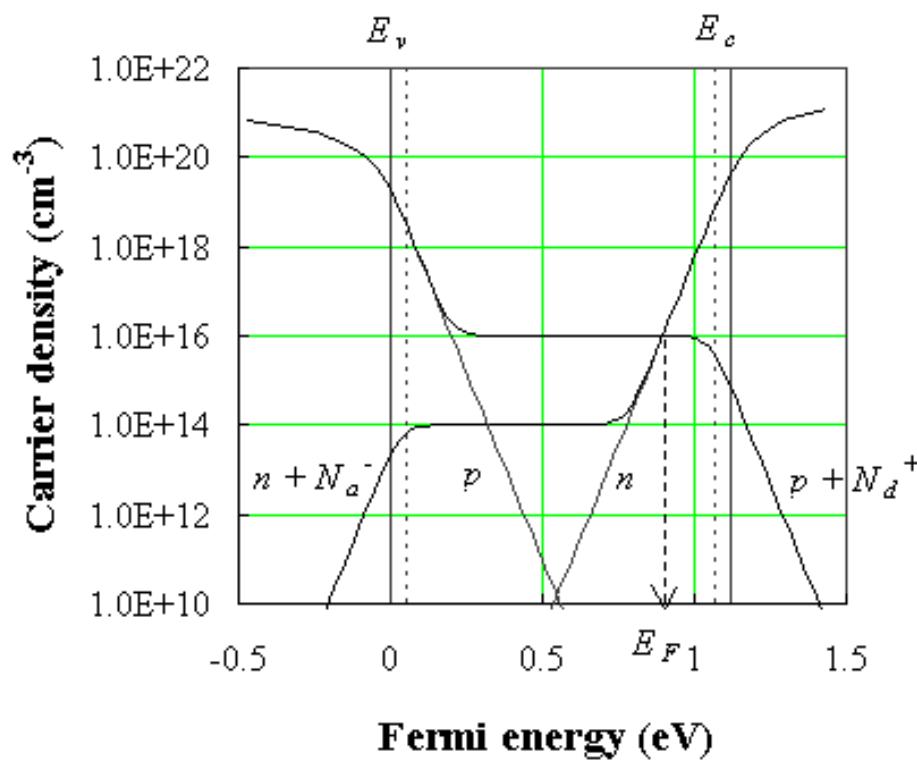


Figure 2.6.7 : Graphical solution of the Fermi energy based on the general analysis. The value of the Fermi energy and the free carrier density is obtained at the intersection of the two curves, which represent the total positive and total negative charge in the semiconductor. N_a equals 10^{16} cm^{-3} and N_d equals 10^{14} cm^{-3} . 

Figure 2.6.7 shows the positive and negative charge densities as well as the electron and hole densities as a function of the Fermi energy. The dotted lines indicate the position of the acceptor and donor energies. The Fermi energy is obtained at the intersection of both curves as indicated by the arrow.

This graphical solution is a very useful tool to explore the Fermi energy as a function of the doping densities, ionization energies and temperature.

Operation of devices over a wide temperature range requires a detailed knowledge of the carrier density as a function of temperature. At intermediate temperatures the carrier density approximately equals the net doping, $|N_a - N_d|$.

Semiconductors, which satisfy this condition, are also called extrinsic semiconductors. The free carrier density increases at high temperatures for which the intrinsic density approaches the net doping density and decreases at low temperatures due to incomplete ionization of the dopants. The carrier density and Fermi energy are shown in Figure 2.6.8 for silicon doped with 10^{16} cm^{-3} donors and 10^{15} cm^{-3} acceptors:

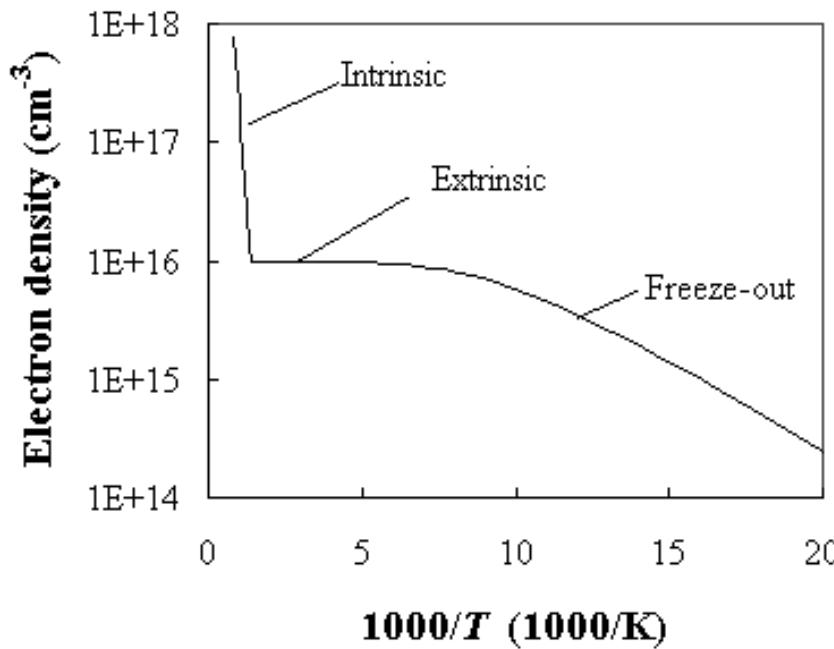


Figure 2.6.8 : Electron density and Fermi energy as a function of temperature in silicon with $N_d = 10^{16} \text{ cm}^{-3}$, $N_a = 10^{14} \text{ cm}^{-3}$ and $E_c - E_d = E_a - E_v = 50 \text{ meV}$. The activation energy at 70 K equals 27.4 meV. 

At high temperatures, the carrier density equals the intrinsic carrier concentration, while at low temperatures the carrier density is dominated by the ionization of the donors.

The temperature dependence is related to an activation energy by fitting the carrier density versus $1/T$ on a semi-logarithmic scale to a straight line of the form $n_o(T) = C \exp(-E_A/kT)$, where C is a constant. At high temperatures this activation energy equals half the bandgap energy or $E_A = E_g/2$.

The temperature dependence at low temperatures is somewhat more complex as it depends on whether or not the material is compensated. Figure 2.6.8 was calculated for silicon containing both donors and acceptors. At 70 K the electron density is below the donor density but still larger than the acceptor density. Under such conditions the activation energy, E_A , equals half of the ionization energy of the donors or $(E_c - E_d)/2$. At lower temperatures where the electron density is lower than the acceptor density, the activation energy equals the ionization energy or $E_c - E_d$. This behavior is explained by the fact that the Fermi energy in compensated material is fixed at the donor energy. The donors levels are always partially empty as electrons are removed from the donor atoms to fill the acceptor energy levels. If the acceptor density is smaller than the electron density - as is true for uncompensated material - the Fermi energy does change with temperature and the activation energy approaches half of the ionization energy.

Lightly doped semiconductors suffer from freeze-out at relatively high temperature. Higher-doped semiconductors freeze-out at lower temperatures. Highly-doped semiconductors do not contain a single donor energy, but rather an impurity band which overlaps with the conduction or valence band. The overlap of the two bands results in free carriers even at zero Kelvin. Degenerately doping a semiconductor therefore eliminates freeze-out effects.

2.6.5. Non-equilibrium carrier densities



Up until now, we have only considered the thermal equilibrium carrier densities, n_o and p_o . However most devices of interest are not in thermal equilibrium. Keep in mind that a constant ambient constant temperature is not a sufficient condition for thermal equilibrium. In fact, applying a non-zero voltage to a device or illuminating it with light will cause a non-equilibrium condition, even if the temperature is constant.

To describe a system that is not in thermal equilibrium we assume that each of the carrier distributions is still in equilibrium with itself. Such assumption is justified on the basis that electrons readily interact with each other and interact with holes only on a much longer time scale. As a result the electron density can still be calculated using the Fermi-Dirac distribution function, but with a different value for the Fermi energy. The total carrier density for a non-degenerate semiconductor is then described by:

$$n = n_o + \delta n = n_i \exp\left(\frac{F_n - E_i}{kT}\right) \quad (2.6.37)$$

Where δn is the excess electron density and F_n is the quasi-Fermi energy for the electrons. Similarly, the hole density can be expressed as:

$$p = p_o + \delta p = n_i \exp\left(\frac{E_i - F_p}{kT}\right) \quad (2.6.38)$$

Where δp is the excess hole density and F_p is the quasi-Fermi energy for the holes.

Example 2.7	A piece of germanium doped with 10^{16} cm^{-3} shallow donors is illuminated with light generating 10^{15} cm^{-3} excess electrons and holes. Calculate the quasi-Fermi energies relative to the intrinsic energy and compare it to the Fermi energy in the absence of illumination.
Solution	<p>The carrier densities when illuminating the semiconductor are:</p> $n = n_o + \delta n = 10^{16} + 10^{15} = 1.1 \times 10^{16} \text{ cm}^{-3}$ $p = p_o + \delta p \approx 10^{15} \text{ cm}^{-3}$ <p>and the quasi-Fermi energies are:</p> $F_n - E_i = kT \ln \frac{n}{n_i} = 0.0259 \times \ln \frac{1.1 \times 10^{16}}{2 \times 10^{13}} = 163 \text{ meV}$ $F_p - E_i = -kT \ln \frac{p}{n_i} = 0.0259 \times \ln \frac{1 \times 10^{15}}{2 \times 10^{13}} = -101 \text{ meV}$ <p>In comparison, the Fermi energy in the absence of light equals</p> $E_F - E_i = kT \ln \frac{n_o}{n_i} = 0.0259 \times \ln \frac{10^{16}}{2 \times 10^{13}} = 161 \text{ meV}$ <p>which is very close to the quasi-Fermi energy of the majority carriers.</p>

Example 1.4 Consider an infinitely long cylinder with charge density \mathbf{r} , dielectric constant ϵ_0 and radius r_0 . What is the electric field in and around the cylinder?

Solution Because of the cylinder symmetry one expects the electric field to be only dependent on the radius, r . Applying Gauss's law one finds:

$$E \cdot A = E 2\pi r L = \frac{Q}{\epsilon_0} = \frac{r \mathbf{p} r^2 L}{\epsilon_0} \text{ for } r < r_0$$

and

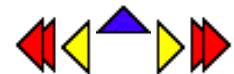
$$E \cdot A = E 2\pi r L = \frac{Q}{\epsilon_0} = \frac{r \mathbf{p} r_0^2 L}{\epsilon_0} \text{ for } r > r_0$$

where a cylinder with length L was chosen to define the surface A , and edge effects were ignored. The electric field then equals:

$$E(r) = \frac{r r}{2\epsilon_0} \text{ for } r < r_0 \text{ and } E(r) = \frac{r r_0^2}{2\epsilon_0 r} \text{ for } r > r_0$$

The electric field increases within the cylinder with increasing radius. The electric field decreases outside the cylinder with increasing radius.

Chapter 7: MOS Field-Effect-Transistors



Equations

$$I_D = -\frac{Q_{inv}WL}{t_r} \quad \bullet(7.3.1)\bullet$$

$$t_r = \frac{L}{v} \quad \bullet(7.3.2)\bullet$$

$$v = \mu \mathcal{E} = \mu \frac{V_{DS}}{L} \quad \bullet(7.3.3)\bullet$$

$$I_D = -\mu Q_{inv} \cdot \frac{W}{L} \cdot V_{DS} \quad \bullet(7.3.4)\bullet$$

$$Q_{inv} = -C_{ox}(V_{GS} - V_T), \text{ for } V_{GS} > V_T \quad \bullet(7.3.5)\bullet$$

$$I_D = \mu C_{ox} \frac{W}{L} (V_{GS} - V_T) V_{DS}, \text{ for } |V_{DS}| \ll (V_{GS} - V_T) \quad \bullet(7.3.6)\bullet$$

$$I_D = \mu C_{ox} \frac{W}{dy} (V_G - V_S - V_C - V_T) dV_C \quad \bullet(7.3.7)\bullet$$

$$\int_0^L I_D dy = \mu C_{ox} W \int_0^{V_{DS}} (V_G - V_S - V_C - V_T) dV_C \quad \bullet(7.3.8)\bullet$$

$$I_D = \mu C_{ox} \frac{W}{L} [(V_{GS} - V_T) V_{DS} - \frac{V^2}{2}], \quad \text{for } V_{DS} < V_{GS} - V_T \quad \bullet(7.3.9)\bullet$$

$$I_{D,sat} = \mu C_{ox} \frac{W}{L} \frac{(V_{GS} - V_T)^2}{2}, \quad \text{for } V_{DS} > V_{GS} - V_T \quad \bullet(7.3.10)\bullet$$

$$I_D = 0, \quad \text{for } V_{GS} < V_T \quad \bullet(7.3.11)\bullet$$

$$g_m \stackrel{\Delta}{=} \left. \frac{\partial I_D}{\partial V_{GS}} \right|_{V_{DS}} \quad \bullet(7.3.12)\bullet$$

$$g_{m,quad} = \mu C_{ox} \frac{W}{L} V_{DS} \quad \bullet(7.3.13)\bullet$$

$$g_{m,sat} = \mu C_{ox} \frac{W}{L} (V_{GS} - V_T) \quad \bullet(7.3.14)\bullet$$

$$g_d \stackrel{\Delta}{=} \left. \frac{\partial I_D}{\partial V_{DS}} \right|_{V_{GS}} \quad \bullet(7.3.15)\bullet$$

$$g_{d,quad} = \mu C_{ox} \frac{W}{L} (V_{GS} - V_T - V_{DS}) \quad \bullet(7.3.16)\bullet$$

$$g_{d,sat} = 0 \quad \bullet(7.3.17)\bullet$$

$$I_{D,sat} = \mu C_{ox} \frac{W}{L} \frac{(V_{GS} - V_T)^2}{2} (1 + \lambda V_{DS}), \quad \text{for } V_{DS} > V_{GS} - V_T \quad \bullet(7.3.18)\bullet$$

$$Q_{inv} = -C_{ox} (V_{GS} - V_T), \quad \text{for } V_{GS} > V_T \quad \bullet(7.3.19)\bullet$$

$$V_T = V_{FB} + V_C + 2\phi_F + \frac{\sqrt{2\varepsilon_s q N_a (2\phi_F + V_{SB} + V_C)}}{C_{ox}} \quad \bullet(7.3.20)\bullet$$

$$I_D = \mu_n C_{ox} \frac{W}{dy} (V_{GS} - V_{FB} - 2\phi_F - V_C - \frac{\sqrt{2\varepsilon_s q N_a (2\phi_F + V_{SB} + V_C)}}{C_{ox}}) dV_C \quad \bullet(7.3.21)\bullet$$

$$\int_0^L I_D dy = \mu_n C_{ox} W \int_0^{V_{DS}} (V_{GS} - V_{FB} - 2\phi_F - V_C) dV_C - \mu_n W \int_0^{V_{DS}} \sqrt{2\varepsilon_s q N_a (2\phi_F + V_{SB} + V_C)} dV_C \quad \bullet(7.3.22)\bullet$$

$$I_D = \frac{\mu_n C_{ox} W}{L} (V_{GS} - V_{FB} - 2\phi_F - \frac{V_{DS}}{2}) V_{DS} - \frac{2}{3} \mu_n \frac{W}{L} \sqrt{2\varepsilon_s q N_a} ((2\phi_F + V_{DB})^{3/2} - (2\phi_F + V_{SB})^{3/2}) \quad \bullet(7.3.23)\bullet$$

$$V_{DS,sat} = V_{GS} - V_{FB} - 2\phi_F - \frac{q N_a \varepsilon_s}{C_{ox}^2} \left(\sqrt{1 + 2 \frac{C_{ox}^2}{q N_a \varepsilon_s} (V_{GB} - V_{FB})} - 1 \right) \quad \bullet(7.3.24)\bullet$$

$$g_{m,sat} = \mu_n C_{ox} \frac{W}{L} [V_{GS} - V_{FB} - 2\phi_F - \frac{q N_a \varepsilon_s}{C_{ox}^2} \left(\sqrt{1 + 2 \frac{C_{ox}^2}{q N_a \varepsilon_s} (V_{GB} - V_{FB})} - 1 \right)] \quad \bullet(7.3.25)\bullet$$

$$g_{m,sat} = \mu_n^* C_{ox} \frac{W}{L} (V_{GS} - V_T) \quad \bullet(7.3.26)\bullet$$

$$\mu_n^* = \mu_n \left(1 - \frac{1}{\sqrt{1 + \frac{2(2\phi_F + V_{SB}) C_{ox}^2}{q N_a \varepsilon_s}}} \right) \quad \bullet(7.3.27)\bullet$$

$$V_T = V_{FB} + 2\phi_F + \frac{\sqrt{2\varepsilon_s q N_a (2\phi_F + V_{SB})}}{C_{ox}} \quad \bullet(7.4.1)\bullet$$

$$V_{FB} = \Phi_{MS} - \frac{Q_f}{C_{ox}} - \frac{1}{C_{ox}} \int_0^{t_{ox}} \frac{x}{x_{ox}} \rho_{ox}(x) dx \quad \bullet(7.4.2)\bullet$$

$$\Phi_{MS} = \Phi_M - \Phi_S = \Phi_M - \left(\chi + \frac{E_g}{2q} + \phi_F \right) \quad \bullet(7.4.3)\bullet$$

$$\phi_F = V_t \ln \frac{N_a}{n_i}, \text{ p - substrate} \quad \bullet(7.4.4)\bullet$$

$$V_T = V_{FB} - |2\phi_F| - \frac{\sqrt{2\varepsilon_s q N_a (|2\phi_F| - V_{SB})}}{C_{ox}} \quad \bullet(7.4.5)\bullet$$

$$V_{FB} = \Phi_{MS} - \frac{Q_f}{C_{ox}} - \frac{1}{C_{ox}} \int_0^{t_{ox}} \frac{x}{x_{ox}} \rho_{ox}(x) dx \quad \bullet(7.4.6)\bullet$$

$$\Phi_{MS} = \Phi_M - \Phi_S = \Phi_M - \left(\chi + \frac{E_g}{2q} - |\phi_F| \right) \quad \bullet(7.4.7)\bullet$$

$$|\phi_F| = V_t \ln \frac{N_d}{n_i}, \text{ n - substrate} \quad \bullet(7.4.8)\bullet$$

$$V_T = V_{FB} + 2\phi_F + \frac{\sqrt{2\varepsilon_s q N_a (2\phi_F + V_{SB})}}{C_{ox}} \quad \bullet(7.4.9)\bullet$$

$$\Delta V_T = \gamma(\sqrt{(2\phi_F + V_{SB})} - \sqrt{2\phi_F}) \quad \bullet(7.4.10)\bullet$$

$$\gamma = \frac{\sqrt{2\varepsilon_s q N_a}}{C_{ox}} \quad \bullet(7.4.11)\bullet$$



•(7.5.1)•

$$KP = \mu C_{ox} \quad •(7.5.2)•$$

$$VTO = V_{FB} + 2\phi_F + \frac{\sqrt{2\epsilon_s q N_a (2\phi_F)}}{C_{ox}} \quad •(7.5.3)•$$

$$GAMMA = \gamma = \frac{\sqrt{2\epsilon_s q N_a}}{C_{ox}} \quad •(7.5.4)•$$



•(7.5.5)•



•(7.5.6)•



•(7.5.7)•



•(7.5.8)•

$$CGSO = CGDO = \frac{\epsilon_{ox} \Delta L}{t_{ox}} \quad •(7.5.9)•$$

$$I_D \propto \exp\left(\frac{V_G - V_T}{V_t}\right) \quad •(7.7.1)•$$

$$n = 1 + \frac{1}{2C_{ox}} \sqrt{\frac{q \varepsilon_s N_a}{\phi_F}} \quad \bullet(7.7.2)\bullet$$

$$Q_d \propto \exp\left(-\frac{\phi_s}{V_t}\right) \quad \bullet(7.7.3)\bullet$$

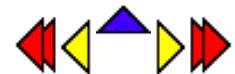
$$V_G = V_{FB} + \phi_s + V_{ox} = V_{FB} + \phi_s + \frac{\sqrt{2q \varepsilon_s \phi_s}}{C_{ox}} \quad \bullet(7.7.4)\bullet$$

$$\frac{dV_G}{d\phi_s} = 1 + \frac{1}{2C_{ox}} \sqrt{\frac{2q \varepsilon_s}{\phi_s}} \cong 1 + \frac{1}{2C_{ox}} \sqrt{\frac{q \varepsilon_s}{2\phi_F}} = n \quad \bullet(7.7.5)\bullet$$

$$I_D \propto Q_d \propto \exp\left(-\frac{\phi_s}{V_t}\right) \propto \exp\left(-\frac{V_G}{nV_t}\right) \quad \bullet(7.7.6)\bullet$$

$$\mu_{surface} \propto \varepsilon^{-1/3} \quad \bullet(7.7.7)\bullet$$

Chapter 1: Review of Modern Physics



Bibliography

1. Quantum Mechanics, H. Kroemer, Prentice Hall, 1994.
2. Electrons in Solids, Third edition, R.H. Bube, Academic Press, 1992.
3. Solid State Electronic Devices, Fifth edition, B. G. Streetman, Prentice Hall, 2000.
4. Quantum Mechanics, Second edition, E. Merzbacher, Wiley, 1970.
5. Quantum Physics of Electronics, S.N. Levine, The Macmillan company, 1965.
6. Thermal Physics, Second edition, C. Kittel and H. Kroemer, Freeman, 1980.

Chapter 1: Review of Modern Physics

1.2. Quantum mechanics

$$E_{ph} = h\mathbf{n} = \frac{hc}{\mathbf{l}} \quad (1.2.1)$$

$$\mathbf{l} = \frac{h}{p} \quad (1.2.2)$$

$$E_{ph} = h\mathbf{n} = \hbar\mathbf{w} \quad (1.2.3)$$

$$u_{\mathbf{w}} = \frac{\hbar}{\mathbf{p}^2 c^3} \frac{\mathbf{w}^3}{\exp(\frac{\hbar\mathbf{w}}{kT}) - 1} \quad (1.2.4)$$

$$E_n = -\frac{13.6 \text{ eV}}{n^2}, \text{ with } n=1, 2, \dots \quad (1.2.5)$$

$$E_{ph} = 13.6 \text{ eV} \left(\frac{1}{j^2} - \frac{1}{i^2} \right), \text{ with } i > j \quad (1.2.6)$$

$$2\mathbf{p} \cdot \mathbf{r} = n\mathbf{l} \quad (1.2.7)$$

$$m \frac{v^2}{r} = \frac{q^2}{4\mathbf{p}\mathbf{e}_0 r^2} \quad (1.2.8)$$

$$a_0 = \frac{\mathbf{e}_0 h^2 n^2}{\mathbf{p} m_0 q^2} \quad (1.2.9)$$

$$E_n = -\frac{m_0 q^4}{8\mathbf{e}_0^2 h^2 n^2}, \text{ with } n=1, 2, \dots \quad (1.2.10)$$

$$V(r) = -\frac{q^2}{4\mathbf{p}\mathbf{e}_0 r} \quad (1.2.11)$$

$$E = K.E. + V = \frac{p^2}{2m} + V(x) \quad (1.2.12)$$

$$E\Psi = \frac{p^2}{2m}\Psi + V(x)\Psi \quad (1.2.13)$$

$$-\hbar^2 \frac{\nabla^2 \Psi}{\Psi x^2} = \hbar^2 k^2 \Psi = p^2 \Psi \quad \text{for } \Psi = e^{i(kx - \mathbf{w}t)} \quad (1.2.14)$$

$$-\frac{\hbar^2}{2m} \frac{d^2 \Psi(x)}{dx^2} + V(x) \Psi(x) = E \Psi(x) \quad (1.2.15)$$

$$E_n = \frac{\hbar^2}{2m^*} \left(\frac{n}{2L_x} \right)^2, \text{ with } n = 1, 2, \dots \quad (1.2.16)$$

1.3. Electromagnetic theory

$$\frac{dE(x)}{dx} = \frac{\mathbf{r}(x)}{\epsilon} \quad (1.3.1)$$

$$E(x_2) - E(x_1) = \int_{x_1}^{x_2} \frac{\mathbf{r}(x)}{\epsilon} dx \quad (1.3.2)$$

$$\vec{\nabla} \vec{E}(x, y, z) = \frac{\mathbf{r}(x, y, z)}{\epsilon} \quad (1.3.3)$$

$$\mathbf{E} \cdot \mathbf{A} = \frac{Q}{\epsilon} \quad (1.3.4)$$

$$\frac{d\mathbf{f}(x)}{dx} = -E(x) \quad (1.3.5)$$

$$\mathbf{f}(x_2) - \mathbf{f}(x_1) = - \int_{x_1}^{x_2} E(x) dx \quad (1.3.6)$$

$$\frac{d^2 \mathbf{f}(x)}{dx^2} = - \frac{\mathbf{r}(x)}{\epsilon} \quad (1.3.7)$$

$$\vec{\nabla} \mathbf{f}(x, y, z) = -\vec{E}(x, y, z) \quad (1.3.8)$$

$$\nabla^2 \mathbf{f}(x, y, z) = - \frac{\mathbf{r}(x, y, z)}{\epsilon} \quad (1.3.9)$$

1.4. Statistical Thermodynamics

$$dU = dQ + dW + \mathbf{m} dN \quad (1.4.1)$$

$$f(E) = \frac{1}{1 + e^{(E - E_F)/kT}} \quad (1.4.2)$$

Principles of Semiconductor Devices



How to use this electronic text

Toolbar

By pointing to the "Contents" tag positioned in the upper left corner, the toolbar emerges. It can be used to quickly gain access to the [table of contents](#) or the contents of any chapter or the appendix. The toolbar can be closed by clicking on the "Contents" tag.

Equation window

The top part of the screen is the [equation window](#). As equations are referenced in the text, they can be displayed in the equation window. The equation window can be [cleared](#) by clicking in the very upper left corner of the screen.

Icons and symbols in the text

Several navigation aids are provided in the text. Pointing at the icons/symbols provides a short description. A more detailed description is provided below:



Link to a print file in PDF format.

The print files enable to print a quality hardcopy. The user should have the Acrobat Reader Plug-in to view and print the files. [Get the plug-in](#)



Link to a QuickTime movie.

The movie files require the QuickTime plugin. [Get the plug-in](#)



Link to a spreadsheet in XLS format.

The spreadsheets are provided with the examples and a selected set of figures. The user should have a recent version of EXCEL or the EXCEL reader.



Link to previous chapter.

From Chapter 1 one reaches the Title page followed by the Appendix and Chapters 7, 6, 5, 4, 3 and 2 before returning to Chapter 1.

When viewing the examples, problems, review questions, bibliography and equation sheet of a particular chapter, one obtains the same item of the previous chapter.



Link to previous section.

When viewing the examples, problems, review questions, bibliography or equation sheet of a particular chapter, one obtains the previous item of the same chapter.



Link to next section.

When viewing the examples, problems, review questions, bibliography or equation sheet of a particular chapter, one obtains the next item of the same chapter.



Link to next chapter.

From Chapter 1 one reaches Chapters 2, 3, 4, 5, 6, 7 followed by the Appendix, the Title page before returning to Chapter 1.

When viewing the examples, problems, review questions, bibliography or equation sheet of a particular chapter, one obtains the same item of the next chapter.



Link to sub sections in the same file.

Except for when viewing an introductory section where it links to the examples, problems, review questions, bibliography and equation sheet of that chapter.



Back to introductory section of the corresponding chapter.

In addition, clicking on a figure or table will provide immediate access to the next figure or table in the same file.

Chapter 1: Review of Modern Physics



Problems

1. Calculate the wavelength of a photon with a photon energy of 2 eV. Also, calculate the wavelength of an electron with a kinetic energy of 2 eV. 
2. Consider a beam of light with a power of 1 Watt and a wavelength of 800 nm. Calculate a) the photon energy of the photons in the beam, b) the frequency of the light wave and c) the number of photons provided by the beam in one second. 
3. Show that the spectral density, u_{ω} (equation 1.2.4) peaks at $E_{\text{ph}} = 2.82 kT$. Note that a numeric iteration is required. 
4. Calculate the peak wavelength of blackbody radiation emitted from a human body at a temperature of 37°C. 
5. Derive equations (1.2.9) and (1.2.10). 
6. What is the width of an infinite quantum well if the second lowest energy of a free electron confined to the well equals 100 meV. 
7. Calculate the lowest three possible energies of an electron in a hydrogen atom in units of electron volt. 
8. Derive the electric field of a proton with charge q as a function of the distance from the proton using Gauss's law. Integrated the electric field to find the potential $\phi(r)$:

$$\mathbf{E}(r) = \frac{q}{4\pi\epsilon_0 r}$$

Treat the proton as a point charge and assume the potential to be zero far away from the proton. 

9. Prove that the probability of occupying an energy level below the Fermi energy equals the probability that an energy level above the Fermi energy and equally far away from the Fermi energy is not occupied. 

Chapter 7: MOS Field-Effect-Transistors



Bibliography

Chapter 1: Review of Modern Physics



Review Questions

1. List three experiments, which can only be explained using quantum mechanics.
2. What is a Rydberg?
3. Name the two primary assumptions of the Bohr model.
4. How do we know that the energy levels in a hydrogen atom are quantized?
5. What two parameters are linked by Gauss's law?
6. What two parameters are linked by Poisson's equation?
7. What is the definition of thermal equilibrium?
8. List the three laws of thermodynamics.
9. Explain in words the meaning of the thermodynamic identity.
10. What is the Fermi function?

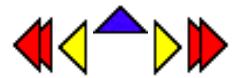
Bibliography

1. Quantum Mechanics, H. Kroemer, Prentice Hall, 1994.
2. Electrons in Solids, Third edition, R.H. Bube, Academic Press, 1992.
3. Solid State Electronic Devices, Fifth edition, B. G. Streetman, Prentice Hall, 2000.
4. Quantum Mechanics, Second edition, E. Merzbacher, Wiley, 1970.
5. Quantum Physics of Electronics, S.N. Levine, The Macmillan company, 1965.
6. Thermal Physics, Second edition, C. Kittel and H. Kroemer, Freeman, 1980.

Bibliography

1. Physics of Semiconductor Devices, Second edition, S. M. Sze, Wiley & Sons, 1981, Chapter 5.
2. Device Electronics for Integrated Circuits, Second edition, R.S. Muller and T. I. Kamins, Wiley & Sons, 1986, Chapter 3.

Chapter 6: MOS Capacitors



Bibliography

Chapter 6: MOS Capacitors

6.3. MOS analysis

$$V_{FB} = \Phi_M - \Phi_S \quad (6.3.1)$$

$$\Phi_M - \Phi_S = \Phi_M - \mathbf{c} - \frac{E_g}{q} - V_t \ln\left(\frac{N_a}{n_i}\right) \quad (6.3.2)$$

$$\Phi_{poly,S} = \Phi_{poly} - \Phi_S = V_t \ln\left(\frac{N_{a,poly}}{N_a}\right) \quad \text{p - type polysilico n gate} \quad (6.3.3)$$

$$\Phi_{poly,S} = \Phi_{poly} - \Phi_S = V_t \ln\left(\frac{n_i^2}{N_{d,poly} N_a}\right) \quad \text{n - type polysilico n gate}$$

$$V_{FB} = \Phi_{MS} - \frac{Q_i}{C_{ox}} - \frac{1}{\epsilon_{ox}} \int_0^{t_{ox}} \mathbf{r}_{ox}(x) x dx \quad (6.3.4)$$

$$\begin{aligned} Q_{inv} &= C_{ox}(V_G - V_T) & \text{for } V_G > V_T \\ Q_{inv} &= 0 & V_G \leq V_T \end{aligned} \quad (6.3.5)$$

$$Q_d = -qN_a x_d \quad (6.3.6)$$

$$E_s = \frac{qN_a x_d}{\epsilon_s} \text{ and } E_{ox} = \frac{qN_a x_d}{\epsilon_{ox}} \quad (6.3.7)$$

$$\mathbf{f}_s = \frac{qN_a x_d^2}{2\epsilon_s} \quad (6.3.8)$$

$$\mathbf{f}_F = V_t \ln \frac{N_a}{n_i} \quad (6.3.9)$$

$$x_d = \sqrt{\frac{2\epsilon_s \mathbf{f}_s}{qN_a}}, \text{ for } 0 \leq \mathbf{f}_s \leq 2\mathbf{f}_F \quad (6.3.10)$$

$$Q_{d,T} = -qN_a x_{d,T} \quad (6.3.11)$$

$$x_{d,T} = \sqrt{\frac{2\epsilon_s (2\mathbf{f}_F)}{qN_a}} \quad (6.3.12)$$

$$Q_M = -(Q_d + Q_{inv}) \quad (6.3.13)$$

$$V_G = V_{FB} + \mathbf{f}_s + \frac{Q_M}{C_{ox}} = V_{FB} + \mathbf{f}_s - \frac{Q_d + Q_{inv}}{C_{ox}} \quad (6.3.14)$$

$$V_G = V_{FB} + \mathbf{f}_s + \frac{\sqrt{2\mathbf{e}_s q N_a \mathbf{f}_s}}{C_{ox}}, \text{ for } 0 \leq \mathbf{f}_s \leq 2\mathbf{f}_F \quad (6.3.15)$$

$$V_G = V_{FB} + 2\mathbf{f}_F + \frac{\sqrt{4\mathbf{e}_s q N_a \mathbf{f}_F}}{C_{ox}} - \frac{Q_{inv}}{C_{ox}} \stackrel{\Delta}{=} V_T - \frac{Q_{inv}}{C_{ox}} \quad (6.3.16)$$

$$V_T = V_{FB} + 2\mathbf{f}_F + \frac{\sqrt{4\mathbf{e}_s q N_a \mathbf{f}_F}}{C_{ox}} \quad (6.3.17)$$

$$C_{LF} = C_{HF} = C_{ox}, \text{ for } V_G \leq V_{FB} \quad (6.3.18)$$

$$C_{LF} = C_{HF} = \frac{1}{\frac{1}{C_{ox}} + \frac{x_d}{\mathbf{e}_s}}, \text{ for } V_{FB} \leq V_G \leq V_T \quad (6.3.19)$$

$$x_d = \sqrt{\frac{2\mathbf{e}_s \mathbf{f}_s}{q N_d}} \quad (6.3.20)$$

$$V_G = V_{FB} + \mathbf{f}_s + \frac{\sqrt{2\mathbf{e}_s q N_a \mathbf{f}_s}}{C_{ox}}, \text{ for } 0 \leq \mathbf{f}_s \leq 2\mathbf{f}_F \quad (6.3.15)$$

$$C_{LF} = C_{ox} \text{ and } C_{HF} = \frac{1}{\frac{1}{C_{ox}} + \frac{x_{d,T}}{\mathbf{e}_s}}, \text{ for } V_G \geq V_T \quad (6.3.21)$$

$$\frac{d^2\mathbf{f}}{dx^2} = \frac{q}{\mathbf{e}_s} (N_a^+ - p) = \frac{q N_a}{\mathbf{e}_s} (1 - e^{-\frac{f}{V_t}}) \equiv \frac{q N_a}{\mathbf{e}_s} \frac{\mathbf{f}}{V_t} \quad (6.3.22)$$

$$\mathbf{f} = \mathbf{f}_s e^{-\frac{x}{L_D}}, \text{ with } L_D = \sqrt{\frac{\mathbf{e}_s V_t}{q N_a}} \quad (6.3.23)$$

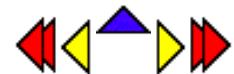
$$C_{s,FB} = \frac{dQ_s}{d\mathbf{f}_s} = \frac{d(\mathbf{e}_s \frac{\mathbf{f}_s}{L_D})}{d\mathbf{f}_s} = \frac{\mathbf{e}_s}{L_D} \quad (6.3.24)$$

$$C_{FB} = \frac{1}{\frac{1}{C_{ox}} + \frac{L_D}{\mathbf{e}_s}} \quad (6.3.25)$$

$$time = \frac{|\mathbf{Q}_{inv}|}{qG(x_{d,dd} + L_n)} = \frac{C_{ox}(V_G - V_T)}{\frac{qn_i}{2\mathbf{t}} \left(\sqrt{\frac{2\mathbf{e}_s \mathbf{f}_{s,dd}}{q N_a}} + \sqrt{\mathbf{m}_n V_t \mathbf{t}_n} \right)} \quad (6.3.26)$$

$$\frac{dV_G}{dt} > \frac{qn_i}{2C_{ox}} \sqrt{\frac{\mathbf{m}_n V_t}{\mathbf{t}_n}} \quad (6.3.27)$$

Chapter 6: MOS Capacitors



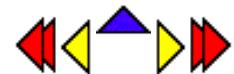
Review Questions

1. Draw an MOS flatband diagram. Indicate the workfunction of the metal and the semiconductor as well as the flatband voltage. Draw it approximately to scale using $\Phi_M = 4.1$ V, $\chi = 4.05$ V, $E_g = 1.12$ eV (silicon) and $N_a = 10^{16}$ cm⁻³.
2. Derive the metal-semiconductor workfunction for n-type and p-type poly-silicon gate structures. (equation [6.3.2](#))
3. Explain why the flatband voltage depends on the charge in the oxide or at the oxide-semiconductor interface.
4. Name the three bias regimes of an MOS capacitor and explain what happens in the semiconductor in each of these bias modes.
5. What is the basic assumption regarding the charge in the inversion layer?
6. What are the assumptions of the MOS capacitor analysis?
7. What is the difference between the high frequency and quasi-static capacitance?
8. Why is the high frequency capacitance constant in inversion?
9. Why does the flatband capacitance not equal the oxide capacitance?
10. What is deep depletion?
11. Why does light illumination affect the capacitance of an MOS structure?
12. Name the non-ideal effects in MOS capacitors. What causes them and how do they affect the MOS characteristics?

Review Questions

1. Describe the motion of electrons and holes in a pnp bipolar transistor biased in the forward active mode with $V_{BC} = 0$.
2. What is the definition of the emitter efficiency? Explain in words and provide the corresponding equation.
3. What is the definition of the base transport factor? Explain in words and provide the corresponding equation.
4. Derive the relation between the current gain and the transport factor.
5. How does recombination in the quasi-neutral base region affect the emitter, base and collector current?
6. How does recombination in the base-emitter depletion region affect the emitter, base and collector current?
7. Explain the four different bias modes of a bipolar transistor.
8. Explain why a transistor can have a current gain larger than one in the common emitter mode. Provide the necessary and sufficient conditions needed to obtain a current gain larger than one.
9. What is the Early effect and how does it affect the transistor characteristics?

Chapter 7: MOS Field-Effect-Transistors



Review Questions

1. What is the difference between a source and a drain of a MOSFET?
2. What is the difference between an n-type and a p-type MOSFET?
3. What is the difference between an enhancement and a depletion MOSFET?
4. Which device has most positive threshold voltage a depletion mode p-type MOSFET or an enhancement mode p-type MOSFET?
5. Which device has the highest drain current in saturation at zero gate voltage, a p-type enhancement MOSFET or an n-type depletion MOSFET?
6. Why is the electron layer of an n-MOSFET called an inversion layer?
7. Why is there no current in a MOSFET when the device is biased in accumulation?
8. What is the difference between the linear and the quadratic MOSFET model?
9. Does the body effect affect CMOS circuits? Explain.
10. How does the oxide thickness and substrate doping affect the threshold voltage of a MOSFET. Plot the threshold voltage as a function of the oxide thickness for different doping densities.
11. What is the advantage of a poly-silicon gate technology?
12. What is channel length modulation?
13. What is punchthrough?
14. Explain the effect of scaling on the different MOSFET parameters as listed in Table [7.7.1](#)

Review Questions

1. Draw an MOS flatband diagram. Indicate the workfunction of the metal and the semiconductor as well as the flatband voltage. Draw it approximately to scale using $\Phi_M = 4.1$ V, $c = 4.05$ V, $E_g = 1.12$ eV (silicon) and $N_a = 10^{16}$ cm⁻³.
2. Derive the metal-semiconductor workfunction for n-type and p-type poly-silicon gate structures. (equation 6.3.2)
3. Explain why the flatband voltage depends on the charge in the oxide or at the oxide-semiconductor interface.
4. Name the three bias regimes of an MOS capacitor and explain what happens in the semiconductor in each of these bias modes.
5. What is the basic assumption regarding the charge in the inversion layer?
6. What are the assumptions of the MOS capacitor analysis?
7. What is the difference between the high frequency and quasi-static capacitance?
8. Why is the high frequency capacitance constant in inversion?
9. Why does the flatband capacitance not equal the oxide capacitance?
10. What is deep depletion?
11. Why does light illumination affect the capacitance of an MOS structure?
12. Name the non-ideal effects in MOS capacitors. What causes them and how do they affect the MOS characteristics?

Chapter 7: MOS Field-Effect-Transistors



Problems



1. Consider an n-type MOSFET, which consists of a 10 nm thick oxide ($\epsilon_r = 3.9$) and has a gate length of 1 micron, a gate width of 20 micron and a threshold voltage of 1.5 Volt. Calculate the resistance of the MOSFET in the linear region as measured between source and drain when applying a gate-source voltage of 3 Volt. What should the gate-source voltage be to double the resistance? The surface mobility of the electrons is $300 \text{ cm}^2/\text{V}\cdot\text{sec}$.
2. Consider an n-type MOSFET with an oxide thickness $t_{\text{ox}} = 20 \text{ nm}$ ($\epsilon_r = 3.9$) and a gate length, $L = 1 \text{ micron}$, a gate width, $W = 10 \text{ micron}$ and a threshold voltage, $V_T = 1 \text{ Volt}$. Calculate the capacitance per unit area of the oxide, C_{OX} , and from it the capacitance of the gate, C_G . Calculate the drain current, I_D , at a gate-source voltage, $V_{\text{GS}} = 3 \text{ Volt}$ and a drain-source voltage, $V_{\text{DS}} = 0.05 \text{ Volt}$. The surface mobility of the electrons $\mu_n = 300 \text{ cm}^2/\text{V}\cdot\text{sec}$. Use the linear model of the MOSFET.
3. A MOSFET ($L = 1 \mu\text{m}$, $t_{\text{ox}} = 15 \text{ nm}$, $V_T = 1 \text{ V}$ and $\mu_n = 300 \text{ cm}^2/\text{V}\cdot\text{sec}$) must provide a current of 20 mA at a drain-source voltage of 0.5 Volt and a gate-source voltage of 5 Volt. How wide should the gate be?
4. A MOSFET ($L = 1 \mu\text{m}$, $t_{\text{ox}} = 10 \text{ nm}$, $V_T = 1 \text{ V}$ and $\mu_n = 300 \text{ cm}^2/\text{V}\cdot\text{sec}$) is to be used as 50 W terminating resistor when applying a gate-source voltage, $V_{\text{GS}} = 5 \text{ Volt}$. How wide should the gate be?
5. The capacitance of an n-type silicon MOSFET is 1 pF. Provided that the oxide thickness is 50 nm and the gatelength is 1 micron, what is the resistance of the MOSFET in the linear regime when biased at a gate voltage which is 5 Volt larger than the threshold voltage? Use a reasonable value for the surface mobility knowing that the bulk mobility equals $1400 \text{ cm}^2/\text{V}\cdot\text{sec}$.
6. Consider a p-channel silicon MOSFET with an aluminum gate.
 - a. Draw the energy band diagram of the MOS structure for $V_G = V_{\text{FB}}$. Indicate the workfunction of the metal and the semiconductor, as well as the electron affinity.
 - b. Draw the field distribution for $V_G = V_T$ (onset of inversion).
 - c. Calculate the depletion layer width and the field in the oxide at the onset of inversion. ($N_d = 10^{16} \text{ cm}^{-3}$, $t_{\text{ox}} = 100 \text{ nm}$, $V_{\text{FB}} = -0.5 \text{ V}$)
7. Calculate the depletion region width within a p-type bulk silicon MOS capacitor with $N_d = 10^{17} \text{ cm}^{-3}$, at the onset of inversion.
8. A silicon p-substrate ($N_a = 10^{16} \text{ cm}^{-3}$) MOSFET with $t_{\text{ox}} = 0.1 \mu\text{m}$, $\epsilon_{\text{ox}}/\epsilon_0 = 3.9$ and a negative interface charge per unit area of -10^{-8} C/cm^2 , has a threshold voltage which is 1 Volt smaller than desired. By what value should one change the oxide thickness to obtain the desired threshold voltage? Should one increase or decrease the oxide thickness?
9. A silicon MOSFET ($n_I = 10^{10} \text{ cm}^{-3}$, $\epsilon_s/\epsilon_0 = 11.9$ and $\epsilon_{\text{ox}}/\epsilon_0 = 3.9$) is scaled by reducing all dimensions by a factor of 2 and by increasing the doping density of the substrate by a factor of 4.

Calculate the ratio of the following parameters of the scaled device relative to that of the original device: (make approximations if necessary)

- The transconductance at $V_{GS} - V_T = 1$ V.
 - The gate capacitance
 - The transit frequency at
 - $V_{GS} - V_T = 1$ V. (Assume that $C_{DS} = 0$)
10. The threshold shift when increasing the reverse bias of the source-bulk diode from 1 Volt to 3 Volt.
 11. The breakdown voltage of the oxide assuming the breakdown field to be constant.
 12. The breakdown voltage of the drain-to-bulk p-n diode assuming the breakdown field to be constant.
13. A silicon p-substrate ($N_a = 10^{16} \text{ cm}^{-3}$) MOSFET with $t_{ox} = 0.1 \mu\text{m}$, $\epsilon_{ox}/\epsilon_0 = 3.9$ and $V_{FB} = -0.2$ V, has a threshold voltage which is 1 Volt smaller than desired. By what value should one change the oxide thickness, t_{ox} , to obtain the desired threshold voltage? Should one increase or decrease the oxide thickness?

Problems

1. Consider an n-type MOSFET, which consists of a 10 nm thick oxide ($\epsilon_r = 3.9$) and has a gate length of 1 micron, a gate width of 20 micron and a threshold voltage of 1.5 Volt. Calculate the resistance of the MOSFET in the linear region as measured between source and drain when applying a gate-source voltage of 3 Volt. What should the gate-source voltage be to double the resistance? The surface mobility of the electrons is $300 \text{ cm}^2/\text{V}\cdot\text{sec}$.
2. Consider an n-type MOSFET with an oxide thickness $t_{ox} = 20 \text{ nm}$ ($\epsilon_r = 3.9$) and a gate length, $L = 1 \text{ micron}$, a gate width, $W = 10 \text{ micron}$ and a threshold voltage, $V_T = 1 \text{ Volt}$. Calculate the capacitance per unit area of the oxide, C_{ox} , and from it the capacitance of the gate, C_G . Calculate the drain current, I_D , at a gate-source voltage, $V_{GS} = 3 \text{ Volt}$ and a drain-source voltage, $V_{DS} = 0.05 \text{ Volt}$. The surface mobility of the electrons $\mu_n = 300 \text{ cm}^2/\text{V}\cdot\text{sec}$. Use the linear model of the MOSFET.
3. A MOSFET ($L = 1 \mu\text{m}$, $t_{ox} = 15 \text{ nm}$, $V_T = 1 \text{ V}$ and $\mu_n = 300 \text{ cm}^2/\text{V}\cdot\text{sec}$) must provide a current of 20 mA at a drain-source voltage of 0.5 Volt and a gate-source voltage of 5 Volt. How wide should the gate be?
4. A MOSFET ($L = 1 \mu\text{m}$, $t_{ox} = 10 \text{ nm}$, $V_T = 1 \text{ V}$ and $\mu_n = 300 \text{ cm}^2/\text{V}\cdot\text{sec}$) is to be used as 50Ω terminating resistor when applying a gate-source voltage, $V_{GS} = 5 \text{ Volt}$. How wide should the gate be?
5. The capacitance of an n-type silicon MOSFET is 1 pF. Provided that the oxide thickness is 50 nm and the gate length is 1 micron, what is the resistance of the MOSFET in the linear regime when biased at a gate voltage which is 5 Volt larger than the threshold voltage? Use a reasonable value for the surface mobility knowing that the bulk mobility equals $1400 \text{ cm}^2/\text{V}\cdot\text{sec}$.
6. Consider a p-channel silicon MOSFET with an aluminum gate.
 - a) Draw the energy band diagram of the MOS structure for $V_G = V_{FB}$. Indicate the workfunction of the metal and the semiconductor, as well as the electron affinity.
 - b) Draw the field distribution for $V_G = V_T$ (onset of inversion).
 - c) Calculate the depletion layer width and the field in the oxide at the onset of inversion. ($N_d = 10^{16} \text{ cm}^{-3}$, $t_{ox} = 100 \text{ nm}$, $V_{FB} = -0.5 \text{ V}$)
7. Calculate the depletion region width within a p-type bulk silicon MOS- capacitor with $N_d = 10^{17} \text{ cm}^{-3}$, at the onset of inversion.
8. A silicon p-substrate ($\psi \equiv N_d = 10^{16} \text{ cm}^{-3}$) MOSFET with $t_{ox} = 0.1 \mu\text{m}$, $\epsilon_{ox}/\epsilon_0 = 3.9$ and a negative interface charge per unit area of -10^{-8} C/cm^2 , has a threshold voltage which is 1 Volt smaller than desired. By what value should one change the oxide thickness to obtain the desired threshold voltage? Should one increase or decrease the oxide thickness?
9. A silicon MOSFET ($n_i = 10^{10} \text{ cm}^{-3}$, $\epsilon_s/\epsilon_0 = 11.9$ and $\epsilon_{ox}/\epsilon_0 = 3.9$) is scaled by reducing all dimensions by a factor of 2 and by increasing the doping density of the substrate by a factor of 4.

Calculate the ratio of the following parameters of the scaled device relative to that of the original device: (make approximations if necessary)

- The transconductance at $V_{GS} - V_T = 1$ V.
 - The gate capacitance
 - The transit frequency at $V_{GS} - V_T = 1$ V. (Assume that $C_{DS} = 0$)
 - The threshold shift when increasing the reverse bias of the source-bulk diode from 1 Volt to 3 Volt.
 - The breakdown voltage of the oxide assuming the breakdown field to be constant.
 - The breakdown voltage of the drain-to-bulk p-n diode assuming the breakdown field to be constant.
10. A silicon p-substrate ($p \equiv N_a = 10^{16} \text{ cm}^{-3}$) MOSFET with $t_{ox} = 0.1 \mu\text{m}$, $\epsilon_{ox}/\epsilon_0 = 3.9$ and $V_{FB} = -0.2$ V, has a threshold voltage which is 1 Volt smaller than desired. By what value should one change the oxide thickness, t_{ox} , to obtain the desired threshold voltage? Should one increase or decrease the oxide thickness?

Problems

1. Consider an aluminum-SiO₂-silicon MOS capacitor ($F_M = 4.1$ V, $\epsilon_{ox}/\epsilon_0 = 3.9$, $\chi = 4.05$ V and $N_a = 10^{17}$ cm⁻³) MOS capacitor with $t_{ox} = 5$ nm.
 - a) Calculate the flatband voltage and threshold voltage.
 - b) Repeat for an n-type silicon substrate with $N_d = 10^{16}$ cm⁻³.
 - c) Repeat with a surface charge of 10^{-7} C/cm²
 - d) Repeat with a charge density in the oxide of 10^{-1} C/cm³
2. A high-frequency capacitance voltage measurement of a silicon MOS structure was fitted by the following expression:
$$C(V_G) = 6 \text{ pF} + 12 \text{ pF}/(1 + \exp(V_G))$$
 - a) Calculate the oxide capacitance per unit area and the oxide thickness. The area of the capacitor is 100 x 100 micron and the relative dielectric constant equals 3.9.
 - b) From the minimum capacitance, calculate the maximum depletion layer width and the substrate doping density.
 - c) Calculate the bulk potential.
 - d) Calculate the flatband capacitance and the flatband voltage.
 - e) Calculate the threshold voltage.
3. An MOS capacitor with an oxide thickness of 20 nm has an oxide capacitance, which is three times larger than the minimum high-frequency capacitance in inversion. Find the substrate doping density.
4. A CMOS gate requires n-type and p-type MOS capacitors with a threshold voltage of 2 and -2 Volt respectively. If the gate oxide is 50 nm what are the required substrate doping densities? Assume the gate electrode is aluminum. Repeat for a p⁺ poly-silicon gate.
5. Consider a p-MOS capacitor (with an n-type substrate) and with an aluminum gate. Find the doping density for which the threshold voltage is 3 times larger than the flat band voltage. $t_{ox} = 25$ nm. Repeat for a capacitor with 10^{11} cm⁻² electronic charges at the oxide-semiconductor interface.
6. A silicon p-MOS capacitor ($N_d = 4 \times 10^{16}$ cm⁻³, $t_{ox} = 40$ nm) is biased halfway between the flatband and threshold voltage. Calculate the applied voltage and the corresponding capacitance

Example 7.1 Calculate the drain current of a silicon nMOSFET with $V_T = 1$ V, $W = 10$ μm , $L = 1$ μm and $t_{ox} = 20$ nm. The device is biased with $V_{GS} = 3$ V and $V_{DS} = 5$ V. Use the quadratic model, a surface mobility of $300 \text{ cm}^2/\text{V}\cdot\text{s}$ and set $V_{BS} = 0$ V. Also calculate the transconductance at $V_{GS} = 3$ V and $V_{DS} = 5$ V and compare it to the output conductance at $V_{GS} = 3$ V and $V_{DS} = 0$ V.

Solution The MOSFET is biased in saturation since $V_{DS} > V_{GS} - V_T$. Therefore the drain current equals:

$$I_D = m_n C_{ox} \frac{W}{L} \frac{(V_{GS} - V_T)^2}{2}$$

$$= 300 \times \frac{3.9 \times 8.85 \times 10^{-14}}{20 \times 10^{-7}} \frac{10}{1} \times \frac{(3-1)^2}{2} = 1.04 \text{ mA}$$

The transconductance equals:

$$g_m = m_n C_{ox} \frac{W}{L} (V_{GS} - V_T)$$

$$= 300 \times \frac{3.9 \times 8.85 \times 10^{-14}}{20 \times 10^{-7}} \frac{10}{1} \times (3-1) = 1.04 \text{ mS}$$

and the output conductance equals:

$$g_d = m_n C_{ox} \frac{W}{L} (V_{GS} - V_T - V_{DS})$$

$$= 300 \times \frac{3.9 \times 8.85 \times 10^{-14}}{20 \times 10^{-7}} \frac{10}{1} \times (3-1-0) = 1.04 \text{ mS}$$

Example 7.2 Repeat example 7.1 using the variable depletion layer model.
Use $V_{FB} = -0.807$ V and $N_a = 10^{17}$ cm⁻³.

Solution To find out whether the MOSFET is biased in saturation, one first calculates the saturation voltage, $V_{D,sat}$:

$$V_{DS,sat} = V_{GS} - V_{FB} - 2f_F$$

$$- \frac{qN_a e_s}{C_{ox}^2} \left\{ \sqrt{1 + 2 \frac{C_{ox}^2}{qN_a e_s} (V_{GB} - V_{FB})} - 1 \right\}$$

$$= 1.39 \text{ V}$$

The drain current is then obtained from:

$$I_D = \frac{m_n C_{ox} W}{L} (V_{GS} - V_{FB} - 2f_F - \frac{V_{DS,sat}}{2}) V_{DS,sat}$$

$$- \frac{2}{3} m_n \frac{W}{L} \sqrt{2e_s q N_a} ((2f_F + V_{DB})^{3/2} - (2f_F + V_{SB})^{3/2})$$

$$= 0.7 \text{ mA}$$

The transconductance equals:

$$g_{m,sat} = m_n C_{ox} \frac{W}{L} [V_{GS} - V_{FB} - 2f_F$$

$$- \frac{qN_a e_s}{C_{ox}^2} \left\{ \sqrt{1 + 2 \frac{C_{ox}^2}{qN_a e_s} (V_{GB} - V_{FB})} - 1 \right\}]$$

$$= 0.52 \text{ mS}$$

corresponding to a modified mobility $m_i^* = 149 \text{ cm}^2/\text{V}\cdot\text{s}$.

The output conductance at $V_{DS} = 0$ V equals:

$$g_d = \left. \frac{I_D}{V_{DS}} \right|_{V_{GS}} = 1.04 \text{ mS}$$

Which is the same as that of example 7.1 since the depletion layer width is constant for $V_{DS} = 0$.

Example 7.3 Calculate the threshold voltage of a silicon nMOSFET when applying a substrate voltage, $V_{BS} = 0, -2.5, -5, -7.5$ and -10 V. The capacitor has a substrate doping $N_a = 10^{17}$ cm $^{-3}$, a 20 nm thick oxide ($\epsilon_{ox} = 3.9 \epsilon_0$) and an aluminum gate ($\Phi_M = 4.1$ V). Assume there is no fixed charge in the oxide or at the oxide-silicon interface.

Solution The threshold voltage at $V_{BS} = -2.5$ V equals:

$$V_T = V_{T0} + \frac{g}{\sqrt{2f_F}} \left(\sqrt{1 + \frac{V_{SB}}{2f_F}} - 1 \right)$$

$$= -0.09 + \frac{0.75}{\sqrt{2 \times 0.42}} \left(\sqrt{1 + \frac{2.5}{2 \times 0.42}} - 1 \right) = 0.73 \text{ V}$$

Where the flatband voltage without substrate bias, V_{T0} , was already calculated in example 6.2. The body effect parameter was obtained from:

$$g = \frac{\sqrt{2\epsilon_s q N_a}}{C_{ox}} = \frac{\sqrt{2 \times 11.9 \times 8.85 \times 10^{-14} \times 1.6 \times 10^{-19} \times 10^{17}}}{3.9 \times 8.85 \times 10^{-14} / 20 \times 10^{-7}}$$

$$= 0.75 \text{ V}^{-1/2}$$

The threshold voltages for the different substrate voltages are listed in the table below.

	$V_{BS} = -2.5$ V	-5 V	-7.5 V	-10 V
V_T	0.73 V	1.26 V	1.68 V	2.04 V

Chapter 5: Bipolar Junction Transistors



5.2. Structure and principle of operation

A bipolar junction transistor consists of two back-to-back p-n junctions, who share a thin common region with width, w_B . Contacts are made to all three regions, the two outer regions called the emitter and collector and the middle region called the base. The structure of an NPN bipolar transistor is shown in Figure 5.2.1 (a). The device is called "bipolar" since its operation involves both types of mobile carriers, electrons and holes.

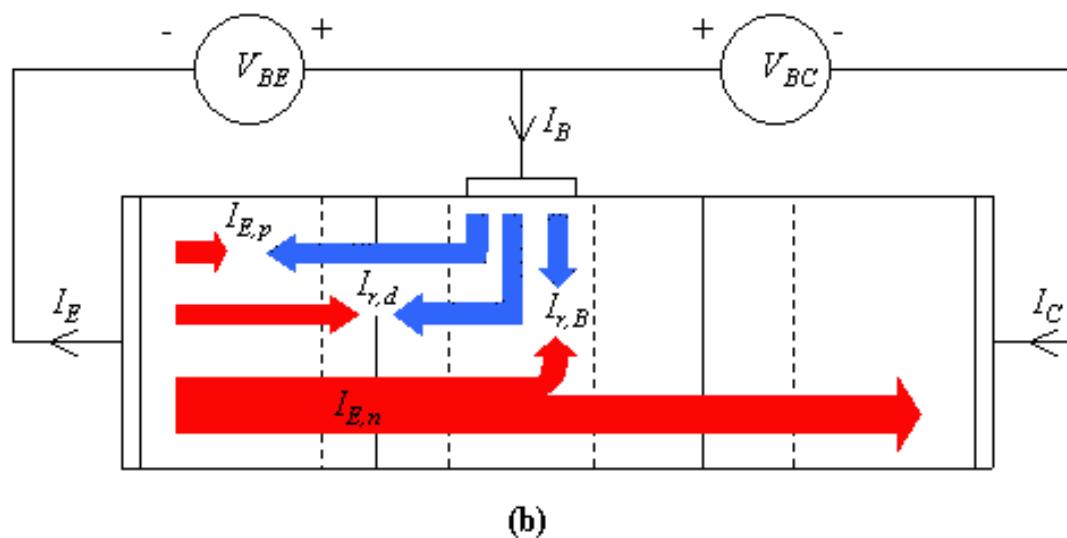
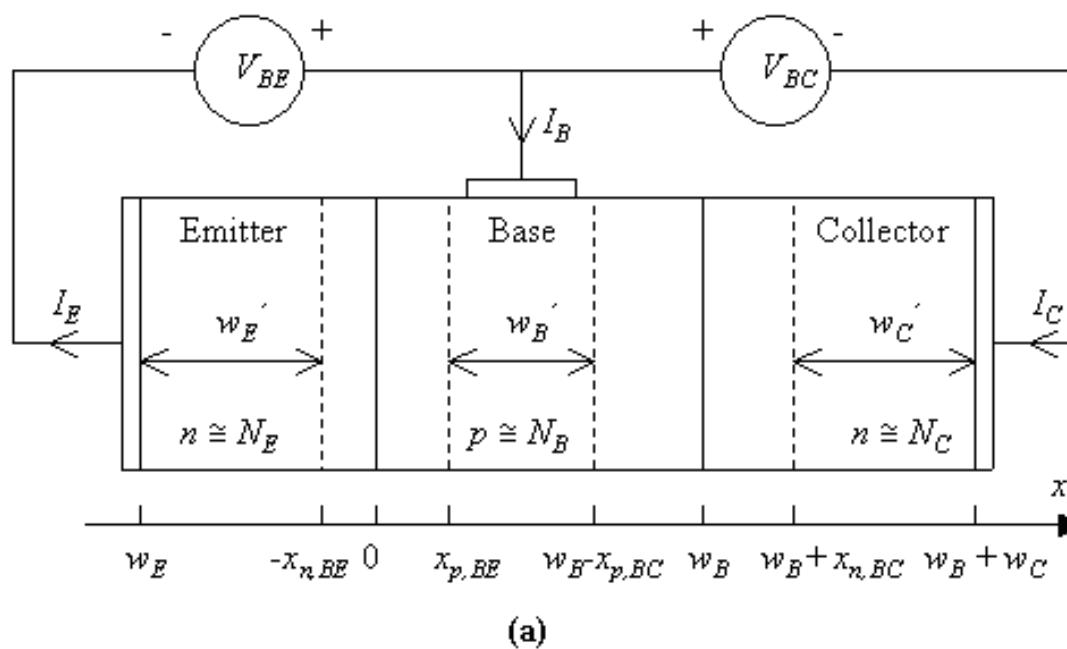


Figure 5.2.1.: (a) Structure and sign convention of an NPN bipolar junction transistor. (b) Electron and hole flow under forward active bias, $V_{BE} > 0$ and $V_{BC} = 0$.

Since the device consists of two back-to-back diodes, there are depletion regions between the quasi-neutral regions. The width of the quasi neutral regions in the emitter, base and collector are indicated with the symbols w_E' , w_B' and w_C' and are calculated from

$$w_E^+ = w_E - x_{n,BE} \quad (5.2.1)$$

$$w_B^+ = w_B - x_{p,BE} - x_{p,BC} \quad (5.2.2)$$

$$w_C^+ = w_C - x_{n,BC} \quad (5.2.3)$$

where the depletion region widths are given by:

$$x_{n,BE} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BE} - V_{BE})}{q}} \frac{N_B}{N_E} \left(\frac{1}{N_B + N_E} \right) \quad (5.2.4)$$

$$x_{p,BE} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BE} - V_{BE})}{q}} \frac{N_E}{N_B} \left(\frac{1}{N_B + N_E} \right) \quad (5.2.5)$$

$$x_{p,BC} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BC} - V_{BC})}{q}} \frac{N_C}{N_B} \left(\frac{1}{N_B + N_C} \right) \quad (5.2.6)$$

$$x_{n,BC} = \sqrt{\frac{2\varepsilon_s(\phi_{i,BC} - V_{BC})}{q}} \frac{N_B}{N_C} \left(\frac{1}{N_B + N_C} \right) \quad (5.2.7)$$

with

$$\phi_{i,BE} = V_t \ln \frac{N_E N_B}{n_i^2} \quad (5.2.8)$$

$$\phi_{i,BC} = V_t \ln \frac{N_C N_B}{n_i^2} \quad (5.2.9)$$

The sign convention of the currents and voltage is indicated on Figure 5.2.1(a). The base and collector current are positive if a positive current goes into the base or collector contact. The emitter current is positive for a current coming out of the emitter contact. This also implies the emitter current, I_E , equals the sum of the base current, I_B , and the collector current, I_C :

$$I_E = I_C + I_B \quad (5.2.10)$$

The base-emitter voltage and the base-collector voltage are positive if a positive voltage is applied to the base contact relative to the emitter and collector respectively.

The operation of the device is illustrated with Figure 5.2.1 (b). We consider here only the forward active bias mode of operation, obtained by forward biasing the base-emitter junction and reverse biasing the base-collector junction. To simplify the discussion further, we also set $V_{CE} = 0$. The corresponding energy band diagram is shown in Figure 5.2.2. Electrons diffuse from the emitter into the base and holes diffuse from the base into the emitter. This carrier diffusion is identical to that in a p-n junction. However, what is different is that the electrons can diffuse as minority carriers through the quasi-neutral region in the base. Once the electrons arrive at the base-collector depletion region, they are swept through the depletion layer due to the electric field. These electrons contribute to the collector current. In addition, there are two more currents, the base recombination current, indicated on Figure 5.2.2 by the vertical arrow, and the base-emitter depletion layer recombination current (not shown).

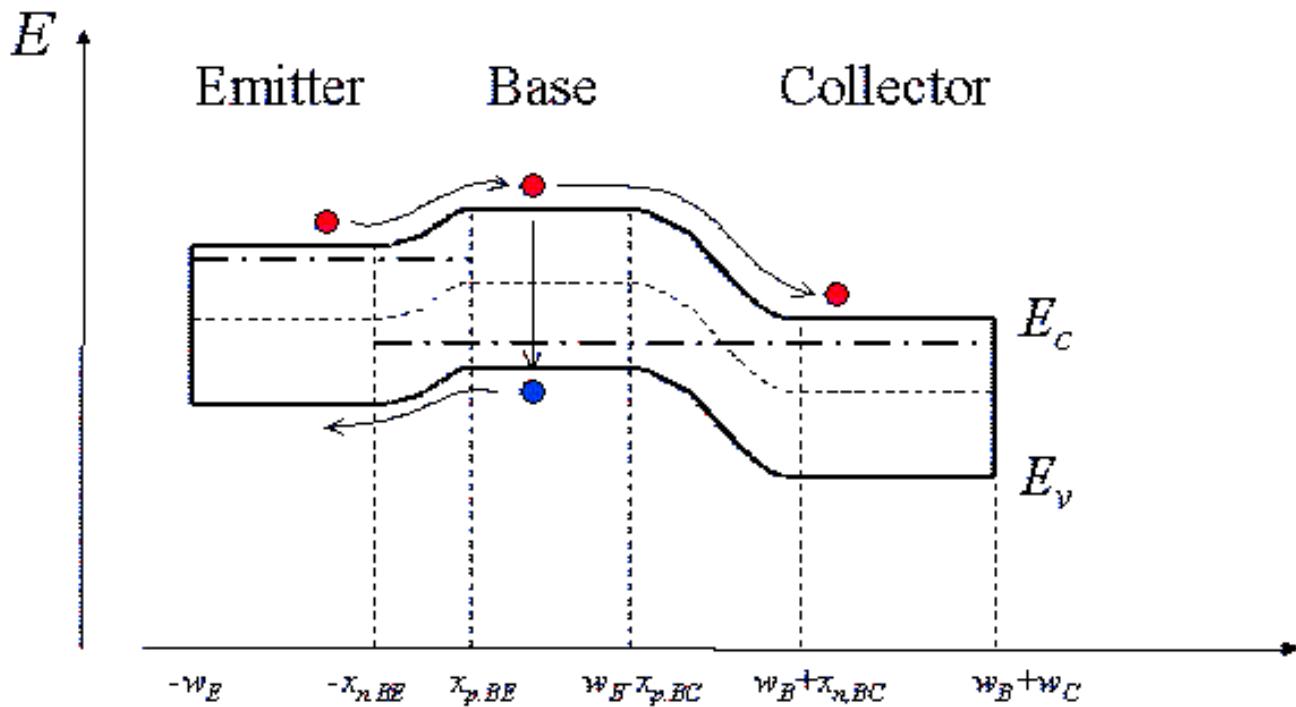


Figure 5.2.2. : Energy band diagram of a bipolar transistor biased in the forward active mode.

The total emitter current is the sum of the electron diffusion current, $I_{E,n}$, the hole diffusion current, $I_{E,p}$ and the base-emitter depletion layer recombination current, $I_{r,d}$.

$$I_E = I_{E,n} + I_{E,p} + I_{r,d} \quad (5.2.11)$$

The total collector current is the electron diffusion current, $I_{E,n}$, minus the base recombination current, $I_{r,B}$.

$$I_C = I_{E,n} - I_{r,B} \quad (5.2.12)$$

The base current is the sum of the hole diffusion current, $I_{E,p}$, the base recombination current, $I_{r,B}$ and the base-emitter depletion layer recombination current, $I_{r,d}$.

$$I_B = I_{E,p} + I_{r,B} + I_{r,d} \quad (5.2.13)$$

The transport factor, α , is defined as the ratio of the collector and emitter current:

$$\alpha = \frac{I_C}{I_E} \quad (5.2.14)$$

Using Kirchoff's current law and the sign convention shown in Figure 5.2.1(a), we find that the base current equals the difference between the emitter and collector current. The current gain, β , is defined as the ratio of the collector and base current and equals:

$$\beta = \frac{I_C}{I_B} = \frac{\alpha}{1 - \alpha} \quad (5.2.15)$$

This explains how a bipolar junction transistor can provide current amplification. If the collector current is almost equal to the emitter current, the transport factor, α , approaches one. The current gain, β , can therefore become much larger than one.

To facilitate further analysis, we now rewrite the transport factor, α , as the product of the emitter efficiency, γ_E , the base transport factor, α_T , and the depletion layer recombination factor, δ_r .

$$\alpha = \alpha_T \gamma_E \delta_r \quad (5.2.16)$$

The emitter efficiency, γ_E , is defined as the ratio of the electron current in the emitter, $I_{E,n}$, to the sum of the electron and hole current diffusing across the base-emitter junction, $I_{E,n} + I_{E,p}$.

$$\gamma_E = \frac{I_{E,n}}{I_{E,n} + I_{E,p}} \quad (5.2.17)$$

The base transport factor, α_T , equals the ratio of the current due to electrons injected in the collector, to the current due to electrons injected in the base.

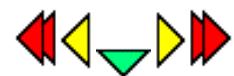
$$\alpha_T = \frac{I_{E,n} - I_{r,B}}{I_{E,n}} \quad (5.2.18)$$

Recombination in the depletion-region of the base-emitter junction further reduces the current gain, as it increases the emitter current without increasing the collector current. The depletion layer recombination factor, δ_r , equals the ratio of the current due to electron and hole diffusion across the base-emitter junction to the total emitter current:

$$\delta_r = \frac{I_E - I_{r,d}}{I_E} \quad (5.2.19)$$

Example 5.1	A bipolar transistor with an emitter current of 1 mA has an emitter efficiency of 0.99, a base transport factor of 0.995 and a depletion layer recombination factor of 0.998. Calculate the base current, the collector current, the transport factor and the current gain of the transistor.
Solution	<p>The transport factor and current gain are:</p> $\alpha = \gamma_E \alpha_T \delta_r = 0.99 \times 0.995 \times 0.998 = 0.983$ <p>and</p> $\beta = \frac{\alpha}{1 - \alpha} = 58.1$ <p>The collector current then equals</p> $I_C = \alpha I_E = 0.983 \text{ mA}$ <p>And the base current is obtained from:</p> $I_B = I_E - I_C = 17 \mu\text{A}$

Chapter 5: Bipolar Junction Transistors



5.3. Ideal transistor model

- [5.3.1. Forward active mode of operation](#)
- [5.3.2. General bias modes of a bipolar transistor](#)
- [5.3.3. The Ebers-Moll model](#)
- [5.3.4. Saturation.](#)

The ideal transistor model is based on the ideal p-n diode model and provides a first-order calculation of the dc parameters of a bipolar junction transistor. To further simplify this model, we will assume that all quasi-neutral regions in the device are much smaller than the minority-carrier diffusion lengths in these regions, so that the "short" diode expressions apply. The use of the ideal p-n diode model implies that no recombination within the depletion regions is taken into account. Such recombination current will be discussed in section [5.4.3](#).

The discussion of the ideal transistor starts with a discussion of the forward active mode of operation, followed by a general description of the four different bias modes, the corresponding Ebers-Moll model and a calculation of the collector-emitter voltage when the device is biased in saturation.

5.3.1. Forward active mode of operation



The forward active mode is obtained by forward-biasing the base-emitter junction. In addition we eliminate the base-collector junction current by setting $V_{BC} = 0$. The minority-carrier distribution in the quasi-neutral regions of the bipolar transistor, as shown in Figure [5.3.1](#), is used to analyze this situation in more detail.

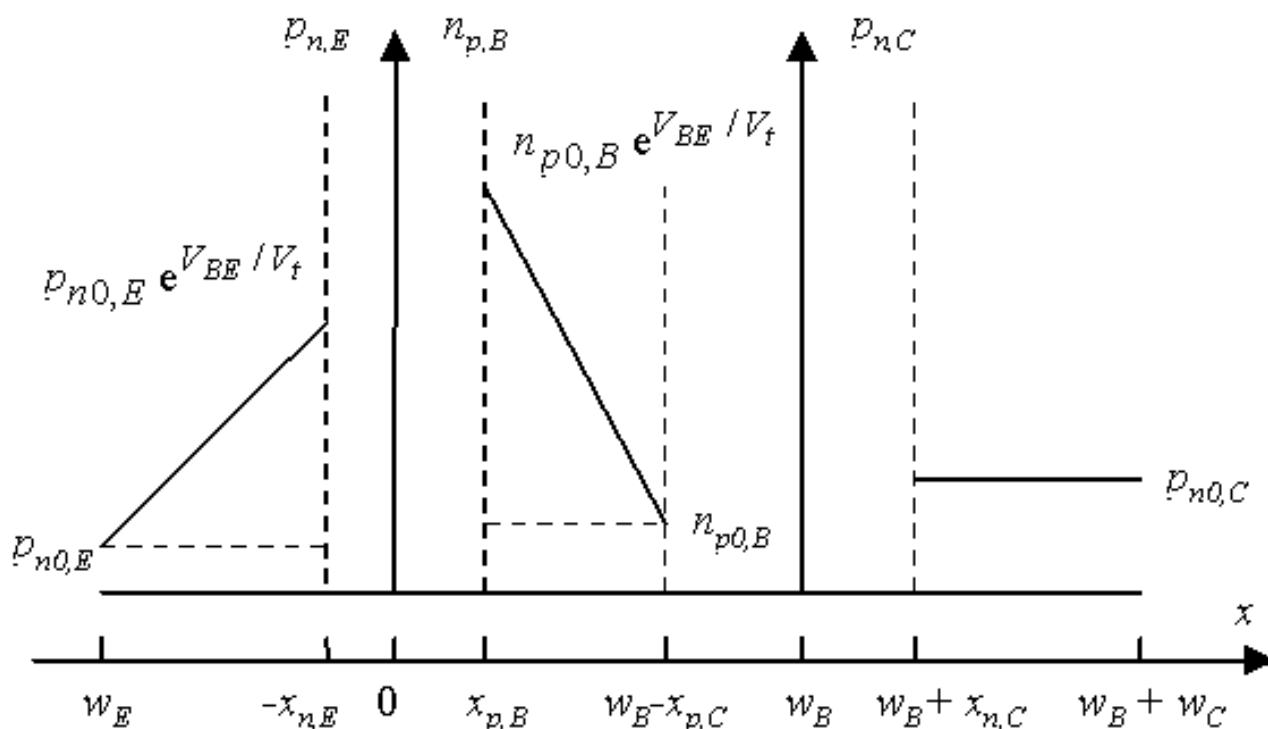


Figure 5.3.1. : Minority-carrier distribution in the quasi-neutral regions of a bipolar transistor (a) Forward active bias mode. (b) Saturation mode.

The values of the minority carrier densities at the edges of the depletion regions are indicated on the Figure 5.3.1. The carrier densities vary linearly between the boundary values as expected when using the assumption that no significant recombination takes place in the quasi-neutral regions. The minority carrier densities on both sides of the base-collector depletion region equal the thermal equilibrium values since V_{BC} was set to zero. While this boundary condition is mathematically equivalent to that of an ideal contact, there is an important difference. The minority carriers arriving at $x = w_B - x_{p,C}$ do not recombine. Instead, they drift through the base-collector depletion region and end up as majority carriers in the collector region.

The emitter current due to electrons and holes are obtained using the "short" diode expressions derived in section 4.4.2.5, yielding:

$$I_{E,n} = q n_i^2 A_E \left(\frac{D_{n,B}}{N_B w_B} \right) \left(\exp\left(\frac{V_{BE}}{V_t}\right) - 1 \right) \quad (5.3.1)$$

and

$$I_{E,p} = q n_i^2 A_E \left(\frac{D_{p,E}}{N_E w_E} \right) \left(\exp\left(\frac{V_{BE}}{V_t}\right) - 1 \right) \quad (5.3.2)$$

It is convenient to rewrite the emitter current due to electrons, $I_{E,n}$, as a function of the total excess minority charge in the base, $\Delta Q_{n,B}$. This charge is proportional to the triangular area in the quasi-neutral base as shown in Figure 5.3.1 a) and is calculated from:

$$\Delta Q_{n,B} = q A_E \int_{x_{p,E}}^{w_B - x_{p,C}} (n_p(x) - n_{p0}) dx \quad (5.3.3)$$

which for a "short" diode becomes:

$$\Delta Q_{n,B} = q A_E \frac{n_i^2}{N_B} \left(\exp\left(\frac{V_{BE}}{V_t}\right) - 1 \right) \frac{w_B}{2} \quad (5.3.4)$$

And the emitter current due to electrons, $I_{E,n}$, simplifies to:

$$I_{E,n} = \frac{\Delta Q_{n,B}}{t_r} \quad (5.3.5)$$

where t_r is the average time the minority carriers spend in the base layer, i.e. the transit time. The emitter current therefore equals the excess minority carrier charge present in the base region, divided by the time this charge spends in the base. This and other similar relations will be used to construct the charge control model of the bipolar junction transistor in section 5.5.2.

A combination of equations (5.3.1), (5.3.4) and (5.3.5) yields the transit time as a function of the quasi-neutral layer width, w_B , and the electron diffusion constant in the base, $D_{n,B}$.

$$t_r = \frac{w_B^2}{2D_{n,B}} \quad (5.3.6)$$

We now turn our attention to the recombination current in the quasi-neutral base and obtain it from the continuity equation:

$$\frac{\partial n_p(x)}{\partial t} = \frac{1}{q} \frac{\partial J_n(x)}{\partial x} - \frac{n_p(x) - n_{p0}}{\tau_n} \quad (5.3.7)$$

In steady state and applied to the quasi-neutral region in the base, the continuity equation yields the base recombination current, $I_{r,B}$:

$$I_{r,B} = qA_E \int_{x_{p,BE}}^{w_B - x_{p,BC}} \frac{n_p(x) - n_{p0}}{\tau_n} dx \quad (5.3.8)$$

which in turn can be written as a function of the excess minority carrier charge, $\Delta Q_{n,B}$, using equation (5.3.3).

$$I_{r,B} = \frac{\Delta Q_{n,B}}{\tau_n} \quad (5.3.9)$$

Next, we need to find the emitter efficiency and base transport factor. The emitter efficiency defined by equation (5.2.17), becomes:

$$\gamma_E = \frac{1}{1 + \frac{D_{p,E} N_B w_B}{D_{n,B} N_E w_E}} \quad (5.3.10)$$

It is typically the emitter efficiency, which limits the current gain in transistors made of silicon or germanium. The long minority-carrier lifetime and the long diffusion lengths in those materials justify the exclusion of recombination in the base or the depletion layer. The resulting current gain, under such conditions, is:

$$\beta \equiv \frac{D_{n,B} N_E w_E}{D_{p,E} N_B w_B}, \quad \text{if } \alpha \equiv \gamma_E \quad (5.3.11)$$

From this equation, we conclude that the current gain can be larger than one if the emitter doping is much larger than the base doping. A typical current gain for a silicon bipolar transistor is 50 - 150.

The base transport factor, as defined in equation (5.2.18), equals:

$$\alpha_T = 1 - \frac{t_r}{\tau_n} = 1 - \frac{w_B^2}{2D_{n,B}\tau_n} \quad (5.3.12)$$

This expression is only valid if the base transport factor is very close to one, since it was derived using the "short-diode" carrier distribution. This base transport factor can also be expressed in function of the diffusion length in the base:

$$\alpha_T = 1 - \frac{1}{2} \left(\frac{w_B}{L_n} \right)^2 \quad (5.3.13)$$

Example 5.2	<p>Consider a pnp bipolar transistor with emitter doping of 10^{18} cm^{-3} and base doping of 10^{17} cm^{-3}. The quasi-neutral region width in the emitter is $1 \mu\text{m}$ and $0.2 \mu\text{m}$ in the base. Use $\mu_n = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $\mu_p = 300 \text{ cm}^2/\text{V}\cdot\text{s}$. The minority carrier lifetime in the base is 10 ns. Calculate the emitter efficiency, the base transport factor, and the current gain of the transistor biased in the forward active mode. Assume there is no recombination in the depletion region.</p>
Solution	<p>The emitter efficiency is obtained from:</p> $\gamma_E = \frac{1}{1 + \frac{D_{p,E} N_B w_B}{D_{n,B} N_E w_E}} = 0.994$ <p>The base transport factor equals:</p> $\alpha_T = 1 - \frac{w_B^2}{2 D_{n,B} \tau_n} = 0.9992$ <p>The current gain then becomes:</p> $\beta = \frac{\alpha}{1 - \alpha} = 147.5$ <p>where the transport factor, α, was calculated as the product of the emitter efficiency and the base transport factor:</p> $\alpha = \gamma_E \alpha_T = 0.994 \times 0.9992 = 0.993$

5.3.2. General bias modes of a bipolar transistor



While the forward active mode of operation is the most useful bias mode when using a bipolar junction transistor as an amplifier, one cannot ignore the other bias modes especially when using the device as a digital switch. All possible bias modes are illustrated with Figure 5.3.2. They are the forward active mode of operation, the reverse active mode of operation, the saturation mode and the cut-off mode.

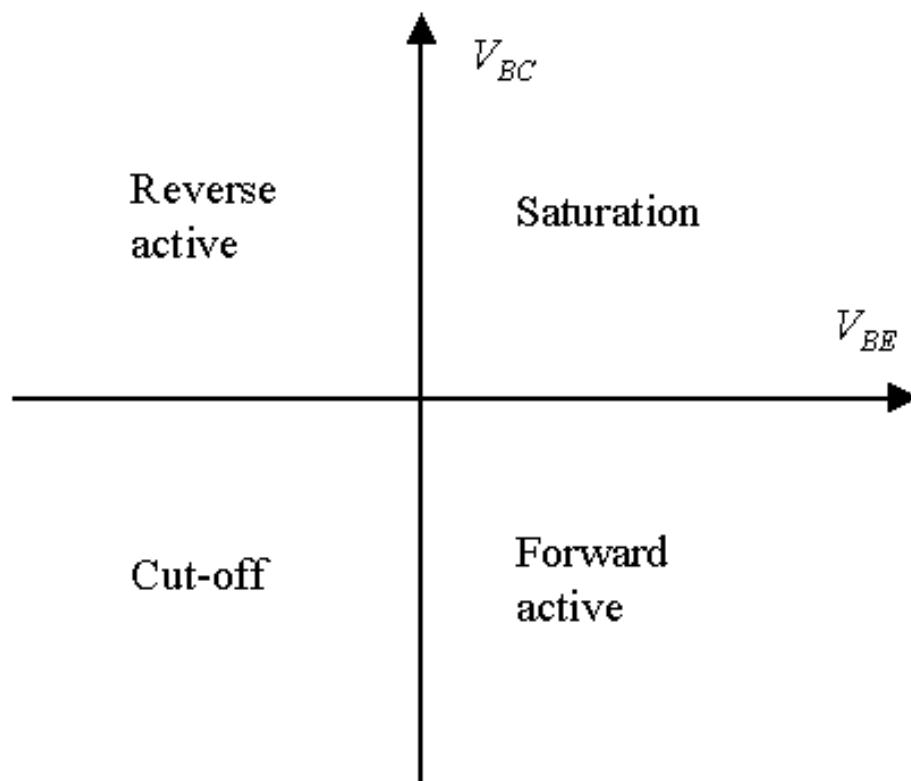


Figure 5.3.2.: Possible bias modes of operation of a bipolar junction transistor.

The forward active mode is the one where we forward bias the base-emitter junction, $V_{BE} > 0$ and reverse bias the base-collector junction, $V_{BC} < 0$. This mode, as discussed in section [5.3.1](#), is the one used in bipolar transistor amplifiers. In bipolar transistor logic circuits, one frequently switches the transistor from the "off" state to the low resistance "on" state. This "off" state is the cut-off mode and the "on" state is the saturation mode. In the cut-off mode, both junctions are reversed biased, $V_{BE} < 0$ and $V_{BC} < 0$, so that very little current goes through the device. This corresponds to the "off" state of the device. In the saturation mode, both junctions are forward biased, $V_{BE} > 0$ and $V_{CB} > 0$. This corresponds to the low resistance "on" state of the transistor.

Finally, there is the reverse active mode of operation. In the reverse active mode, we reverse the function of the emitter and the collector. We reverse bias the base-emitter junction and forward bias the base-collector junction, or $V_{BE} < 0$ and $V_{BC} > 0$. In this mode, the transistor has an emitter efficiency and base transport factor as described by equations [\(5.3.10\)](#) and [\(5.3.12\)](#), where we replace the emitter parameters by the collector parameters. Most transistors, however, have poor emitter efficiency under reverse active bias since the collector doping density is typically much less than the base doping density to ensure high base-collector breakdown voltages. In addition, the collector-base area is typically larger than the emitter-base area, so that even fewer electrons make it from the collector into the emitter.

Having described the forward active mode of operation, there remains the saturation mode, which needs further discussion. Cut-off requires little further analysis, while the reverse active mode of operation is analogous to the forward active mode with the added complication that the areas of the base-emitter and base-collector junction, A_E and A_C , differ. The Ebers-Moll model describes all of these bias modes.

5.3.3. The Ebers-Moll model



The Ebers-Moll model is an ideal model for a bipolar transistor, which can be used, in the forward active mode of operation, in the reverse active mode, in saturation and in cut-off. This model is the predecessor of today's computer simulation models and contains only the "ideal" diode currents.

The model contains two diodes and two current sources as shown in Figure 5.3.3. The two diodes represent the base-emitter and base-collector diodes. The current sources quantify the transport of minority carriers through the base region. These are current sources depend on the current through each diode. The parameters $I_{E,s}$, $I_{C,s}$, α_F and α_R are the saturation currents of the base-emitter and base collector diode and the forward and reverse transport factors.

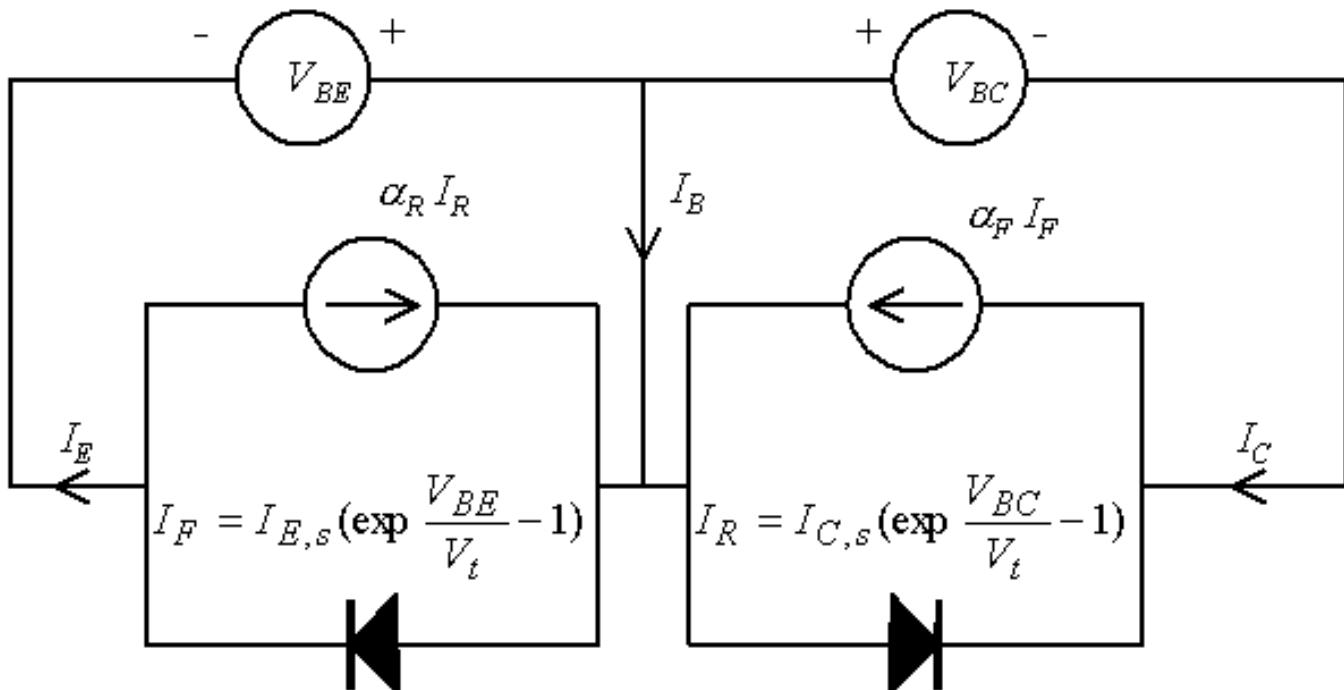


Figure 5.3.3 : Equivalent circuit for the Ebers-Moll model of an NPN bipolar junction transistor

Using the parameters identified in Figure 5.3.3, we can relate the emitter, base and collector current to the forward and reverse currents and transport factors, yielding:

$$I_E = I_F - \alpha_R I_R \quad (5.3.14)$$

$$I_B = (1 - \alpha_F) I_F + (1 - \alpha_R) I_R \quad (5.3.15)$$

$$I_C = -I_R + \alpha_F I_F \quad (5.3.16)$$

The Ebers-Moll parameters are related by the following equation:

$$I_{E,s} \alpha_F = I_{C,s} \alpha_R \quad (5.3.17)$$

This relationship is also referred as the reciprocity relation and can be derived by examining the minority carrier current through the base. For the specific case where the base-emitter and base-collector voltage are the same and the base doping is uniform, there can be no minority carrier diffusion in the base so that:

$$I_F(V_{BE}) \alpha_F = I_R(V_{BC} = V_{BE}) \alpha_R \quad (5.3.18)$$

from which the reciprocity relation is obtained.

The forward- and reverse-bias transport factors are obtained by measuring the current gain in the forward active and reverse active mode of operation. The saturation currents $I_{E,s}$ and $I_{C,s}$ are obtained by measuring the base-emitter (base-collector) diode saturation current while shorting the base-collector (base-emitter) diode.



5.3.4. Saturation.

In the low resistance "on" state of a bipolar transistor, one finds that the voltage between the collector and emitter is less than the forward bias voltage of the base-emitter junction. Typically the "on" state voltage of a silicon BJT is 100 mV and the forward bias voltage is 700 mV. Therefore, the base-collector junction is forward biased. Using the Ebers-Moll model, we can calculate the "on" voltage from:

$$V_{CE,sat} = V_{BE} - V_{BC} = V_t \ln \left\{ \frac{I_F}{I_R} \frac{I_{C,s}}{I_{E,s}} \right\} \quad (5.3.19)$$

and using equations (5.3.15), (5.3.16) and the reciprocity relation (5.3.17), one obtains:

$$V_{CE,sat} = V_t \ln \left\{ \frac{1 + \frac{I_C}{I_B} (1 - \alpha_R)}{\alpha_R \left[1 - \frac{I_C}{I_B} \frac{(1 - \alpha_F)}{\alpha_F} \right]} \right\} \quad (5.3.20)$$

Saturation also implies that a large amount of minority carrier charge is accumulated in the base region. As a transistor is switched from saturation to cut-off, this charge initially remains in the base and a collector current will remain until this charge is removed by recombination. This causes an additional delay before the transistor is turned off. Since the carrier lifetime can be significantly longer than the base transit time, the turn-off delay causes a large and undesirable asymmetry between turn-on and turn-off time. Saturation is therefore avoided in high-speed bipolar logic circuits. Two techniques are used to reduce the turn-off delay: 1) adding a Schottky diode in parallel to the base-collector junction and 2) using an emitter-coupled circuit configuration. Both approaches avoid biasing the transistor in the saturation mode. The Schottky diode clamps the base-collector voltage at a value, which is slightly lower than the turn-on voltage of the base-collector diode. An emitter-coupled circuit is biased with a current source, which can be designed so that the collector voltage cannot be less than the base voltage.

Example 5.3	Calculate the saturation voltage of a bipolar transistor biased with a base current of 1 mA and a collector current of 10 mA. Use $\alpha_R = 0.993$ and $\alpha_F = 0.2$.
Solution	<p>The saturation voltage equals:</p> $V_{CE,sat} = V_t \ln \left\{ \frac{1 + \frac{I_C}{I_B} (1 - \alpha_R)}{\alpha_R \left[1 - \frac{I_C}{I_B} \frac{(1 - \alpha_F)}{\alpha_F} \right]} \right\} = 0.1 \text{ V}$

Chapter 5: Bipolar Junction Transistors



5.4. Non-ideal effects

5.4.1. Base-width modulation

5.4.2. Recombination in the depletion region

A variety of effects occur in bipolar transistors, which are not included in the ideal transistor model. These include the base-width modulation effects and the current due to recombination in the depletion layers. Both are described next.

5.4.1. Base-width modulation



As the voltages applied to the base-emitter and base-collector junctions are changed, the depletion layer widths and the quasi-neutral regions vary as well. This causes the collector current to vary with the collector-emitter voltage as illustrated in Figure 5.4.1.

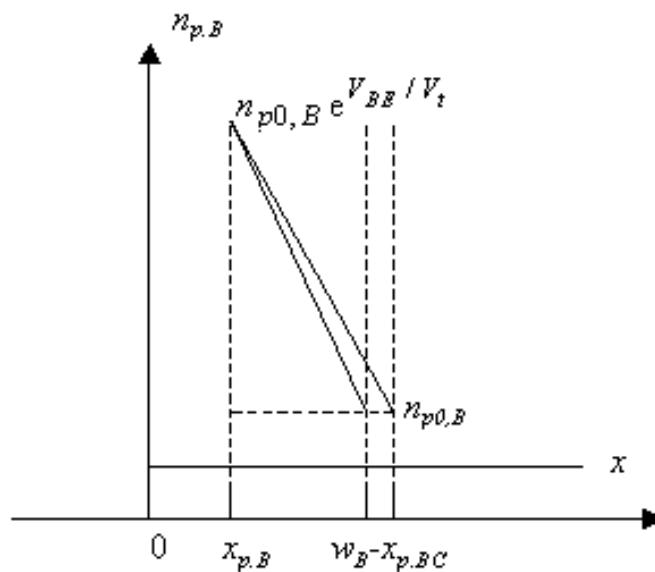


Figure 5.4.1. : Variation of the minority-carrier distribution in the base quasi-neutral region due to a variation of the base-collector voltage.

A variation of the base-collector voltage results in a variation of the quasi-neutral width in the base. The gradient of the minority-carrier density in the base therefore changes, yielding an increased collector current as the collector-base current is increased. This effect is referred to as the Early effect. The Early effect is observed as an increase in the collector current with increasing collector-emitter voltage as illustrated with Figure 5.4.2. The Early voltage, V_A , is obtained by drawing a line tangential to the transistor I - V characteristic at the point of interest. The Early voltage equals the horizontal distance between the point chosen on the I - V characteristics and the intersection between the tangential line and the horizontal axis. It is indicated on the figure by the horizontal arrow.

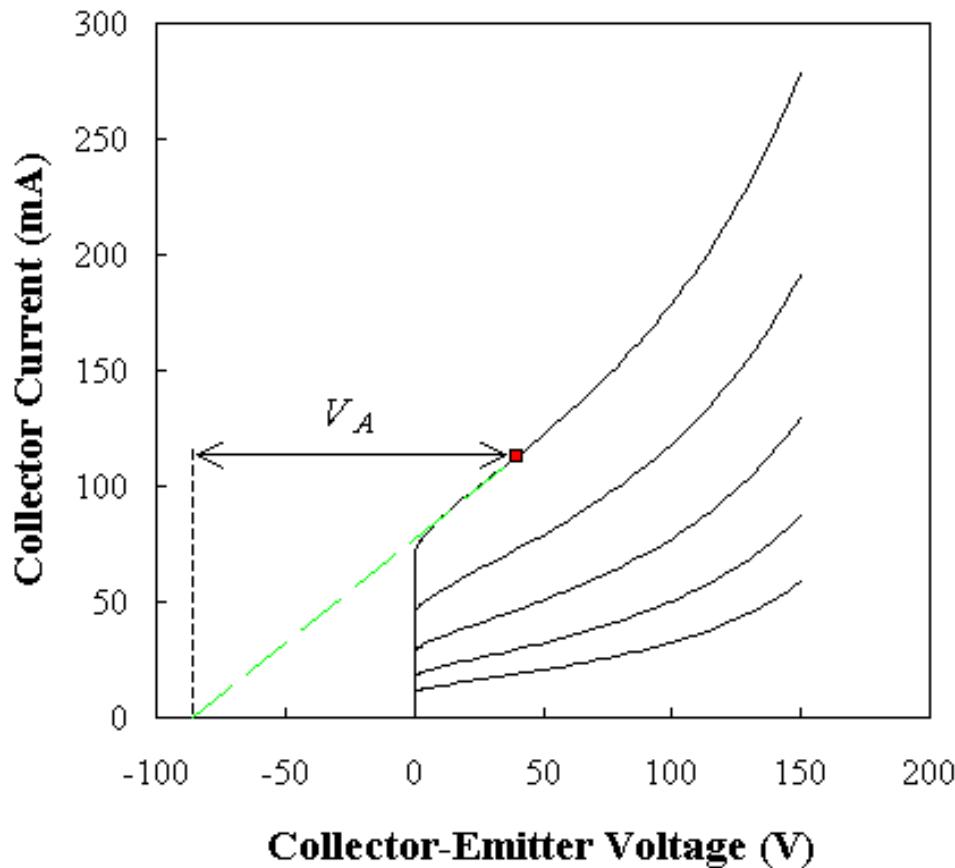


Figure 5.4.2. : Collector current increase with an increase of the collector-emitter voltage due to the Early effect. The Early voltage, V_A , is also indicated on the figure.

The change of the collector current when changing the collector-emitter voltage is primarily due to the variation of the base-collector voltage, since the base-emitter junction is forward biased and a constant base current is applied. The collector current depends on the base-collector voltage since the base-collector depletion layer width varies, which also causes the quasi-neutral width, w_B' , in the base to vary. This variation can be calculated for a piece-wise uniformly-doped transistor using the ideal transistor mode as described by equations (5.2.10) and (5.3.1):

$$\frac{dI_C}{dV_{CE}} \cong -\frac{dI_C}{dV_{BC}} = \frac{I_C}{w_B'} \frac{dw_B'}{dV_{BC}} \quad (5.4.1)$$

This variation can be expressed by the Early voltage, V_A , which quantifies what voltage variation would result in zero collector current.

$$\frac{dI_C}{dV_{CE}} \cong \frac{I_C}{|V_A|} \quad (5.4.2)$$

It can be shown that the Early voltage also equals the majority carrier charge in the base, Q_B , divided by the base-collector junction capacitance, $C_{j,BC}$:

$$|V_A| = \frac{Q_B}{C_{j,BC}} = \frac{\frac{qA_C N_B w_B}{\epsilon_s A_C}}{x_{p,BC} + x_{n,BC}} \quad (5.4.3)$$

In addition to the Early effect, there is a less pronounced effect due to the variation of the base-emitter voltage, which changes the ideality factor of the collector current. However, the effect at the base-emitter junction is much smaller since the base-emitter junction capacitance is larger and the base-emitter voltage variation is very limited since the junction is forward biased. The effect does lead to a variation of the ideality factor, n , given by:

$$n = \frac{1}{V_t \frac{d \ln I_C}{dV_{BE}}} \approx 1 + \frac{V_t}{Q_B} C_{j,BE} \quad (5.4.4)$$

Example 5.4	Consider a bipolar transistor with a base doping of 10^{17} cm^{-3} and a quasi-neutral base width of $0.2 \mu\text{m}$. Calculate the Early voltage and collector current ideality factor given that the base-emitter capacitance and the base-collector capacitance are 0.2 nF and 0.2 pF . The collector area equals 10^{-4} cm^2 .
Solution	<p>The Early voltage equals:</p> $ V_A = \frac{Q_B}{C_{j,BC}} = \frac{\frac{qA_C N_B w_B}{\epsilon_s A_C}}{x_{p,BC} + x_{n,BC}} = 160 \text{ V}$ <p>The saturation voltage equals:</p> $n \approx 1 + \frac{V_t}{Q_B} C_{j,BE} = 1.16$

An extreme case of base-width modulation is punchthrough. As the collector-emitter voltage is increased, the quasi-neutral width of the base decreases, so that eventually it becomes zero. The collector current becomes very large and no longer depends on the voltage applied to the base. This mode of operation is undesirable since most performance characteristics degrade as one approaches punchthrough. The rapid increase of the collector current at the punchthrough voltage can cause the destruction of the transistor due to excessive power dissipation. Punchthrough is therefore one of the possible breakdown modes of a bipolar transistor.

5.4.2. Recombination in the depletion region



So far, we have ignored the recombination in the depletion region. As in a p-n diode, the recombination in the depletion region causes an additional diode current. We can identify this contribution to the current because of the different voltage dependence as described in section 4.4.4. An example is shown in Figure 5.4.3. Shown are the collector and base current of a silicon bipolar transistor, biased in the forward active mode of operation with $V_{BC} = -12 \text{ V}$, as a function of the base-emitter voltage. This type of plot is also called a Gummel plot.

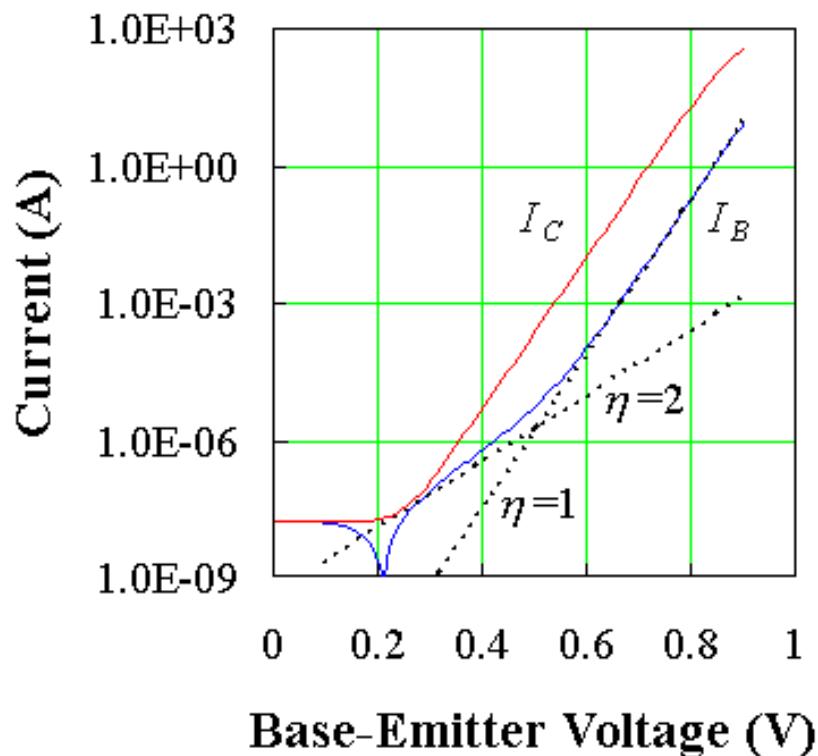


Figure 5.4.3 : Gummel plot: Collector current (top curve) and base current (bottom curve) of a silicon bipolar transistor versus the base-emitter voltage. 

The current due to recombination in the depletion region can be observed as an additional base current between $V_{BE} = 0.2$ and 0.4 V. The collector current does not include this additional current, since recombination in the depletion region does not affect the flow of electrons through the base.

Chapter 2: Semiconductor Fundamentals



Review Questions



1. Why do solids occur in the form of a crystal?
2. How do we classify the different crystals?
3. How many Bravais lattices are there in two dimensions? How many in three dimensions?
4. List the three cubic bravais lattices.
5. How do you explain that the allowed energies for electrons in solids are restricted to energy bands? Why are these bands separated energy band gaps? Why are the energies not discreet as in an atom. Why are they not continuous, as is the case for a free electron?
6. How does the conductivity of a solid depend on whether the energy bands are completely filled, partially filled or empty? How does the existence of overlapping bandgaps affect the conductivity?
7. Why does a completely filled band not contribute the conductivity of a solid?
8. Explain physically why the bandgap of a semiconductor decreases with temperature.
9. What are holes? Carefully justify your definition.
10. What is a state?
11. How many states are there in 1 micron sized cube for which an electron has a kinetic energy less than 1 eV? Treat the electron as a free electron confined to a box with infinite potential walls.
12. What is the physical meaning of the Fermi energy?
13. What is the value of the Fermi function at an energy, which is $3kT$ larger/lower than the Fermi energy?
14. What is the basic assumption used in statistical thermodynamics when calculating the probability distribution functions?
15. What are the two boundary conditions used to find the possible ways to fill energy levels with electrons.
16. How does a boson differ from a Fermion? Name two bosons.
17. List the assumptions made to obtain equations (2.6.12).
18. What is an intrinsic semiconductors? What is the hole density in an intrinsic semiconductor?

19. Why is the product of the electron and hole density in a non-degenerate semiconductor constant rather than for instance the sum? This relationship is also referred to as the mass-action law. Why?
20. Define a non-degenerate semiconductor. Why do we need this concept?
21. What is the difference between a doped semiconductor and an extrinsic semiconductor?
22. What assumptions are made when deriving equations (2.6.29) and (2.6.30)?
23. Describe the temperature dependence of the carrier density in a semiconductor. Identify the three regions and explain what happens by indicating the filled and empty states on an energy band diagram. Do this for n-type, p-type and compensated material.
24. Name the two transport mechanisms in semiconductors.
25. Describe the microscopic behavior of electrons and holes in a semiconductor.
26. Define the mobility.
27. Explain why the mobility in a semiconductor depends on the doping density.
28. Define the resistivity and conductivity of a semiconductor.
29. Explain why the velocity in a semiconductor is limited.
30. What is the driving force, which causes diffusion?
31. Explain the relation between the mean free path, the scattering time and the thermal velocity.
32. List three recombination-generation mechanisms.
33. Explain why the net recombination rate as described by the simple model depends on the excess carrier density.
34. Describe the continuity equation in words.
35. What assumptions are made to obtain the diffusion equations (2.9.9) and (2.9.10) from the continuity equations (2.9.3) and (2.9.4)?
36. What is the diffusion length and how does it relate to the diffusion constant and the minority carrier lifetime?
37. What is the drift-diffusion model?

Bibliography

1. Physical Properties of crystals, J.F. Nye, Oxford University Press, 1979.
2. Crystal physics, H.J. Juretschke, W.A. Benjamin, Inc., 1974
3. Introduction to Solid State Physics, C. Kittel, Wiley & sons,
4. Electrons and holes in semiconductors, W. Shockley, D. Van Nostrand Company, 1959
5. Electrons in Solids Third edition, R.H. Bube, Academic Press, 1992.
6. Solid State Electronic Devices, Fourth edition, B. G. Streetman, Prentice Hall, 1995
7. Solid State Theory, W. A. Harrison, Dover publications, 1974
8. Thermal Physics, second edition, C. Kittel and H. Kroemer, Freeman, 1980

Bibliography

Review Questions

1. Why do solids occur in the form of a crystal?
2. How do we classify the different crystals?
3. How many Bravais lattices are there in two dimensions? How many in three dimensions?
4. List the three cubic bravais lattices.
5. How do you explain that the allowed energies for electrons in solids are restricted to energy bands? Why are these bands separated energy band gaps? Why are the energies not discreet as in an atom. Why are they not continuous, as is the case for a free electron?
6. How does the conductivity of a solid depend on whether the energy bands are completely filled, partially filled or empty? How does the existence of overlapping bandgaps affect the conductivity?
7. Why does a completely filled band not contribute the conductivity of a solid?
8. Explain physically why the bandgap of a semiconductor decreases with temperature.
9. What are holes? Carefully justify your definition.
10. What is a state?
11. How many states are there in 1 micron sized cube for which an electron has a kinetic energy less than 1 eV? Treat the electron as a free electron confined to a box with infinite potential walls.
12. What is the physical meaning of the Fermi energy?
13. What is the value of the Fermi function at an energy, which is $3kT$ larger/lower than the Fermi energy?
14. What is the basic assumption used in statistical thermodynamics when calculating the probability distribution functions?
15. What are the two boundary conditions used to find the possible ways to fill energy levels with electrons.
16. How does a boson differ from a Fermion? Name two bosons.
17. List the assumptions made to obtain equations (2.6.12).
18. What is an intrinsic semiconductors? What is the hole density in an intrinsic semiconductor?
19. Why is the product of the electron and hole density in a non-degenerate semiconductor constant rather than for instance the sum? This relationship is also referred to as the mass-action law. Why?
20. Define a non-degenerate semiconductor. Why do we need this concept?
21. What is the difference between a doped semiconductor and an extrinsic semiconductor?

22. What assumptions are made when deriving equations (2.6.29) and (2.6.30)?
23. Describe the temperature dependence of the carrier density in a semiconductor. Identify the three regions and explain what happens by indicating the filled and empty states on an energy band diagram. Do this for n-type, p-type and compensated material.
24. Name the two transport mechanisms in semiconductors.
25. Describe the microscopic behavior of electrons and holes in a semiconductor.
26. Define the mobility.
27. Explain why the mobility in a semiconductor depends on the doping density.
28. Define the resistivity and conductivity of a semiconductor.
29. Explain why the velocity in a semiconductor is limited.
30. What is the driving force, which causes diffusion?
31. Explain the relation between the mean free path, the scattering time and the thermal velocity.
32. List three recombination-generation mechanisms.
33. Explain why the net recombination rate as described by the simple model depends on the excess carrier density.
34. Describe the continuity equation in words.
35. What assumptions are made to obtain the diffusion equations (2.9.9) and (2.9.10) from the continuity equations (2.9.3) and (2.9.4)?
36. What is the diffusion length and how does it relate to the diffusion constant and the minority carrier lifetime?
37. What is the drift-diffusion model?

Problems

1. Calculate the packing density of the body centered cubic, the face centered cubic and the diamond lattice, listed in example 2.1 p 28.
2. At what temperature does the energy bandgap of silicon equal exactly 1 eV?
3. Prove that the probability of occupying an energy level below the Fermi energy equals the probability that an energy level above the Fermi energy and equally far away from the Fermi energy is not occupied.
4. At what energy (in units of kT) is the Fermi function within 1 % of the Maxwell-Boltzmann distribution function? What is the corresponding probability of occupancy?
5. Calculate the Fermi function at 6.5 eV if $E_F = 6.25$ eV and $T = 300$ K. Repeat at $T = 950$ K assuming that the Fermi energy does not change. At what temperature does the probability that an energy level at $E = 5.95$ eV is empty equal 1 %.
6. Calculate the effective density of states for electrons and holes in germanium, silicon and gallium arsenide at room temperature and at 100 °C. Use the effective masses for density of states calculations.
7. Calculate the intrinsic carrier density in germanium, silicon and gallium arsenide at room temperature (300 K). Repeat at 100 °C. Assume that the energy bandgap is independent of temperature and use the room temperature values.
8. Calculate the position of the intrinsic energy level relative to the midgap energy

$$E_{midgap} = (E_c + E_v)/2$$

in germanium, silicon and gallium arsenide at 300 K. Repeat at $T = 100$ °C.

9. Calculate the electron and hole density in germanium, silicon and gallium arsenide if the Fermi energy is 0.3 eV above the intrinsic energy level. Repeat if the Fermi energy is 0.3 eV below the conduction band edge. Assume that $T = 300$ K.
10. The equations (2.6.34) and (2.6.35) derived in section 2.6 are only valid for non-degenerate semiconductors (i.e. $E_v + 3kT < E_F < E_c - 3kT$). Where exactly in the derivation was the assumption made that the semiconductor is non-degenerate?
11. A silicon wafer contains 10^{16} cm⁻³ electrons. Calculate the hole density and the position of the intrinsic energy and the Fermi energy at 300 K. Draw the corresponding band diagram to scale, indicating the conduction and valence band edge, the intrinsic energy level and the Fermi energy level. Use $n_i = 10^{10}$ cm⁻³.
12. A silicon wafer is doped with 10^{13} cm⁻³ shallow donors and 9×10^{12} cm⁻³ shallow acceptors. Calculate the electron and hole density at 300 K. Use $n_i = 10^{10}$ cm⁻³.
13. The resistivity of a silicon wafer at room temperature is 5 Ωcm. What is the doping density? Find all possible solutions.
14. How many phosphorus atoms must be added to decrease the resistivity of n-type silicon at room temperature from 1 Ωcm to 0.1 Ωcm. Make sure you include the doping dependence of

the mobility. State your assumptions.

15. A piece of n-type silicon ($N_d = 10^{17} \text{ cm}^{-3}$) is uniformly illuminated with green light ($\lambda = 550 \text{ nm}$) so that the power density in the material equals 1 mW/cm^2 . a) Calculate the generation rate of electron-hole pairs using an absorption coefficient of 10^4 cm^{-1} . b) Calculate the excess electron and hole density using the generation rate obtained in (a) and a minority carrier lifetime due to Shockley-Read-Hall recombination of 0.1 ms. c) Calculate the electron and hole quasi-Fermi energies (relative to E_i) based on the excess densities obtained in (b).
16. A piece of intrinsic silicon is instantaneously heated from 0 K to room temperature (300 K). The minority carrier lifetime due to Shockley-Read-Hall recombination in the material is 1 ms. Calculate the generation rate of electron-hole pairs immediately after reaching room temperature. ($E_t = E_i$). If the generation rate is constant, how long does it take to reach thermal equilibrium?
17. Calculate the conductivity and resistivity of intrinsic silicon. Use $n_i = 10^{10} \text{ cm}^{-3}$, $\mu_n = 1400 \text{ cm}^2/\text{V-sec}$ and $\mu_p = 450 \text{ cm}^2/\text{V-sec}$.
18. Consider the problem of finding the doping density which results the maximum possible resistivity of silicon at room temperature. ($n_i = 10^{10} \text{ cm}^{-3}$, $\mu_n = 1400 \text{ cm}^2/\text{V-sec}$ and $\mu_p = 450 \text{ cm}^2/\text{V-sec}$.)

Should the silicon be doped at all or do you expect the maximum resistivity when dopants are added?

If the silicon should be doped, should it be doped with acceptors or donors (assume that all dopant are shallow).

Calculate the maximum resistivity, the corresponding electron and hole density and the doping density.

19. The electron density in silicon at room temperature is twice the intrinsic density. Calculate the hole density, the donor density and the Fermi energy relative to the intrinsic energy. Repeat for $n = 5 n_i$ and $n = 10 n_i$. Also repeat for $p = 2 n_i$, $p = 5 n_i$ and $p = 10 n_i$, calculating the electron and acceptor density as well as the Fermi energy relative to the intrinsic energy level.
20. What photon energy (in electron volt) corresponds to a wavelength of 1 micron? What wavelength corresponds to a photon energy of 1 eV?
21. 1 billion photons with a wavelength of 0.3 micron hit a detector every second. How large is the incident power?
22. The expression for the Bohr radius can also be applied to the hydrogen-like atom consisting of an ionized donor and the electron provided by the donor. Modify the expression for the Bohr radius so that it applies to this hydrogen-like atom. Calculate the Bohr radius of an electron orbiting around the ionized donor in silicon. ($\epsilon_r = 11.9$ and $m_e^* = 0.26 m_0$)
23. Calculate the density of electrons per unit energy (in electron volt) and per unit area (per

cubic centimeter) at 1 eV above the band minimum. Assume that $m_e^* = 1.08 m_0$

24. Calculate the probability that an electron occupies an energy level which is $3kT$ below the Fermi energy. Repeat for an energy level which is $3kT$ above the Fermi energy.
25. Calculate and plot as a function of energy the product of the probability that an energy level is occupied with the probability that that same energy level is not occupied. Assume that the Fermi energy is zero and that $kT = 1$ eV
26. The effective mass of electrons in silicon is $0.26 m_0$ and the effective mass of holes is $0.36 m_0$. If the scattering time is the same for both carrier types, what is the ratio of the electron mobility and the hole mobility.
27. Electrons in silicon carbide have a mobility of $1000 \text{ cm}^2/\text{V}\cdot\text{sec}$. At what value of the electric field do the electrons reach a velocity of $3 \times 10^7 \text{ cm/s}$? Assume that the mobility is constant and independent of the electric field. What voltage is required to obtain this field in a 5 micron thick region? How much time do the electrons need to cross the 5 micron thick region?
28. A piece of silicon has a resistivity which is specified by the manufacturer to be between 2 and 5 Ohm cm. Assuming that the mobility of electrons is $1400 \text{ cm}^2/\text{V}\cdot\text{sec}$ and that of holes is $450 \text{ cm}^2/\text{V}\cdot\text{sec}$, what is the minimum possible carrier density and what is the corresponding carrier type? Repeat for the maximum possible carrier density.
29. A silicon wafer has a 2 inch diameter and contains 10^{14} cm^{-3} electrons with a mobility of $1400 \text{ cm}^2/\text{V}\cdot\text{sec}$. How thick should the wafer be so that the resistance between the front and back surface equals 0.1 Ohm.
30. The electron mobility in germanium is $1000 \text{ cm}^2/\text{V}\cdot\text{sec}$. If this mobility is due to impurity and lattice scattering and the mobility due to lattice scattering only is $1900 \text{ cm}^2/\text{V}\cdot\text{sec}$, what is the mobility due to impurity scattering only?

Problem 2.1 Calculate the packing density of the body centered cubic, the face centered cubic and the diamond lattice, listed in example 2.1 p 28.

Solution The packing density is calculated as in example 2.1 p 28 and obtained from:

$$\frac{\text{Volume of atoms}}{\text{Volume of the unit cell}} = \frac{\frac{4}{3} \mathbf{p} r^3}{a^3}$$

The correct radius and number of atoms per unit cell should be used.

A **body centered cubic** lattice contains an additional atom in the middle and therefore contains two atoms per unit cell. The atoms touch along the body diagonal, which equals $\sqrt{3} a$. The radius is one quarter of the body diagonal.

A **face centered cubic** lattice contains six additional atoms in the center of all six faces of the cube. Since only half of the atoms is within the cube the total number of atoms per unit cell equals four. The atoms touch along the diagonal of the faces of the cube, which equals $\sqrt{2} a$. The radius is one quarter of the diagonal.

The **diamond lattice** contains two face centered cubic lattice so that the total number of atoms per unit cell equals twice that of the face centered lattice, namely eight. The atoms touch along the body diagonal, where two atoms are one quarter of the body diagonal apart or $\sqrt{3} a / 4$. The radius equals half the distance between the two atoms.

The radius, number of atoms per unit cell and the packing density are summarized in the table below.

	Radius	Atoms/ unit cell	Packing density
Simple cubic	$\frac{a}{2}$	1	$\frac{\mathbf{p}}{6} = 52\%$
Body centered cubic	$\frac{\sqrt{3} a}{4}$	2	$\frac{\mathbf{p} \sqrt{3}}{8} = 68\%$
Face centered cubic	$\frac{\sqrt{2} a}{4}$	4	$\frac{\mathbf{p} \sqrt{2}}{6} = 74\%$
Diamond	$\frac{\sqrt{3} a}{8}$	8	$\frac{\mathbf{p} \sqrt{3}}{16} = 34\%$

Problem 2.2 At what temperature does the energy bandgap of silicon equal exactly 1 eV?

Solution The energy bandgap is obtained from:

$$E_g(T) = E_g(0 \text{ K}) - \frac{\mathbf{a}T^2}{T + \mathbf{b}}$$
$$= 1.166 - \frac{0.473 \times 10^{-3} \times T^2}{T + 636} = 1.0 \text{ eV}$$

This quadratic equation can be solved yielding:

$$T = \frac{E_g(0 \text{ K}) - E_g(T)}{2\mathbf{a}}$$
$$+ \sqrt{\left(\frac{E_g(0 \text{ K}) - E_g(T)}{2\mathbf{a}}\right)^2 + \frac{\mathbf{b}(E_g(0 \text{ K}) - E_g(T))}{\mathbf{a}}} = 679 \text{ K}$$

which is consistent with Figure 2.3.5

Problem 2.3 (same as 1.9) Prove that the probability of occupying an energy level below the Fermi energy equals the probability that an energy level above the Fermi energy and equally far away from the Fermi energy is not occupied.

Solution The probability that an energy level with energy ΔE **below** the Fermi energy E_F is occupied can be rewritten as:

$$\begin{aligned} f(E_F - \Delta E) &= \frac{1}{1 + \exp \frac{E_F - \Delta E - E_F}{kT}} = \frac{\exp \frac{\Delta E}{kT}}{\exp \frac{\Delta E}{kT} + 1} \\ &= 1 - \frac{1}{\exp \frac{\Delta E}{kT} + 1} = 1 - \frac{1}{1 + \exp \frac{E_F + \Delta E - E_F}{kT}} = 1 - f(E_F + \Delta E) \end{aligned}$$

so that it also equals the probability that an energy level with energy ΔE **above** the Fermi energy, E_F , is **not** occupied.

Problem 2.4 At what energy (in units of kT) is the Fermi function within 1 % of the Maxwell-Boltzmann distribution function? What is the corresponding probability of occupancy?

Solution The Fermi function can be approximated by the Maxwell-Boltzmann distribution function with an approximate error of 1 % if:

$$\frac{f_{MB} - f_{FB}}{f_{FD}} = 0.01, \text{ or } \frac{1}{f_{FD}} = \frac{1.01}{f_{MB}}$$

using $x = (E - E_F)/kT$, this condition can be rewritten as:

$$1 + \exp(-x) = 1.01 \exp(-x)$$

from which one finds $x = \ln(100) = 4.605$ so that

$$E = E_F + 4.605 kT \text{ and } f_{FD}(E_F + 4.605 kT) = 0.0099$$

Problem 2.5 Calculate the Fermi function at 6.5 eV if $E_F = 6.25$ eV and $T = 300$ K. Repeat at $T = 950$ K assuming that the Fermi energy does not change. At what temperature does the probability that an energy level at $E = 5.95$ eV is empty equal 1 %.

Solution The Fermi function at 300 K equals:

$$f(6.5 \text{ eV}) = \frac{1}{1 + \exp\left(\frac{6.5 - 6.25}{0.02586}\right)} = 6.29 \times 10^{-5}$$

The Fermi function at 950 K equals:

$$f(6.5 \text{ eV}) = \frac{1}{1 + \exp\left(\frac{6.5 - 6.25}{0.0818}\right)} = 0.045$$

The probability that the Fermi function equals 1 % implies:

$$f(5.95 \text{ eV}) = 0.99 = \frac{1}{1 + \exp\left(\frac{5.95 - 6.25}{kT/q}\right)}$$

resulting in

$$T = \frac{-0.3q/k}{\ln\left(\frac{1}{0.99} - 1\right)} = 484.7 \text{ }^{\circ}\text{C}$$

Problem 2.6 Calculate the effective density of states for electrons and holes in germanium, silicon and gallium arsenide at room temperature and at 100 °C. Use the effective masses for density of states calculations.

Solution The effective density of states in the conduction band for germanium equals:

$$\begin{aligned}
 N_c &= 2 \left(\frac{2\mathbf{p} m_e^* k T}{h^2} \right)^{3/2} \\
 &= 2 \left(\frac{2\mathbf{p} 0.55 \times 9.11 \times 10^{-31} \times 1.38 \times 10^{-23} \times 300}{(6.626 \times 10^{-34})^2} \right)^{3/2} \\
 &= 1.02 \times 10^{25} \text{ m}^{-3} = 1.02 \times 10^{19} \text{ cm}^{-3}
 \end{aligned}$$

where the effective mass for density of states was used (Appendix 3). Similarly one finds the effective densities for silicon and gallium arsenide and those of the valence band, using the effective masses listed below:

	Germanium	Silicon	Gallium Arsenide
m_e/m_0	0.55	1.08	0.067
$N_c (\text{cm}^{-3})$	1.02×10^{19}	2.82×10^{19}	4.35×10^{17}
m_e/m_0	0.37	0.81	0.45
$N_v (\text{cm}^{-3})$	5.64×10^{18}	1.83×10^{19}	7.57×10^{18}

The effective density of states at 100 °C (372.15 K) are obtain from:

$$N_c(T) = N_c(300 \text{ K}) \left(\frac{T}{300} \right)^{3/2}$$

yielding:

	Germanium	Silicon	Gallium Arsenide
$N_c (\text{cm}^{-3})$	1.42×10^{19}	3.91×10^{19}	6.04×10^{17}
$N_v (\text{cm}^{-3})$	7.83×10^{18}	2.54×10^{19}	1.05×10^{18}

Problem 2.7 Calculate the intrinsic carrier density in germanium, silicon and gallium arsenide at room temperature (300 K). Repeat at 100 °C. Assume that the energy bandgap is independent of temperature and use the room temperature values.

Solution The intrinsic carrier density is obtained from:

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

where both effective densities of states are also temperature dependent. Using the solution of Problem 2.6 one obtains:

$T = 300 \text{ K}$	Germanium	Silicon	Gallium Arsenide
$n_i (\text{cm}^{-3})$	2.16×10^{13}	8.81×10^9	1.97×10^6
$T = 100^\circ\text{C}$	Germanium	Silicon	Gallium Arsenide
$n_i (\text{cm}^{-3})$	3.67×10^{14}	8.55×10^{11}	6.04×10^8

Problem 2.8 Calculate the position of the intrinsic energy level relative to the midgap energy

$E_{midgap} = (E_c + E_v)/2$
in germanium, silicon and gallium arsenide at 300 K. Repeat at $T = 100^\circ\text{C}$.

Solution The intrinsic energy level relative to the midgap energy is obtained from:

$$E_i - E_{midgap} = \frac{3}{4}kT \ln \frac{m_h^*}{m_e^*}$$

where the effective masses are the effective masses for density of states calculations as listed in the table below.

The corresponding values of the intrinsic level relative to the midgap energy are listed as well.

	Germanium	Silicon	Gallium arsenide
m_e^*/m_0	0.55	1.08	0.067
m_h^*/m_0	0.37	0.81	0.45
$T = 300\text{ K}$	-7.68 meV	-5.58 meV	36.91 meV
$T = 100\text{ C}$	-9.56 meV	-6.94 meV	45.92 meV

Problem 2.9 Calculate the electron and hole density in germanium, silicon and gallium arsenide if the Fermi energy is 0.3 eV above the intrinsic energy level. Repeat if the Fermi energy is 0.3 eV below the conduction band edge. Assume that $T = 300$ K.

Solution The electron density, n , can be calculated from the Fermi energy using:

$$n = n_i \exp \frac{E_F - E_i}{kT} = n_i \exp \left(\frac{0.3}{0.02586} \right)$$

and the corresponding hole density equals:

$$p = n_i^2 / n$$

the resulting values are listed in the table below.

If the Fermi energy is 0.3 eV below the conduction band edge, one obtains the carrier densities using:

$$n = N_c \exp \frac{E_F - E_c}{kT} = N_c \exp \left(\frac{-0.3}{0.02586} \right)$$

and the corresponding hole density equals:

$$p = n_i^2 / n$$

the resulting values are listed in the table below.

	Germanium	Silicon	Gallium Arsenide
n_i (cm ⁻³)	2.03×10^{13}	1.45×10^{10}	2.03×10^6
N_c (cm ⁻³)	1.02×10^{19}	6.62×10^{19}	4.37×10^{17}
$E_F - E_i$ = 0.3 eV	n (cm ⁻³) 2.24×10^{18}	1.60×10^{15}	2.23×10^{11}
	p (cm ⁻³) 1.48×10^8	1.32×10^5	18.4
$E_F - E_i$ = - 0.3 eV	n (cm ⁻³) 9.27×10^{13}	6.02×10^{14}	3.97×10^{12}
	p (cm ⁻³) 4.45×10^{12}	3.50×10^5	1.04

Problem 2.10 The equations (2.6.34) and (2.6.35) derived in section 2.6 are only valid for non-degenerate semiconductors (i.e. $E_v + 3kT < E_F < E_c - 3kT$). Where exactly in the derivation was the assumption made that the semiconductor is non-degenerate?

Solution Equations (2.6.34) and (2.6.35) were derived using charge neutrality and the mass action law. Of those two assumptions, the use of the mass action law implies that the semiconductor is non-degenerate.

The mass action law was derived using (2.6.12) and (2.6.13). These equations, representing a closed form solution for the thermal equilibrium carrier densities as a function of the Fermi energy, were in turn obtained by solving the Fermi integral and assuming that:

$$E_v + 3kT < E_F < E_c - 3kT$$

i.e. that the Fermi energy must be at least $3kT$ away from either bandedge and within the bandgap.

Problem 2.11 A silicon wafer contains 10^{16} cm^{-3} electrons. Calculate the hole density and the position of the intrinsic energy and the Fermi energy at 300 K. Draw the corresponding band diagram to scale, indicating the conduction and valence band edge, the intrinsic energy level and the Fermi energy level. Use $n_i = 10^{10} \text{ cm}^{-3}$.

Solution The hole density is obtained using the mass action law:

$$p = \frac{n_i^2}{n} = \frac{10^{20}}{10^{16}} = 10^4 \text{ cm}^{-3}$$

The position of the intrinsic energy relative to the midgap energy equals:

$$E_i - \frac{E_c + E_v}{2} = -\frac{3}{4} kT \ln\left(\frac{m_h^*}{m_e^*}\right) = -\frac{3}{4} \times 0.0258 \times \ln \frac{0.81}{1.08} = -5.58 \text{ meV}$$

The position of the Fermi energy relative to the intrinsic energy equals:

$$E_F - E_i = kT \ln\left(\frac{N_d}{n_i}\right) = 0.0258 \times \ln \frac{10^{16}}{10^{10}} = 357 \text{ meV}$$

Problem 2.12 A silicon wafer is doped with 10^{13} cm^{-3} shallow donors and $9 \times 10^{12} \text{ cm}^{-3}$ shallow acceptors. Calculate the electron and hole density at 300 K. Use $n_i = 10^{10} \text{ cm}^{-3}$.

Solution Since there are more donors than acceptors, the resulting material is n-type and the electron density equals the difference between the donor and acceptor density or:

$$n = N_d - N_a = 10^{13} - 9 \times 10^{12} = 10^{12} \text{ cm}^{-3}$$

The hole density is obtained by applying the mass action law:

$$p = \frac{n_i^2}{n} = \frac{10^{20}}{10^{12}} = 10^8 \text{ cm}^{-3}$$

Problem 2.13 The resistivity of a silicon wafer at room temperature is $5 \Omega\text{cm}$. What is the doping density? Find all possible solutions.

Solution Starting with a initial guess that the conductivity is due to electrons with a mobility of $1400 \text{ cm}^2/\text{V}\cdot\text{s}$, the corresponding doping density equals:

$$N_d \cong n = \frac{1}{q\mathbf{m}_n \mathbf{r}} = \frac{1}{1.6 \times 10^{-19} \times 1400 \times 5} = 8.9 \times 10^{14} \text{ cm}^{-3}$$

The mobility corresponding to this doping density equals

$$\mathbf{m}_n = \mathbf{m}_{\min} + \frac{\mathbf{m}_{\max} - \mathbf{m}_{\min}}{1 + \left(\frac{N_d}{N_r}\right)^a} = 1366 \text{ cm}^2/\text{V}\cdot\text{s}$$

Since the calculated mobility is not the same as the initial guess, this process must be repeated until the assumed mobility is the same as the mobility corresponding to the calculated doping density, yielding:

$$N_d = 9.12 \times 10^{14} \text{ cm}^{-3} \text{ and } \mathbf{m}_n = 1365 \text{ cm}^2/\text{V}\cdot\text{s}$$

For p-type material one finds:

$$N_a = 2.56 \times 10^{15} \text{ cm}^{-3} \text{ and } \mathbf{m}_p = 453 \text{ cm}^2/\text{V}\cdot\text{s}$$

Problem 2.14 How many phosphorus atoms must be added to decrease the resistivity of n-type silicon at room temperature from $1 \Omega\text{-cm}$ to $0.1 \Omega\text{-cm}$. Make sure you include the doping dependence of the mobility. State your assumptions.

Solution Starting with a initial guess that the conductivity is due to electrons with a mobility of $1400 \text{ cm}^2/\text{V}\cdot\text{s}$, the corresponding doping density corresponding to the initial resistivity of $1 \Omega\text{-cm}$ equals:

$$N_d \approx n = \frac{1}{q\mathbf{m}_h \mathbf{r}} = \frac{1}{1.6 \times 10^{-19} \times 1400 \times 1} = 4.46 \times 10^{15} \text{ cm}^{-3}$$

The mobility corresponding to this doping density equals

$$\mathbf{m}_h = \mathbf{m}_{\min} + \frac{\mathbf{m}_{\max} - \mathbf{m}_{\min}}{1 + \left(\frac{N_d}{N_r}\right)^a} = 1274 \text{ cm}^2/\text{V}\cdot\text{s}$$

Since the calculated mobility is not the same as the initial guess, this process must be repeated until the assumed mobility is the same as the mobility corresponding to the calculated doping density, yielding:

$$N_{d,initial} = 4.94 \times 10^{15} \text{ cm}^{-3} \text{ and } \mathbf{m}_h = 1265 \text{ cm}^2/\text{V}\cdot\text{s}$$

Repeating this procedure for a resistivity of $0.1 \Omega\text{-cm}$ one find the final doping density to be

$$N_{d,final} = 8.08 \times 10^{16} \text{ cm}^{-3} \text{ and } \mathbf{m}_h = 772 \text{ cm}^2/\text{V}\cdot\text{s}$$

The added density of phosphorous atoms therefore equals

$$N_{d, added} = 4.94 \times 10^{15} - = 7.59 \times 10^{16} \text{ cm}^{-3}$$

Problem 2.18 Consider the problem of finding the doping density, which results in the maximum possible resistivity of silicon at room temperature. ($n_i = 10^{10} \text{ cm}^{-3}$, $\mathbf{m}_h = 1400 \text{ cm}^2/\text{V-sec}$ and $\mathbf{m}_p = 450 \text{ cm}^2/\text{V-sec}$.)

Should the silicon be doped at all or do you expect the maximum resistivity when dopants are added?

If the silicon should be doped, should it be doped with acceptors or donors (assume that all dopant are shallow).

Calculate the maximum resistivity, the corresponding electron and hole density and the doping density.

Solution Since the mobility of electrons is larger than that of holes, one expects the resistivity to initially decrease as acceptors are added to intrinsic silicon.

The maximum resistivity (or minimum conductivity) is obtained from:

$$\frac{d\mathbf{S}}{dn} = q \frac{d(\mathbf{m}_h n + \mathbf{m}_p p)}{dn} = q \frac{d(\mathbf{m}_h n + \mathbf{m}_p n_i^2 / n)}{dn} = 0$$

which yields:

$$n = \sqrt{\frac{\mathbf{m}_p}{\mathbf{m}_h}} n_i = 0.57 n_i = 5.7 \times 10^9 \text{ cm}^{-3}$$

The corresponding hole density equals $p = 1.76 n_i = 1.76 \times 10^9 \text{ cm}^{-3}$ and the amount of acceptors one needs to add equals $N_a = 1.20 n_i = 1.20 \times 10^9 \text{ cm}^{-3}$. The maximum resistivity equals:

$$\mathbf{r}_{\max} = \frac{1}{q(\mathbf{m}_h n + \mathbf{m}_p p)} = \frac{1}{qn_i 1587} = 394 \text{ k}\Omega\text{cm}$$

Problem 2.20 The expression for the Bohr radius can also be applied to the hydrogen-like atom consisting of an ionized donor and the electron provided by the donor. Modify the expression for the Bohr radius so that it applies to this hydrogen-like atom. Calculate the resulting radius of an electron orbiting around the ionized donor in silicon. ($\epsilon_r = 11.9$ and $m_e^* = 0.26 m_0$)

Solution The Bohr radius is obtained from:

$$a_0 = \frac{e_0 h^2 n^2}{p m_0 q^2}$$

However since the electron travel through silicon one has to replace the permittivity of vacuum with the dielectric constant of silicon and the free electron mass with the effective mass for conductivity calculations so that:

$$a_{0, \text{donor in silicon}} = a_0 \frac{\epsilon_r}{m_e^* / m_0} = 529 \times \frac{11.9}{0.26} \text{ pm} = 2.42 \text{ nm}$$

Problem 2.25 Electrons in silicon carbide have a mobility of $1000 \text{ cm}^2/\text{V}\cdot\text{sec}$. At what value of the electric field do the electrons reach a velocity of $3 \times 10^7 \text{ cm/s}$? Assume that the mobility is constant and independent of the electric field. What voltage is required to obtain this field in a 5 micron thick region? How much time do the electrons need to cross the 5 micron thick region?

Solution The electric field is obtained from the mobility and the velocity:

$$E = \frac{m}{v} = \frac{1400}{3 \times 10^7} = 30 \text{ kV/cm}$$

Combined with the length one finds the applied voltage.

$$V = E L = 30,000 \times 5 \times 10^{-4} = 15 \text{ V}$$

The transit time equals the length divided by the velocity:

$$t_r = L/v = 5 \times 10^{-4} / 3 \times 10^7 = 16.7 \text{ ps}$$

Problem 2.26 A piece of silicon has a resistivity which is specified by the manufacturer to be between 2 and 5 Ohm cm. Assuming that the mobility of electrons is $1400 \text{ cm}^2/\text{V}\cdot\text{sec}$ and that of holes is $450 \text{ cm}^2/\text{V}\cdot\text{sec}$, what is the minimum possible carrier density and what is the corresponding carrier type? Repeat for the maximum possible carrier density.

Solution The minimum carrier density is obtained for the highest resistivity and the material with the highest carrier mobility, i.e. the n-type silicon.

The minimum carrier density therefore equals:

$$n = \frac{1}{q m_n r_{\max}} = \frac{1}{1.6 \times 10^{19} \times 1400 \times 5} = 8.92 \times 10^{14} \text{ cm}^{-3}$$

The maximum carrier density is obtained for the lowest resistivity and the material with the lowest carrier mobility, i.e. the p-type silicon.

The maximum carrier density therefore equals:

$$p = \frac{1}{q m_p r_{\min}} = \frac{1}{1.6 \times 10^{19} \times 450 \times 2} = 6.94 \times 10^{15} \text{ cm}^{-3}$$

Problem 2.27 A silicon wafer has a 2-inch diameter and contains 10^{14} cm^{-3} electrons with a mobility of $1400 \text{ cm}^2/\text{V}\cdot\text{sec}$. How thick should the wafer be so that the resistance between the front and back surface equals 0.1 Ohm?

Solution The resistance is given by

$$R = r \frac{L}{A}$$

Where A is the area of the wafer and L the thickness, so that the wafer thickness equals:

$$L = \frac{RA}{r} = \frac{0.1 \times \rho \times (2.54)^2}{44.6} = 0.455 \text{ mm}$$

The resistivity, r , was obtained from:

$$r = \frac{1}{q m_n n} = \frac{1}{1.6 \times 10^{19} \times 1400 \times 10^{14}} = 44.6 \Omega\text{-cm}$$

Problem 2.29 A piece of n-type silicon is doped with 10^{17} cm^{-3} shallow donors. Calculate the density of electrons per unit energy at $kT/2$ above the conduction band edge. $T = 300 \text{ K}$. Calculate the electron energy for which the density of electrons per unit energy has a maximum. What is the corresponding probability of occupancy at that maximum?

Solution

The density of electrons per unit energy at a given energy equals:

$$n(E) = g_c(E)f(E)$$

where

$$g_c(E) = \frac{8\sqrt{2}p m^{3/2}}{h^3} \sqrt{E - E_c} = 1.05 \text{ cm}^{-3}\text{J}^{-1}$$

and

$$f(E) = \frac{1}{1 + \exp \frac{E - E_F}{kT}} = 9.14 \times 10^{-4}$$

The position of the Fermi energy is calculated from the doping density:

$$E_F - E_c = kT \ln \frac{n}{N_c} = kT \ln \frac{N_d}{N_c} = -168 \text{ meV}$$

This last equation is only valid if the semiconductor is non-degenerate, which is a justifiable assumption since the electron density is much smaller than the effective density of states. The Fermi function then becomes:

$$f(E) = \frac{1}{1 + \exp \frac{E - E_F}{kT}} \cong \exp \frac{E_F - E}{kT}$$

And the density of electrons per unit energy can then be further simplified to:

$$n(E) = g_c(E)f(E) \cong \frac{8\sqrt{2}p m^{3/2}}{h^3} \sqrt{E - E_c} \exp \frac{E_F - E}{kT}$$

The maximum is obtained by setting the derivative equal to zero:

$$\frac{dn(E)}{dE} = 0$$

This result in:

$$0 = \frac{1}{2} \frac{1}{\sqrt{E - E_c}} \exp \frac{E_F - E}{kT} - \frac{1}{kT} \sqrt{E - E_c} \exp \frac{E_F - E}{kT}$$

Which can be solved to yield:

$$E = E_{\max} = E_c + kT/2$$

The corresponding probability of occupancy equals the value of the Fermi function calculated above.

Problem 2.30 Phosphorous donor atoms with a concentration of 10^{16} cm^{-3} are added to a piece of silicon. Assume that the phosphorous atoms are distributed homogeneously throughout the silicon. The atomic weight of phosphorous is 31.

- a) What is the sample resistivity at 300 K?
 - b) What proportion by weight does the donor impurity comprise? The density of silicon is 2.33 gram/cm³.
 - c) If $10^{17} \text{ atoms cm}^{-3}$ of boron are included in addition to phosphorous, and distributed uniformly, what is the resulting resistivity and type (i.e., p- or n-type material)?
 - d) Sketch the energy-band diagram under the condition of part c) and show the position of the Fermi energy relative to the valence band edge.
-

Solution

- a) The electron mobility in the silicon equals:

$$\mathbf{m}_n = \mathbf{m}_{\min} + \frac{\mathbf{m}_{\max} - \mathbf{m}_{\min}}{a} = 68.5 + \frac{1414 - 68.5}{1 + \left(\frac{N}{N_r}\right)^{0.711}} \frac{10^{16}}{1 + \left(\frac{10^{16}}{9.2 \times 10^{16}}\right)}$$

$$= 1184 \text{ cm}^2/\text{V-s}$$

$$\mathbf{r} = \frac{1}{\mathbf{s}} = \frac{1}{q(\mathbf{m}_n n + \mathbf{m}_p p)} \approx \frac{1}{q \mathbf{m}_n n} = \frac{1}{1.6 \times 10^{-19} \times 1184 \times 10^{16}} = 0.53 \Omega\text{-cm}$$

b)
$$\frac{\frac{\text{weight}}{\text{volume}}_P}{\frac{\text{weight}}{\text{volume}}_{P+Si}} \approx \frac{Mm_A N_d}{\text{density of Si}}$$

$$= \frac{31 \times 1.6 \times 10^{27} \times 10^3 \times 10^{16}}{2.928} = 2.1 \times 10^{-7}$$

- c) The semiconductor is p-type since $N_a > N_d$

The hole density is obtained from:

$$p = \frac{N_a^+ - N_d^-}{2} + \sqrt{\left(\frac{N_a^+ - N_d^-}{2}\right)^2 + n_i^2}$$
$$= \frac{9 \times 10^{16}}{2} + \sqrt{\left(\frac{9 \times 10^{16}}{2}\right)^2 + (10^{10})^2} = 9 \times 10^{16} \text{ cm}^{-3}$$

and the mobility is calculated from the sum of the donor and acceptor densities

$$\mathbf{m}_p = \mathbf{m}_{\min} + \frac{\mathbf{m}_{\max} - \mathbf{m}_{\min}}{a} =$$
$$1 + \left(\frac{N}{N_r}\right)^{0.719}$$
$$44.9 + \frac{470.5 - 44.9}{1 + \left(\frac{11 \times 10^{16}}{2.23 \times 10^{17}}\right)} = 310.6 \text{ cm}^2/\text{V-s}$$

leading to the conductivity of the material:

$$\mathbf{r} = \frac{1}{\mathbf{s}} = \frac{1}{q(\mathbf{m}_n n + \mathbf{m}_p p)} \approx \frac{1}{q \mathbf{m}_p p} =$$
$$\frac{1}{1.6 \times 10^{-19} \times 310.6 \times 9 \times 10^{16}} = 0.22 \Omega\text{-cm}$$

d)
$$E_F - E_v = kT \ln \frac{N_v}{p} = 0.0259 \ln \frac{1.04 \times 10^{19}}{9 \times 10^{16}} = 123 \text{ meV}$$

Problem 2.31 Find the equilibrium electron and hole concentrations and the location of the Fermi energy relative to the intrinsic energy in silicon at 27 °C, if the silicon contains the following concentrations of shallow dopants.

- a) $1 \times 10^{16} \text{ cm}^{-3}$ boron atoms
- b) $3 \times 10^{16} \text{ cm}^{-3}$ arsenic atoms and $2.9 \times 10^{16} \text{ cm}^{-3}$ boron atoms.

Solution a) Boron atoms are acceptors, therefore $N_a = 10^{16} \text{ cm}^{-3}$. Since these are shallow acceptors and the material is not compensated, degenerate or close to intrinsic, the hole density equals the acceptor density:

$$p \approx 10^{16} \text{ cm}^{-3}$$

Using the mass action law we then find the electron density

$$n = n_i^2/p = 1 \times 10^4 \text{ cm}^{-3}$$

The Fermi energy is then obtained from:

$$E_F - E_i = kT \ln \frac{n}{n_i} = 0.0259 \ln \frac{10^4}{10^{10}} = -357 \text{ meV}$$

b) Arsenic atoms are donors, therefore $N_d = 3 \times 10^{16} \text{ cm}^{-3}$ and $N_a = 2.9 \times 10^{16} \text{ cm}^{-3}$. Since these are shallow acceptors and the material is not degenerate or close to intrinsic, the electron density approximately equals the difference between the donor and acceptor density

$$n \approx N_d - N_a = 10^{15} \text{ cm}^{-3}$$

Using the mass action law we then find the hole density

$$p = n_i^2/n = 1 \times 10^5 \text{ cm}^{-3}$$

The Fermi energy is then obtained from:

$$E_F - E_i = kT \ln \frac{n}{n_i} = 0.0259 \ln \frac{10^{15}}{10^{10}} = 298 \text{ meV}$$

Problem 2.32 The electron concentration in a piece of lightly doped, *n*-type silicon at room temperature varies linearly from 10^{17} cm^{-3} at $x = 0$ to $6 \times 10^{16} \text{ cm}^{-3}$ at $x = 2 \mu\text{m}$. Electrons are supplied to keep this concentration constant with time. Calculate the electron current density in the silicon if no electric field is present. Assume $m_i = 1000 \text{ cm}^2/\text{V}\cdot\text{s}$ and $T = 300 \text{ K}$.

Solution The diffusion current is obtained from:

$$J_n = qD_n \frac{dn}{dx} = 1.6 \times 10^{-19} \times 25.8 \times \frac{10^{17} - 6 \times 10^{16}}{2 \times 10^{-4}} = 828 \text{ A/cm}^2$$

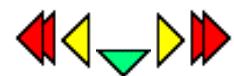
where the diffusion constant D_n is obtained from:

$$D_n = m_i \times V_t = 1000 \times 0.0258 = 25.8 \text{ cm}^2$$

Review Questions

1. What is the difference between a source and a drain of a MOSFET?
2. What is the difference between an n-type and a p-type MOSFET?
3. What is the difference between an enhancement and a depletion MOSFET?
4. Which device has most positive threshold voltage a depletion mode p-type MOSFET or an enhancement mode p-type MOSFET?
5. Which device has the highest drain current in saturation at zero gate voltage, a p-type enhancement MOSFET or an n-type depletion MOSFET?
6. Why is the electron layer of an n-MOSFET called an inversion layer?
7. Why is there no current in a MOSFET when the device is biased in accumulation?
8. What is the difference between the linear and the quadratic MOSFET model?
9. Does the body effect affect CMOS circuits? Explain.
10. How does the oxide thickness and substrate doping affect the threshold voltage of a MOSFET. Plot the threshold voltage as a function of the oxide thickness for different doping densities.
11. What is the advantage of a poly-silicon gate technology?
12. What is channel length modulation?
13. What is punchthrough?
14. Explain the effect of scaling on the different MOSFET parameters as listed in Table 7.7.1

Chapter 7: MOS Field-Effect-Transistors



7.7. Advanced MOSFET issues

- [7.7.1. Channel length modulation](#)
- [7.7.2. Drain induced barrier lowering](#)
- [7.7.3. Punch through](#)
- [7.7.4. Sub-threshold current](#)
- [7.7.5. Field dependent mobility](#)
- [7.7.6. Avalanche breakdown and parasitic bipolar action](#)
- [7.7.7. Velocity saturation](#)
- [7.7.8. Oxide Breakdown](#)
- [7.7.9. Scaling](#)

7.7.1. Channel length modulation



Channel length modulation in a MOSFET is caused by the increase of the depletion layer width at the drain as the drain voltage is increased. This leads to a shorter channel length and an increased drain current. An example is shown in Figure [7.7.1](#). The channel-length-modulation effect typically increases in small devices with low-doped substrates. An extreme case of channel length modulation is punch through where the channel length reduces to zero. Proper scaling can reduce channel length modulation, namely by increasing the doping density as the gate length is reduced.

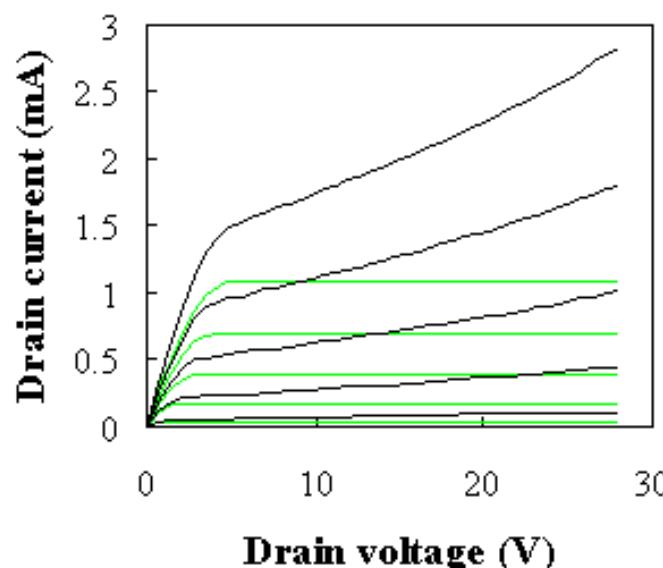


Figure 7.7.1: Current-Voltage characteristics of a MOSFET with and without channel length modulation. ($N_d = 10^{17}$ cm $^{-3}$, $L = 1 \mu\text{m}$)



7.7.2. Drain induced barrier lowering

Drain induced barrier lowering (DIBL) is the effect a voltage of the drain has on the output conductance and measured threshold voltage. This effect occurs in devices where only the gate length is reduced without properly scaling the other dimensions. It is observed as a variation of the measured threshold voltage with reduced gate length. The threshold variation is caused by the increased current with increased drain voltage as the applied drain voltage controls the inversion layer charge at the drain, thereby competing with the gate voltage. This effect is due to the two-dimensional field distribution at the drain end and can typically be eliminated by properly scaling the drain and source depths while increasing the substrate doping density.



7.7.3. Punch through

Punch through in a MOSFET is an extreme case of channel length modulation where the depletion layers around the drain and source regions merge into a single depletion region. The field underneath the gate then becomes strongly dependent on the drain-source voltage, as is the drain current. Punch through causes a rapidly increasing current with increasing drain-source voltage. This effect is undesirable as it increases the output conductance and limits the maximum operating voltage of the device.



7.7.4. Sub-threshold current

The basic assumption of the MOS capacitor analysis in section [6.3.2](#) is that no inversion layer charge exists below the threshold voltage. This leads to zero current below threshold. The actual sub-threshold current is not zero but reduces exponentially below the threshold voltage as:

$$I_D \propto \exp\left(\frac{V_G - V_T}{V_t}\right) \quad (7.7.1)$$

with

$$n = 1 + \frac{1}{2C_{ox}} \sqrt{\frac{q \epsilon_s N_a}{\phi_F}} \quad (7.7.2)$$

The sub-threshold behavior is critical for dynamic circuits since one needs to ensure that no charge leaks through transistors biased below threshold.

7.7.4.1. Derivation of the sub-threshold ideality factor

The charge density below threshold can be expressed as:

$$Q_d \propto \exp\left(\frac{\phi_s}{V_t}\right) \quad (7.7.3)$$

Where the surface potential, ϕ_s , is related to the gate voltage by:

$$V_G = V_{FB} + \phi_s + V_{ox} = V_{FB} + \phi_s + \frac{\sqrt{2q\epsilon_s\phi_s}}{C_{ox}} \quad (7.7.4)$$

The gate voltage, V_G , is therefore related to the surface potential, ϕ_s , by:

$$\frac{dV_G}{d\phi_s} = 1 + \frac{1}{2C_{ox}} \sqrt{\frac{2q\epsilon_s}{\phi_s}} \cong 1 + \frac{1}{2C_{ox}} \sqrt{\frac{q\epsilon_s}{2\phi_F}} = n \quad (7.7.5)$$

Where the surface potential below threshold was approximated to its value, $2\phi_F$, at threshold. The sub-threshold current therefore equals:

$$I_D \propto Q_d \propto \exp\left(\frac{\phi_s}{V_t}\right) \propto \exp\left(\frac{V_G}{nV_t}\right) \quad (7.7.6)$$

7.7.5. Field dependent mobility



The mobility in the inversion layer is distinctly lower than in bulk material. This is due to the fact the electron wavefunction extends into the oxide and the carrier mobility is lowered due to the lower mobility in the oxide. Higher electric fields at the surface - as typically obtained in scaled down devices - push the electron wavefunction even more into the oxide yielding a field dependent mobility. The mobility at the surface, $\mu_{surface}$, varies with the electric field, \mathcal{E} , in the following way:

$$\mu_{surface} \propto \mathcal{E}^{-1/3} \quad (7.7.7)$$

7.7.6. Avalanche breakdown and parasitic bipolar action



As the electric field in the channel is increased, avalanche breakdown occurs in the channel at the drain. This avalanche breakdown increases the current as in a p-n diode (see section 4.5.3 and 2.8). In addition, there is parasitic bipolar action taking place. Holes generated by the avalanche breakdown move from drain to source underneath the inversion layer. This hole current forward biases the source-bulk p-n diode so that now also electrons are injected as minority carriers into the p-type substrate underneath the inversion layer. These electrons arrive at the drain and again create more electron-hole pairs through avalanche multiplication. The positive feedback between the avalanche breakdown and the parasitic bipolar action results in breakdown at lower drain voltage.

7.7.7. Velocity saturation



As devices are reduced in size, the electric field typically also increases and the carriers in the channel have an increased velocity. However at high fields there is no longer a linear relation between the electric field and the velocity as the velocity gradually saturates reaching the saturation velocity. This velocity saturation is caused by the increased scattering rate of highly energetic electrons, primarily due to optical phonon emission. This effect increases the transit time of carriers through the channel. In sub-micron MOSFETs one finds that the average electron velocity is larger than in bulk material so that velocity saturation is not quite as much of a restriction as initially thought.



7.7.8. Oxide Breakdown

As the gate-oxide is scaled down, breakdown of the oxide and oxide reliability becomes more of a concern. Higher fields in the oxide increase the tunneling of carriers from the channel into the oxide. These carriers slowly degrade the quality of the oxide and lead over time to failure of the oxide. This effect is referred to as time dependent destructive breakdown (TDDB).

A simple reduction of the power supply voltage has been used to eliminate this effect. However as gate oxides approach a thickness of 1.5 - 3 nm, carrier tunneling becomes less dependent on the applied electric field so that this problem will require more attention.

Oxides other than silicon dioxide have been considered as alternate oxides and are typically referred to as high-k dielectrics. These oxides have a larger dielectric constant so that the same gate capacitance can be obtained with a thicker oxide. The challenge is to obtain the same stability, reliability and breakdown voltage as silicon dioxide. Oxides of interest include Al_2O_3 , ZrO and TiO .

7.7.9. Scaling



The reduction of the dimensions of a MOSFET has been has dramatic during the last three decades. Starting at a minimum feature length of 10 μm in 1970 the gate length was gradually reduced to 0.15 μm minimum feature size in 2000, resulting in a 13% reduction per year. Proper scaling of MOSFET however requires not only a size reduction of the gate length and width. It also requires a reduction of all other dimensions including the gate/source and gate/drain alignment, the oxide thickness and the depletion layer widths. Scaling of the depletion layer widths also implies scaling of the substrate doping density.

Two types of scaling are common: constant field scaling and constant voltage scaling. Constant field scaling yields the largest reduction in the power delay product of a single transistor. However, it requires a reduction in the power supply voltage as one decreases the minimum feature size. Constant voltage scaling does not have this problem and is therefore the preferred scaling method since it provides voltage compatibility with older circuit technologies. The disadvantage of constant voltage scaling is that the electric field increases as the minimum feature length is reduced. This leads to velocity saturation, mobility degradation, increased leakage currents and lower breakdown voltages.

The scaling of MOSFET device parameters is illustrated by Table 7.7.1 where constant field, constant voltage and constant voltage scaling in the presence of velocity saturation are compared.

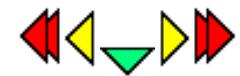
Parameter	Symbol	Constant Field Scaling	Constant Voltage Scaling	Constant Voltage Scaling with velocity saturation
Gate length	L	$1/\alpha$	$1/\alpha$	$1/\alpha$
Gate width	W	$1/\alpha$	$1/\alpha$	$1/\alpha$
Field	\mathcal{E}	1	α	α
Oxide thickness	t_{ox}	$1/\alpha$	$1/\alpha$	$1/\alpha$
Substrate doping	N_a	α^2	α^2	α^2
Gate capacitance	C_G	$1/\alpha$	$1/\alpha$	$1/\alpha$
Oxide capacitance	C_{ox}	α	α	α
Transit time	t_r	$1/\alpha^2$	$1/\alpha^2$	$1/\alpha$
Transit frequency	f_r	α	α^2	α
Voltage	V	$1/\alpha$	1	1
Current	I	$1/\alpha$	α	1
Power	P	$1/\alpha^2$	α	1
Power-delay	$P \triangleq t$	$1/\alpha^3$	$1/\alpha$	$1/\alpha$

Table 7.7.1 : Comparison of the effect of scaling on MOSFET device parameters. Compared are constant field scaling, constant voltage scaling and constant voltage scaling in the presence of velocity overshoot.

Review Questions

1. List three experiments, which can only be explained using quantum mechanics.
2. What is a Rydberg?
3. Name the two primary assumptions of the Bohr model.
4. How do we know that the energy levels in a hydrogen atom are quantized?
5. What two parameters are linked by Gauss's law?
6. What two parameters are linked by Poisson's equation?
7. What is the definition of thermal equilibrium?
8. List the three laws of thermodynamics.
9. Explain in words the meaning of the thermodynamic identity.
10. What is the Fermi function?

Chapter 7: MOS Field-Effect-Transistors



7.6. MOSFET Circuits and Technology

- [7.6.1. MOSFET fabrication process](#)
- [7.6.2. Poly-silicon gate technology](#)
- [7.6.3. CMOS](#)
- [7.6.4. MOSFET Memory](#)

MOSFET circuit technology has dramatically changed over the last three decades. Starting with a ten-micron pMOS process with an aluminum gate and a single metallization layer around 1970, the technology has evolved into a tenth-micron self-aligned-gate CMOS process with up to five metallization levels. The transition from dopant diffusion to ion implantation, from thermal oxidation to oxide deposition, from a metal gate to a poly-silicon gate, from wet chemical etching to dry etching and more recently from aluminum (with 2% copper) wiring to copper wiring has provided vastly superior analog and digital CMOS circuits.

7.6.1. MOSFET fabrication process



The MOSFET fabrication process has evolved dramatically over the years. Around 1970, pMOS circuits with aluminum gate metal and wiring were dominant. The corresponding steps of a typical pMOSFET fabrication process steps are listed in Table 7.6.1.

Lithography step	Process step	Process
1	Field oxide growth Oxide etch Source-drain diffusion	Thermal oxidation HF etch Boron diffusion
2	Oxide etch Gate oxide growth	HF etch Thermal oxidation
3	Via hole etch	HF etch
4	Aluminum metal deposition Aluminum etch Contact anneal and surface state reduction	Evaporation Wet chemical etch Furnace anneal in H_2/N_2

Table 7.6.1:

pMOS process steps

The primary problem at the time was threshold voltage control. Positively charged ions in the oxide decreased the threshold voltage of the devices. p-type MOSFETs were therefore the device of choice despite the lower hole mobility, since they would still be enhancement-type devices even when charge was present. Circuits were still operational at somewhat higher power supply voltages.

Thermal oxidation of the silicon in an oxygen or water vapor atmosphere provided a quality gate oxide with easily controlled thickness. The same process was also used to provide a high-temperature mask for the diffusion process and a passivation and isolation layer.

The oxide was easily removed in hydrofluoric acid (HF), without removing the underlying silicon.

Aluminum was evaporated over the whole wafer and then etched yielding both the gate metal and the metal wiring connecting the devices. A small amount of copper (~2%) was added to make the aluminum more resistant to electromigration. Electromigration is the movement of atoms due to the impact with the electrons carrying the current through the wire. This effect can cause open circuits and is therefore a well-known reliability problem. It typically occurs at points where the local current density is very high, in narrow wires, at corners or when crossing an oxide step. The addition of a small amount of copper provides a more rigid structure within the aluminum and eliminates the effect.

A metal anneal in nitrogen/hydrogen (N_2/H_2) ambient was used to improve the metal-semiconductor contact and to reduce the surface state density at the semiconductor/gate-oxide interface.

Since then the fabrication process has changed as illustrated with Table 7.6.2. Most changes were introduced to provide superior performance, better reliability and higher yield. The most important change has been the reduction of the gate length. A gate length reduction provides a shorter transit time and hence a faster device. In addition, a gate length reduction is typically linked to a reduction of the minimum feature size and therefore yields smaller transistors as well as a larger number of transistors on a chip with a given size. As the technology improved, it was also possible to make larger chips, so that the number of transistors per chip increased even faster. At the same time the wafer size was increased to accommodate the larger chips while reducing the loss due to partial chips at the wafer periphery and to reduce the cost per chip as more chips can be accommodated on a single wafer.

The other changes can be split into process improvements and circuit improvements. The distinction is at times artificial, as circuit improvements typically require new processes.

The key circuit improvement is the use of CMOS circuits, containing both nMOS and pMOS transistors. Early on, the pMOS devices were replaced with nMOS transistors because of the better electron mobility. Enhancement-mode loads were replaced for a while by resistor loads and then depletion-mode loads yielding faster logic circuits with larger operating margins. Analog circuits benefited in similar ways. The use of complementary circuits was first introduced by RCA but did not immediately catch on since the logic circuits were somewhat slower and larger than the then-dominant nMOS depletion logic. It was only when the number of transistors per chip became much larger that the inherent advantages of CMOS circuits, namely the lower power dissipation and larger operating margins became highly desirable. By now the CMOS technology is the dominant technology in the IC industry as the ten-fold reduction of power dissipation largely outweighs the 30%-50% speed reduction and size increase.

The process improvements can in turn be split into those aimed at improving the circuit performance and those improving the manufacturability and reliability. Again the split is somewhat artificial but it is beneficial to understand what factors affect the process changes. The latter group includes CVD deposition, ion implantation, RIE etching, sputtering, planarization and deuterium annealing. The process changes, which improve the circuit performance, are the self-aligned poly-silicon gate process, the silicide gate cap, LOCOS isolation, multilevel wiring and copper wiring.

The self-aligned poly-silicon gate process was introduced before CMOS and marked the beginning of modern day MOSFETs. The self-aligned structure, as further discussed in section 7.6.2, is obtained by using the gate as the mask for the source-drain implant. Since the crystal damage caused by the high-energy ions must be annealed at high temperature (~800 °C), an aluminum gate could no longer be used. Doped poly-silicon was found to be a very convenient gate material as it withstands the high anneal temperature and can be oxidized just like silicon. The self-aligned process lowers the parasitic capacitance between gate and drain and therefore improves the high-frequency performance and switching time. The addition of a silicide layer on top of the gate reduces the gate resistance while still providing a quality implant mask. The self-aligned process also reduced the transistor size and hence increased the density. The field oxide was replaced by a local oxidation isolation structure (LOCOS), where a Si_3N_4 layer is used to prevent the oxidation in the MOSFET region. The oxide provides an implant mask and contact hole mask yielding an even more compact device.

Multilevel wiring is a necessity when one increases the number of transistors per chip since the number of wires increases with the square of the number of transistors and the average wire length increase linearly with the chip size. While multilevel wiring simply consists of a series of metal wiring levels separated by insulators, the multilevel wiring has increasingly become a bottleneck in the fabrication of high-performance circuits. Planarization techniques, as discussed below, and the introduction of copper instead of aluminum-based metals have further increased the wiring density and lowered the wiring resistance.

Initial process and process parameters	Current process and process parameters
10 μm gate length	0.1 μm gate length
1 inch wafers	300 mm wafers
2 x 2 mm chips	1 x 2 cm chips
Thermal oxidation	CVD deposition
Field oxide isolation	LOCOS isolation, trench isolation
Wet chemical etching	Reactive ion etching (RIE)
Diffusion	Ion implantation
PMOS	nMOS, CMOS
Enhancement load, resistor load	Depletion load, complementary load
Aluminum gate	Poly-silicon/Silicide self-aligned gate
Evaporated aluminum wiring with 2% copper	Sputtered copper
One or two metal wiring levels without planarization	Up to six wiring levels with planarization and tungsten plugs
Metal evaporation	Sputtering
Hydrogen anneal	Deuterium anneal

Table 7.6.2: MOS process changes and improvements

Chemical vapor deposition (CDV) of insulating layers is now used instead of thermal oxidation since it does not consume the underlying silicon. Also because there is no limit to the obtainable thickness and since materials other than SiO_2 (for instance Si_3N_4) can be deposited. CDV deposition is also frequently used to deposit refractory metals such as tungsten.

Ion implantation has replaced diffusion because of the superior control and uniformity. Dry etching including plasma etching, reactive ion etching (RIE) and ion beam etching has replaced wet chemical etching. These etch processes provide better etch rate uniformity and control as well as pronounced anisotropic etching. The high etch rate selectivity of wet chemical etching is not obtained with these dry etch techniques, but are well compensated by the better uniformity.

Sputtering of metals has completely replaced evaporation. Sputtering typically provides better adhesion and thickness control. It is also better suited to deposit refractory metals and silicides.

Planarization is the process by which the top surface of the wafer is planarized after each step. The purpose of this planarization process is to provide a flat surface, so that fine-line lithography can be performed at all stages of the fabrication process. The planarization enables high-density multi-layer wiring levels.

Deuterium anneal is a recent modification of the standard hydrogen anneal, which passivates the surface states. The change to deuterium was prompted because it is a heavier isotope of hydrogen. It chemically acts the same way but is less likely to be knocked out of place by the energetic carriers in the inversion layer. The use of deuterium therefore reduces the increase of the surface state density due to hot-electron impact.



7.6.2. Poly-silicon gate technology

An early improvement of the technology was obtained by using a poly-silicon gate, yielding a self-aligned structure which is both compact and has better performance. The poly-silicon gate is used as a mask during the implantation so that the source and drain regions are self-aligned with respect to the gate. This self-alignment structure reduces the device size. In addition, it eliminates the large overlap capacitance between gate and drain, while maintaining a continuous inversion layer between source and drain.

A further improvement of this technique is the use of a low-doped drain (LDD) structure. As an example we consider the structure shown in Figure 7.6.1. Here a first shallow implant is used to contact the inversion layer underneath the gate. The shallow implant causes only a small overlap between the gate and source/drain regions. After adding a sidewall to the gate a second deep implant is added to the first one. This deep implant has a low sheet resistance and adds a minimal series resistance. The combination of the two implants therefore yields a minimal overlap capacitance and low access resistance.

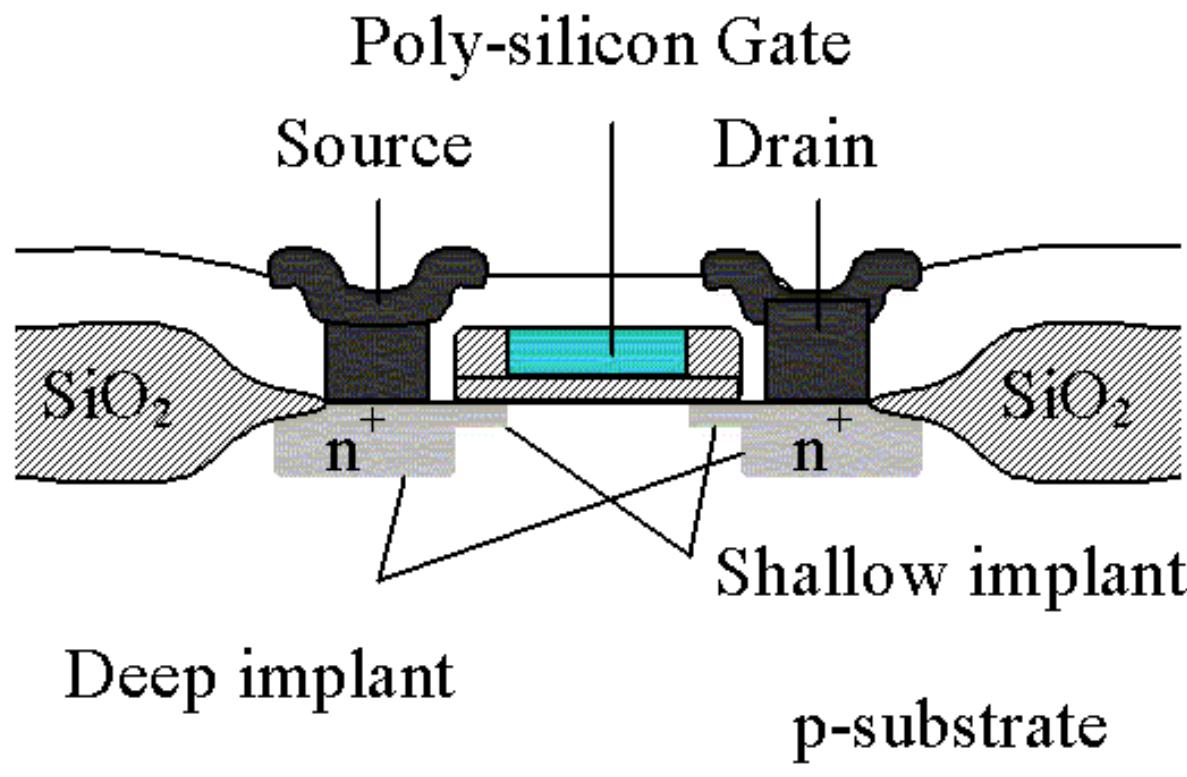


Figure 7.6.1: Cross-sectional view of a self-aligned poly-silicon gate transistor with LOCOS isolation

Shown is also the local oxidation isolation (LOCOS). Typically, there would also be an additional field and channel implant. The field implant increases the doping density under the oxide and thereby increases the threshold voltage of the parasitic transistor formed by the metal wiring on top of the isolation oxide. The channel implant provides an adjustment of the threshold voltage of the transistors. The use of a poly-silicon gate has the disadvantage that the sheet resistance of the gate is much larger than that of a metal gate. This leads to high RC time-constants of long poly-silicon lines. These long RC time-constants are reduced by using silicides (WSi, TaSi, CoSi etc.) instead or on top of poly-silicon. Also by using the poly-silicon only as gate material and not as a wiring level one can further eliminate such RC time delays.



7.6.3. CMOS

Complementary Metal-Oxide-Silicon circuits require an nMOS and pMOS transistor technology on the same substrate. To this end, an n-type well is provided in the p-type substrate. Alternatively one can use a p-well or both an n-type and p-type well in a low-doped substrate. The gate oxide, poly-silicon gate and source-drain contact metal are typically shared between the pMOS and nMOS technology, while the source-drain implants must be done separately. Since CMOS circuits contain pMOS devices, which are affected by the lower hole mobility, CMOS circuits are not faster than their all-nMOS counter parts. Even when scaling the size of the pMOS devices so that they provide the same current, the larger pMOS device has a higher capacitance.

The CMOS advantage is that the output of a CMOS inverter can be as high as the power supply voltage and as low as ground. This large voltage swing and the steep transition between logic high and low yield large operation margins and therefore also a high circuit yield. In addition, there is no power dissipation in either logic state. Instead the power dissipation occurs only when a transition is made between logic states. CMOS circuits are therefore not faster than nMOS circuits but are more suited for very/ultra large-scale integration (VLSI/ULSI).

CMOS circuits have one property, which is very undesirable, namely latchup. Latchup occurs when four alternating p and n-type regions are brought in close proximity. Together they form two bipolar transistors, one npn and one pnp transistor. The base of each transistor is connected to the collector of the other, forming a cross-coupled thyristor-like combination. As a current is applied to the base of one transistor, the current is amplified by the transistor and provided as the base current of the other one. If the product of the current gain of both transistors is larger than unity, the current through both devices increases until the series resistances of the circuit limits the current. Latchup therefore results in excessive power dissipation and faulty logic levels in the gates affected. In principle, this effect can be eliminated by separating n-type and p-type device. A more effective and less space-consuming solution is the use of trenches, which block the minority carrier flow. A deep and narrow trench is etched between all n-type and p-type wells, passivated and refilled with an insulating layer.



7.6.4. MOSFET Memory

MOSFET memory is an important application of MOSFETs. Memory chips contain the largest number of devices per unit area since the transistors are arranged in a very dense regular structure. The generic structure of a memory chip is shown in Figure 7.6.2.

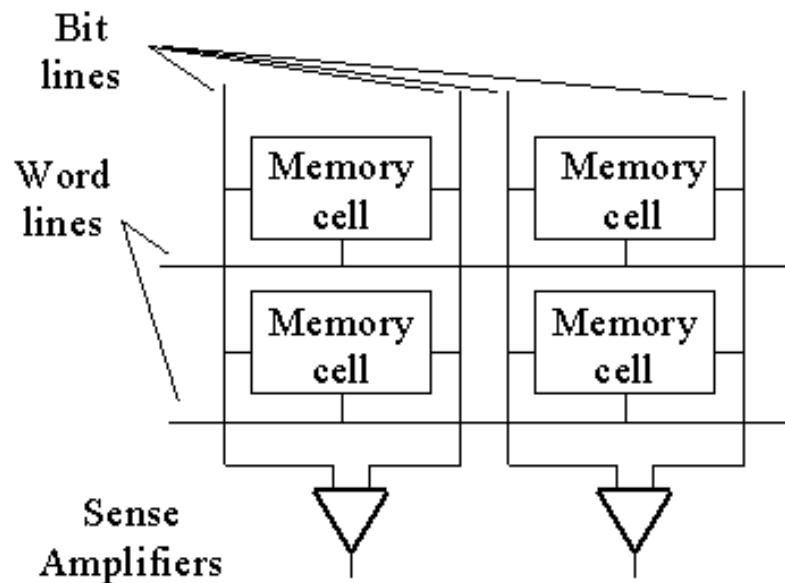


Figure 7.6.2: Arrangement of memory cells into an array

A two dimensional array of memory cells, which store a single bit, are connected through a series of word lines and bit lines. One row of cells is activated by changing the voltage on the corresponding word line. The information is then stored in the cell by applying voltages to the bit lines. During a read operation, the information is retrieved by sensing the voltage on the bit lines with a sense amplifier. A possible implementation of a static random access memory (SRAM) is shown in Figure 7.6.3.

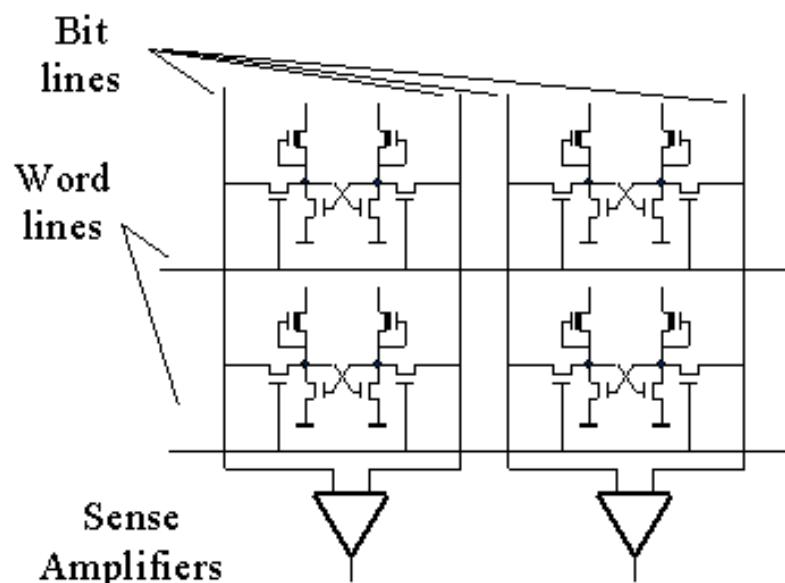


Figure 7.6.3: Static random access memory (SRAM) using a six-transistor cell.

The memory cell consists of a flip-flop and the cells are accessed through two pass transistors connected to the bit lines and controlled by the word line. Depletion mode transistors are shown here as load devices. A common alternate load is an amorphous silicon resistor.

A simpler cell leading to denser memory chips is the dynamic random access memory shown in Figure 7.6.4.

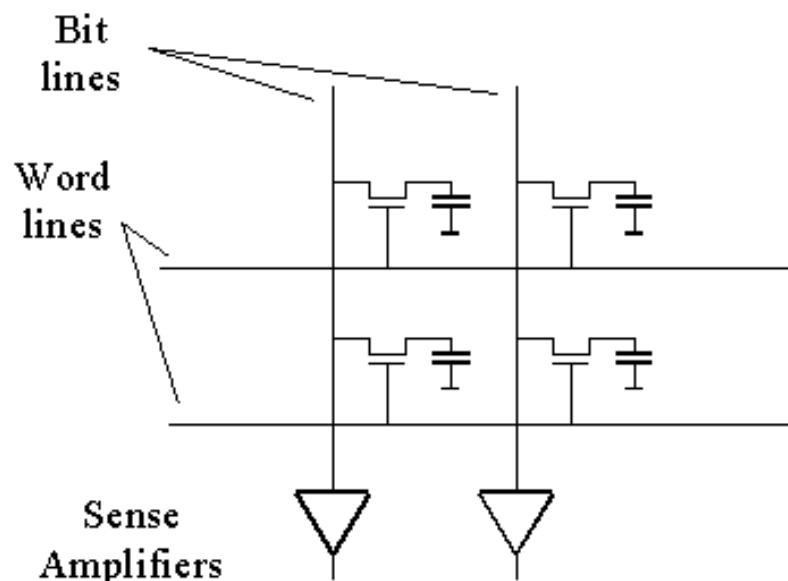


Figure 7.6.4: Dynamic random access memory (DRAM) using a one-transistor cell.

The dynamic cell contains only a single transistor and capacitor. The cell is called dynamic since the information is stored as charge on the capacitor. This charge slowly leaks away so that the cell needs to be refreshed periodically. The reading process is also destructive since the storage capacitor is discharged as a voltage is applied to the word line. Therefore, one has to rewrite the information into all the cells of a given row after reading a single cell from that row. Despite these restrictions, dynamic memory chips represent the largest section of the memory market. The advantage of a higher storage density outweighs all other considerations. Process advances such as the use of a vertical trench, have further increased the density of dynamic memory chips.

As an example we now consider the dynamic memory cell shown in Figure 7.6.5. Shown are the top view and cross-sectional view. The figure illustrates how compact the cell can be by using the gate as the word line of the array and by using a trench capacitor. Also note that the drain of the transistor and one side of the capacitor are merged into one n-type region. The bit lines shown in the top view are placed next to the transistor for clarity. Actual memory cells have the bit lines on top of the transistors as shown in the cross-sectional view. More recent memory cells even have the transistor buried in the trench together with the capacitor.

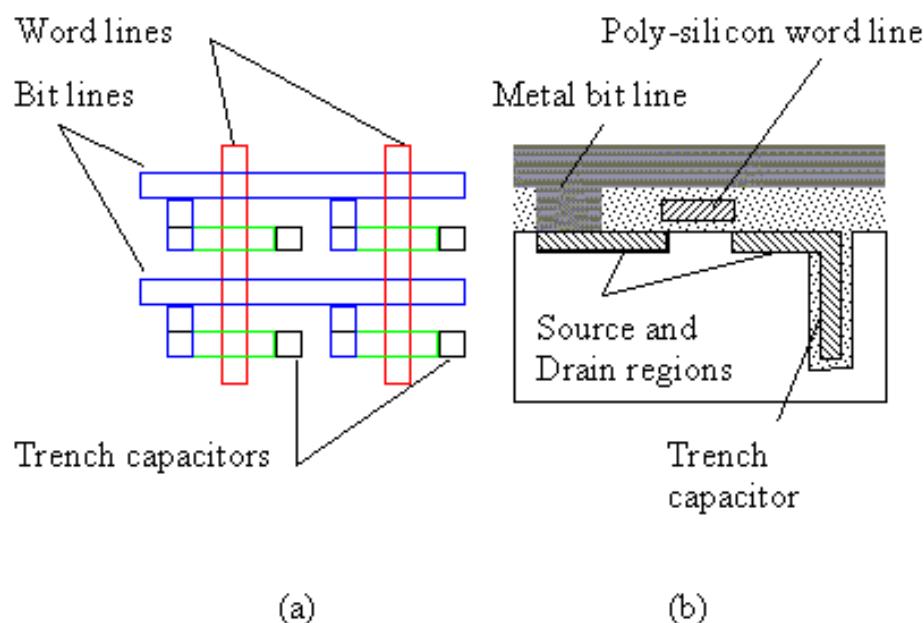


Figure 7.6.5: Dynamic random access memory (DRAM) using a one-transistor cell. (a) top view of four cells and (b) cross-sectional view of one cell.

A critical issue when scaling dynamic memory circuits is the capacitance of the storage capacitor. Scaling of all dimensions would yield a smaller value of the capacitor. However, larger arrays, made possible by scaling the device size, require a larger capacitance. After all, the critical operation in a dynamic memory is the read-out. During read-out, the memory capacitor is connected to the bit line and the charge is now distributed between the memory cell capacitance, the bit line capacitance and the parasitic capacitance of all the devices connected to the bit line. The remaining voltage on the bit line therefore depends on the ratio of the cell capacitance to that of the bit line and connected elements. In large memory chips the voltage would become unacceptably low if the memory capacitance would be scaled down with all other device dimensions. Instead the capacitance of the memory capacitor is kept almost constant from one generation to the next at a value around 1 fF. This value corresponds to the storage of 25,000 electrons at a voltage of 5 V and results in a bit line voltage of a few millivolts.

Chapter 6: MOS Capacitors



6.3. MOS analysis

- [6.3.1. Flatband voltage calculation](#)
- [6.3.2. Inversion layer charge](#)
- [6.3.3. Full depletion analysis](#)
- [6.3.4. MOS Capacitance](#)

6.3.1. Flatband voltage calculation



If there is no charge present in the oxide or at the oxide-semiconductor interface, the flat band voltage simply equals the difference between the gate metal workfunction, Φ_M , and the semiconductor workfunction, Φ_S .

$$V_{FB} = \Phi_M - \Phi_S \quad (6.3.1)$$

The workfunction is the voltage required to extract an electron from the Fermi energy to the vacuum level. This voltage is between three and five Volt for most metals. It should be noted that the actual value of the workfunction of a metal deposited onto silicon dioxide is not exactly the same as that of the metal in vacuum. Figure 6.3.1 provides experimental values for the workfunction of different metals as obtained from a measurement of a MOS capacitor as a function of the measured workfunction in vacuum. The same data is also listed in Table 6.3.1.

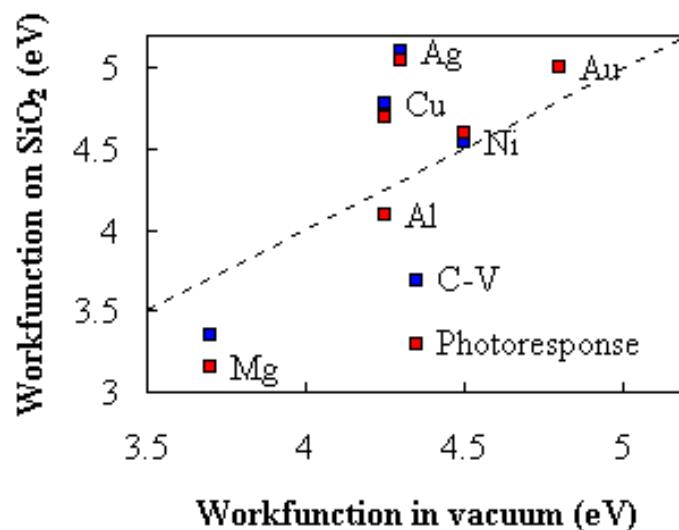


Figure 6.3.1.: Workfunction of Magnesium (Mg), Aluminum (Al), Copper (Cu), Silver (Ag), Nickel (Ni) and Gold (Au) obtained from I-V and C-V measurements on MOS structures as a function of the workfunction of those metals measured in vacuum.

	Ag	Al	Au	Cr	Mg	Ni
Φ_M (in vacuum)	4.3	4.25	4.8	4.25	3.7	4.5
Φ_M (on SiO_2) ($C-V$)	5.1	4.1	5	4.7	3.35	4.55

Table 6.3.1: Workfunction of selected metals as measured in vacuum and as obtained from a $C-V$ measurement on an MOS structure.

The workfunction of a semiconductor, Φ_S , requires some more thought since the Fermi energy varies with the doping type as well as with the doping concentration. This workfunction equals the sum of the electron affinity in the semiconductor, χ , the difference between the conduction band energy and the intrinsic energy divided by the electronic charge and the bulk potential as expressed by the following equation:

$$\Phi_M - \Phi_S = \Phi_M - \chi - \frac{E_g}{2q} - V_t \ln\left(\frac{N_a}{n_i}\right) \quad (6.3.2)$$

For MOS structures with a highly doped poly-silicon gate one must also calculate the workfunction of the gate based on the bulk potential of the poly-silicon.

$$\begin{aligned} \Phi_{poly,S} &= V_t \ln\left(\frac{N_{a,poly}}{N_a}\right) && \text{p - type polysilicon gate} \\ \Phi_{poly,S} &= V_t \ln\left(\frac{n_i^2}{N_{d,poly} N_a}\right) && \text{n - type polysilicon gate} \end{aligned} \quad (6.3.3)$$

Where $N_{a,poly}$ and $N_{d,poly}$ are the acceptor and donor density of the p-type and n-type poly-silicon gate respectively.

The flatband voltage of real MOS structures is further affected by the presence of charge in the oxide or at the oxide-semiconductor interface. The flatband voltage still corresponds to the voltage, which, when applied to the gate electrode, yields a flat energy band in the semiconductor. The charge in the oxide or at the interface changes this flatband voltage. For a charge, Q_i , located at the interface between the oxide and the semiconductor, and a charge density, ρ_{ox} , distributed within the oxide, the flatband voltage is given by:

$$V_{FB} = \Phi_{MS} - \frac{Q_i}{C_{ox}} - \frac{1}{\epsilon_{ox}} \int_0^{t_{ox}} \rho_{ox}(x) x dx \quad (6.3.4)$$

where the second term is the voltage across the oxide due to the charge at the oxide-semiconductor interface and the third term is due to the charge density in the oxide.

The actual calculation of the flatband voltage is further complicated by the fact that charge can move within the oxide. The charge at the oxide-semiconductor interface due to surface states also depends on the position of the Fermi energy.

Since any additional charge affects the flatband voltage and thereby the threshold voltage, great care has to be taken during fabrication to avoid the incorporation of charged ions as well as creation of surface states.

Example 6.1	Calculate the flatband voltage of a silicon nMOS capacitor with a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$ and an aluminum gate ($\Phi_M = 4.1 \text{ V}$). Assume there is no fixed charge in the oxide or at the oxide-silicon interface.
-------------	---

Solution

The flatband voltage equals the work function difference since there is no charge in the oxide or at the oxide-semiconductor interface.

$$V_{FB} = \Phi_{MS} = \Phi_M - \chi - \frac{E_g}{2q} - V_t \ln \frac{N_a}{n_i}$$

$$= 4.1 - 4.05 - 0.56 - 0.026 \times \ln \frac{10^{17}}{10^{10}} = -0.93 \text{ V}$$

The flatband voltages for nMOS and pMOS capacitors with an aluminum or a poly-silicon gate are listed in the table below.

	Aluminum	p^+ poly	n^+ poly
nMOS	-0.93 eV	0.14 eV	-0.98 eV
pMOS	-0.09 eV	0.98 eV	-0.14 eV

6.3.2. Inversion layer charge



The basis assumption as needed for the derivation of the MOSFET models is that the inversion layer charge is proportional with the applied voltage. In addition, the inversion layer charge is zero at and below the threshold voltage as described by:

$$Q_{inv} = C_{ox}(V_G - V_T) \quad (6.3.5)$$

$$Q_{inv} = 0$$

The linear proportionality can be explained by the fact that a gate voltage variation causes a charge variation in the inversion layer. The proportionality constant between the charge and the applied voltage is therefore expected to be the gate oxide capacitance. This assumption also implies that the inversion layer charge is located exactly at the oxide-semiconductor interface.

Because of the energy band gap of the semiconductor separating the electrons from the holes, the electrons can only exist if the p-type semiconductor is first depleted. The voltage at which the electron inversion-layer forms is referred to as the threshold voltage.

To justify this assumption we now examine a comparison of a numeric solution with equation (6.3.4) as shown in Figure 6.3.2.

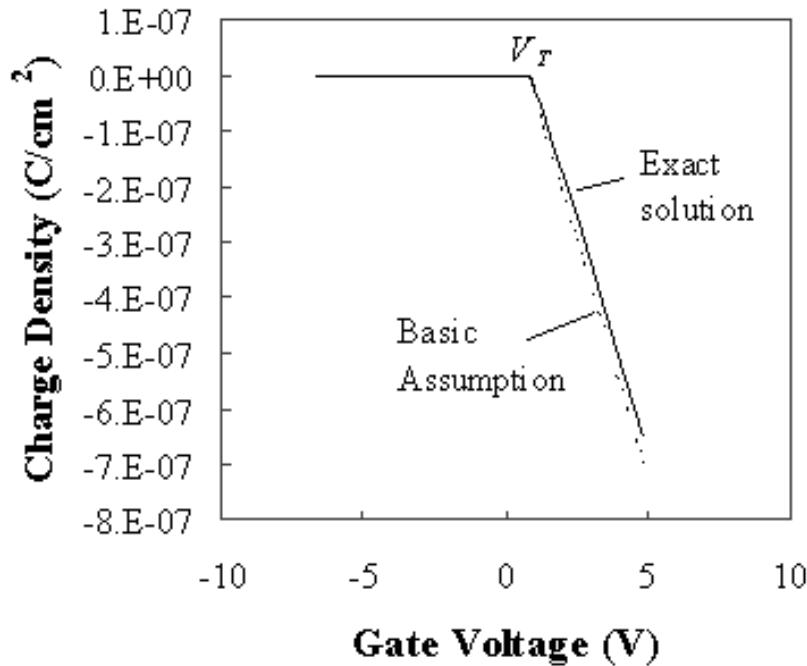


Figure 6.3.2.: Charge density due to electrons in the inversion layer of an MOS capacitor. Compared are the analytic solution (solid line) and equation (6.3.5) (dotted line) for $N_a = 10^{17} \text{ cm}^{-3}$ and $t_{\text{ox}} = 20 \text{ nm}$. 

While there is a clear difference between the curves, the difference is small. We will therefore use our basic assumption when deriving the different MOSFET models since it dramatically simplifies the derivation, be it while losing some accuracy.

6.3.3. Full depletion analysis

We now derive the MOS parameters at threshold with the aid of Figure 6.3.3. To simplify the analysis we make the following assumptions: 1) we assume that we can use the full depletion approximation and 2) we assume that the inversion layer charge is zero below the threshold voltage. Beyond the threshold voltage we assume that the inversion layer charge changes linearly with the applied gate voltage.

The derivation starts by examining the charge per unit area in the depletion layer, Q_d . As can be seen in Figure 6.3.3 (a), this charge is given by:

$$Q_d = -qN_a x_d \quad (6.3.6)$$

Where x_d is the depletion layer width and N_a is the acceptor density in the substrate. Integration of the charge density then yields the electric field distribution shown in Figure 6.3.3 (b). The electric field in the semiconductor at the interface, \mathcal{E}_s , and the field in the oxide equal, \mathcal{E}_{ox} :

$$\mathcal{E}_s = \frac{qN_a x_d}{\epsilon_s} \quad (6.3.7)$$

The electric field changes abruptly at the oxide-semiconductor interface due to the difference in the dielectric constant. At a silicon/SiO₂ interface the field in the oxide is about three times larger since the dielectric constant of the oxide ($\epsilon_{ox} = 3.9 \epsilon_0$) is about one third that of silicon ($\epsilon_s = 11.9 \epsilon_0$). The electric field in the semiconductor changes linearly due to the constant doping density and is zero at the edge of the depletion region.

The potential shown in Figure 6.3.3 (c) is obtained by integrating the electric field. The potential at the surface, ϕ_s , equals:

$$\phi_s = \frac{q N_a x_d^2}{2 \epsilon_s} \quad (6.3.8)$$

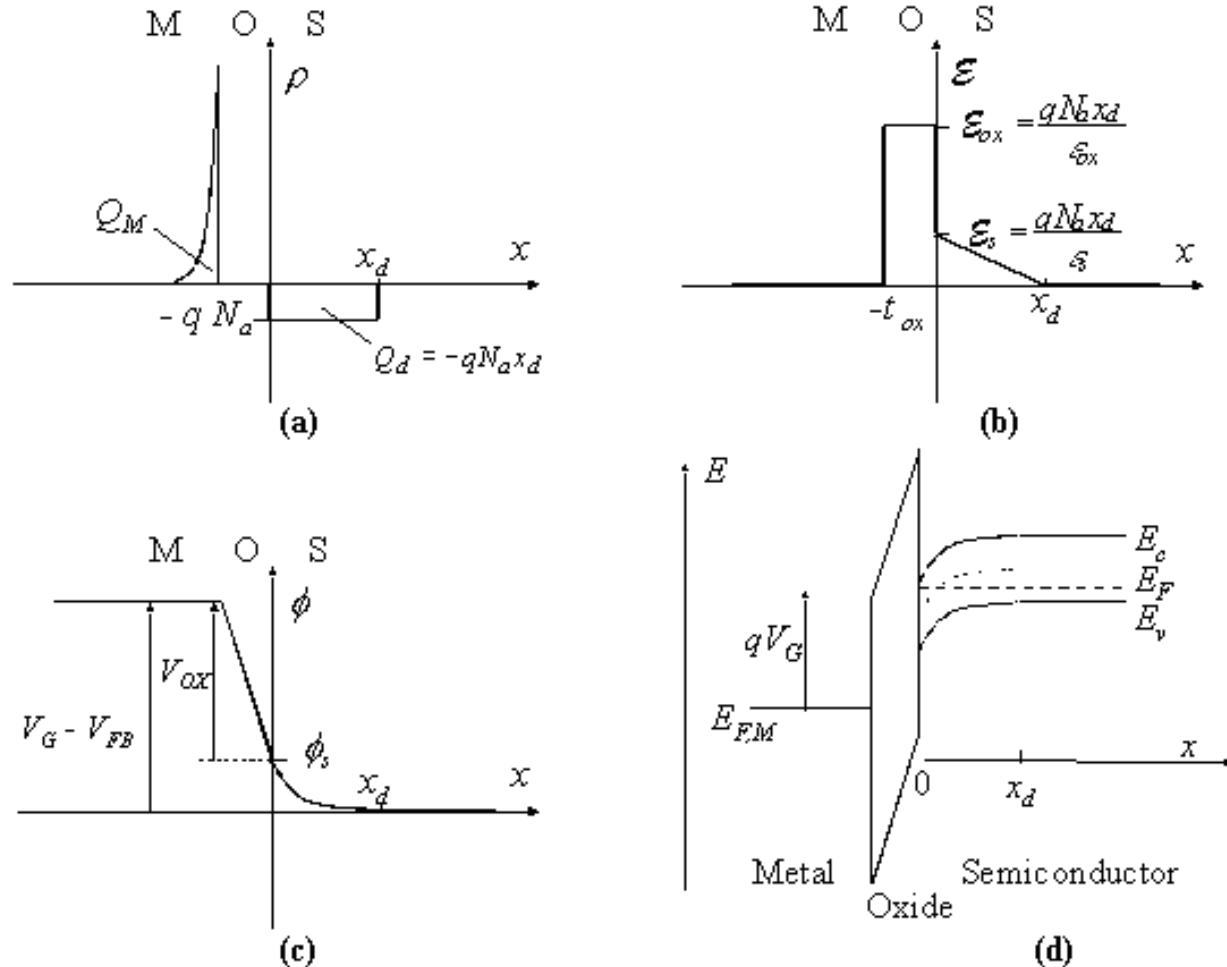


Figure 6.3.3: Electrostatic analysis of an MOS structure. Shown are (a) the charge density, (b) the electric field, (c) the potential and (d) the energy band diagram for an n-MOS structure biased in depletion.

The calculated field and potential is only valid in depletion. In accumulation, there is no depletion region and the full depletion approximation does not apply. In inversion, there is an additional charge in the inversion layer, Q_{inv} . This charge increases gradually as the gate voltage is increased. However, this charge is only significant once the electron density at the surface exceeds the hole density in the substrate, N_a . We therefore define the threshold voltage as the gate voltage for which the electron density at the surface equals N_a . This corresponds to the situation where the total potential across the surface equals twice the bulk potential, ϕ_F .

$$\phi_F = V_t \ln \frac{N_a}{n_i} \quad (6.3.9)$$

The depletion layer in depletion is therefore restricted to this potential range:

$$x_d = \sqrt{\frac{2 \epsilon_s \phi}{q N_a}}, \text{ for } 0 \leq \phi_s \leq 2 \phi_F \quad (6.3.10)$$

For a surface potential larger than twice the bulk potential, the inversion layer charge change increases exponentially with the surface potential. Consequently, an increased gate voltage yields an increased voltage across the oxide while the surface potential remains almost constant. We will therefore assume that the surface potential and the depletion layer width at threshold equal those in inversion. The corresponding expressions for the depletion layer charge at threshold, $Q_{d,T}$, and the depletion layer width at threshold, $x_{d,T}$, are:

$$Q_{d,T} = -q N_a x_{d,T} \quad (6.3.11)$$

$$x_{d,T} = \sqrt{\frac{2 \epsilon_s (2 \phi_F)}{q N_a}} \quad (6.3.12)$$

Beyond threshold, the total charge in the semiconductor has to balance the charge on the gate electrode, Q_M , or:

$$Q_M = -(Q_d + Q_{inv}) \quad (6.3.13)$$

where we define the charge in the inversion layer as a quantity which needs to be determined but should be consistent with our basic assumption. This leads to the following expression for the gate voltage, V_G :

$$V_G = V_{FB} + \phi_s + \frac{Q_M}{C_{ox}} = V_{FB} + \phi_s - \frac{Q_d + Q_{inv}}{C_{ox}} \quad (6.3.14)$$

In depletion, the inversion layer charge is zero so that the gate voltage becomes:

$$V_G = V_{FB} + \phi_s + \frac{\sqrt{2 \epsilon_s q N_a \phi_s}}{C_{ox}}, \text{ for } 0 \leq \phi_s \leq 2 \phi_F \quad (6.3.15)$$

while in inversion this expression becomes:

$$V_G = V_{FB} + 2 \phi_F + \frac{\sqrt{4 \epsilon_s q N_a \phi_F}}{C_{ox}} - \frac{Q_{inv}}{C_{ox}} \stackrel{\Delta}{=} V_T - \frac{Q_{inv}}{C_{ox}} \quad (6.3.16)$$

the third term in (6.3.16) states our basic assumption, namely that any change in gate voltage beyond the threshold requires a change of the inversion layer charge. From the second equality in equation (6.3.16), we then obtain the threshold voltage or:

$$V_T = V_{FB} + 2 \phi_F + \frac{\sqrt{4 \epsilon_s q N_a \phi_F}}{C_{ox}} \quad (6.3.17)$$

Example 6.2	Calculate the threshold voltage of a silicon nMOS capacitor with a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$, a 20 nm thick oxide ($\epsilon_{ox} = 3.9 \epsilon_0$) and an aluminum gate ($\Phi_M = 4.1 \text{ V}$). Assume there is no fixed charge in the oxide or at the oxide-silicon interface.
-------------	---

Solution

The threshold voltage equals:

$$\begin{aligned}
 V_T &= V_{FB} + 2\phi_F + \frac{\sqrt{4\epsilon_s q N_a \phi_F}}{C_{ox}} \\
 &= -0.93 + 2 \times 0.42 \\
 &+ \frac{\sqrt{4 \times 11.9 \times 8.85 \times 10^{-14} \times 1.6 \times 10^{-19} \times 10^{17} \times 0.42}}{3.9 \times 8.85 \times 10^{-14} / 20 \times 10^{-7}} \\
 &= -0.09 \text{ V}
 \end{aligned}$$

Where the flatband voltage was already calculated in example [6.1](#). The threshold voltage voltages for nMOS and pMOS capacitors with an aluminum or a poly-silicon gate are listed in the table below.

	Aluminum	p ⁺ poly	n ⁺ poly
nMOS	-0.09 eV	0.98 eV	-0.14 eV
pMOS	-0.93 eV	0.14 eV	-0.98 eV

6.3.4. MOS Capacitance



[6.3.4.1. Simple capacitance model](#)

[6.3.4.2. Calculation of the flat band capacitance](#)

[6.3.4.3. Deep depletion capacitance](#)

[6.3.4.4. Experimental results and comparison with theory](#)

[6.3.4.5. Non-Ideal effects in MOS capacitors](#)

Capacitance voltage measurements of MOS capacitors provide a wealth of information about the structure, which is of direct interest when one evaluates an MOS process. Since the MOS structure is simple to fabricate, the technique is widely used.

To understand capacitance-voltage measurements one must first be familiar with the frequency dependence of the measurement. This frequency dependence occurs primarily in inversion since a certain time is needed to generate the minority carriers in the inversion layer. Thermal equilibrium is therefore not immediately obtained.

The low frequency or quasi-static measurement maintains thermal equilibrium at all times. This capacitance is the ratio of the change in charge to the change in gate voltage, measured while the capacitor is in equilibrium. A typical measurement is performed with an electrometer, which measured the charge added per unit time as one slowly varies the applied gate voltage.

The high frequency capacitance is obtained from a small-signal capacitance measurement at high frequency. The bias voltage on the gate is varied slowly to obtain the capacitance versus voltage. Under such conditions, one finds that the charge in the inversion layer does not change from the equilibrium value corresponding to the applied DC voltage. The high frequency capacitance therefore reflects only the charge variation in the depletion layer and the (rather small) movement of the inversion layer charge.

In this section, we first derive the simple capacitance model, which is based on the full depletion approximation and our basic assumption. The comparison with the exact low frequency capacitance will reveal that the largest error occurs at the flatband voltage. We therefore derive the exact flatband capacitance using the linearized Poisson's equation. Then we discuss the full exact analysis followed by a discussion of deep depletion as well as the non-ideal effects in MOS capacitors.

6.3.4.1. Simple capacitance model

The capacitance of an MOS capacitor is obtained using the same assumptions as those listed in section [6.3.3](#). The MOS structure is treated as a series connection of two capacitors: the capacitance of the oxide and the capacitance of the depletion layer.

In accumulation, there is no depletion layer. The remaining capacitor is the oxide capacitance, so that the capacitance equals:

$$C_{LF} = C_{HF} = C_{ox}, \text{ for } V_G \leq V_{FB} \quad (6.3.18)$$

In depletion, the MOS capacitance is obtained from the series connection of the oxide capacitance and the capacitance of the depletion layer, or:

$$C_{LF} = C_{HF} = \frac{1}{\frac{1}{C_{ox}} + \frac{x_d}{\epsilon_s}}, \text{ for } V_{FB} \leq V_G \leq V_T \quad (6.3.19)$$

where x_d is the variable depletion layer width which is calculated from:

$$x_d = \sqrt{\frac{2\epsilon_s \phi_s}{qN_d}} \quad (6.3.20)$$

In order to find the capacitance corresponding to a specific value of the gate voltage we also need to use the relation between the potential across the depletion region and the gate voltage, given by:

$$V_G = V_{FB} + \phi_s + \frac{\sqrt{2\epsilon_s q N_d \phi_s}}{C_{ox}}, \text{ for } 0 \leq \phi_s \leq 2\phi_F \quad (6.3.15)$$

In inversion, the capacitance becomes independent of the gate voltage. The low frequency capacitance equals the oxide capacitance since charge is added to and removed from the inversion layer. The high frequency capacitance is obtained from the series connection of the oxide capacitance and the capacitance of the depletion layer having its maximum width, $x_{d,T}$. The capacitances are given by:

$$C_{LF} = C_{ox} \text{ and } C_{HF} = \frac{1}{\frac{1}{C_{ox}} + \frac{x_{d,T}}{\epsilon_s}}, \text{ for } V_G \geq V_T \quad (6.3.21)$$

The capacitance of an MOS capacitor as calculated using the simple model is shown in Figure [6.3.4](#). The dotted lines represent the simple model while the solid line corresponds to the low frequency capacitance as obtained from the exact analysis.

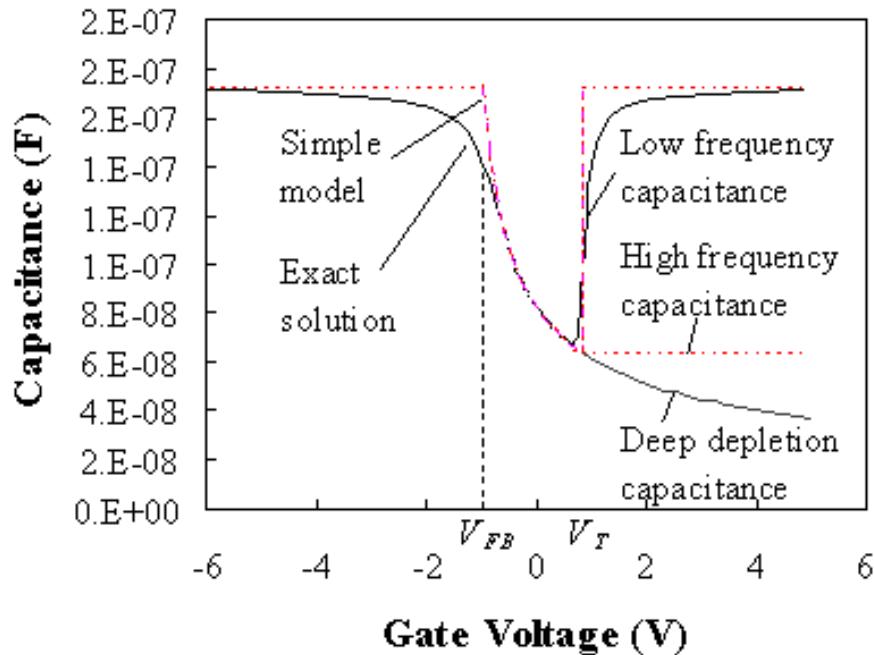


Figure 6.3.4 : Low frequency capacitance of an nMOS capacitor. Shown are the exact solution for the low frequency capacitance (solid line) and the low and high frequency capacitance obtained with the simple model (dotted lines). $N_a = 10^{17} \text{ cm}^{-3}$ and $t_{\text{ox}} = 20 \text{ nm}$. 

6.3.4.2. Calculation of the flat band capacitance

The simple model predicts that the flatband capacitance equals the oxide capacitance. However, the comparison with the exact solution of the low frequency capacitance as shown in Figure 6.3.4 reveals that the error can be substantial. The reason for this is that we have ignored any charge variation in the semiconductor. We will therefore now derive the exact flatband capacitance.

To derive the flatband capacitance including the charge variation in the semiconductor, we first linearize Poisson's equation. Since the potential across the semiconductor at flatband is zero, we expect the potential to be small as we vary the gate voltage around the flatband voltage. Poisson's equation can then be simplified to:

$$\frac{d^2 \phi}{dx^2} = \frac{q}{\epsilon_s} (N_a^+ - p) = \frac{qN_a}{\epsilon_s} (1 - e^{-\frac{\phi}{V_t}}) \cong \frac{qN_a}{\epsilon_s} \frac{\phi}{V_t} \quad (6.3.22)$$

Charge due to ionized donors or electrons were eliminated, since neither are present in a p-type semiconductor around flatband. The linearization is obtained by replacing the exponential function by the first two terms of its Taylor series expansion. The solution to this equation is:

$$\phi = \phi_s e^{-\frac{x}{L_D}}, \text{ with } L_D = \sqrt{\frac{\epsilon_s V_t}{qN_a}} \quad (6.3.23)$$

Where ϕ_s is the potential at the surface of the semiconductor and L_D is called the Debye length. The solution of the potential enables the derivation of the capacitance of the semiconductor under flatband conditions, or:

$$C_{s,FB} = \frac{dQ_s}{d\phi_s} = \frac{d(\epsilon_s \frac{\phi_s}{L_D})}{d\phi_s} = \frac{\epsilon_s}{L_D} \quad (6.3.24)$$

The flatband capacitance of the MOS structure at flatband is obtained by calculating the series connection of the oxide capacitance and the capacitance of the semiconductor, yielding:

$$C_{FB} = \frac{1}{\frac{1}{C_{ox}} + \frac{L_D}{\epsilon_s}} \quad (6.3.25)$$

Example 6.3	Calculate the oxide capacitance, the flatband capacitance and the high frequency capacitance in inversion of a silicon nMOS capacitor with a substrate doping $N_a = 10^{17} \text{ cm}^{-3}$, a 20 nm thick oxide ($\epsilon_{ox} = 3.9 \epsilon_0$) and an aluminum gate ($\Phi_M = 4.1 \text{ V}$).
-------------	---

Solution	<p>The oxide capacitance equals:</p> $C_{ox} = \frac{\epsilon_{ox}}{t_{ox}} = \frac{3.9 \times 8.85 \times 10^{-14}}{2 \times 10^{-6}} = 173 \text{ nF/cm}^2$ <p>The flatband capacitance equals:</p> $C_{FB} = \frac{1}{\frac{1}{C_{ox}} + \frac{L_D}{\epsilon_s}}$ $= \frac{1}{\frac{1}{173 \times 10^{-9}} + \frac{1.3 \times 10^{-6}}{11.9 \times 8.85 \times 10^{-14}}} = 142 \text{ nF/cm}^2$ <p>where the Debye length is obtained from:</p> $L_D = \sqrt{\frac{\epsilon_s V_t}{q N_a}} = \sqrt{\frac{11.9 \times 8.85 \times 10^{-14} \times 0.0259}{1.6 \times 10^{-19} \times 10^{17}}} = 13 \text{ nm}$ <p>The high frequency capacitance in inversion equals:</p> $C_{HF,inv} = \frac{1}{\frac{1}{C_{ox}} + \frac{x_{d,T}}{\epsilon_s}}$ $= \frac{1}{\frac{1}{173 \times 10^{-9}} + \frac{1.05 \times 10^{-5}}{11.9 \times 8.85 \times 10^{-14}}} = 63 \text{ nF/cm}^2$
----------	---

and the depletion layer width at threshold equals:

$$x_{d,T} = \sqrt{\frac{2 \varepsilon_s (2 \phi_F)}{q N_a}}$$

$$= \sqrt{\frac{2 \times 11.9 \times 8.85 \times 10^{-14} \times 2 \times 0.419}{1.6 \times 10^{-19} \times 10^{17}}} = 105 \text{ nm}$$

The bulk potential, ϕ_F , was already calculated in example 6.1

6.3.4.3. Deep depletion capacitance

Deep depletion occurs in an MOS capacitor when measuring the high-frequency capacitance while sweeping the gate voltage "quickly". Quickly means that the gate voltage must be changed fast enough so that the structure is not in thermal equilibrium. One then observes that, when ramping the voltage from flatband to threshold and beyond, the inversion layer is not or only partially formed. This occurs since the generation of minority carriers can not keep up with the amount needed to form the inversion layer. The depletion layer therefore keeps increasing beyond its maximum thermal equilibrium value, $x_{d,T}$ resulting in a capacitance which further decreases with voltage.

The time required to reach thermal equilibrium can be estimated by taking the ratio of the total charge in the inversion layer to the thermal generation rate of minority carriers. A complete analysis should include both the surface generation rate as well as generation in the depletion layer and the quasi-neutral region. A good approximation is obtained by considering only the generation rate in the depletion region in deep depletion, $x_{d,dd}$. This yields the following equation:

$$\text{time} = \frac{|\mathcal{Q}_{inv}|}{qG(x_{d,dd} + L_n)} = \frac{C_{ox}(V_G - V_T)}{\frac{qn_i}{2\tau} \left(\sqrt{\frac{2 \varepsilon_s \phi_{s,dd}}{q N_a}} + \sqrt{\mu_n V_t \tau} \right)} \quad (6.3.26)$$

where the generation in the depletion layer was assumed to be constant. The rate of change required to observe deep depletion is then obtained from:

$$\frac{dV_G}{dt} > \frac{qn_i}{2C_{ox}} \sqrt{\frac{\mu_n V_t}{\tau}} \quad (6.3.27)$$

This equation predicts that deep depletion is less likely at higher ambient temperature, since the intrinsic carrier density n_i increases exponentially with temperature. The intrinsic density also decreases exponentially with the energy bandgap. Therefore, MOS structures made with wide bandgap materials (for instance 6H-SiC for which $E_g = 3 \text{ eV}$), have an extremely pronounced deep depletion effect.

In silicon MOS capacitors, one finds that the occurrence of deep depletion can be linked to the minority carrier lifetime. Structures with a long (0.1 ms) lifetime require a few seconds to reach thermal equilibrium which results in a pronounced deep depletion effect at room temperature, structures with a short (1 ms) lifetime do not show this effect.

Carrier generation due to light will increase the generation rate beyond the thermal generation rate, which we assumed above and reduce the time needed to reach equilibrium. Deep depletion measurements are therefore done in the dark.

6.3.4.4. Experimental results and comparison with theory

As an example, we show below the measured low frequency (quasi-static) and high frequency capacitance-voltage curves of an MOS capacitor. The capacitance was measured in the presence of ambient light as well as in the dark as explained in Figure 6.3.5.

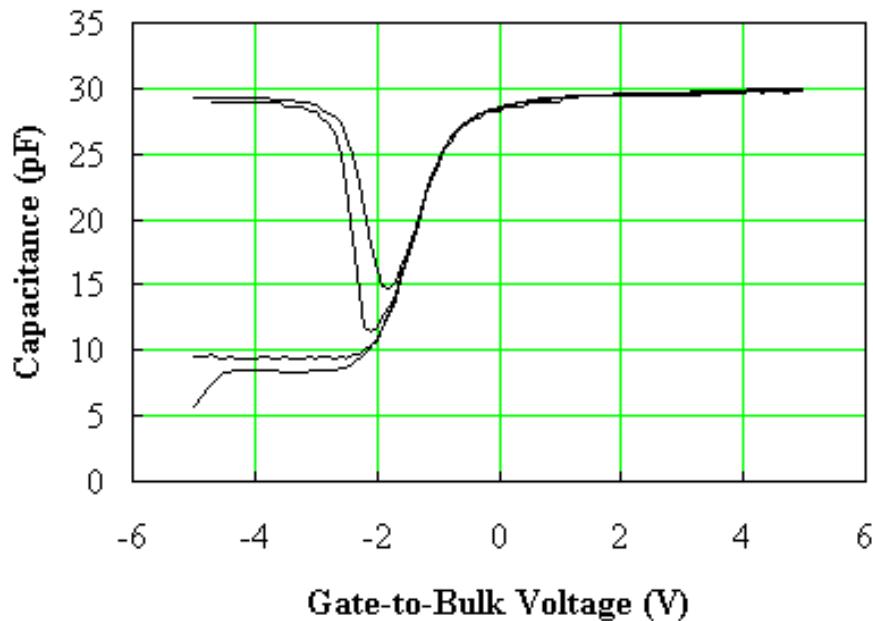


Figure 6.3.5 : Low frequency (quasi-static) and high frequency capacitance measurement of a pMOS capacitor. Shown are, from top to bottom, the low frequency capacitance measured in the presence of ambient light (top curve), the low frequency capacitance measured in the dark, the high frequency capacitance measured in the presence of ambient light and the high frequency capacitance measured in the dark (bottom curve). All curves were measured from left to right. The MOS parameters are $N_a - N_d = 4 \times 10^{15} \text{ cm}^{-3}$ and $t_{\text{ox}} = 80 \text{ nm}$. The device area is 0.0007 cm^2 .

Figure 6.3.5 illustrates some of the issues when measuring the capacitance of an MOS capacitor. First, one should measure the devices in the dark. The presence of light causes carrier generation in the semiconductor, which affects the measured capacitance. In addition, one must avoid the deep depletion effects such as the initial linearly varying capacitance of the high frequency capacitance measured in the dark on the above figure (bottom curve). The larger the carrier lifetime, the slower the voltage is to be changed to avoid deep depletion.

The low frequency measured is compared to the theoretical value in Figure 6.3.6. The high frequency capacitance measured in the presence of light is also shown on the figure. The figure illustrates the agreement between experiment and theory. A comparison of the experimental low (rather than high) frequency capacitance with theory is somewhat easier to carry out. The low frequency capacitance is easier to calculate while the measurement tends to be less sensitive to deep depletion effects.

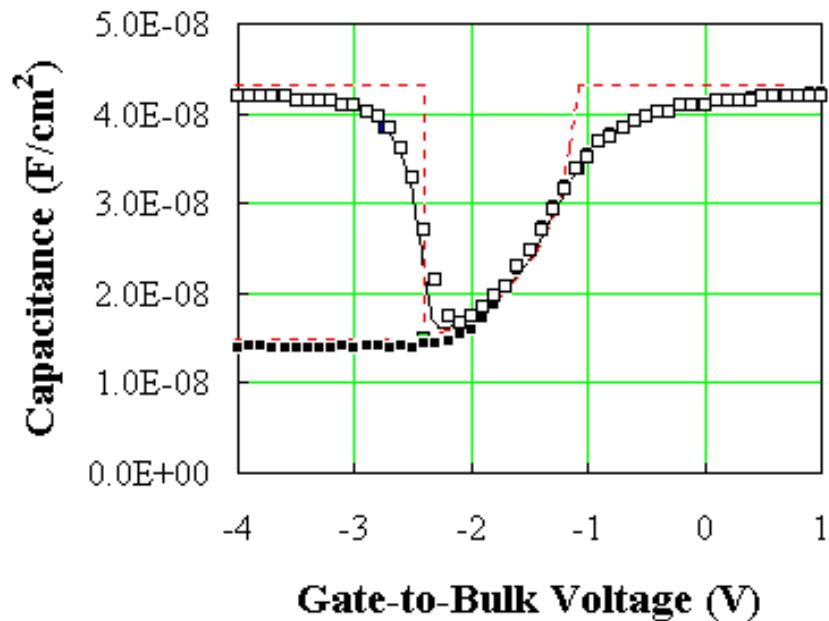


Figure 6.3.6: Comparison of the theoretical low frequency capacitance (solid line) and the experimental data (open squares) obtained in the dark. Fitting parameters are $N_a - N_d = 3.95 \times 10^{15} \text{ cm}^{-3}$ and $t_{\text{ox}} = 80 \text{ nm}$.

6.3.4.5. Non-Ideal effects in MOS capacitors

Non-ideal effects in MOS capacitors include fixed charge, mobile charge and charge in surface states. All three types of charge can be identified by performing a capacitance-voltage measurement.

Fixed charge in the oxide simply shifts the measured curve. A positive fixed charge at the oxide-semiconductor interface shifts the flatband voltage by an amount, which equals the charge divided by the oxide capacitance. The shift reduces linearly as one reduces the position of the charge relative to the gate electrode and becomes zero if the charge is located at the metal-oxide interface. A fixed charge is caused by ions, which are incorporated in the oxide during growth or deposition.

The flatband voltage shift due to mobile charge is described by the same equation as that due to fixed charge. However, the measured curves differ since a positive gate voltage causes any negative mobile charge to move away from the gate electrode, while a negative voltage attracts the charge towards the gate. This causes the curve to shift towards the applied voltage. One can recognize mobile charge by the hysteresis in the high frequency capacitance curve when sweeping the gate voltage back and forth. Sodium ions incorporated in the oxide of silicon MOS capacitors are known to yield mobile charges. It is because of the high sensitivity of MOS structures to a variety of impurities that the industry carefully controls the purity of the water and the chemicals used.

Charge due to electrons occupying surface states also yields a shift in flatband voltage. However as the applied voltage is varied, the Fermi energy at the oxide-semiconductor interface changes also and affects the occupancy of the surface states. The interface states cause the transition in the capacitance measurement to be less abrupt. The combination of the low frequency and high frequency capacitance allows calculating the surface state density. This method provides the surface state density over a limited (but highly relevant) range of energies within the bandgap. Measurements on n-type and p-type capacitors at different temperatures provide the surface state density throughout the bandgap.

Problems

1. Calculate the wavelength of a photon with a photon energy of 2 eV. Also, calculate the wavelength of an electron with a kinetic energy of 2 eV.
2. Consider a beam of light with a power of 1 Watt and a wavelength of 800 nm. Calculate a) the photon energy of the photons in the beam, b) the frequency of the light wave and c) the number of photons provided by the beam in one second.
3. Show that the spectral density, u_w (equation 1.2.4) peaks at $E_{ph} = 2.82 kT$. Note that a numeric iteration is required.
4. Calculate the peak wavelength of blackbody radiation emitted from a human body at a temperature of 37°C.
5. Derive equations (1.2.9) and (1.2.10).
6. What is the width of an infinite quantum well if the second lowest energy of a free electron confined to the well equals 100 mV?
7. Calculate the lowest three possible energies of an electron in a hydrogen atom in units of electron volt. Identify all possible electron energies between the lowest energy and -2 eV.
8. Derive the electric field of a proton with charge q as a function of the distance from the proton using Gauss's law. Integrated the electric field to find the potential $f(r)$:

$$f(r) = \frac{q}{4\pi \epsilon_s r}$$

Treat the proton as a point charge and assume the potential to be zero far away from the proton.

9. Prove that the probability of occupying an energy level below the Fermi energy equals the probability that an energy level above the Fermi energy and equally far away from the Fermi energy is not occupied.

Problem 1.1 Calculate the wavelength of a photon with a photon energy of 2 eV. Also, calculate the wavelength of an electron with a kinetic energy of 2 eV.

Solution The wavelength of a 2 eV photon equals:

$$l = \frac{hc}{E_{ph}} = \frac{6.626 \times 10^{-34} \text{ Js} \times 3 \times 10^8 \text{ m/s}}{1.602 \times 10^{-19} \text{ C} \times 2 \text{ eV}} = 0.62 \text{ } \mu\text{m}$$

where the photon energy (2 eV) was first converted to Joules by multiplying with the electronic charge.

The wavelength of an electron with a kinetic energy of 2 eV is obtained by calculating the deBroglie wavelength:

$$l = \frac{h}{p} = \frac{6.626 \times 10^{-34} \text{ Js}}{7.62 \times 10^{-25} \text{ kg m/s}} = 0.87 \text{ nm}$$

Where the momentum of the particle was calculated from the kinetic energy:

$$p = \sqrt{2mE} = \sqrt{2 \times 9.11 \times 10^{-31} \text{ kg} \times 1.6 \times 10^{-19} \text{ C} \times 2 \text{ eV}} = 7.64 \times 10^{-25} \text{ kg m/s}$$

Problem 1.2 Consider a beam of light with a power of 1 Watt and a wavelength of 800 nm. Calculate a) the photon energy of the photons in the beam, b) the frequency of the light wave and c) the number of photons provided by the beam in one second.

Solution The photon energy is calculated from the wavelength as:

$$E_{ph} = \frac{hc}{\lambda} = \frac{6.626 \times 10^{-34} \text{ Js} \times 3 \times 10^8 \text{ m/s}}{800 \times 10^{-9} \text{ m}} = 2.48 \times 10^{-19} \text{ J}$$

or in electron Volt:

$$E_{ph} = \frac{2.48 \times 10^{-19} \text{ J}}{1.602 \times 10^{-19} \text{ C}} = 1.55 \text{ eV}$$

The frequency then equals:

$$n = \frac{E_{ph}}{h} = \frac{2.48 \times 10^{-19} \text{ J}}{6.626 \times 10^{-34} \text{ Js}} = 375 \text{ THz}$$

And the number of photons equals the ratio of the optical power and the energy per photon:

$$\# \text{ photons} = \frac{1 \text{ Watt}}{E_{ph}} = \frac{1 \text{ Watt}}{2.48 \times 10^{-19} \text{ J}} = 4 \times 10^{18}$$

Problem 1.3 Show that the spectral density, u_w (equation 1.2.4) peaks at $E_{ph} = 2.82 kT$. Note that a numeric iteration is required.

Solution The spectral density, u_w , can be rewritten as a function of $x = \frac{\hbar w}{kT}$

$$u_w = \frac{k^3 T^3}{\hbar^2 p^2 c^3} \frac{x^3}{\exp(x) - 1}$$

The maximum of this function is obtained if its derivative is zero or:

$$\frac{du_w}{dx} = \frac{3x^2}{\exp(x) - 1} - \frac{x^3 \exp(x)}{(\exp(x) - 1)^2} = 0$$

Therefore x must satisfy:

$$3 - 3\exp(-x) = x$$

This transcendental equation can be solved starting with an arbitrary positive value of x . A repeated calculation of the left hand side using this value and the resulting new value for x quickly converges to $x_{\max} = 2.82144$. The maximum spectral density therefore occurs at:

$$E_{ph,\max} = x_{\max} kT = 2.82144 \text{ kT}$$

Problem 1.4 Calculate the peak wavelength of blackbody radiation emitted from a human body at a temperature of 37°C.

Solution The peak wavelength is obtained through the peak energy:

$$I_{\max} = \frac{hc}{E_{ph, \max}} = \frac{hc}{2.82kT}$$

$$I_{\max} = \frac{6.626 \times 10^{-34} \times 3 \times 10^8}{2.82 \times 1.38 \times 10^{-23} \times 310.15} = 1.65 \times 10^{-5} \text{ m} = 16.5 \text{ } \mu\text{m}$$

Where the temperature was first converted to units Kelvin.

Problem 1.5 Derive equations (1.2.9) and (1.2.10). Calculate the total energy as the sum of the kinetic and potential energy.

Solution The derivation starts by setting the centrifugal force equal to the electrostatic force:

$$m_0 \frac{v^2}{r} = \frac{p^2}{m_0 r} = \frac{q^2}{4p e_0 r^2}$$

where the velocity, v , is expressed as a function of the momentum, p .

The momentum in turn is calculated as a function of the deBroglie wavelength and the wavelength must be an integer fraction of the length of the circular orbit

$$\frac{p^2}{mr} = \frac{h^2}{mr l^2} = \frac{h^2 n^2}{mr 4p^2 r^2} = \frac{q^2}{4p e_0 r^2}$$

The corresponding radius equals the Bohr radius, a_0 :

$$a_0 = \frac{e_0 h^2 n^2}{p m_0 q^2}$$

The corresponding energies are obtained by adding the kinetic and potential energy:

$$E_n = \frac{p^2}{2m_0} - \frac{q^2}{4p e_0 a_0} = -\frac{m_0 q^4}{8e_0^2 h^2 n^2}, \text{ with } n = 1, 2, \dots$$

Note that the potential energy equals the potential of a proton multiplied with the electron charge, $-q$.

Problem 1.6 What is the width of an infinite quantum well if the second lowest energy of a free electron confined to the well equals 100 meV?

Solution The second lowest energy is calculated from

$$E_2 = \frac{h^2}{2m^*} \left(\frac{2}{2L_x} \right)^2 = 1.6 \times 10^{-20} \text{ J}$$

One can therefore solve for the width, L_x , of the well, yielding:

$$L_x = \frac{h}{\sqrt{2m^* E_2}} = \frac{6.626 \times 10^{-34}}{\sqrt{2 \times 9.11 \times 10^{-31} \times 1.6 \times 10^{-20}}} = 3.88 \text{ nm}$$

Problem 1.7 Calculate the lowest three possible energies of an electron in a hydrogen atom in units of electron volt. Identify all possible electron energies between the lowest energy and -2 eV.

Solution The three lowest electron energies in a hydrogen atom can be calculated from:

$$E_n = -\frac{13.6 \text{ eV}}{n^2}, \text{ with } n = 1, 2, \text{ and } 3$$

resulting in:

$$E_1 = -13.6 \text{ eV}, E_2 = -3.4 \text{ eV} \text{ and } E_3 = -1.51 \text{ eV}$$

The second lowest energy, E_2 , is the only one between the lowest energy, E_1 , and -2 eV.

Problem 1.8 Derive the electric field of a proton with charge q as a function of the distance from the proton using Gauss's law. Integrate the electric field to find the potential $\mathbf{f}(r)$:

$$\mathbf{f}(r) = \frac{q}{4\pi \epsilon_0 r}$$

Treat the proton as a point charge and assume the potential to be zero far away from the proton.

Solution Using a sphere with radius, r , around the charged proton as a surface where the electric field, E , is constant, one can apply Gauss's law:

$$E(r)4\pi r^2 = \frac{q}{\epsilon_0}$$

so that

$$E(r) = \frac{q}{4\pi r^2 \epsilon_0}$$

The potential is obtained by integrating this electric field from to Resulting in:

$$\mathbf{f}(r) - \mathbf{f}(\infty) = - \int_{\infty}^r \frac{q}{4\pi \epsilon_0 r^2} dr = \frac{q}{4\pi \epsilon_0 r}$$

where the potential at infinity was set to zero.

Problem 1.9 Prove that the probability of occupying an energy level below the Fermi energy equals the probability that an energy level above the Fermi energy and equally far away from the Fermi energy is not occupied.

Solution The probability that an energy level with energy ΔE **below** the Fermi energy E_F is occupied can be rewritten as:

$$\begin{aligned} f(E_F - \Delta E) &= \frac{1}{1 + \exp \frac{E_F - \Delta E - E_F}{kT}} = \frac{\exp \frac{\Delta E}{kT}}{\exp \frac{\Delta E}{kT} + 1} \\ &= 1 - \frac{1}{\exp \frac{\Delta E}{kT} + 1} = 1 - \frac{1}{1 + \exp \frac{E_F + \Delta E - E_F}{kT}} = 1 - f(E_F + \Delta E) \end{aligned}$$

so that it also equals the probability that an energy level with energy ΔE **above** the Fermi energy, E_F , is **not** occupied.

Chapter 6: MOS Capacitors



6.2. Structure and principle of operation

[6.2.1. Flatband diagram](#)

[6.2.2. Accumulation](#)

[6.2.3. Depletion](#)

[6.2.4. Inversion](#)

The MOS capacitor consists of a Metal-Oxide-Semiconductor structure as illustrated by Figure [6.2.1](#). Shown is the semiconductor substrate with a thin oxide layer and a top metal contact, also referred to as the gate. A second metal layer forms an Ohmic contact to the back of the semiconductor, also referred to as the bulk. The structure shown has a p-type substrate. We will refer to this as an n-type MOS capacitor since the inversion layer as discussed below contains electrons

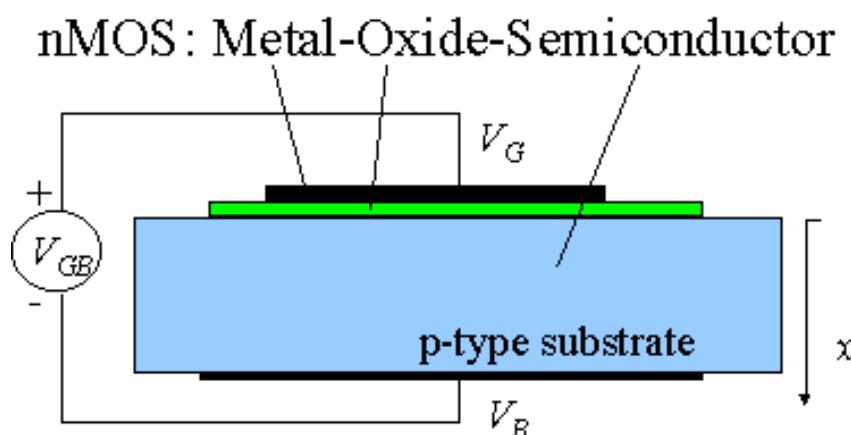


Figure 6.2.1:

To understand the different bias modes of an MOS capacitor we now consider three different bias voltages. One below the flatband voltage, V_{FB} , a second between the flatband voltage and the threshold voltage, V_T , and finally one larger than the threshold voltage. These bias regimes are called the accumulation, depletion and inversion mode of operation. These three modes as well as the charge distributions associated with each of them are shown in Figure [6.2.2](#).

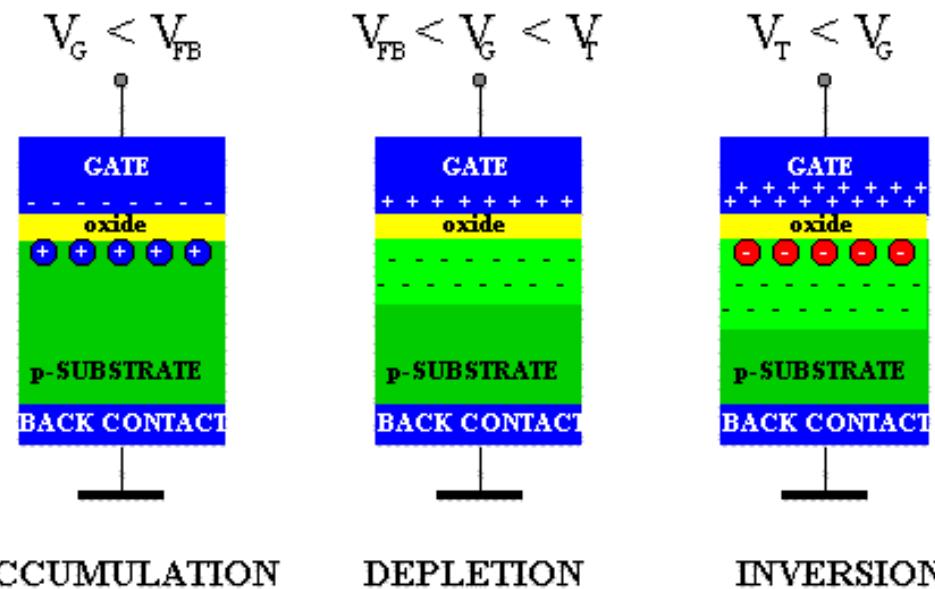


Figure 6.2.2.: Charges in a Metal-Oxide-Semiconductor structure under accumulation, depletion and inversion conditions

Accumulation occurs typically for negative voltages where the negative charge on the gate attracts holes from the substrate to the oxide-semiconductor interface. Depletion occurs for positive voltages. The positive charge on the gate pushes the mobile holes into the substrate. Therefore, the semiconductor is depleted of mobile carriers at the interface and a negative charge, due to the ionized acceptor ions, is left in the space charge region. The voltage separating the accumulation and depletion regime is referred to as the flatband voltage, V_{FB} . Inversion occurs at voltages beyond the threshold voltage. In inversion, there exists a negatively charged inversion layer at the oxide-semiconductor interface in addition to the depletion-layer. This inversion layer is due to minority carriers, which are attracted to the interface by the positive gate voltage.

The energy band diagram of an n-MOS capacitor biased in inversion is shown in Figure 6.2.3. The oxide is characterized as a semiconductor with a very large bandgap, which blocks any flow of carriers between the semiconductor and the gate metal. The band-bending in the semiconductor is consistent with the presence of a depletion layer. At the semiconductor-oxide interface, the Fermi energy is close to the conduction band edge as expected when a high density of electrons is present. An interesting point to note is that as the oxide behaves as an ideal insulator, the semiconductor is in thermal equilibrium even when a voltage is applied to the gate. The presence of an electric field does not automatically lead to a non-equilibrium condition, as was also the case for a p-n diode with zero bias.

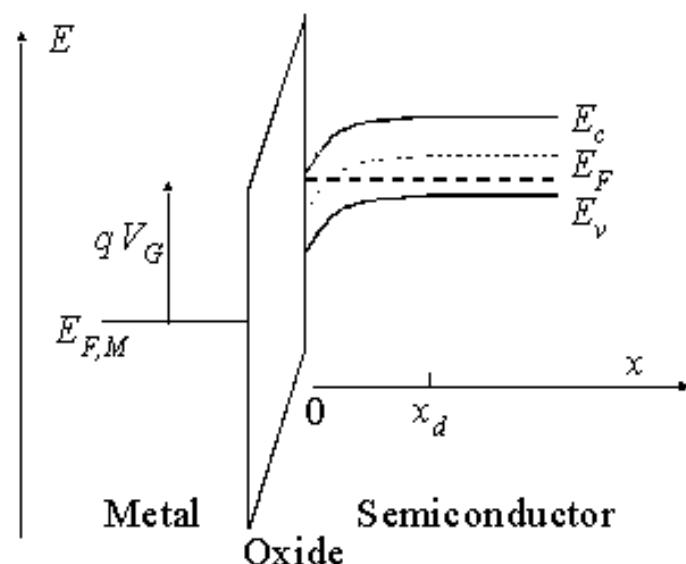


Figure 6.2.3:

Energy band diagram of an MOS structure biased in inversion.

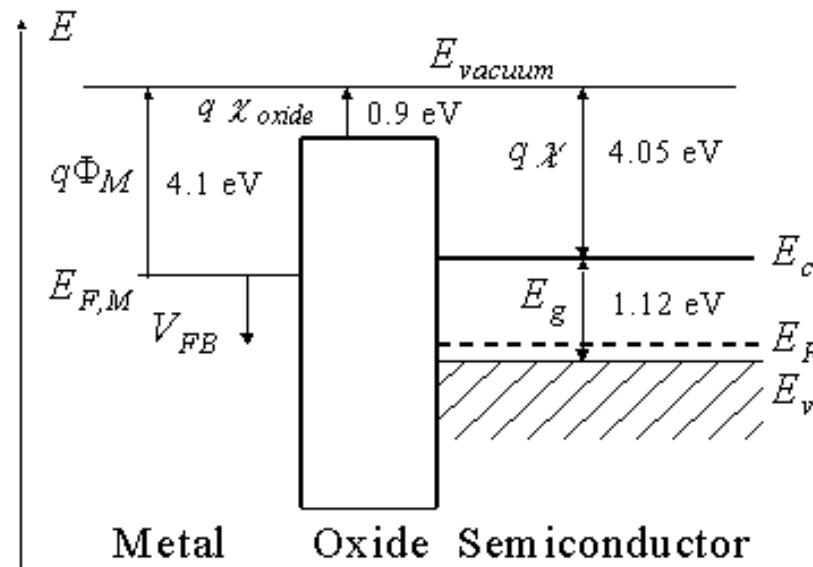
We will discuss in the next sections the four modes of operation of an MOS structure: Flatband, Depletion, Inversion and Accumulation. Flatband conditions exist when no charge is present in the semiconductor so that the silicon energy band is flat. Initially we will assume that this occurs at zero gate bias. Later we will consider the actual flat band voltage in more detail. Surface depletion occurs when the holes in the substrate are pushed away by a positive gate voltage. A more positive voltage also attracts electrons (the minority carriers) to the surface, which form the so-called inversion layer. Under negative gate bias, one attracts holes from the p-type substrate to the surface, yielding accumulation.

6.2.1. Flatband diagram



The flatband diagram is by far the easiest energy band diagram. The term flatband refers to fact that the energy band diagram of the semiconductor is flat, which implies that no charge exists in the semiconductor. The flatband diagram of an aluminum-silicon dioxide-silicon MOS structure is shown in Figure 6.2.4. Note that a voltage, V_{FB} , must be applied to obtain this flat band diagram. Indicated on the figure is also the work function of the aluminum gate, Φ_M , the electron affinity of the oxide, χ_{oxide} , and that of silicon, χ , as well as the bandgap energy of silicon, E_g . The bandgap energy of the oxide is quoted in the literature to be between 8 and 9 electron volt. The reader should also realize that the oxide is an amorphous material and the use of semiconductor parameters for such material can justifiably be questioned.

The flat band voltage is obtained when the applied gate voltage equals the workfunction difference between the gate metal and the semiconductor. If there is also a fixed charge in the oxide and/or at the oxide-silicon interface, the expression for the flatband voltage must be modified accordingly.

**Figure 6.2.4:**

Flatband energy diagram of a metal-oxide-semiconductor (MOS) structure consisting of an aluminum metal, silicon dioxide and silicon.

6.2.2. Accumulation



Accumulation occurs when one applies a voltage, which is less than the flatband voltage. The negative charge on the gate attracts holes from the substrate to the oxide-semiconductor interface. Only a small amount of band bending is needed to build up the accumulation charge so that almost all of the potential variation is within the oxide.



6.2.3. Depletion

As a more positive voltage than the flatband voltage is applied, a negative charge builds up in the semiconductor. Initially this charge is due to the depletion of the semiconductor starting from the oxide-semiconductor interface. The depletion layer width further increases with increasing gate voltage.



6.2.4. Inversion

As the potential across the semiconductor increases beyond twice the bulk potential, another type of negative charge emerges at the oxide-semiconductor interface: this charge is due to minority carriers, which form a so-called inversion layer. As one further increases the gate voltage, the depletion layer width barely increases further since the charge in the inversion layer increases exponentially with the surface potential.

Chapter 6: MOS Capacitors



6.4. MOS capacitor technology

The fabrication of the oxide of an MOS structure is one of the critical steps when fabricating MOSFETs. This is in part due to the need for an ideal oxide-semiconductor interface with low surface-state density but also because of the extremely thin oxides that are currently used for sub-micron MOSFETs. Two techniques are commonly used to form silicon dioxide. One involves the oxidation of the silicon yielding a thermal oxide. The other technique relies on the deposition of SiO_2 using a chemical vapor-deposition (CVD) process.

The thermal oxidation of silicon is obtained by heating the wafer in an oxygen or water vapor ambient. Typical temperatures range from 800 to 1200°C. The oxidation of a silicon surface also occurs at room temperature but the resulting 3 nm layer of oxide limits any further oxidation. At high temperatures, oxygen or water molecules can diffuse through the oxide so that further oxidation takes place. The oxidation in oxygen ambient is called a dry oxidation. The one in water vapor is a wet oxidation. The thermal oxidation provides a high quality interface and oxide. It is used less these days because of the high process temperatures.

The deposition of SiO_2 using a CVD process is one where two gases, silane and oxygen, react to form silicon dioxide, which then sublimes onto any solid surface. The wafers are heated to 200 - 400°C yielding high quality oxides. The lower process temperature and the quality of the deposited layers make CVD deposition the preferred method to fabricate MOS oxides.

Chapter 7: MOS Field-Effect-Transistors



7.5. MOSFET SPICE MODEL

The SPICE model of a MOSFET includes a variety of parasitic circuit elements and some process related parameters in addition to the elements previously discussed in this chapter. The syntax of a MOSFET incorporates the parameters a circuit designer can change as shown below:

MOSFET syntax

```
M <name> <drain node> <gate node> <source node> <bulk/substrate node>
+ [L=][W=][AD=][AS=]
+ [PD=][PS=][NRD=][NRS=]
+ [NRG=][NRB=]
```

where L is the gate length, W the gate width, AD the drain area, AS the source area

PD is the drain perimeter, PS is the source perimeter

Example:

M1 3 2 1 0 NMOS L=1u W=6u

```
.MODEL NFET NMOS (LEVEL=2 L=1u W=1u VTO=-1.44 KP=8.64E-6
+ NSUB=1E17 TOX=20n)
```

where M1 is one specific transistor in the circuit, while the transistor model "NFET" uses the built-in model NFET to specify the process and technology related parameters of the MOSFET. A list of SPICE parameters and their relation to the parameters discussed in this text is provided in Table 7.5.1.

SPICE variable	Equation
TOX	$TOX = t_{ox}$
KP	$KP = \mu C_{ox}$
VTO	$VTO = V_{FB} + 2\sqrt{\frac{2\epsilon_s q N_a}{C_{ox}}}$
GAMMA	$GAMMA = \gamma = \frac{\sqrt{2\epsilon_s q N_a}}{C_{ox}}$
NSUB	$NSUB = N_d \text{ or } N_a$
U0	$U0 = \mu$
LAMBDA	$LAMBDA = \lambda$
VMAX	$VMAX = v_{sat}$

Table 7.5.1:

SPICE parameters and corresponding equations

In addition there are additional parameters, which can be specified to further enhance the accuracy of the model, such as:

LD, lateral diffusion (length)
 RD, drain ohmic resistance
 RG, gate ohmic resistance
 IS, bulk p-n saturation current
 CBD, bulk-drain zero-bias p-n capacitance
 CGSO/CGDO, gate-source/drain overlap capacitance/channel width
 XJ, metallurgical junction depth
 WD, lateral diffusion (width)
 RS, source ohmic resistance
 RB, bulk ohmic resistance
 JS, bulk p-n saturation current/area
 CBS, bulk-source zero-bias p-n capacitance

The gate-source/drain overlap capacitance per channel width is obtained from:

$$CGSO = CGDO = \frac{\varepsilon_{ox} \Delta L}{t_{ox}}$$

Where ΔL is the overlap between the gate and the source/drain region. The corresponding equivalent circuit is provided in Figure 7.5.1.

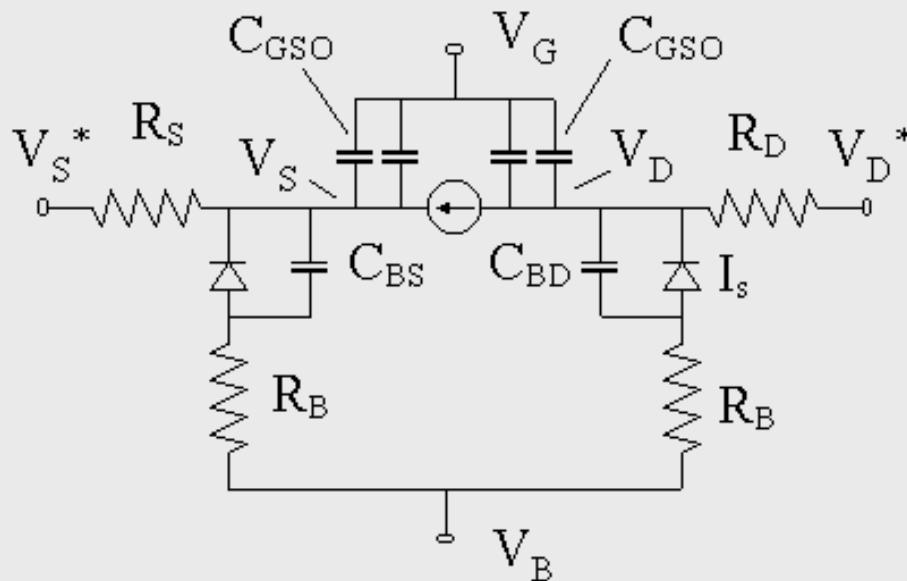


Figure 7.5.1 :

Large signal model of a MOSFET

Principles of Semiconductor Devices



Table of Contents

[Short Table of Contents](#)

[Title Page](#)
[Table of Contents](#)
[Help](#)
[Introduction](#)

Chapter 1: Review of Modern Physics

1. [Introduction](#)
2. [Quantum Mechanics](#)
3. [Electromagnetic theory](#)
4. [Statistical Thermodynamics](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 2: Semiconductor Fundamentals

1. [Introduction](#)
2. [Crystals and crystal structures](#)
3. [Energy bands](#)
4. [Density of States](#)
5. [Distribution Functions](#)
6. [Carrier Densities](#)
7. [Carrier Transport](#)
8. [Recombination and Generation](#)
9. [Continuity Equation](#)
10. [Drift-Diffusion Model](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation sheet](#)

Chapter 3: Metal-Semiconductor Junctions

1. [Introduction](#)
2. [Structure and principle of operation](#)
3. [Electrostatic analysis](#)
4. [Schottky diode current](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 4: p-n Junctions

1. [Introduction](#)
2. [Structure and principle of operation](#)
3. [Electrostatic analysis](#)
4. [p-n diode current](#)
5. [Breakdown](#)
6. [Optoelectronic devices](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 5: Bipolar Transistors

1. [Introduction](#)
2. [Structure](#)
3. [Ideal BJT model](#)
4. [Non-ideal effects](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 6: MOS Capacitors

1. [Introduction](#)
2. [Structure](#)
3. [Analysis](#)
4. [Technology](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 7: MOS Field-Effect-Transistors

1. [Introduction](#)
2. [Structure and principle of operation](#)
3. [MOSFET analysis](#)
4. [Threshold voltage](#)
5. [MOSFET SPICE model](#)
6. [MOSFET circuits and technology](#)
7. [Advance MOSFET issues](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Appendix

1. List of symbols

[By symbol](#)

[By name](#)

[Extended List of symbols](#)

2. [Physical Constants](#)

3. [Material Parameters](#)

4. [Prefixes](#)

5. [Units](#)

6. [Greek Alphabet](#)

7. [Periodic Table](#)

[Quick access](#)

© B. Van Zeghbroeck, 2001

Equation Window

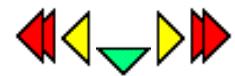
Principles of Semiconductor Devices

B. Van Zeghbroeck

Review Questions

1. What is a flatband diagram?
2. Define the barrier height of a metal-semiconductor junction. Can the barrier height be negative? Explain.
3. Define the built-in potential. Also provide an equation and state the implicit assumption(s).
4. Name three possible reasons why a measured barrier height can differ from the value calculated using equations (3.2.1) or (3.2.2).
5. How does the energy band diagram of a metal-semiconductor junction change under forward and reverse bias? How does the depletion layer width change with bias?
6. What is the full depletion approximation? Why do we need the full depletion approximation?
7. What mechanism(s) cause(s) current in a metal-semiconductor junction?

Chapter 7: MOS Field-Effect-Transistors



7.4. Threshold voltage

7.4.1. Threshold voltage calculation

7.4.2. The substrate bias effect

In this section we summarize the calculation of the threshold voltage and discuss the dependence of the threshold voltage on the bias applied to the substrate, called the substrate bias effect.

7.4.1. Threshold voltage calculation



The threshold voltage equals the sum of the flatband voltage, twice the bulk potential and the voltage across the oxide due to the depletion layer charge, or:

$$V_T = V_{FB} + 2\phi_F + \frac{\sqrt{2\epsilon_s q N_a (2\phi_F + V_{SB})}}{C_{ox}} \quad (7.4.1)$$

where the flatband voltage, V_{FB} , is given by:

$$V_{FB} = \Phi_{MS} - \frac{Q_f}{C_{ox}} - \frac{1}{C_{ox}} \int_0^{t_{ox}} \frac{x}{x_{ox}} \rho_{ox}(x) dx \quad (7.4.2)$$

With

$$\Phi_{MS} = \Phi_M - \Phi_S = \Phi_M - \left(\chi + \frac{E_g}{2q} + \phi_F \right) \quad (7.4.3)$$

and

$$\phi_F = V_t \ln \frac{N_a}{n_i}, \text{ p - substrate} \quad (7.4.4)$$

The threshold voltage of a p-type MOSFET with an n-type substrate is obtained using the following equations:

$$V_T = V_{FB} - |2\phi_F| - \frac{\sqrt{2\epsilon_s q N_d (|2\phi_F| - V_{SB})}}{C_{ox}} \quad (7.4.5)$$

where the flatband voltage, V_{FB} , is given by:

$$V_{FB} = \Phi_{MS} - \frac{Q_f}{C_{ox}} - \frac{1}{C_{ox}} \int_0^{t_{ox}} \frac{x}{x_{ox}} \rho_{ox}(x) dx \quad (7.4.6)$$

With

$$\Phi_{MS} = \Phi_M - \Phi_S = \Phi_M - \left(\chi + \frac{E_g}{2q} - |\phi_F| \right) \quad (7.4.7)$$

and

$$|\phi_F| = V_f \ln \frac{N_d}{n_i}, \text{ n - substrate} \quad (7.4.8)$$

The threshold voltage dependence on the doping density is illustrated with Figure 7.4.1 for both n-type and p-type MOSFETs with an aluminum gate metal.

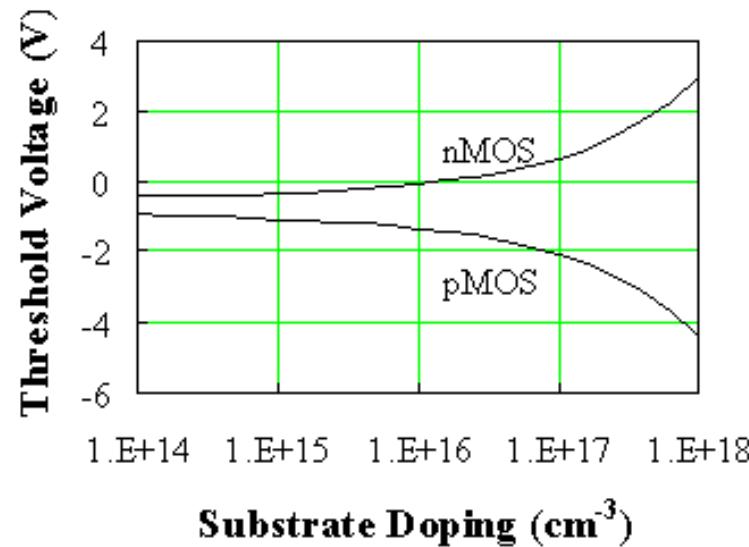


Figure 7.4.1 : Threshold voltage of n-type (upper curve) and p-type (lower curve) MOSFETs versus substrate doping density. 

The threshold of both types of devices is slightly negative at low doping densities and differs by 4 times the absolute value of the bulk potential. The threshold of nMOSFETs increases with doping while the threshold of pMOSFETs decreases with doping in the same way. A variation of the flatband voltage due to oxide charge will cause a reduction of both thresholds if the charge is positive and an increase if the charge is negative.

7.4.2. The substrate bias effect



The voltage applied to the back contact affects the threshold voltage of a MOSFET. The voltage difference between the source and the bulk, V_{BS} changes the width of the depletion layer and therefore also the voltage across the oxide due to the change of the charge in the depletion region. This results in a modified expression for the threshold voltage, as given by:

$$V_T = V_{FB} + 2\phi_F + \frac{\sqrt{2\varepsilon_s q N_a (2\phi_F + V_{SB})}}{C_{ox}} \quad (7.4.9)$$

The threshold difference due to an applied source-bulk voltage can therefore be expressed by:

$$\Delta V_T = \gamma(\sqrt{(2\phi_F + V_{SB})} - \sqrt{2\phi_F}) \quad (7.4.10)$$

Where γ is the body effect parameter given by:

$$\gamma = \frac{\sqrt{2\varepsilon_s q N_a}}{C_{ox}} \quad (7.4.11)$$

The variation of the threshold voltage with the applied bulk-to-source voltage can be observed by plotting the transfer curve for different bulk-to-source voltages. The expected characteristics, as calculated using the quadratic model and the variable depletion layer model, are shown in Figure 7.4.2.

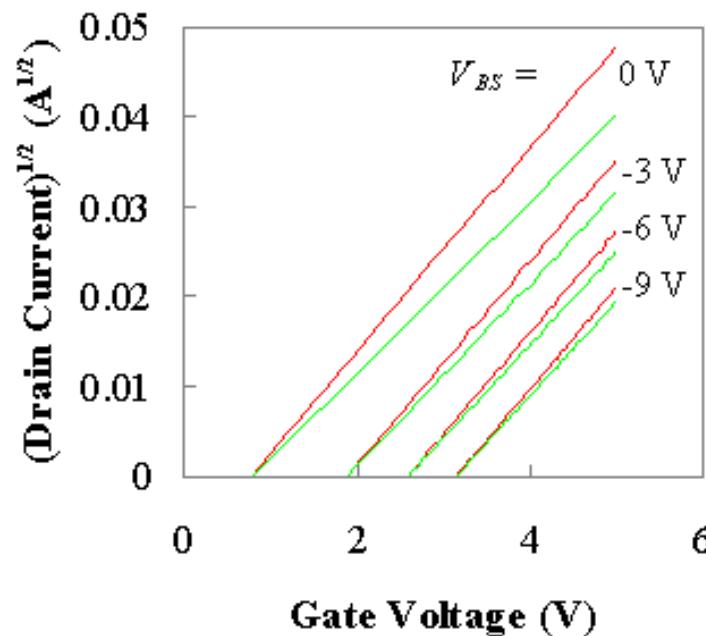


Figure 7.4.2 : Square root of I_D versus the gate-source voltage as calculated using the quadratic model (upper curves) and the variable depletion layer model (lower curves). 

A first observation is that the threshold shift is the same for both models. When biasing the device at the threshold voltage, drain saturation is obtained at zero drain-to-source voltage so that the depletion layer width is constant along the channel. As the drain-source voltage at saturation is increased, there is an increasing difference between the drain current as calculated with each model. The difference however reduces as a more negative bulk-source voltage is applied. This is due to the larger depletion layer width, which reduces the relative variation of the depletion layer charge along the channel.

Example 7.3	Calculate the threshold voltage of a silicon nMOSFET when applying a substrate voltage, $V_{BS} = 0, -2.5, -5, -7.5$ and -10 V. The capacitor has a substrate doping $N_a = 10^{17}$ cm $^{-3}$, a 20 nm thick oxide ($\epsilon_{ox} = 3.9 \epsilon_0$) and an aluminum gate ($\Phi_M = 4.1$ V). Assume there is no fixed charge in the oxide or at the oxide-silicon interface.
Solution	<p>The threshold voltage at $V_{BS} = -2.5$ V equals:</p> $V_T = V_{T0} + \frac{\gamma}{\sqrt{2\phi_F}} \left(\sqrt{1 + \frac{V_{SB}}{2\phi_F}} - 1 \right)$ $= -0.09 + \frac{0.75}{\sqrt{2 \times 0.42}} \left(\sqrt{1 + \frac{2.5}{2 \times 0.42}} - 1 \right) = 0.73 \text{ V}$ <p>Where the flatband voltage without substrate bias, V_{T0}, was already calculated in example 6.2. The body effect parameter was obtained from:</p>

$$\gamma = \frac{\sqrt{2 \epsilon_3 q N_a}}{C_{ox}} = \frac{\sqrt{2 \times 11.9 \times 8.85 \times 10^{-14} \times 1.6 \times 10^{-19} \times 10^{17}}}{3.9 \times 8.85 \times 10^{-14} / 20 \times 10^{-7}} \\ = 0.75 \text{ V}^{-1/2}$$

The threshold voltages for the different substrate voltages are listed in the table below.

	$V_{BS} = -2.5 \text{ V}$	-5 V	-7.5 V	-10 V
V_T	0.73 V	1.26 V	1.68 V	2.04 V

Principles of Semiconductor Devices



Table of Contents

Chapter 1: Review of Modern Physics

Chapter 2: Semiconductor Fundamentals

Chapter 3: Metal-Semiconductor Junctions

Chapter 4: p-n Junctions

Chapter 5: Bipolar Transistors

Chapter 6: MOS Capacitors

Chapter 7: MOS Field-Effect-Transistors

Appendix

© B. Van Zeghbroeck, 2001

Principles of Semiconductor Devices



Table of Contents

[Title Page](#)
[Table of Contents](#)
[Help](#)
[Introduction](#)

Chapter 1: Review of Modern Physics

Chapter 2: Semiconductor Fundamentals

Chapter 3: Metal-Semiconductor Junctions

Chapter 4: p-n Junctions

Chapter 5: Bipolar Transistors

Chapter 6: MOS Capacitors

Chapter 7: MOS Field-Effect-Transistors

Appendix

© B. Van Zeghbroeck, 2001

Principles of Semiconductor Devices



Table of Contents

Chapter 1: Review of Modern Physics

1. [Introduction](#)
2. [Quantum Mechanics](#)
3. [Electromagnetic theory](#)
4. [Statistical Thermodynamics](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 2: Semiconductor Fundamentals

Chapter 3: Metal-Semiconductor Junctions

Chapter 4: p-n Junctions

Chapter 5: Bipolar Transistors

Chapter 6: MOS Capacitors

Chapter 7: MOS Field-Effect-Transistors

Appendix

© B. Van Zeghbroeck, 2001

Principles of Semiconductor Devices



Table of Contents

Chapter 1: Review of Modern Physics

Chapter 2: Semiconductor Fundamentals

1. [Introduction](#)
2. [Crystals and crystal structures](#)
3. [Energy bands](#)
4. [Density of States](#)
5. [Distribution Functions](#)
6. [Carrier Densities](#)
7. [Carrier Transport](#)
8. [Recombination and Generation](#)
9. [Continuity Equation](#)
10. [Drift-Diffusion Model](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation sheet](#)

Chapter 3: Metal-Semiconductor Junctions

Chapter 4: p-n Junctions

Chapter 5: Bipolar Transistors

Chapter 6: MOS Capacitors

Chapter 7: MOS Field-Effect-Transistors

Appendix

© B. Van Zeghbroeck, 2001

Chapter 2: Semiconductor Fundamentals



2.4 Density of states

2.4.1. Calculation of the density of states

Before we can calculate the density of carriers in a semiconductor, we have to find the number of available states at each energy. The number of electrons at each energy is then obtained by multiplying the number of states with the probability that a state is occupied by an electron. Since the number of energy levels is very large and dependent on the size of the semiconductor, we will calculate the number of states per unit energy and per unit volume.

2.4.1 Calculation of the density of states



The density of states in a semiconductor equals the density per unit volume and energy of the number of solutions to Schrödinger's equation. We will assume that the semiconductor can be modeled as an infinite quantum well in which electrons with effective mass, m^* , are free to move. The energy in the well is set to zero. The semiconductor is assumed a cube with side L . This assumption does not affect the result since the density of states per unit volume should not depend on the actual size or shape of the semiconductor.

The solutions to the wave equation (equation [1.2.14](#)) where $V(x) = 0$ are sine and cosine functions:

$$\Psi = A \sin(k_x x) + B \cos(k_x x) \quad (2.4.1)$$

Where A and B are to be determined. The wavefunction must be zero at the infinite barriers of the well. At $x = 0$ the wavefunction must be zero so that only sine functions can be valid solutions or B must equal zero. At $x = L$, the wavefunction must also be zero yielding the following possible values for the wavenumber, k_x .

$$k_x = \frac{n\pi}{L}, n = 1, 2, 3, \dots \quad (2.4.2)$$

This analysis can now be repeated in the y and z direction. Each possible solution corresponds to a cube in k -space with size $n\pi/L$ as indicated on Figure [2.4.1](#).

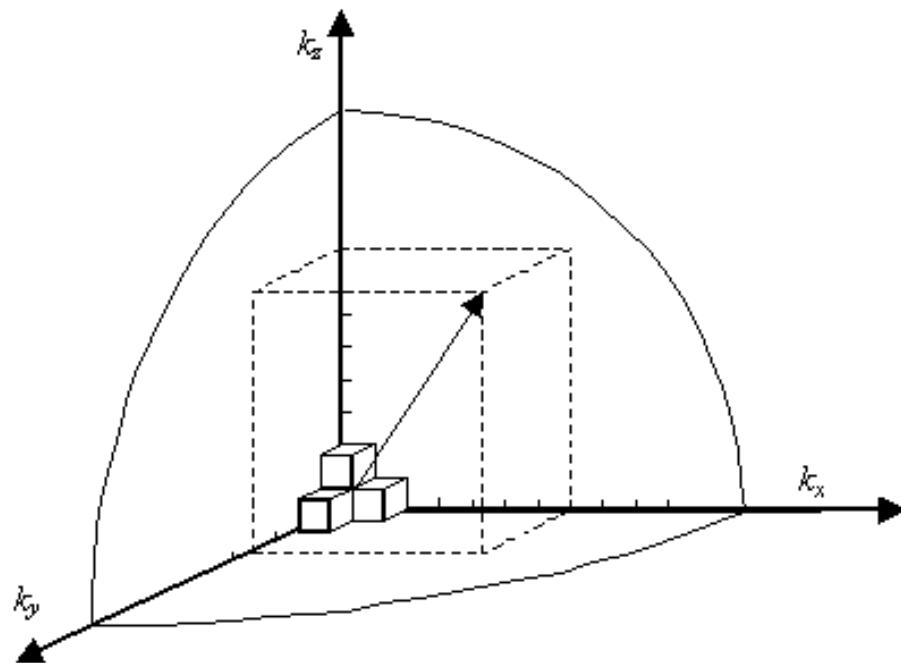


Figure 2.4.1: Calculation of the number of states with wavenumber less than k .

The total number of solutions with a different value for k_x , k_y and k_z and with a magnitude of the wavevector less than k is obtained by calculating the volume of one eighth of a sphere with radius k and dividing it by the volume corresponding to a

single solution, $\left(\frac{\pi}{L}\right)^3$, yielding:

$$N = 2 \times \frac{1}{8} \times \left(\frac{L}{\pi}\right)^3 \times \frac{4}{3} \times \pi \times k^3 \quad (2.4.3)$$

A factor of two is added to account for the two possible spins of each solution. The density per unit energy is then obtained using the chain rule:

$$\frac{dN}{dE} = \frac{dN}{dk} \frac{dk}{dE} = \left(\frac{L}{\pi}\right)^3 \pi k^2 \frac{dk}{dE} \quad (2.4.4)$$

The kinetic energy E of a particle with mass m^* is related to the wavenumber, k , by:

$$E(k) = \frac{\hbar^2 k^2}{2m^*} \quad (2.4.5)$$

And the density of states per unit volume and per unit energy, $g(E)$, becomes:

$$g(E) = \frac{1}{L^3} \frac{dN}{dE} = \frac{8\pi\sqrt{2}}{\hbar^3} m^{*3/2} \sqrt{E}, \text{ for } E \geq 0 \quad (2.4.6)$$

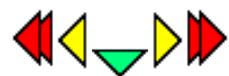
The density of states is zero at the bottom of the well as well as for negative energies.

The same analysis also applies to electrons in a semiconductor. The effective mass takes into account the effect of the periodic potential on the electron. The minimum energy of the electron is the energy at the bottom of the conduction band, E_c , so that the density of states for electrons in the conduction band is given by:

$$g_c(E) = \frac{1}{L^3} \frac{dN}{dE} = \frac{8\pi\sqrt{2}}{\hbar^3} m^{*3/2} \sqrt{E - E_c}, \text{ for } E \geq E_c \quad (2.4.7)$$

Example 2.3	Calculate the number of states per unit energy in a 100 by 100 by 10 nm piece of silicon ($m^* = 1.08 m_0$) 100 meV above the conduction band edge. Write the result in units of eV ⁻¹ .
Solution	<p>The density of states equals:</p> $g(E) = \frac{8 \pi \sqrt{2} m^*}{h^3} \frac{3/2}{\sqrt{E - E_c}}$ $= \frac{8 \pi \sqrt{2} (1.08 \times 9.1 \times 10^{-31})^{3/2}}{(6.626 \times 10^{-34})^3} \sqrt{0.1 \times 1.6 \times 10^{-19}}$ $= 1.51 \times 10^{56} \text{ m}^{-3} \text{J}^{-1}$ <p>So that the total number of states per unit energy equals:</p> $g(E)V = 1.51 \times 10^{56} \times 10^{-22} \text{ J}^{-1} = 2.41 \times 10^5 \text{ eV}^{-1}$

Chapter 2: Semiconductor Fundamentals



2.5 Carrier distribution functions

[2.5.1. Fermi-Dirac distribution function](#)

[2.5.2. Example](#)

[2.5.3. Impurity distribution functions](#)

[2.5.4. Other distribution functions and comparison](#)

The distribution or probability density functions describe the probability with which one can expect particles to occupy the available energy levels in a given system. Of particular interest is the probability density function of electrons, called the Fermi function. The derivation of such probability density functions belongs in a statistical thermodynamics course. However, given the importance of the Fermi distribution function, we will carefully examine an example as well as the characteristics of this function. Other distribution functions such as the impurity distribution functions, the Bose-Einstein distribution function and the Maxwell Boltzmann distribution are also provided.

2.5.1 Fermi-Dirac distribution function

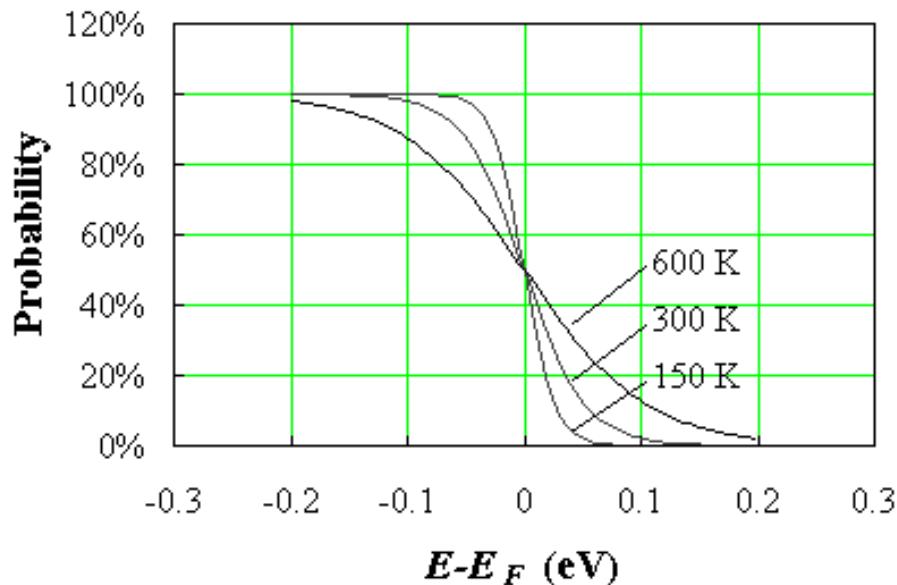


The Fermi-Dirac distribution function, also called Fermi function, provides the probability of occupancy of energy levels by Fermions. Fermions are half-integer spin particles, which obey the Pauli exclusion principle. The Pauli exclusion principle postulates that only one Fermion can occupy a single quantum state. Therefore, as Fermions are added to an energy band, they will fill the available states in an energy band just like water fills a bucket. The states with the lowest energy are filled first, followed by the next higher ones. At absolute zero temperature ($T = 0$ K), the energy levels are all filled up to a maximum energy, which we call the Fermi level. No states above the Fermi level are filled. At higher temperature, one finds that the transition between completely filled states and completely empty states is gradual rather than abrupt.

Electrons are Fermions. Therefore, the Fermi function provides the probability that an energy level at energy, E , in thermal equilibrium with a large system, is occupied by an electron. The system is characterized by its temperature, T , and its Fermi energy, E_F . The Fermi function is given by:

$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}} \quad (2.5.1)$$

This function is plotted in Figure [2.5.1](#) for different temperatures.

**Figure 2.5.1 :**

The Fermi function at three different temperatures. 

The Fermi function has a value of one for energies, which are more than a few times kT below the Fermi energy. It equals 1/2 if the energy equals the Fermi energy and decreases exponentially for energies which are a few times kT larger than the Fermi energy. While at $T = 0$ K the Fermi function equals a step function, the transition is more gradual at finite temperatures and more so at higher temperatures.

2.5.2 Example



To better understand the origin of distribution functions, we now consider a specific system with equidistant energy levels at 0.5, 1.5, 2.5, 3.5, 4.5, 5.5, ... eV. Each energy level can contain two electrons. Since electrons are indistinguishable from each other, no more than two electrons (with opposite spin) can occupy a given energy level. This system contains 20 electrons.

The minimum energy of this system corresponds to the situation where all 20 electrons occupy the ten lowest energy levels without placing more than 2 in any given level. This situation occurs at $T = 0$ K and the total energy equals 100 eV.

Since we are interested in a situation where the temperature is not zero, we arbitrarily set the total energy at 106 eV, which is 6 eV more than the minimum possible energy of this system. This ensures that the thermal energy is not zero so that the system must be at a non-zero temperature.

There are 24 possible and different configurations, which satisfy these particular constraints. Eight of those configurations are shown in Figure 2.5.2, where the filled circles represent the electrons:

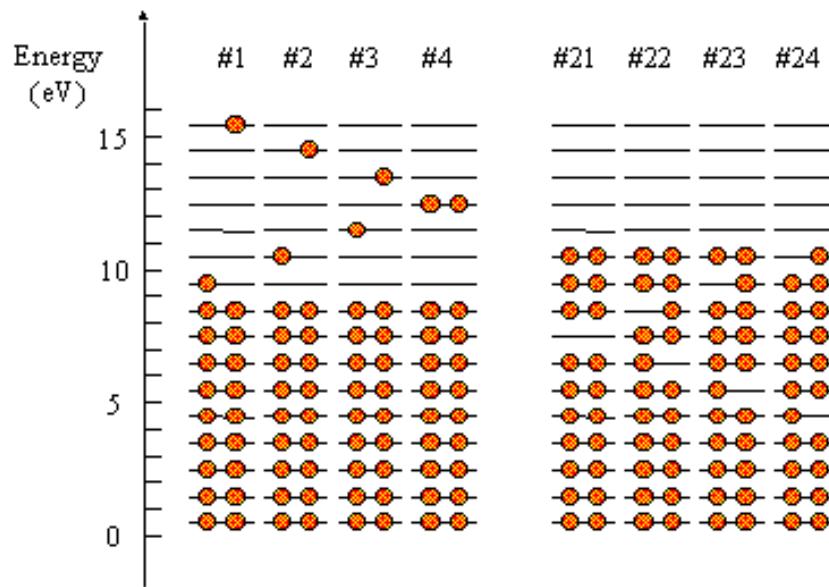


Figure 2.5.2 : Eight of the 24 possible configurations in which 20 electrons can be placed having a total energy of 106 eV.

We now apply the basic postulate of statistical thermodynamics, namely that all possible configurations are equally likely to occur. The expected configuration therefore equals the average occupancy of all possible configurations.

The average occupancy of each energy level taken over all (and equally probable) 24 configurations is compared in Figure 2.5.3 to the Fermi-Dirac distribution function. A best fit was obtained using a Fermi energy of 9.998 eV and $kT = 1.447$ eV or $T = 16,800$ K. The agreement is surprisingly good considering the small size of this system.

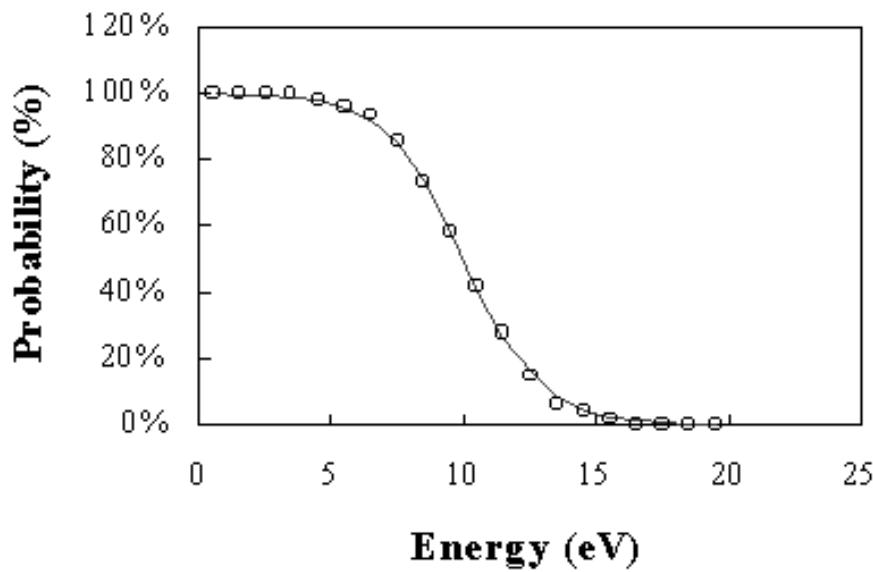


Figure 2.5.3 : Probability versus energy averaged over the 24 possible configurations (circles) fitted with a Fermi-Dirac function (solid line) using $kT = 1.447$ eV and $E_F = 9.998$ eV. 

Based on the construction of the distribution function in this example, one would expect the distribution function to be dependent on the density of states. This is the case for small systems. However, for large systems and for a single energy level in thermal equilibrium with a larger system, the distribution function no longer depends on the density of states. This is very fortunate, since it dramatically simplifies the carrier density calculations. One should also keep in mind that the Fermi energy for a particular system as obtained in section 2.6 does depend on the density of states.



2.5.3 Impurity distribution functions

The distribution function of impurities differs from the Fermi-Dirac distribution function although the particles involved are Fermions. The difference is due to the fact that an ionized donor energy level still contains one electron, which can have either spin (spin up or spin down). The donor energy level cannot be empty since this would leave a doubly positively charged atom, which would have an energy different from the donor energy. The distribution function for donors therefore differs from the Fermi function and is given by:

$$f_{\text{donor}}(E_d) = \frac{1}{1 + \frac{1}{2} e^{(E_d - E_F)/kT}} \quad (2.5.2)$$

The distribution function for acceptors differs also because of the different possible ways to occupy the acceptor level. The neutral acceptor contains no electrons. The ionized acceptor contains one electron, which can have either spin, while the doubly negatively charged state is not allowed since this would require a different energy. This restriction would yield a factor of 2 in front of the exponential term. In addition, one finds that most commonly used semiconductors have a two-fold degenerate valence band, which causes this factor to increase to four, yielding:

$$f_{\text{acceptor}}(E_a) = \frac{1}{1 + 4 e^{(E_a - E_F)/kT}} \quad (2.5.3)$$



2.5.4 Other distribution functions and comparison

Other distribution functions include the Bose-Einstein distribution and the Maxwell-Boltzmann distribution. These are briefly discussed below and compared to the Fermi-Dirac distribution function.

The Bose-Einstein distribution function applies to bosons. Bosons are particles with integer spin and include photons, phonons and a large number of atoms. Bosons do not obey the Pauli exclusion principle so that any number can occupy one energy level. The Bose-Einstein distribution function is given by:

$$f_{BE}(E) = \frac{1}{e^{(E-E_F)/kT} - 1} \quad (2.5.4)$$

This function is only defined for $E > E_F$.

The Maxwell Boltzmann applies to non-interacting particles, which can be distinguished from each other. This distribution function is also called the classical distribution function since it provides the probability of occupancy for non-interacting particles at low densities. Atoms in an ideal gas form a typical example of such particles. The Maxwell-Boltzmann distribution function is given by:

$$f_{MB}(E) = \frac{1}{e^{(E-E_F)/kT}} \quad (2.5.5)$$

A plot of the three distribution functions, the Fermi-Dirac distribution, the Maxwell-Boltzmann distribution and the Bose-Einstein distribution is shown in Figure [2.5.4](#).

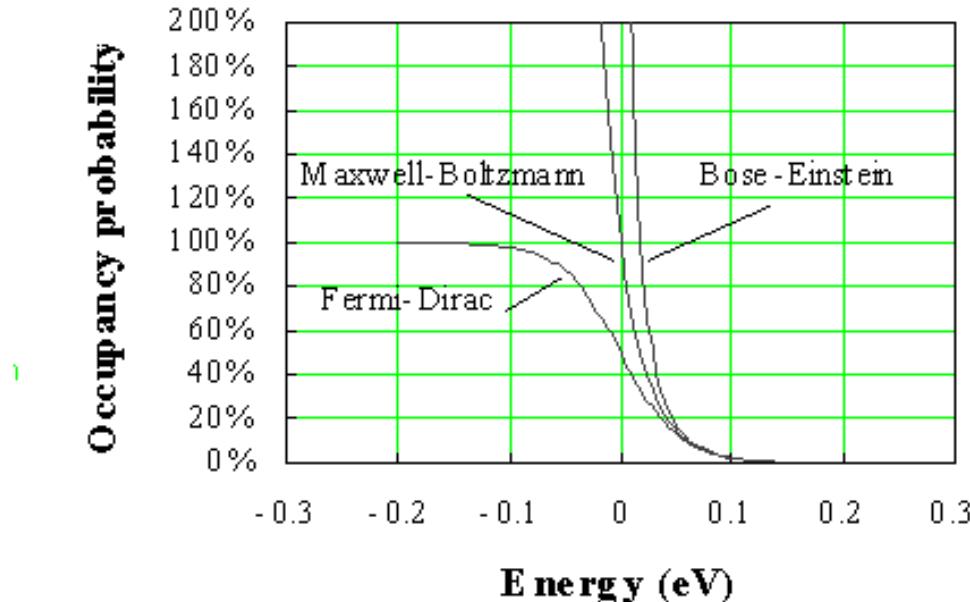


Figure 2.5.4 : Probability of occupancy versus energy of the Fermi-Dirac, the Bose-Einstein and the Maxwell-Boltzmann distribution. The Fermi energy, E_F , is assumed to be zero. 

All three functions are almost equal for large energies (more than a few kT beyond the Fermi energy). The Fermi-Dirac distribution reaches a maximum of 100% for energies, which are a few kT below the Fermi energy, while the Bose-Einstein distribution diverges at the Fermi energy and has no validity for energies below the Fermi energy.

Principles of Semiconductor Devices



Table of Contents

Chapter 1: Review of Modern Physics

Chapter 2: Semiconductor Fundamentals

Chapter 3: Metal-Semiconductor Junctions

1. [Introduction](#)
2. [Structure and principle of operation](#)
3. [Electrostatic analysis](#)
4. [Schottky diode current](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 4: p-n Junctions

Chapter 5: Bipolar Transistors

Chapter 6: MOS Capacitors

Chapter 7: MOS Field-Effect-Transistors

Appendix

© B. Van Zeghbroeck, 2001

Principles of Semiconductor Devices



Table of Contents

Chapter 1: Review of Modern Physics

Chapter 2: Semiconductor Fundamentals

Chapter 3: Metal-Semiconductor Junctions

Chapter 4: p-n Junctions

1. [Introduction](#)
2. [Structure and principle of operation](#)
3. [Electrostatic analysis](#)
4. [p-n diode current](#)
5. [Breakdown](#)
6. [Optoelectronic devices](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 5: Bipolar Transistors

Chapter 6: MOS Capacitors

Chapter 7: MOS Field-Effect-Transistors

Appendix

© B. Van Zeghbroeck, 2001

Principles of Semiconductor Devices



Table of Contents

Chapter 1: Review of Modern Physics

Chapter 2: Semiconductor Fundamentals

Chapter 3: Metal-Semiconductor Junctions

Chapter 4: p-n Junctions

Chapter 5: Bipolar Transistors

1. [Introduction](#)
2. [Structure](#)
3. [Ideal BJT model](#)
4. [Non-ideal effects](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 6: MOS Capacitors

Chapter 7: MOS Field-Effect-Transistors

Appendix

© B. Van Zeghbroeck, 2001

Principles of Semiconductor Devices



Table of Contents

Chapter 1: Review of Modern Physics

Chapter 2: Semiconductor Fundamentals

Chapter 3: Metal-Semiconductor Junctions

Chapter 4: p-n Junctions

Chapter 5: Bipolar Transistors

Chapter 6: MOS Capacitors

1. [Introduction](#)
2. [Structure](#)
3. [Analysis](#)
4. [Technology](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Chapter 7: MOS Field-Effect-Transistors

Appendix

© B. Van Zeghbroeck, 2001

Principles of Semiconductor Devices



Table of Contents

Chapter 1: Review of Modern Physics

Chapter 2: Semiconductor Fundamentals

Chapter 3: Metal-Semiconductor Junctions

Chapter 4: p-n Junctions

Chapter 5: Bipolar Transistors

Chapter 6: MOS Capacitors

Chapter 7: MOS Field-Effect-Transistors

1. [Introduction](#)
2. [Structure and principle of operation](#)
3. [MOSFET analysis](#)
4. [Threshold voltage](#)
5. [MOSFET SPICE model](#)
6. [MOSFET circuits and technology](#)
7. [Advance MOSFET issues](#)

[Examples](#)- [Problems](#)- [Review Questions](#)- [Bibliography](#)- [Equation Sheet](#)

Appendix

© B. Van Zeghbroeck, 2001

Chapter 7: MOS Field-Effect-Transistors



7.2. Structure and principle of operation

A top view of the same MOSFET is shown in Figure 7.2.1, where the gate length, L , and gate width, W , are identified. Note that the gate length does not equal the physical dimension of the gate, but rather the distance between the source and drain regions underneath the gate. The overlap between the gate and the source/drain region is required to ensure that the inversion layer forms a continuous conducting path between the source and drain region. Typically this overlap is made as small as possible in order to minimize its parasitic capacitance.



Figure 7.2.1 : Top view of an n-type Metal-Oxide-Semiconductor- Field-Effect-Transistor (MOSFET)

The flow of electrons from the source to the drain is controlled by the voltage applied to the gate. A positive voltage applied to the gate attracts electrons to the interface between the gate dielectric and the semiconductor. These electrons form a conducting channel between the source and the drain called the inversion layer. No gate current is required to maintain the inversion layer at the interface since the gate oxide blocks any carrier flow. The net result is that the current between drain and source is controlled by the voltage, which is applied to the gate.

The typical current versus voltage (I - V) characteristics of a MOSFET are shown in Figure 7.2.2.

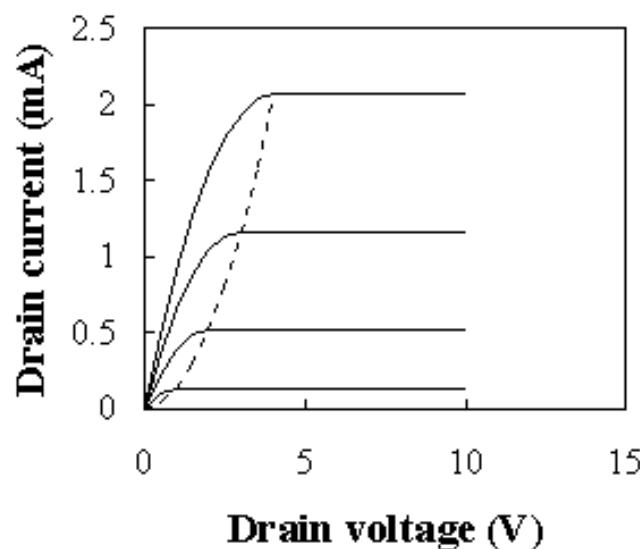


Figure 7.2.2 : I - V characteristics of an n-type MOSFET with $V_G = 5$ V (top curve), 4 V, 3 V and 2 V (bottom curve)

NOTE: We will primarily discuss the n-type or n-channel MOSFET in this chapter. This type of MOSFET is fabricated on a p-type semiconductor substrate. The complementary MOSFET is the p-type or p-channel MOSFET. It contains p-type source and drain regions in an n-type substrate. The inversion layer is formed when holes are attracted to the interface by a negative gate voltage. While the holes still flow from source to drain, they result in a negative drain current. CMOS circuits require both n-type and p-type devices.

Chapter 7: MOS Field-Effect-Transistors



7.3. MOSFET analysis

7.3.1. The linear model

7.3.2. The quadratic model

7.3.3. The variable depletion layer model

In this section, we present three different models for the MOSFET, the linear model, the quadratic model and the variable depletion layer model. The linear model correctly predicts the MOSFET behavior for small drain-source voltages, where the MOSFET acts as a variable resistor. The quadratic model includes the voltage variation along the channel between source and drain. This model is most commonly used despite the fact that the variation of the depletion layer charge is ignored. The variable depletion layer model is more complex as it does include the variation of the depletion layer along the channel.

7.3.1. The linear model



The linear model describes the behavior of a MOSFET biased with a small drain-to-source voltage. As the name suggests, the MOSFET, as described by the linear model, acts as a linear device, more specifically a linear resistor whose resistance can be modulated by changing the gate-to-source voltage. In this regime, the MOSFET can be used as a switch for analog signals or as an analog multiplier.

The general expression for the drain current equals the total charge in the inversion layer divided by the time the carriers need to flow from the source to the drain:

$$I_D = -\frac{Q_{inv}WL}{t_r} \quad (7.3.1)$$

where Q_{inv} is the inversion layer charge per unit area, W is the gate width, L is the gate length and t_r is the transit time. If the velocity of the carriers is constant between source and drain, the transit time equals:

$$t_r = \frac{L}{v} \quad (7.3.2)$$

where the velocity, v , equals the product of the mobility and the electric field:

$$v = \mu E = \mu \frac{V_{DS}}{L} \quad (7.3.3)$$

The constant velocity also implies a constant electric field so that the field equals the drain-source voltage divided by the gate length. This leads to the following expression for the drain current:

$$I_D = -\mu C_{inv} \cdot \frac{W}{L} \cdot V_{DS} \quad (7.3.4)$$

We now assume that the charge density in the inversion layer is constant between source and drain. We also assume that the basic assumption described in section 6.3.2 applies, namely that the charge density in the inversion layer is given by the product of the capacitance per unit area and the gate-to-source voltage minus the threshold voltage:

$$Q_{inv} = -C_{ox}(V_{GS} - V_T), \text{ for } V_{GS} > V_T \quad (7.3.5)$$

The inversion layer charge is zero if the gate voltage is lower than the threshold voltage. Replacing the inversion layer charge density in the expression for the drain current yields the linear model:

$$I_D = \mu C_{ox} \frac{W}{L} (V_{GS} - V_T) V_{DS}, \text{ for } |V_{DS}| \ll (V_{GS} - V_T) \quad (7.3.6)$$

Note that the capacitance in the above equations is the gate oxide capacitance per unit area. Also note that the drain current is zero if the gate-to-source voltage is less than the threshold voltage. The linear model is only valid if the drain-to-source voltage is much smaller than the gate-to-source voltage minus the threshold voltage. This insures that the velocity, the electric field and the inversion layer charge density is indeed constant between the source and the drain.

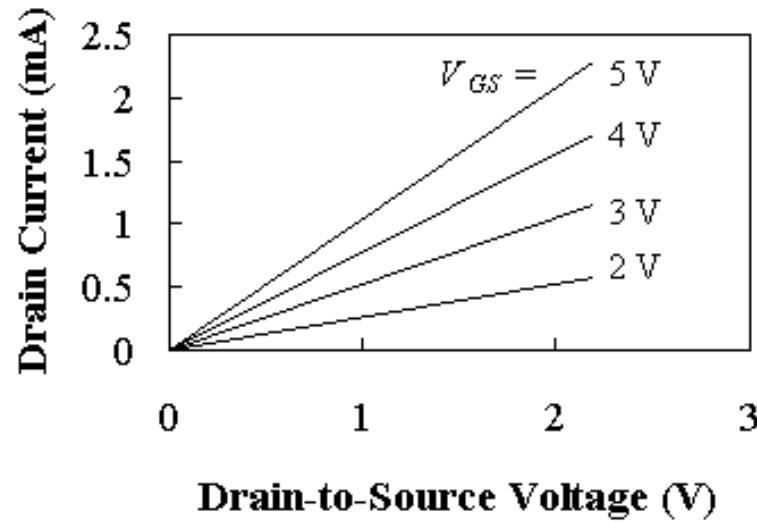


Figure 7.3.1 : Linear I - V characteristics of a MOSFET with $V_T = 1$ V. ($\mu_n = 300$ cm 2 /V-s, $W/L = 5$ and $t_{ox} = 20$ nm).



The figure illustrates the behavior of the device in the linear regime: While there is no drain current if the gate voltage is less than the threshold voltage, the current increases with gate voltage once it is larger than the threshold voltage. The slope of the curves equals the conductance of the device, which increases linearly with the applied gate voltage. The figure clearly illustrates the use of a MOSFET as a voltage-controlled resistor.

7.3.2. The quadratic model



The quadratic model uses the same assumptions as the linear model except that the inversion layer charge density is allowed to vary in the channel between the source and the drain.

The derivation is based on the fact that the current at each point in the channel is constant. The current can also be related to the local channel voltage.

Considering a small section within the device with width dy and channel voltage $V_C + V_S$ one can still use the linear model described by equation (7.3.6), yielding:

$$I_D = \mu C_{ox} \frac{W}{dy} (V_G - V_S - V_C - V_T) dV_C \quad (7.3.7)$$

where the drain-source voltage is replaced by the change in channel voltage over a distance dy , namely dV_C . Both sides of the equation can be integrated from the source to the drain, so that y varies from 0 to the gate length, L , and the channel voltage V_C varies from 0 to the drain-source voltage, V_{DS} .

$$\int_0^L I_D dy = \mu C_{ox} W \int_0^{V_{DS}} (V_G - V_S - V_C - V_T) dV_C \quad (7.3.8)$$

Using the fact that the DC drain current is constant throughout the device one obtains the following expression:

$$I_D = \mu C_{ox} \frac{W}{L} [(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2}], \quad \text{for } V_{DS} < V_{GS} - V_T \quad (7.3.9)$$

The drain current first increases linearly with the applied drain-to-source voltage, but then reaches a maximum value. According to the above equation the current would even decrease and eventually become negative. The charge density at the drain end of the channel is zero at that maximum and changes sign as the drain current decreases. As explained in section 6.2, the change in the inversion layer does go to zero and reverses its sign as holes are accumulated at the interface. However, these holes cannot contribute to the drain current since the reversed-biased p-n diode between the drain and the substrate blocks any flow of holes into the drain. Instead the current reaches its maximum value and maintains that value for higher drain-to-source voltages. A depletion layer located at the drain end of the gate accommodates the additional drain-to-source voltage. This behavior is referred to as drain current saturation.

Drain current saturation therefore occurs when the drain-to-source voltage equals the gate-to-source voltage minus the threshold voltage. The value of the drain current is then given by the following equation:

$$I_{D,sat} = \mu C_{ox} \frac{W}{L} \frac{(V_{GS} - V_T)^2}{2}, \quad \text{for } V_{DS} > V_{GS} - V_T \quad (7.3.10)$$

The quadratic model explains the typical current-voltage characteristics of a MOSFET, which are normally plotted for different gate-to-source voltages. An example is shown in Figure 7.3.2. The saturation occurs to the right of the dotted line which is given by $I_D = \mu C_{ox} W/L V_{DS}^2$.

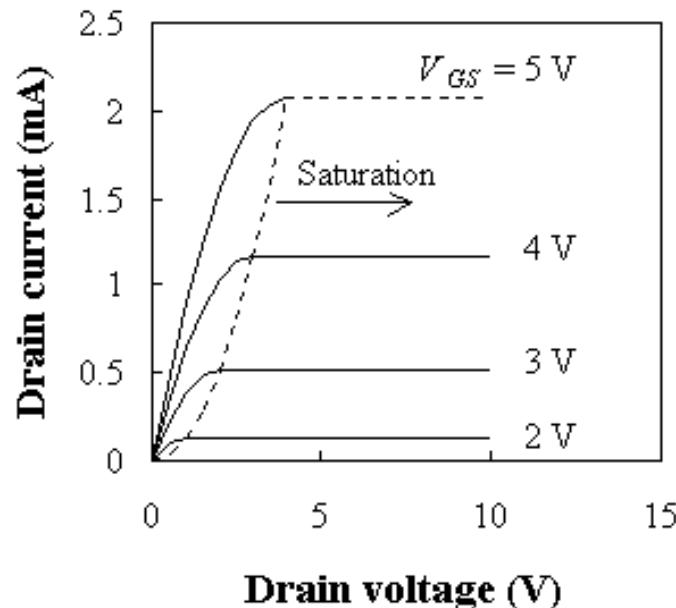


Figure 7.3.2: Current-Voltage characteristics of an n-type MOSFET as obtained with the quadratic model. The dotted line separates the quadratic region of operation on the left from the saturation region on the right. 

The drain current is again zero if the gate voltage is less than the threshold voltage.

$$I_D = 0, \text{ for } V_{GS} < V_T \quad (7.3.11)$$

For negative drain-source voltages, the transistor is in the quadratic regime and is described by equation (7.3.9). However, it is possible to forward bias the drain-bulk p-n junction. A complete circuit model should therefore also include the p-n diodes between the source, the drain and the substrate.

The quadratic model can be used to calculate some of the small signal parameters, namely the transconductance, g_m and the output conductance, g_d .

The transconductance quantifies the drain current variation with a gate-source voltage variation while keeping the drain-source voltage constant, or:

$$g_m = \frac{\Delta I_D}{\Delta V_{GS}} \Big|_{V_{DS}} \quad (7.3.12)$$

The transconductance in the quadratic region is given by:

$$g_{m,quad} = \mu C_{ox} \frac{W}{L} V_{DS} \quad (7.3.13)$$

which is proportional to the drain-source voltage for $V_{DS} < V_{GS} - V_T$. In saturation, the transconductance is constant and equals:

$$g_{m,sat} = \mu C_{ox} \frac{W}{L} (V_{GS} - V_T) \quad (7.3.14)$$

The output conductance quantifies the drain current variation with a drain-source voltage variation while keeping the gate-source voltage constant, or:

$$g_d = \frac{\Delta I_D}{\Delta V_{DS}} \Big|_{V_{GS}} \quad (7.3.15)$$

The output conductance in the quadratic region decreases with increasing drain-source voltage:

$$g_{d,quad} = \mu C_{ox} \frac{W}{L} (V_{GS} - V_T - V_{DS}) \quad (7.3.16)$$

and becomes zero as the device is operated in the saturated region:

$$g_{d,sat} = 0 \quad (7.3.17)$$

Example 7.1	<p>Calculate the drain current of a silicon nMOSFET with $V_T = 1$ V, $W = 10$ μm, $L = 1$ μm and $t_{ox} = 20$ nm. The device is biased with $V_{GS} = 3$ V and $V_{DS} = 5$ V. Use the quadratic model, a surface mobility of 300 $\text{cm}^2/\text{V}\cdot\text{s}$ and set $V_{BS} = 0$ V.</p> <p>Also calculate the transconductance at $V_{GS} = 3$ V and $V_{DS} = 5$ V and compare it to the output conductance at $V_{GS} = 3$ V and $V_{DS} = 0$ V.</p>
-------------	--

Solution	<p>The MOSFET is biased in saturation since $V_{DS} > V_{GS} - V_T$.</p> <p>Therefore the drain current equals:</p> $I_D = \mu_n C_{ox} \frac{W}{L} \frac{(V_{GS} - V_T)^2}{2}$ $= 300 \times \frac{3.9 \times 8.85 \times 10^{-14}}{20 \times 10^{-7}} \frac{10}{1} \times \frac{(3-1)^2}{2} = 1.04 \text{ mA}$ <p>The transconductance equals:</p> $g_m = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_T)$ $= 300 \times \frac{3.9 \times 8.85 \times 10^{-14}}{20 \times 10^{-7}} \frac{10}{1} \times (3-1) = 1.04 \text{ mS}$ <p>and the output conductance equals:</p> $g_d = \mu_n C_{ox} \frac{W}{L} (V_{GS} - V_T - V_{DS})$ $= 300 \times \frac{3.9 \times 8.85 \times 10^{-14}}{20 \times 10^{-7}} \frac{10}{1} \times (3-1-0) = 1.04 \text{ mS}$
----------	---

The measured drain current in saturation is not constant as predicted by the quadratic model. Instead it increases with drain-source voltage due to channel length modulation, drain induced barrier lowering or two-dimensional field distributions, as discussed in section 7.7.1. A simple empirical model, which considers these effects, is given by:

$$I_{D,sat} = \mu C_{ox} \frac{W}{L} \frac{(V_{GS} - V_T)^2}{2} (1 + \lambda V_{DS}), \quad \text{for } V_{DS} > V_{GS} - V_T \quad (7.3.18)$$

Where λ is a fitting parameter.

7.3.3. The variable depletion layer model



The variable depletion layer model includes the variation of the charge in the depletion layer between the source and drain. This variation is caused by the voltage variation along the channel. The inversion layer charge is still given by:

$$Q_{inv} = -C_{ox} (V_{GS} - V_T), \quad \text{for } V_{GS} > V_T \quad (7.3.19)$$

where we now include the implicit dependence of the threshold voltage on the charge in the depletion region, or:

$$V_T = V_{FB} + V_C + 2\phi_F + \frac{\sqrt{2\epsilon_s q N_a (2\phi_F + V_{SB} + V_C)}}{C_{ox}} \quad (7.3.20)$$

The voltage V_C is the difference between the voltage within the channel and the source voltage. We can now apply the linear model to a small section at a distance y from the source and with a thickness dy . The voltage at that point equals $V_C + V_S$ while the voltage across that section equals dV_C . This results in the following expression for the drain current, I_D :

$$I_D = \mu_n C_{ox} \frac{W}{dy} (V_{GS} - V_{FB} - 2\phi_F - V_C - \frac{\sqrt{2\epsilon_s q N_a (2\phi_F + V_{SB} + V_C)}}{C_{ox}}) dV_C \quad (7.3.21)$$

Both sides of the equation can be integrated from the source to the drain with y varying from 0 to the gate length, L , and the channel voltage, V_C varying from 0 to the drain-source voltage, V_{DS} .

$$\int_0^L I_D dy = \mu_n C_{ox} W \int_0^{V_{DS}} (V_{GS} - V_{FB} - 2\phi_F - V_C) dV_C - \mu_n W \int_0^{V_{DS}} \sqrt{2\epsilon_s q N_a (2\phi_F + V_{SB} + V_C)} dV_C \quad (7.3.22)$$

Integration along the channel yields the following drain current:

$$I_D = \frac{\mu_n C_{ox} W}{L} (V_{GS} - V_{FB} - 2\phi_F - \frac{V_{DS}}{2}) V_{DS} - \frac{2}{3} \mu_n \frac{W}{L} \sqrt{2\epsilon_s q N_a} ((2\phi_F + V_{DS})^{3/2} - (2\phi_F + V_{SB})^{3/2}) \quad (7.3.23)$$

The current-voltage characteristics as obtained with the above equation are shown in Figure 7.3.3, together with those obtained with the quadratic model. Again, it was assumed that the drain current saturates at its maximum value, since a positive inversion layer charge can not exist in an n-type MOSFET. The drain voltage at which saturation occurs is given by:

$$V_{DS,sat} = V_{GS} - V_{FB} - 2\phi_F - \frac{q N_a \epsilon_s}{C_{ox}^2} \left(\sqrt{1 + 2 \frac{C_{ox}^2}{q N_a \epsilon_s} (V_{GS} - V_{FB})} - 1 \right) \quad (7.3.24)$$

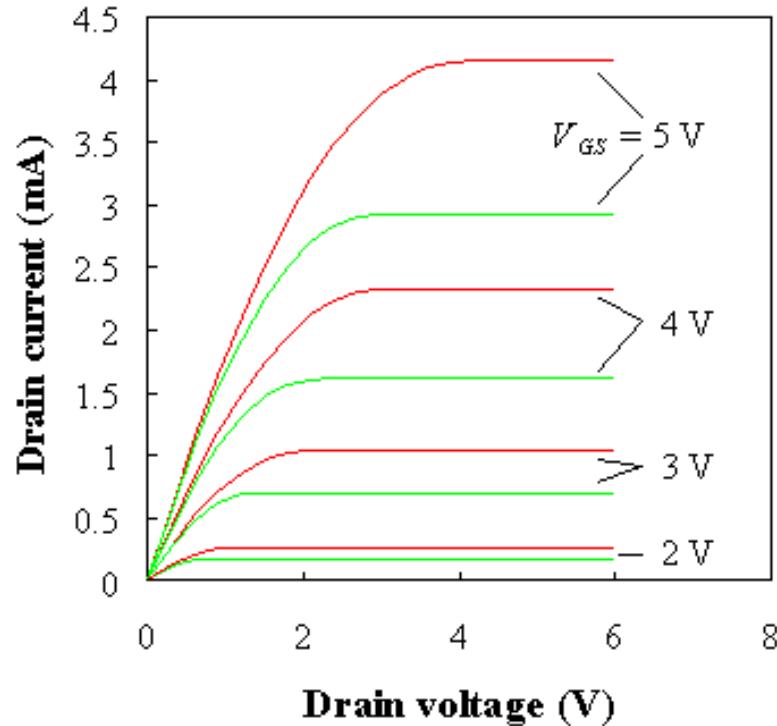


Figure 7.3.3 : Comparison of the quadratic model (upper curves) and the variable depletion layer model (lower curves). 

The figure shows a clear difference between the two models: the quadratic model yields a larger drain current compared to the more accurate variable depletion layer charge model. The transconductance is still given by equation (7.3.13), which combined with the saturation voltage (equation (7.3.24)) yields:

$$g_{m, \text{sat}} = \mu_n C_{ox} \frac{W}{L} [V_{GS} - V_{FB} - 2\phi_F - \frac{qN_a \varepsilon_s}{C_{ox}^2} \{ \sqrt{1 + 2 \frac{C_{ox}^2}{qN_a \varepsilon_s} (V_{GS} - V_{FB})} - 1 \}] \quad (7.3.25)$$

This transconductance is almost linearly dependent on V_{GS} , so that it can still be written in the form of equation (7.3.10) with a modified mobility μ_n^* :

$$g_{m, \text{sat}} = \mu_n^* C_{ox} \frac{W}{L} (V_{GS} - V_T) \quad (7.3.26)$$

Where μ_n^* equals:

$$\mu_n^* = \mu_n \left(1 - \frac{1}{\sqrt{1 + \frac{2(2\phi_F + V_{SB})C_{ox}^2}{qN_a \varepsilon_s}}} \right) \quad (7.3.27)$$

The term under the square root depends on the ratio of the oxide capacitance to the depletion layer capacitance at the onset of inversion. Since this ratio is larger than one in most transistors, the modified mobility is 10% to 40% smaller than the actual mobility. This effective mobility can also be used with the quadratic model, yielding a simple but reasonably accurate model for the MOSFET.

Example 7.2	Repeat example 7.1 using the variable depletion layer model. Use $V_{FB} = -0.807$ V and $N_a = 10^{17}$ cm $^{-3}$.
Solution	<p>To find out whether the MOSFET is biased in saturation, one first calculates the saturation voltage, $V_{D,sat}$:</p> $V_{DS,sat} = V_{GS} - V_{FB} - 2\phi_F$ $= \frac{qN_a \epsilon_s}{C_{ox}^2} \left(\sqrt{1 + 2 \frac{C_{ox}^2}{qN_a \epsilon_s} (V_{GS} - V_{FB})} - 1 \right)$ $= 1.39 \text{ V}$ <p>The drain current is then obtained from:</p> $I_D = \frac{\mu_n C_{ox} W}{L} (V_{GS} - V_{FB} - 2\phi_F - \frac{V_{DS,sat}}{2}) V_{DS,sat}$ $= \frac{2}{3} \mu_n \frac{W}{L} \sqrt{2 \epsilon_s q N_a} ((2\phi_F + V_{DB})^{3/2} - (2\phi_F + V_{SB})^{3/2})$ $= 0.7 \text{ mA}$ <p>The transconductance equals:</p> $g_{m,sat} = \mu_n C_{ox} \frac{W}{L} [V_{GS} - V_{FB} - 2\phi_F$ $- \frac{qN_a \epsilon_s}{C_{ox}^2} \left(\sqrt{1 + 2 \frac{C_{ox}^2}{qN_a \epsilon_s} (V_{GS} - V_{FB})} - 1 \right)]$ $= 0.52 \text{ mS}$ <p>corresponding to a modified mobility $\mu_n^* = 149$ cm2/V-s. The output conductance at $V_{DS} = 0$ V equals:</p> $g_d = \left. \frac{\partial I_D}{\partial V_{DS}} \right _{V_{GS}} = 1.04 \text{ mS}$ <p>Which is the same as that of example 7.1 since the depletion layer width is constant for $V_{DS} = 0$.</p>

Appendix 1: Symbol Index

[A](#) [B](#) [C](#) [D](#) [E](#) [F](#) [G](#) [H](#) [I](#) [J](#) [K](#) [L](#) [M](#) [N](#) [O](#) [P](#) [Q](#) [R](#) [S](#) [T](#) [U](#) [V](#) [W](#) [X](#) [Y](#) [Z](#)

Description	Symbol	MKS Units
-------------	--------	-----------

- A -

Acceptor doping density	N_a	m^{-3}
Acceptor energy	E_a	Joule
Applied voltage	V_a	V
Area	A	m^2

- B -

Barrier height	ϕ_B	V
Base doping density	N_B	m^{-3}
Base voltage	V_B	V
Base width	w_B	m
Body effect parameter	γ	$\text{V}^{1/2}$
Boltzmann's constant	k	J/K
Built-in potential of a p-n diode or Schottky diode	ϕ_i	V
Bulk potential	ϕ_F	V

- C -

Capacitance per unit area	C	F/m^2
Charge	Q	C
Charge density per unit area at threshold in the depletion layer of an MOS structure	$Q_{d,T}$	C/m^2
Charge density per unit area in the depletion layer of an MOS structure	Q_d	C/m^2
Charge density per unit volume	ρ	C/m^3
Charge density per unit volume in the oxide	ρ_{ox}	C/m^3
Collector doping density	N_C	m^{-3}
Collector voltage	V_C	V

Collector width	w_C	m
Conduction band energy of a semiconductor	E_c	Joule
Conductivity	σ	$\Omega^{-1}m^{-1}$
Current	I	A
Current density	J	A/m^2
Current gain	β	

- D -

Density of states in the conduction band per unit energy and per unit volume	$g_c(E)$	$m^{-3}J^{-1}$
Density of states in the valence band per unit energy and per unit volume	$g_v(E)$	$m^{-3}J^{-1}$
Depletion layer width	w	m
Depletion layer width in a p-type semiconductor	x_p	m
Depletion layer width in an MOS structure	x_d	m
Depletion layer width in an MOS structure at threshold	$x_{d,T}$	m
Depletion layer width in an n-type semiconductor	x_n	m
Dielectric constant of the oxide	ϵ_{ox}	F/m
Dielectric constant of the semiconductor	ϵ_s	F/m
Distribution function (probability density function)	$f(E)$	
Donor doping density	N_d	m^{-3}
Donor energy	E_d	Joule
Doping density	N	
Drain voltage	V_D	V

- E -

Effective density of states in the conduction band	N_c	m^{-3}
Effective density of states in the valence band	N_v	m^{-3}
Effective mass of electrons	m_e^*	kg
Effective mass of holes	m_h^*	kg
Electric field	E	V/m
Electron affinity of the semiconductor	χ	V
Electron current density	J_n	A/m^2
Electron density	n	m^{-3}

Electron density in thermal equilibrium	n_0	m^{-3}
Electron density per unit energy and per unit volume	$n(E)$	m^{-3}
Electron diffusion constant	D_n	m^2/s
Electron diffusion length	L_n	m
Electron energy in vacuum	E_{vacuum}	Joule
Electron generation rate	G_n	$\text{m}^{-3}\text{s}^{-1}$
Electron lifetime	τ_n	s
Electron mobility	μ_n	$\text{m}^2/\text{V}\cdot\text{s}$
Electron recombination rate	R_n	$\text{m}^{-3}\text{s}^{-1}$
electronic charge	q	C
Emitter doping density	N_E	m^{-3}
Emitter efficiency	γ_E	
Emitter voltage	V_E	V
Emitter width	w_E	m
Energy	E	Joule
Energy bandgap of a semiconductor	E_g	Joule
Excess electron charge density in the base	$\Delta Q_{n,B}$	C/m^2
Excess electron density	δ_n	m^{-3}
Excess hole density	δ_p	m^{-3}

- F -

Fermi energy (thermal equilibrium)	E_F	Joule
Flatband capacitance per unit area of a MOS structure	C_{FB}	F/m^2
Flatband voltage	V_{FB}	V
Free electron mass	m_0	kg

- G -

Gate voltage	V_G	V
--------------	-------	------------

- H -

Hole current density	J_p	A/m^2
Hole density	p	m^{-3}

Hole density in an n-type semiconductor	p_n	m^{-3}
Hole density in thermal equilibrium	p_0	m^{-3}
Hole density per unit energy	$p(E)$	m^{-3}
Hole diffusion constant	D_p	m^2/s
Hole diffusion length	L_p	m
Hole generation rate	G_p	$\text{m}^{-3}\text{s}^{-1}$
Hole lifetime	τ_p	s
Hole mobility	μ_p	$\text{m}^2/\text{V}\cdot\text{s}$
Hole recombination rate	R_p	$\text{m}^{-3}\text{s}^{-1}$

- I -

Interface charge density per unit area	Q_I	C/m^2
Intrinsic carrier density	n_I	m^{-3}
Intrinsic carrier density	n_I	m^{-3}
Intrinsic Fermi energy	E_I	Joule
Ionized acceptor density	N_a^-	m^{-3}
Ionized donor density	N_d^+	m^{-3}

- J -

Junction capacitance per unit area	C_j	F/m^2
Junction depth	x_j	m

- K -

Length	L	m

- M -

Mass	m	kg
Mean free path	l	m

- N -

Net recombination rate of electrons	U_n	$\text{m}^{-3}\text{s}^{-1}$
Net recombination rate of holes	U_p	$\text{m}^{-3}\text{s}^{-1}$

- O -

Oxide capacitance per unit area	C_{ox}	F/m^2
Oxide thickness	t_{ox}	m

- P -

Plank's constant	h	Js
Position	x	m
Potential	ϕ	V
Potential at the semiconductor surface	ϕ_s	V

- Q -

Quasi-Fermi energy of electrons	F_n	Joule
Quasi-Fermi energy of holes	F_p	Joule

- R -

Reduced Plank's constant ($= h /2\pi$)	\hbar	Js
Resistance	R	Ohm
Resistivity	ρ	Ωm

- S -

Speed of light in vacuum	c	m/s
--------------------------	-----	--------------

- T -

Temperature	T	Kelvin
Thermal velocity	v_{th}	m/s
Thermal voltage	V_t	V
Thickness	t	m
Threshold voltage of an MOS structure	V_T	V

Transport factor

 α

- U -

- V -

Valence band energy of a semiconductor

 E_v

Joule

Velocity

 v

m/s

- W -

Width of a p-type region

 w_p

m

Width of an n-type region

 w_n

m

Workfunction difference between the metal and the semiconductor

 Φ_{MS}

V

Workfunction of the metal

 Φ_M

V

Workfunction of the semiconductor

 Φ_S

V

- X -

- Y -

- Z -

Appendix:



Appendix 1: Extended List of Symbols

Symbol	Description	MKS Units
a	Acceleration	m/s^2
a_0	Bohr radius	m
A	Area	m^2
A^*	Richardson constant	m/s
A_C	Collector area	m^2
A_E	Emitter area	m^2
b	Bimolecular recombination constant	m^3/s
c	Speed of light in vacuum	m/s
C	Capacitance per unit area	F/m^2
C_D	Diffusion capacitance per unit area	F/m^2
C_{DS}	Drain-source capacitance	F
C_{FB}	Flatband capacitance per unit area of a MOS structure	F/m^2
C_G	Gate capacitance	F
C_{GS}	Gate-source capacitance	F
C_{GD}	Gate-drain capacitance	F
C_{HF}	High-frequency capacitance per unit area of a MOS structure	F/m^2
C_j	Junction capacitance per unit area	F/m^2
C_{LF}	Low-frequency (quasi-static) capacitance per unit area of a MOS structure	F/m^2
C_M	Miller capacitance	F
C_{ox}	Oxide capacitance per unit area	F/m^2
C_s	Semiconductor capacitance per unit area	F/m^2
D_n	Electron diffusion constant	m^2/s
D_p	Hole diffusion constant	m^2/s
E	Energy	Joule
E	Electric field	V/m
E_0	Lowest energy in a one-dimensional quantum well	Joule
E_a	Acceptor energy	Joule
E_{br}	Breakdown field	V/m
E_c	Conduction band energy of a semiconductor	Joule
E_d	Donor energy	Joule
E_F	Fermi energy (thermal equilibrium)	Joule
$E_{F,n}$	Fermi energy in an n-type semiconductor	Joule
$E_{F,p}$	Fermi energy in a p-type semiconductor	Joule

E_g	Energy bandgap of a semiconductor	Joule
E_i	Intrinsic Fermi energy	Joule
E_n	n^{th} quantized energy	Joule
E_{ph}	Photon energy	Joule
E_t	Trap energy	Joule
E_v	Valence band energy of a semiconductor	Joule
E_{vacuum}	Electron energy in vacumm	Joule
$F(E)$	Distribution function (probability density function)	
$f_{BE}(E)$	Bose-Einstein distribution function	
$f_{FD}(E)$	Fermi-Dirac distribution function	
$f_{MB}(E)$	Maxwell-Boltzmann distribution function	
F	Force	Newton
F_n	Quasi-Fermi energy of electrons	Joule
F_p	Quasi-Fermi energy of holes	Joule
$g(E)$	Density of states per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
$g(E)$	Density of states in the conduction band per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
$g(E)$	Density of states in the valence band per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
g_d	Output conductance of a MOSFET	S
g_m	Transconductance of a MOSFET	S
G	Carrier generation rate	$\text{m}^{-3}\text{s}^{-1}$
G_n	Electron generation rate	$\text{m}^{-3}\text{s}^{-1}$
G_p	Hole generation rate	$\text{m}^{-3}\text{s}^{-1}$
h	Plank's constant	Js
\square	Reduced Plank's ($= h / 2\pi$)	Js
I	Current	A
I_B	Base current of a bipolar transistor	A
I_C	Collector current of a bipolar transistor	A
I_D	Drain current of a MOSFET	A
I_E	Emitter current of a bipolar transistor	A
I_F	Forward active current of a bipolar transistor	A
$I_{D,sat}$	Drain current of a MOSFET in saturation	A
I_{ph}	Photo current	A
I_r	Recombination current	A
I_R	Reverse active current of a bipolar transistor	A
I_s	Saturation current	A
I_{sc}	Short circuit current of a solar cell	A
J	Current density	A/m^2
J_n	Electron current density	A/m^2
J_p	Hole current density	A/m^2

k	Boltzmann's constant wavenumber	J/K m^{-1}
l	Mean free path	m
L	Length	m
L_D	Debye length	m
L_n	Electron diffusion length	m
L_p	Hole diffusion length	m
L_x	Hole diffusion length	m
m	Mass	kg
m^*	Effective mass	kg
m_A	Atomic mass	kg
m_0	Free electron mass	kg
m_e^*	Effective mass of electrons	kg
m_h^*	Effective mass of holes	kg
M	Proton mass Multiplication factor	kg
	Electron density	m^{-3}
n	Integer Refractive index Ideality factor	
n_i	Intrinsic carrier density	m^{-3}
$n(E)$	Electron density per unit energy and per unit volume	m^{-3}
n_0	Electron density in thermal equilibrium	m^{-3}
n_i	Intrinsic carrier density	m^{-3}
n_n	Electron density in an n-type semiconductor	m^{-3}
n_{n0}	Thermal equilibrium electron density in an n-type semiconductor	m^{-3}
n_p	Electron density in a p-type semiconductor	m^{-3}
n_{p0}	Thermal equilibrium electron density in a p-type semiconductor	m^{-3}
N	Number of particles Doping density	
N_a	Acceptor doping density	m^{-3}
N_a^-	Ionized acceptor density	m^{-3}
N_A	Avogadro's number	
N_B	Base doping density	m^{-3}
N_c	Effective density of states in the conduction band	m^{-3}
N_C	Collector doping density	m^{-3}
N_d	Donor doping density	m^{-3}
N_d^+	Ionized donor density	m^{-3}
N_E	Emitter doping density	m^{-3}
N_{ss}	Surface state density	m^{-2}

N_t	Recombination trap density	m^{-2}
N_v	Effective density of states in the valence band	m^{-3}
	Hole density	m^{-3}
p	Particle momentum	kgm/s
	Pressure	Nm^{-2}
$p(E)$	Hole density per unit energy	m^{-3}
p_0	Hole density in thermal equilibrium	m^{-3}
p_n	Hole density in an n-type semiconductor	m^{-3}
p_{n0}	Thermal equilibrium hole density in an n-type semiconductor	m^{-3}
p_p	Hole density in a p-type semiconductor	m^{-3}
p_{p0}	Thermal equilibrium hole density in a p-type semiconductor	m^{-3}
q	electronic charge	C
Q	Heat	Joule
	Charge	C
Q_B	Majority carrier charge density in the base	C/m^2
Q_d	Charge density per unit area in the depletion layer of an MOS structure	C/m^2
$Q_{d,T}$	Charge density per unit area at threshold in the depletion layer of an MOS structure	C/m^2
Q_i	Interface charge density per unit area	C/m^2
Q_{inv}	Inversion layer charge density per unit area	C/m^2
Q_M	Charge density per unit area in a metal	C/m^2
Q_n	Charge density per unit area in the depletion layer of an n-type region	C/m^2
Q_p	Charge density per unit area in the depletion layer of a p-type region	C/m^2
Q_{ss}	Surface state charge density per unit area	C/m^2
r_e	Emitter resistance	Ohm
r_π	Base resistance	Ohm
R	The Rydberg constant	J
	Resistance	Ohm
R_n	Electron recombination rate	$\text{m}^{-3}\text{s}^{-1}$
R_p	Hole recombination rate	$\text{m}^{-3}\text{s}^{-1}$
R_s	Sheet resistance	Ohm
s	Spin	
S	Entropy	J/K
t	Thickness	m
t_{ox}	Oxide thickness	m
T	Temperature	Kelvin
	Kinetic energy	Joule
u_ω	Spectral density	$\text{Jm}^{-3}\text{s}^{-1}$
U	Total energy	Joule
U_A	Auger recombination rate	$\text{m}^{-3}\text{s}^{-1}$

U_{b-b}	Band-to-band recombination rate	$\text{m}^{-3}\text{s}^{-1}$
U_n	Net recombination rate of electrons	$\text{m}^{-3}\text{s}^{-1}$
U_p	Net recombination rate of holes	$\text{m}^{-3}\text{s}^{-1}$
U_{SHR}	Shockley-Read-Hall recombination rate	$\text{m}^{-3}\text{s}^{-1}$
v	Velocity	m/s
v_R	Richardson velocity	m/s
v_{sat}	Saturation velocity	m/s
v_{th}	Thermal velocity	m/s
V	Potential energy	Joule
	Volume	m^3
V_a	Applied voltage	V
V_A	Early voltage	V
V_{br}	Breakdown voltage	V
V_B	Base voltage	V
V_{BE}	Base-emitter voltage	V
V_{BC}	Base-collector voltage	V
V_C	Collector voltage	V
V_{CE}	Collector-emitter voltage	V
V_D	Drain voltage	V
V_{DS}	Drain-source voltage	V
$V_{DS,sat}$	Drain-source saturation voltage	V
V_E	Emitter voltage	V
V_{FB}	Flatband voltage	V
V_G	Gate voltage	V
V_{GS}	Gate-source voltage	V
V_{oc}	Open circuit voltage of a solar cell	V
V_t	Thermal voltage	V
V_T	Threshold voltage of an MOS structure	V
w	Depletion layer width	m
w_B	Base width	m
w_C	Collector width	m
w_E	Emitter width	m
w_n	Width of an n-type region	m
w_p	Width of a p-type region	m
W	Work	Joule
x	Position	m
x_d	Depletion layer width in an MOS structure	m
$x_{d,T}$	Depletion layer width in an MOS structure at threshold	m
x_j	Junction depth	m
x_n	Depletion layer width in an n-type semiconductor	m

x_p	Depletion layer width in a p-type semiconductor	m
α	Absorption coefficient	m^{-1}
	Transport factor	
α_F	Forward active transport factor	
α_n	Ionization rate coefficient for electrons	m^{-1}
α_R	Reverse active transport factor	
α_T	Base transport factor	
β	Current gain	
γ	Body effect parameter	$\text{V}^{1/2}$
γ_E	Emitter efficiency	
Γ_n	Auger coefficient for electrons	m^6s^{-1}
Γ_p	Auger coefficient for holes	m^6s^{-1}
δ_n	Excess electron density	m^{-3}
δ_p	Excess hole density	m^{-3}
δ_R	Depletion layer recombination factor	
$\Delta Q_{n,B}$	Excess electron charge density in the base	C/m^2
ϵ	Dielectric constant	F/m
ϵ_0	Permittivity of vacuum	F/m
ϵ_{ox}	Dielectric constant of the oxide	F/m
ϵ_s	Dielectric constant of the semiconductor	F/m
μ_0	Permeability of vacuum	H/m
Θ	Tunnel probability	
λ	Wavelength	m
μ	Electro-chemical potential	Joule
μ_n	Electron mobility	$\text{m}^2/\text{V}\cdot\text{s}$
μ_p	Hole mobility	$\text{m}^2/\text{V}\cdot\text{s}$
ν	Frequency	Hz
ρ	Charge density per unit volume	C/m^3
	Resistivity	Ωm
ρ_{ox}	Charge density per unit volume in the oxide	C/m^3
σ	Conductivity	$\Omega^{-1}\text{m}^{-1}$
τ	Scattering time	s
τ_n	Electron lifetime	s
τ_p	Hole lifetime	s
ϕ	Potential	V
ϕ_B	Barrier height	V
ϕ_F	Bulk potential	V
ϕ_i	Built-in potential of a p-n diode or Schottky diode	V
ϕ_s	Potential at the semiconductor surface	V
Φ	Flux	$\text{m}^{-2}\text{s}^{-1}$

Φ_M	Workfunction of the metal	V
Φ_{MS}	Workfunction difference between the metal and the semiconductor	V
Φ_S	Workfunction of the semiconductor	V
χ	Electron affinity of the semiconductor	V
Ψ	Wavefunction	
ω	Radial frequency	rad/s

Appendix:



Appendix 2: Physical Constants

Avogadro's number	N_A	6.022 x 10 ²³ atoms per mole
Bohr radius	a_0	52.9177 picometer 0.529177 Angstrom
Boltzmann's constant	k	1.38 x 10 ⁻²³ Joule/Kelvin 8.62 x 10 ⁻⁵ electron Volt/Kelvin
Electronic charge	q	1.602 x 10 ⁻¹⁹ Coulomb
Free electron rest mass	m_0	9.11 x 10 ⁻³¹ kilogram 5.69 x 10 ⁻¹⁶ eV s ² cm ⁻²
Permeability of free space	μ_0	4π x 10 ⁻⁷ Henry/meter
Permittivity of free space	ϵ_0	8.854 x 10 ⁻¹² Farad/meter 8.854 x 10 ⁻¹⁴ Farad/centimeter
Planck's constant	h	6.625 x 10 ⁻³⁴ Joule second 4.134 x 10 ⁻¹⁵ electron Volt second
Reduced Planck's constant	\hbar	1.054 x 10 ⁻³⁴ Joule second
Proton rest mass	M	1.67 x 10 ⁻²⁷ Kilogram
Rydberg constant	R	2.17991 x 10 ⁻¹⁸ Joule 13.6058 electron Volt
Speed of light in vacuum	c	2.998 x 10 ⁸ meter/second 2.998 x 10 ¹⁰ centimeter/second
Thermal voltage (at T = 300 K)	$V_t = \frac{kT}{q}$	25.86 milliVolt

Appendix:



Appendix 3: Material Parameters

Name	Symbol	Germanium	Silicon	Gallium Arsenide
Bandgap energy at 300 K	E_g (eV)	0.66	1.12	1.424
Breakdown Field	E_{br} (V/cm)	10^5	3×10^5 *	4×10^5
Density	(g/cm ³)	5.33	2.33	5.32
Effective density of states in the conduction band at 300 K	N_c (cm ⁻³)	1.02×10^{19}	2.82×10^{19}	4.35×10^{17}
Effective density of states in the valence band at 300 K	N_v (cm ⁻³)	5.65×10^{18}	1.83×10^{19}	7.57×10^{18}
Intrinsic concentration at 300 K	n_i (cm ⁻³)	2.8×10^{13}	1.0×10^{10}	2.0×10^6
Effective mass for density of states calculations				
Electrons	m_e^* / m_0	0.55	1.08	0.067
Holes	m_h^* / m_0	0.37	0.81	0.45
Electron affinity	χ (V)	4.0	4.05	4.07
Lattice constant	a (pm)	564.613	543.095	565.33
Mobility at 300 K (undoped)				
Electrons	μ_n (cm ² /V-s)	3900	1400 [†]	8800
Holes	μ_p (cm ² /V-s)	1900	450 [†]	400
Relative dielectric constant	ϵ_s / ϵ_0	16	11.9	13.1
Thermal conductivity at 300 K	χ (W/cmK)	0.6	1.5	0.46

Refractive index at 632.8 nm wavelength	n	5.441	3.875	3.856
		- i 0.785	- i 0.0181	- i 0.196

*See also section 4.5.1: Breakdown field in silicon at 300 K

†See also section 2.7.2: Mobility of doped silicon at 300 K



Appendix:

Appendix 4: Prefixes

deci	(d)	10^{-1}	deka	(da)	10^1
centi	(c)	10^{-2}	hecto	(h)	10^2
milli	(m)	10^{-3}	kilo	(k)	10^3
micro	(μ)	10^{-6}	mega	(M)	10^6
nano	(n)	10^{-9}	giga	(G)	10^9
pico	(p)	10^{-12}	tera	(T)	10^{12}
femto	(f)	10^{-15}	peta	(P)	10^{15}
atto	(a)	10^{-18}	exa	(X)	10^{18}

Appendix:



Appendix 5: Units

MKS Units

Unit	Variable	Symbol	MKSA units
Ampere	Current	I	fundamental MKSA unit
Coulomb	Charge	Q	Ampere second
Farad	Capacitance	C	Coulomb Volt ⁻¹ = Joule Volt ⁻²
Henry	Inductance	L	Weber Ampere ⁻¹ = Tesla meter ² Ampere ⁻¹
Joule	Energy	E	Newton meter = kilogram meter ² second ⁻²
Kelvin	Temperature	T	fundamental MKSA unit
kilogram	Mass	m	fundamental MKSA unit
meter	Length	L	fundamental MKSA unit
Newton	Force	F	kilogram meter second ⁻²
Ohm	Resistance	Ω	Volt Ampere ⁻¹ = Volt ² Joule ⁻¹ second ⁻¹
Pascal	Pressure	P	Newton meter ⁻²
second	time	t	fundamental MKSA unit
Siemens	conductance	G	Ampere Volt ⁻¹ = Joule second Volt ⁻²
Tesla	Magnetic field	B	Newton Ampere ⁻¹ meter ⁻¹ = Joule Ampere ⁻¹ meter ⁻² = Volt second meter ⁻²
Volt	potential	f	Joule coulomb ⁻¹
Watt	Power	P	Joule second ⁻¹
Weber	Magnetic flux	Φ	Tesla meter ² = Volt second

non-MKS Units

Electron Volt (Unit of energy) = 1.602×10^{-19} Joule

moles/liter (Unit of concentration) = 6.022×10^{20} cm⁻³

Degrees Centigrade (Unit of Temperature) = $273.16 + T$ (in Centigrade) Kelvin

Inch (Unit of Length) = 2.54 cm

mil or milli inch (Unit of Length) = 25.4 micrometer

A or Angstrom (Unit of Length) = 0.1 nm



Appendix:

Appendix 6: The greek alphabet

α	A	alpha	ν	N	nu
β	B	beta	ξ	Ξ	xi
γ	Γ	gamma	o	Ω	omicron
δ	Δ	delta	π	Π	pi
ϵ	E	epsilon	ρ	P	rho
ζ	Z	zeta	σ	Σ	sigma
η	H	eta	τ	T	tau
θ	Θ	theta	υ	Υ	upsilon
ι	I	iota	ϕ	Φ	phi
κ	K	kappa	ξ	Ξ	chi
λ	Λ	lambda	ψ	Ψ	psi
μ	M	mu	ω	Ω	omega

Appendix:



Appendix 7: Periodic Table

Periodic Table of the Elements

The Periodic Table is organized into groups (Groups IA through VIIA, and Group 0) and blocks (s-block, p-block, d-block, and f-block). The table includes element symbols, atomic numbers, and group/period labels. A red border highlights the transition elements (Groups IIIA through VIIA and the f-block). A green border highlights the lanthanide (Ce to Lu) and actinide (Th to Lr) series.

IA		IIA		Metals																		Non metals		VIII B			
1 H		3 Li	4 Be	d Transition Elements																		5 B	6 C	7 N	8 O	9 F	2 He
11 Na	12 Mg			IIIA		IVA		VA		VIA		VIIA		VIII A		VIII A		VIII A		IB		IIB		10 Ne			
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr										
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe										
55 Cs	56 Ba	57 La*	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn										
87 Fr	88 Ra	89 Ac**	104 Unq	105 Unp	106 Uns	f Transition Elements																					
* Lanthanides		58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu												
** Actinides		90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr												

Appendix:



Quick Access

[Title](#) [Page](#) [Table of Contents](#) [Help](#) [Introduction](#)

Chapter 1:	1.1. 1.2. 1.3. 1.4.	1.ex 1.p 1.r 1.b 1.eq
Chapter 2:	2.1. 2.2. 2.3. 2.4. 2.5. 2.6. 2.7. 2.8. 2.9. 2.10.	2.ex 2.p 2.r 2.b 2.eq
Chapter 3:	3.1. 3.2. 3.3. 3.4.	3.ex 3.p 3.r 3.b 3.eq
Chapter 4:	4.1. 4.2. 4.3. 4.4. 4.5. 4.6.	4.ex 4.p 4.r 4.b 4.eq
Chapter 5:	5.1. 5.2. 5.3. 5.4.	5.ex 5.p 5.r 5.b 5.eq
Chapter 6:	6.1. 6.2. 6.3. 6.4.	6.ex 6.p 6.r 6.b 6.eq
Chapter 7:	7.1. 7.2. 7.3. 7.4. 7.5. 7.6. 7.7.	7.ex 7.p 7.r 7.b 7.eq
Appendix:	A.1 A.2 A.3 A.4 A.5 A.6 A.7 A.8	

Appendix 2: Physical constants

Avogadro's number	N_A	6.022×10^{23} atoms per mole
Bohr radius	a_0	52.9177 picometer
Boltzmann's constant	k	0.529177 Angstrom
Electronic charge	q	1.38×10^{-23} Joule/Kelvin
Free electron rest mass	m_0	8.62×10^{-5} electron Volt/Kelvin
Permeability of vacuum	μ_0	1.602×10^{-19} Coulomb
Permittivity of vacuum	ϵ_0	9.11×10^{-31} kilogram
Planck's constant	h	5.69×10^{-16} eV s ² cm ⁻²
Reduced Planck's constant	\hbar	$4\pi \times 10^{-7}$ Henry/meter
Proton rest mass	M	8.854×10^{-12} Farad/meter
Rydberg constant	R	8.854×10^{-14} Farad/centimeter
Speed of light in vacuum	c	6.625×10^{-34} Joule second
Thermal voltage (at T = 300 K)	$V_T = \frac{kT}{q}$	4.134×10^{-15} electron Volt second
		1.054×10^{-34} Joule second
		1.67×10^{-27} Kilogram
		2.17991×10^{-18} Joule
		13.6058 electron Volt
		2.998×10^8 meter/second
		2.998×10^{10} centimeter/second

Appendix 4: Prefixes

deci	(d)	10^{-1}	deka	(da)	10^1
centi	(c)	10^{-2}	hecto	(h)	10^2
milli	(m)	10^{-3}	kilo	(k)	10^3
micro	(μ)	10^{-6}	mega	(M)	10^6
nano	(n)	10^{-9}	giga	(G)	10^9
pico	(p)	10^{-12}	tera	(T)	10^{12}
femto	(f)	10^{-15}	peta	(P)	10^{15}
atto	(a)	10^{-18}	exa	(X)	10^{18}

Appendix 3: Material Parameters

Name	Symbol	Germanium	Silicon	Gallium Arsenide
Bandgap energy at 300 K	E_g (eV)	0.66	1.12	1.424
Breakdown Field	E_{br} (V/cm)	10^5	3×10^5 *	4×10^5
Density	(g/cm ³)	5.33	2.33	5.32
Effective density of states in the conduction band at 300 K	N_c (cm ⁻³)	1.02×10^{19}	2.82×10^{19}	4.35×10^{17}
Effective density of states in the valence band at 300 K	N_v (cm ⁻³)	5.65×10^{18}	1.83×10^{19}	7.57×10^{18}
Intrinsic concentration at 300 K	n_i (cm ⁻³)	2.8×10^{13}	1.0×10^{10}	2.0×10^6
Effective mass for density of states calculations				
Electrons	m_e^* / m_0	0.55	1.08	0.067
Holes	m_h^* / m_0	0.37	0.81	0.45
Electron affinity	χ (V)	4.0	4.05	4.07
Lattice constant	a (pm)	564.613	543.095	565.33
Mobility at 300 K (undoped)				
Electrons	m (cm ² /V-s)	3900	1400 [†]	8800
Holes	m (cm ² /V-s)	1900	450 [†]	400
Relative dielectric constant	ϵ_s / ϵ_0	16	11.9	13.1
Thermal conductivity at 300 K	χ (W/cmK)	0.6	1.5	0.46
Refractive index at 632.8 nm wavelength	n	5.441	3.875	3.856
		- i 0.785	- i 0.0181	- i 0.196

*See also section 4.5.1: Breakdown field in silicon at 300 K

[†]See also section 2.7.2: Mobility of doped silicon at 300 K

Principles of Semiconductor Devices



Table of Contents

Chapter 1: Review of Modern Physics

Chapter 2: Semiconductor Fundamentals

Chapter 3: Metal-Semiconductor Junctions

Chapter 4: p-n Junctions

Chapter 5: Bipolar Transistors

Chapter 6: MOS Capacitors

Chapter 7: MOS Field-Effect-Transistors

Appendix

1. List of symbols

[By symbol](#)

[By name](#)

[Extended List of symbols](#)

2. [Physical Constants](#)

3. [Material Parameters](#)

4. [Prefixes](#)

5. [Units](#)

6. [Greek Alphabet](#)

7. [Periodic Table](#)

[Quick access](#)

- A -

Acceptor doping density	N_a	m^{-3}
Acceptor energy	E_a	Joule
Applied voltage	V_a	V
Area	A	m^2

- B -

Barrier height	ϕ_B	V
Base doping density	N_B	m^{-3}
Base voltage	V_B	V
Base width	w_B	m
Body effect parameter	γ	$\text{V}^{1/2}$
Boltzmann's constant	k	J/K
Built-in potential of a p-n diode or Schottky diode	ϕ_i	V
Bulk potential	ϕ_F	V

- C -

Capacitance per unit area	C	F/m^2
Charge	Q	C
Charge density per unit area at threshold in the depletion layer of an MOS structure	$Q_{d,T}$	C/m^2
Charge density per unit area in the depletion layer of an MOS structure	Q_d	C/m^2
Charge density per unit volume	ρ	C/m^3
Charge density per unit volume in the oxide	ρ_{ox}	C/m^3
Collector doping density	N_C	m^{-3}
Collector voltage	V_C	V
Collector width	w_C	m
Conduction band energy of a semiconductor	E_c	Joule
Conductivity	σ	$\Omega^{-1}\text{m}^{-1}$
Current	I	A
Current density	J	A/m^2
Current gain	β	

- D -

Density of states in the conduction band per unit energy and per unit volume	$g_c(E)$	$\text{m}^{-3}\text{J}^{-1}$
Density of states in the valence band per unit energy and per unit volume	$g_v(E)$	$\text{m}^{-3}\text{J}^{-1}$
Depletion layer width	w	m
Depletion layer width in a p-type semiconductor	x_p	m
Depletion layer width in an MOS structure	x_d	m
Depletion layer width in an MOS structure at threshold	$x_{d,T}$	m
Depletion layer width in an n-type semiconductor	x_n	m
Dielectric constant of the oxide	ϵ_{ox}	F/m
Dielectric constant of the semiconductor	ϵ_s	F/m
Distribution function (probability density function)	$f(E)$	
Donor doping density	N_d	m^{-3}
Donor energy	E_d	Joule
Doping density	N	
Drain voltage	V_D	V

- E -

Effective density of states in the conduction band	N_c	m^{-3}
Effective density of states in the valence band	N_v	m^{-3}
Effective mass of electrons	m_e^*	kg
Effective mass of holes	m_h^*	kg
Electric field	E	V/m
Electron affinity of the semiconductor	χ	V
Electron current density	J_n	A/m^2
Electron density	n	m^{-3}
Electron density in thermal equilibrium	n_0	m^{-3}
Electron density per unit energy and per unit volume	$n(E)$	m^{-3}
Electron diffusion constant	D_n	m^2/s
Electron diffusion length	L_n	m
Electron energy in vacuum	E_{vacuum}	Joule
Electron generation rate	G_n	$\text{m}^{-3}\text{s}^{-1}$

Electron lifetime	τ_n	s
Electron mobility	μ_n	$\text{m}^2/\text{V}\cdot\text{s}$
Electron recombination rate	R_n	$\text{m}^{-3}\text{s}^{-1}$
electronic charge	q	C
Emitter doping density	N_E	m^{-3}
Emitter efficiency	γ_E	
Emitter voltage	V_E	V
Emitter width	w_E	m
Energy	E	Joule
Energy bandgap of a semiconductor	E_g	Joule
Excess electron charge density in the base	$\Delta Q_{n,B}$	C/m^2
Excess electron density	δ_n	m^{-3}
Excess hole density	δ_p	m^{-3}

- F -

Fermi energy (thermal equilibrium)	E_F	Joule
Flatband capacitance per unit area of a MOS structure	C_{FB}	F/m^2
Flatband voltage	V_{FB}	V
Free electron mass	m_0	kg

- G -

Gate voltage	V_G	V
--------------	-------	---

- H -

Hole current density	J_p	A/m^2
Hole density	p	m^{-3}
Hole density in an n-type semiconductor	p_n	m^{-3}
Hole density in thermal equilibrium	p_0	m^{-3}
Hole density per unit energy	$p(E)$	m^{-3}
Hole diffusion constant	D_p	m^2/s
Hole diffusion length	L_p	m
Hole generation rate	G_p	$\text{m}^{-3}\text{s}^{-1}$

Hole lifetime	τ_p	s
Hole mobility	μ_p	$\text{m}^2/\text{V}\cdot\text{s}$
Hole recombination rate	R_p	$\text{m}^{-3}\text{s}^{-1}$

- I -

Interface charge density per unit area	Q_I	C/m^2
Intrinsic carrier density	n_I	m^{-3}
Intrinsic carrier density	n_I	m^{-3}
Intrinsic Fermi energy	E_I	Joule
Ionized acceptor density	N_a^-	m^{-3}
Ionized donor density	N_d^+	m^{-3}

- J -

Junction capacitance per unit area	C_j	F/m^2
Junction depth	x_j	m

- K -

- L -

Length	L	m
--------	-----	---

- M -

Mass	m	kg
Mean free path	l	m

- N -

Net recombination rate of electrons	U_n	$\text{m}^{-3}\text{s}^{-1}$
Net recombination rate of holes	U_p	$\text{m}^{-3}\text{s}^{-1}$

- O -

Oxide capacitance per unit area	C_{ox}	F/m^2
---------------------------------	----------	-----------------------

Oxide thickness

 t_{ox} m

- P -

Plank's constant

 h Js

Position

 x m

Potential

 ϕ V

Potential at the semiconductor surface

 ϕ_s V

- Q -

Quasi-Fermi energy of electrons

 F_n Joule

Quasi-Fermi energy of holes

 F_p Joule

- R -

Reduced Plank's constant ($= h /2\pi$) \hbar Js

Resistance

 R Ohm

Resistivity

 ρ Ω m

- S -

Speed of light in vacuum

 c m/s

- T -

Temperature

 T Kelvin

Thermal velocity

 v_{th} m/s

Thermal voltage

 V_t V

Thickness

 t m

Threshold voltage of an MOS structure

 V_T V

Transport factor

 α

- U -

Valence band energy of a semiconductor

 E_v Joule

Velocity	v	m/s
----------	-----	-----

- W -

Width of a p-type region	w_p	m
Width of an n-type region	w_n	m
Workfunction difference between the metal and the semiconductor	Φ_{MS}	V
Workfunction of the metal	Φ_M	V
Workfunction of the semiconductor	Φ_S	V

- X -

- Y -

- Z -

Appendix 1: List of symbols

Symbol	Description	MKS Units
a	Acceleration	m/s^2
a_0	Bohr radius	m
A	Area	m^2
A^*	Richardson constant	m/s
A_C	Collector area	m^2
A_E	Emitter area	m^2
b	Bimolecular recombination constant	m^3/s
c	Speed of light in vacuum	m/s
C	Capacitance per unit area	F/m^2
C_D	Diffusion capacitance per unit area	F/m^2
C_{DS}	Drain-source capacitance	F
C_{FB}	Flatband capacitance per unit area of a MOS structure	F/m^2
C_G	Gate capacitance	F
C_{GS}	Gate-source capacitance	F
C_{GD}	Gate-drain capacitance	F
C_{HF}	High-frequency capacitance per unit area of a MOS structure	F/m^2
C_j	Junction capacitance per unit area	F/m^2
C_{LF}	Low-frequency (quasi-static) capacitance per unit area of a MOS structure	F/m^2
C_M	Miller capacitance	F
C_{ox}	Oxide capacitance per unit area	F/m^2
C_s	Semiconductor capacitance per unit area	F/m^2
D_n	Electron diffusion constant	m^2/s
D_p	Hole diffusion constant	m^2/s
E	Energy	Joule

E	Electric field	V/m
E_0	Lowest energy in a one-dimensional quantum well	Joule
E_a	Acceptor energy	Joule
E_{br}	Breakdown field	V/m
E_c	Conduction band energy of a semiconductor	Joule
E_d	Donor energy	Joule
E_F	Fermi energy (thermal equilibrium)	Joule
$E_{F,n}$	Fermi energy in an n-type semiconductor	Joule
$E_{F,p}$	Fermi energy in a p-type semiconductor	Joule
E_g	Energy bandgap of a semiconductor	Joule
E_i	Intrinsic Fermi energy	Joule
E_n	n^{th} quantized energy	Joule
E_{ph}	Photon energy	Joule
E_t	Trap energy	Joule
E_v	Valence band energy of a semiconductor	Joule
E_{vacuum}	Electron energy in vacumm	Joule
$F(E)$	Distribution function (probability density function)	
$f_{BE}(E)$	Bose-Einstein distribution function	
$f_{FD}(E)$	Fermi-Dirac distribution function	
$f_{MB}(E)$	Maxwell-Boltzmann distribution function	
F	Force	Newton
F_n	Quasi-Fermi energy of electrons	Joule
F_p	Quasi-Fermi energy of holes	Joule
$g(E)$	Density of states per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
$g(E)$	Density of states in the conduction band per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
$g(E)$	Density of states in the valence band per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
g_d	Output conductance of a MOSFET	S

g_m	Transconductance of a MOSFET	S
G	Carrier generation rate	$\text{m}^{-3}\text{s}^{-1}$
G_n	Electron generation rate	$\text{m}^{-3}\text{s}^{-1}$
G_p	Hole generation rate	$\text{m}^{-3}\text{s}^{-1}$
h	Plank's constant	Js
\hbar	Reduced Plank's ($= h / 2\pi$)	Js
I	Current	A
I_B	Base current of a bipolar transistor	A
I_C	Collector current of a bipolar transistor	A
I_D	Drain current of a MOSFET	A
I_E	Emitter current of a bipolar transistor	A
I_F	Forward active current of a bipolar transistor	A
$I_{D,sat}$	Drain current of a MOSFET in saturation	A
I_{ph}	Photo current	A
I_r	Recombination current	A
I_R	Reverse active current of a bipolar transistor	A
I_s	Saturation current	A
I_{sc}	Short circuit current of a solar cell	A
J	Current density	A/m^2
J_n	Electron current density	A/m^2
J_p	Hole current density	A/m^2
k	Boltzmann's constant wavenumber	J/K m^{-1}
l	Mean free path	m
L	Length	m
L_D	Debye length	m
L_n	Electron diffusion length	m

L_p	Hole diffusion length	m
L_x	Hole diffusion length	m
m	Mass	kg
m^*	Effective mass	kg
m_A	Atomic mass	kg
m_0	Free electron mass	kg
m_e^*	Effective mass of electrons	kg
m_h^*	Effective mass of holes	kg
M	Proton mass	kg
M	Multiplication factor	
	Electron density	m^{-3}
n	Integer	
n	Refractive index	
	Ideality factor	
n_i	Intrinsic carrier density	m^{-3}
$n(E)$	Electron density per unit energy and per unit volume	m^{-3}
n_0	Electron density in thermal equilibrium	m^{-3}
n_i	Intrinsic carrier density	m^{-3}
n_n	Electron density in an n-type semiconductor	m^{-3}
n_{n0}	Thermal equilibrium electron density in an n-type semiconductor	m^{-3}
n_p	Electron density in a p-type semiconductor	m^{-3}
n_{p0}	Thermal equilibrium electron density in a p-type semiconductor	m^{-3}
N	Number of particles	
	Doping density	
N_a	Acceptor doping density	m^{-3}
N_a^-	Ionized acceptor density	m^{-3}
N_A	Avogadro's number	

N_B	Base doping density	m^{-3}
N_c	Effective density of states in the conduction band	m^{-3}
N_C	Collector doping density	m^{-3}
N_d	Donor doping density	m^{-3}
N_d^+	Ionized donor density	m^{-3}
N_E	Emitter doping density	m^{-3}
N_{ss}	Surface state density	m^{-2}
N_t	Recombination trap density	m^{-2}
N_v	Effective density of states in the valence band	m^{-3}
	Hole density	m^{-3}
p	Particle momentum	kgm/s
	Pressure	Nm^{-2}
$p(E)$	Hole density per unit energy	m^{-3}
p_0	Hole density in thermal equilibrium	m^{-3}
p_n	Hole density in an n-type semiconductor	m^{-3}
p_{no}	Thermal equilibrium hole density in an n-type semiconductor	m^{-3}
p_p	Hole density in a p-type semiconductor	m^{-3}
p_{p0}	Thermal equilibrium hole density in a p-type semiconductor	m^{-3}
q	electronic charge	C
Q	Heat	Joule
	Charge	C
Q_B	Majority carrier charge density in the base	C/m^2
Q_d	Charge density per unit area in the depletion layer of an MOS structure	C/m^2
$Q_{d,T}$	Charge density per unit area at threshold in the depletion layer of an MOS structure	C/m^2
Q_i	Interface charge density per unit area	C/m^2
Q_{inv}	Inversion layer charge density per unit area	C/m^2

Q_M	Charge density per unit area in a metal	C/m^2
Q_n	Charge density per unit area in the depletion layer of an n-type region	C/m^2
Q_p	Charge density per unit area in the depletion layer of a p-type region	C/m^2
Q_{ss}	Surface state charge density per unit area	C/m^2
r_e	Emitter resistance	Ohm
r_p	Base resistance	Ohm
R	The Rydberg constant	J
	Resistance	Ohm
R_n	Electron recombination rate	$\text{m}^{-3}\text{s}^{-1}$
R_p	Hole recombination rate	$\text{m}^{-3}\text{s}^{-1}$
R_s	Sheet resistance	Ohm
s	Spin	
S	Entropy	J/K
t	Thickness	m
t_{ox}	Oxide thickness	m
T	Temperature	Kelvin
	Kinetic energy	Joule
u_w	Spectral density	$\text{Jm}^{-3}\text{s}^{-1}$
U	Total energy	Joule
U_A	Auger recombination rate	$\text{m}^{-3}\text{s}^{-1}$
U_{b-b}	Band-to-band recombination rate	$\text{m}^{-3}\text{s}^{-1}$
U_n	Net recombination rate of electrons	$\text{m}^{-3}\text{s}^{-1}$
U_p	Net recombination rate of holes	$\text{m}^{-3}\text{s}^{-1}$
U_{SHR}	Shockley-Read-Hall recombination rate	$\text{m}^{-3}\text{s}^{-1}$
v	Velocity	m/s
v_R	Richardson velocity	m/s
v_{sat}	Saturation velocity	m/s

v_{th}	Thermal velocity	m/s
V	Potential energy	Joule
	Volume	m^3
V_a	Applied voltage	V
V_A	Early voltage	V
V_{br}	Breakdown voltage	V
V_B	Base voltage	V
V_{BE}	Base-emitter voltage	V
V_{BC}	Base-collector voltage	V
V_C	Collector voltage	V
V_{CE}	Collector-emitter voltage	V
V_D	Drain voltage	V
V_{DS}	Drain-source voltage	V
$V_{DS,sat}$	Drain-source saturation voltage	V
V_E	Emitter voltage	V
V_{FB}	Flatband voltage	V
V_G	Gate voltage	V
V_{GS}	Gate-source voltage	V
V_{oc}	Open circuit voltage of a solar cell	V
V_t	Thermal voltage	V
V_T	Threshold voltage of an MOS structure	V
w	Depletion layer width	m
w_B	Base width	m
w_C	Collector width	m
w_E	Emitter width	m
w_n	Width of an n-type region	m
w_p	Width of a p-type region	m

W	Work	Joule
x	Position	m
x_d	Depletion layer width in an MOS structure	m
$x_{d,T}$	Depletion layer width in an MOS structure at threshold	m
x_j	Junction depth	m
x_n	Depletion layer width in an n-type semiconductor	m
x_p	Depletion layer width in a p-type semiconductor	m
a	Absorption coefficient	m^{-1}
a	Transport factor	
a_F	Forward active transport factor	
a_n	Ionization rate coefficient for electrons	m^{-1}
a_R	Reverse active transport factor	
a_T	Base transport factor	
b	Current gain	
g	Body effect parameter	$\text{V}^{1/2}$
g_E	Emitter efficiency	
G_n	Auger coefficient for electrons	$\text{m}^6 \text{s}^{-1}$
G_p	Auger coefficient for holes	$\text{m}^6 \text{s}^{-1}$
d_n	Excess electron density	m^{-3}
d_p	Excess hole density	m^{-3}
d_R	Depletion layer recombination factor	
$DQ_{n,B}$	Excess electron charge density in the base	C/m^2
ϵ	Dielectric constant	F/m
ϵ_0	Permittivity of vacuum	F/m
ϵ_{ox}	Dielectric constant of the oxide	F/m

ϵ_s	Dielectric constant of the semiconductor	F/m
μ_0	Permeability of vacuum	H/m
Q	Tunnel probability	
λ	Wavelength	m
m	Electro-chemical potential	Joule
m_e	Electron mobility	$\text{m}^2/\text{V}\cdot\text{s}$
m_h	Hole mobility	$\text{m}^2/\text{V}\cdot\text{s}$
ν	Frequency	Hz
r	Charge density per unit volume	C/m^3
r_{ox}	Resistivity	Ωm
r_{ox}	Charge density per unit volume in the oxide	C/m^3
s	Conductivity	$\Omega^{-1}\text{m}^{-1}$
t	Scattering time	s
t_n	Electron lifetime	s
t_p	Hole lifetime	s
f	Potential	V
f_B	Barrier height	V
f_F	Bulk potential	V
f_i	Built-in potential of a p-n diode or Schottky diode	V
f_s	Potential at the semiconductor surface	V
F	Flux	$\text{m}^{-2}\text{s}^{-1}$
F_M	Workfunction of the metal	V
F_{MS}	Workfunction difference between the metal and the semiconductor	V
F_S	Workfunction of the semiconductor	V
c	Electron affinity of the semiconductor	V
Y	Wavefunction	

w Radial frequency rad/s

Appendix 5: Units

MKS Units

Ampere (Unit of current) = fundamental MKSA unit

Coulomb (Unit of charge) = Ampere second

Farad (Unit of capacitance) = Coulomb Volt⁻¹ = Joule Volt⁻²

Henry (Unit of inductance) = Weber Ampere⁻¹ = Tesla meter² Ampere⁻¹

Joule (Unit of energy) = Newton meter = kilogram meter² second⁻²

Kelvin (Unit of temperature) = fundamental MKSA unit

kilogram (Unit of mass) = fundamental MKSA unit

meter (Unit of length) = fundamental MKSA unit

Newton (Unit of force) = kilogram meter second⁻²

Ohm (Unit of resistance) = Volt Ampere⁻¹ = Volt² Joule⁻¹ second⁻¹

Pascal (Unit of pressure) = Newton meter⁻²

second (Unit of time) = fundamental MKSA unit

Siemens (Unit of conductance) = Ampere Volt⁻¹ = Joule second Volt⁻²

Tesla (Unit of magnetic field) = Newton Ampere⁻¹ meter⁻¹ = Joule Ampere⁻¹ meter⁻² = Volt second meter⁻²

Volt (Unit of potential) = Joule coulomb⁻¹

Watt (Unit of power) = Ampere Volt = Joule second⁻¹

Weber (Unit of magnetic flux) = Tesla meter² = Volt second

non-MKS Units

Electron Volt (Unit of energy) = 1.602×10^{-19} Joule

moles/liter (Unit of concentration) = 6.022×10^{20} cm⁻³

Appendix:



Appendix 8: Numeric answers to selected problems

[Spreadsheet](#) with predefined physical constants and material parameters

[Chapter 1](#)

1. **0.62 μm , 0.87 nm**
2. **1.55 eV, 375 THz, 4×10^{18}**
4. **16.5 μm**
6. **3.88 nm**
7. **-13.6 eV, -3.4 eV, -1.51 eV**

[Chapter 2](#)

1. **52.36%, 68.02%, 74.05%, 34.01%**
2. **406.3°C**
4. **0.0099**
5. **6.29×10^{-5} , 0.045, 484.7°C**
6. **1.02×10^{19} , 5.65×10^{18} ; 2.82×10^{19} , 1.83×10^{19} ; 4.37×10^{17} , $7.57 \times 10^{18} \text{ cm}^{-3}$**
 1.42×10^{19} , 7.83×10^{18} ; 3.91×10^{19} , 2.54×10^{19} ; 6.04×10^{17} , $1.05 \times 10^{18} \text{ cm}^{-3}$
7. **2.16×10^{13} , 8.81×10^9 , 1.97×10^6 ; 3.67×10^{14} , 8.55×10^{11} , $6.04 \times 10^8 \text{ cm}^{-3}$**
8. **-7.68 meV, -5.58 meV, 36.91 meV
-9.56 meV, -6.94 meV, 45.92 meV**
9. **2.24×10^{18} , 1.48×10^8 ; 1.60×10^{15} , 1.32×10^5 ; $2.23 \times$**

- 10¹¹, 18.4 cm⁻³**
9.27 x 10¹³, 4.45 x 10¹²; 6.02 x 10¹⁴, 3.50 x 10⁵; 3.97 x 10¹², 1.04 cm⁻³
- 11. 10⁴ cm⁻³, 0.357 eV**
- 12. 10¹² cm⁻³, 10⁸ cm⁻³**
- 13. 9.17 x 10¹⁴, 9.15 x 10¹⁴, 2.76 x 10¹⁵ cm⁻³**
- 14. 7.59 x 10¹⁶ cm⁻³**
- 15. 2.77 x 10¹⁹ cm⁻³s⁻¹, 2.77 x 10¹⁵ cm⁻³, 2.77 x 10¹⁵ cm⁻³, 417 meV, -324 meV**
- 16. 10¹³ cm⁻³s⁻¹, 10⁻³ s**
- 17. 2.96 x 10⁻⁶ cm/Ohm, 337 kOhm-cm**
- 18. 393 kOhm-cm, 5.67 x 10⁹, 1.76 x 10¹⁰, 1.20 x 10¹⁰ cm⁻³**
- 19. 18, 42, 60 meV**
- 20. 2.42 nm**
- 21. 2.42 x 10²¹ cm⁻³eV⁻¹**
- 22. 4.74 %, 95.26 %**
- 24. 1.39**
- 25. 30 kV/cm, 15 V, 16.7 ps**
- 26. 8.92 x 10¹⁴ cm⁻³, 6.94 x 10¹⁵ cm⁻³**
- 27. 0.455 mm**
- 28. 2111 cm²/V-s**
- 29. 1.05 cm⁻³J⁻¹, 9.14 x 10⁻⁴**
- 30. 0.53 Ω-cm, 2.1 x 10⁻⁷, 0.22 Ω-cm, 123 meV**
- 31. 10¹⁶ cm⁻³, 10⁴ cm⁻³, -357 meV
 10¹⁵ cm⁻³, 10⁵ cm⁻³, 298 meV**
- 32. 828 A/cm²**

[Chapter 4](#)[Chapter 5](#)[Chapter 6](#)

1. **-0.927, 0.149, -0.153, 0.632, -1.071, 0.005, -0.963, 0.113 V** 
2. **180 nF/cm², 19.2 nm, 176 nm, $3.3 \times 10^{16} \text{ cm}^{-3}$, 388 mV, 13 pF, -328 mV, 965 mV**
3. **$7.2 \times 10^{16} \text{ cm}^{-3}$ **
4. **7.62×10^{16} , 2.31×10^{16} , 1.49×10^{14} , $7.51 \times 10^{14} \text{ cm}^{-3}$**
5. **1.38×10^{14} , $7.69 \times 10^{14} \text{ cm}^{-3}$**
6. **-1.11 V, 48.4 nF/cm²**

[Chapter 7](#)

1. **322 Ohm, 2.25 V **
2. **173 nF/cm², 17.3 fF, 51.8 μA**
3. **154 μm **
4. **48.3 μm**
5. **0.62 Ohm **
6. **0.31 μm , 46.6 kV/cm**
7. **105 nm**
8. **158 nm**
10. **170 nm**

Appendix 7: Periodic Table of the Elements

IA	IIA	IIIA	IVA	VA	VIA	VIIA	VIII A	VIII A	IB	II B	III B	IV B	VB	VI B	VII B	VIII B	
1 H																2 He	
PERIODIC TABLE OF THE ELEMENTS																	
3 Li	4 Be																
11 Na	12 Mg																
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	57 La ⁺	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	89 Ac ^{**}	104 Unq	105 Unp	106 Uns												
f Transition Elements																	
* Lanthanides		58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu		
** Actinides		90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No		103 Lr	

Appendix 7: Periodic Table of the Elements

1 H	2 He	<h1>Periodic Table of the Elements</h1>															
3 Li	4 Be																
11 Na	12 Mg	d Transition Elements															
19 K	20 Ca	21 Sc	22 Ti	23 V	24 Cr	25 Mn	26 Fe	27 Co	28 Ni	29 Cu	30 Zn	31 Ga	32 Ge	33 As	34 Se	35 Br	36 Kr
37 Rb	38 Sr	39 Y	40 Zr	41 Nb	42 Mo	43 Tc	44 Ru	45 Rh	46 Pd	47 Ag	48 Cd	49 In	50 Sn	51 Sb	52 Te	53 I	54 Xe
55 Cs	56 Ba	57 La*	72 Hf	73 Ta	74 W	75 Re	76 Os	77 Ir	78 Pt	79 Au	80 Hg	81 Tl	82 Pb	83 Bi	84 Po	85 At	86 Rn
87 Fr	88 Ra	89 Ac**	104 Unq	105 Unp	106 Uns	f Transition Elements											
* Lanthanides		58 Ce	59 Pr	60 Nd	61 Pm	62 Sm	63 Eu	64 Gd	65 Tb	66 Dy	67 Ho	68 Er	69 Tm	70 Yb	71 Lu		
** Actinides		90 Th	91 Pa	92 U	93 Np	94 Pu	95 Am	96 Cm	97 Bk	98 Cf	99 Es	100 Fm	101 Md	102 No	103 Lr		

Germanium - [Ar] 3d¹⁰ 4s² 4p²

Atomic number 32

Atomic mass 72.59

Melting point 937.4° C

Boiling point 2830° C

Density 5.32 g/cm³

Electronegativity 1.8

Mohs Hardness

Crystal structure Diamond

More parameters can be found in [Appendix 3](#)

Silicon - [Ne] 3s² 3p²

Atomic number 14

Atomic mass 28.07

Melting point 1410° C

Boiling point 2355° C

Density 2.33 g/cm³

Electronegativity 1.8

Mohs Hardness 7.0

Crystal structure Diamond

More parameters can be found in [Appendix 3](#)

Aluminum - [Ne] 3s² 3p¹

Atomic number 13

Atomic mass 26.98

Melting point 660° C

Boiling point 2467° C

Density 2.70 g/cm³

Electronegativity 1.5

Mohs Hardness 2.9

Crystal structure Face centered cubic

Workfunction 4.1 V

Carbon - [He] 2s² 2p²

Atomic number 6

Atomic mass 12.01

Melting point 3727° C

Boiling point 4830° C

Density 2.26 g/cm³

Electronegativity 1.8

Mohs Hardness 7.0

Crystal structure Diamond

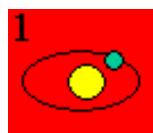
Helium - 1s²

Atomic number	2
Atomic mass	4.0026
Melting point	
Boiling point @ 1 atm.	- 268.94° C (4.2 K)
Density	0.179 mg/cm ³
Electronegativity	
Mohs Hardness	
Crystal structure	

Hydrogen - 1s¹

Atomic number	1
Atomic mass	1.008
Melting point	
Boiling point @ 1 atm.	- 252.8° C (20.35 K)
Density	0.082 mg/cm ³
Electronegativity	
Mohs Hardness	
Crystal structure	

Principles of Semiconductor Devices



Chapter 1: Review of Modern Physics

1. [Introduction](#)
2. [Quantum Mechanics](#)
3. [Electromagnetic theory](#)
4. [Statistical Thermodynamics](#)

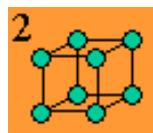
[Examples - Problems](#)

[Review Questions](#)

[Bibliography](#)

[Equation Sheet](#)

Principles of Semiconductor Devices



Chapter 2: Semiconductor Fundamentals

1. [Introduction](#)
2. [Crystals and crystal structures](#)
3. [Energy bands](#)
4. [Density of States](#)
5. [Distribution Functions](#)
6. [Carrier Densities](#)
7. [Carrier Transport](#)
8. [Recombination and Generation](#)
9. [Continuity Equation](#)
10. [Drift-Diffusion Model](#)

[Examples - Problems](#)

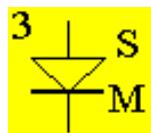
[Review Questions](#)

[Bibliography](#)

[Glossary](#)

[Equation sheet](#)

Principles of Semiconductor Devices



Chapter 3: Metal-Semiconductor Junctions

1. [Introduction](#)
2. [Structure and principle of operation](#)
3. [Electrostatic analysis](#)
4. [Schottky diode current](#)

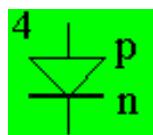
[Examples - Problems](#)

[Review Questions](#)

[Bibliography](#)

[Equation Sheet](#)

Principles of Semiconductor Devices



Chapter 4: p-n Junctions

1. [Introduction](#)
2. [Structure and principle of operation](#)
3. [Electrostatic analysis](#)
4. [p-n diode current](#)
5. [Breakdown](#)
6. [Optoelectronic devices](#)

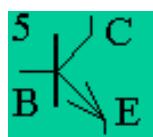
[Examples - Problems](#)

[Review Questions](#)

[Bibliography](#)

[Equation Sheet](#)

Principles of Semiconductor Devices



Chapter 5: Bipolar Transistors

1. [Introduction](#)
2. [Structure](#)
3. [Ideal BJT model](#)
4. [Non-ideal effects](#)

[Examples - Problems](#)

[Review Questions](#)

[Bibliography](#)

[Equation Sheet](#)

Principles of Semiconductor Devices



Chapter 6: MOS Capacitors

1. [Introduction](#)
2. [Structure](#)
3. [Analysis](#)
4. [Technology](#)

[Examples - Problems](#)

[Review Questions](#)

[Bibliography](#)

[Equation Sheet](#)

Principles of Semiconductor Devices



Chapter 7: MOS Field-Effect-Transistors

1. [Introduction](#)
2. [Structure and principle of operation](#)
3. [MOSFET analysis](#)
4. [Threshold voltage](#)
5. [MOSFET SPICE model](#)
6. [MOSFET circuits and technology](#)
7. [Advance MOSFET issues](#)

[Examples - Problems](#)

[Review Questions](#)

[Bibliography](#)

[Equation Sheet](#)

Principles of Semiconductor Devices



Appendix

1. List of symbols

[By symbol](#)

[By name](#)

[Extended List of symbols](#)

2. [Physical Constants](#)

3. [Material Parameters](#)

4. [Prefixes](#)

5. [Units](#)

6. [Greek Alphabet](#)

7. [Periodic Table](#)

8. [Solutions to Selected Problems](#)

[Quick access](#)

-
- Problem 6.1 Consider an aluminum-SiO₂-silicon MOS capacitor ($F_M = 4.1$ V, $\epsilon_{ox}/\epsilon_0 = 3.9$, $\chi = 4.05$ V and $N_a = 10^{17}$ cm⁻³) MOS capacitor with $t_{ox} = 5$ nm.
- a) Calculate the flatband voltage and threshold voltage.
 - b) Repeat for an n-type silicon substrate with $N_d = 10^{16}$ cm⁻³.
 - c) Repeat with a surface charge of 10^{-7} C/cm².
 - d) Repeat with a charge density in the oxide of 10^{-1} C/cm³.
-

Solution

The work function difference of the nMOS capacitor equals

$$\begin{aligned}\Phi_M - \Phi_S &= \Phi_M - \mathbf{c} - \frac{E_g}{2q} - V_t \ln\left(\frac{N_a}{n_i}\right) \\ &= 4.1 - 4.05 - 0.56 - 0.42 = -0.93 = V_{FB}\end{aligned}$$

Since no charge is present in the oxide or at the interface, the flat band voltage equals the work function difference.

The threshold voltage equals:

$$\begin{aligned}V_T &= V_{FB} + 2\mathbf{f}_F + \frac{\sqrt{4\mathbf{e}_s q N_a \mathbf{f}_F}}{C_{ox}} \\ &= -0.93 + 2 \times 0.42 \\ &+ \frac{\sqrt{4 \times 11.9 \times 8.85 \times 10^{-14} \times 1.6 \times 10^{-19} \times 10^{17} \times 0.42}}{3.9 \times 8.85 \times 10^{-14} / 5 \times 10^{-7}} \\ &= 0.15 \text{ V}\end{aligned}$$

For the pMOS capacitor one finds similarly:

$$\begin{aligned}\Phi_M - \Phi_S &= \Phi_M - \mathbf{c} - \frac{E_g}{2q} + V_t \ln\left(\frac{N_d}{n_i}\right) \\ &= 4.1 - 4.05 - 0.56 + 0.36 = -0.15 = V_{FB}\end{aligned}$$

and

$$\begin{aligned}V_T &= V_{FB} - 2\mathbf{f}_F - \frac{\sqrt{4\mathbf{e}_s q N_d \mathbf{f}_F}}{C_{ox}} \\ &= -0.15 - 2 \times 0.36 \\ &- \frac{\sqrt{4 \times 11.9 \times 8.85 \times 10^{-14} \times 1.6 \times 10^{-19} \times 10^{16} \times 0.36}}{3.9 \times 8.85 \times 10^{-14} / 5 \times 10^{-7}} \\ &= -0.94 \text{ V}\end{aligned}$$

In the presence of 10^{-7} C/cm^2 interface charge the flat band voltage of the nMOS capacitor becomes:

$$V_{FB} = \Phi_{MS} - \frac{Q_i}{C_{ox}} = -0.93 - 0.145 = -1.07$$

The threshold voltage shift by the same amount yielding:

$$V_T = 0.15 - 0.145 = -0.005$$

In the presence of 10^{-1} C/cm^3 throughout the oxide the flat band voltage of the nMOS capacitor becomes:

$$V_{FB} = \Phi_{MS} - \frac{1}{\mathbf{e}_{ox}} \int_0^{t_{ox}} \mathbf{r}_{ox}(x) x dx = -0.93 - 0.03 = -0.96$$

The threshold voltage shift by the same amount yielding:

$$V_T = 0.149 - 0.036 = 0.113 \text{ V}$$

Problem 6.3 An MOS capacitor with an oxide thickness of 20 nm has an oxide capacitance which is three times larger than the minimum high-frequency capacitance in inversion. Find the substrate doping density.

Solution The high frequency capacitance in inversion equals:

$$C_{ox} = 3C_{HF,inv} = \frac{3}{\frac{1}{C_{ox}} + \frac{x_{d,T}}{\epsilon_s}}$$

from which one finds $x_{d,T} = 2 \epsilon_s / C_{ox}$
so that:

$$N_a = \frac{f_F}{q\epsilon_s} C_{ox}^2$$

where the bulk potential, f_F , also depends on the doping density, N_a

Iterating, starting with $f_F = 0.4$, yields the following values for the doping density:

$$N_a = 7.06 \times 10^{16}, 7.20 \times 10^{16} \text{ and } 7.21 \times 10^{16} \text{ cm}^{-3}$$

Problem 7.1 Consider an n-type MOSFET which consists of a 10 nm thick oxide ($\epsilon_r = 3.9$) and has a gate length of 1 micron, a gate width of 20 micron and a threshold voltage of 1.5 Volt. Calculate the resistance of the MOSFET in the linear region as measured between source and drain when applying a gate-source voltage of 3 Volt. What should the gate-source voltage be to double the resistance? The surface mobility of the electrons is $300 \text{ cm}^2/\text{V}\cdot\text{sec}$.

Solution The resistance of a MOSFET in the linear region equals

$$\begin{aligned} R_{MOSFET} &= \frac{V_{DS}}{I_D} = \frac{L}{mC_{ox}W(V_{GS} - V_T)} \\ &= \frac{10^{-4}}{300 \times \frac{3.9 \times 8.85 \times 10^{-14}}{10^{-6}} \times 20 \times 10^{-4} \times (V_{GS} - V_T)} \\ &= 322 \text{ Ohm} \end{aligned}$$

To double the resistance one has to half $V_{GS} - V_T$ so that $V_{GS} = 2.25 \text{ V}$

Problem 7.3 A n-type MOSFET ($L = 1 \mu\text{m}$, $t_{ox} = 15 \text{ nm}$, $V_T = 1 \text{ V}$ and $\mu_n = 300 \text{ cm}^2/\text{V}\cdot\text{sec}$) must provide a current of 20 mA at a drain-source voltage of 0.5 Volt and a gate-source voltage of 5 Volt. How wide should the gate be?

Solution The MOSFET is not in saturation so that the gate width can be obtained from:

$$\begin{aligned} W &= \frac{I_D L}{m C_{ox} (V_{GS} - V_T - \frac{V_{DS}}{2}) V_{DS}} \\ &= \frac{0.02 \times 10^{-4}}{300 \times \frac{3.9 \times 8.85 \times 10^{-14}}{1.5 \times 10^{-6}} \times (5 - 1 - 0.25) \times 0.5} \\ &= 154 \text{ micron} \end{aligned}$$

Problem 7.5

The capacitance of an n-type silicon MOSFET is 1 pF. Provided that the oxide thickness is 50 nm and the gatelength is 1 micron, what is the resistance of the MOSFET in the linear regime when biased at a gate voltage, which is 5 Volt larger than the threshold voltage? Use a reasonable value for the surface mobility knowing that the bulk mobility equals $1400 \text{ cm}^2/\text{V}\cdot\text{sec}$.

Solution

The width of the gate is obtained from $C_G = C_{ox}WL$

So that the width equals:

$$W = \frac{C_G}{C_{ox}L} = \frac{10^{-12}}{\frac{3.9 \times 8.85 \times 10^{-14}}{5 \times 10^{-6}} \times 10^{-4}} = 1.45 \text{ mm}$$

The resistance of the MOSFET then equals:

$$\begin{aligned} R_{MOSFET} &= \frac{V_{DS}}{I_D} = \frac{L}{mC_{ox}W(V_{GS} - V_T)} \\ &= \frac{10^{-4}}{400 \times \frac{3.9 \times 8.85 \times 10^{-14}}{5 \times 10^{-6}} \times 0.145 \times 5} \\ &= 5 \text{ Ohm} \end{aligned}$$

The mobility was chosen to be $400 \text{ cm}^2/\text{V}\cdot\text{s}$, which is about half the mobility in bulk material doped with 10^{17} cm^{-3} acceptors.

Appendix 1: List of Symbols

Symbol	Description	MKS Units
A	Area	m^2
c	Speed of light in vacuum	m/s
C	Capacitance per unit area	F/m^2
C_{FB}	Flatband capacitance per unit area of a MOS structure	F/m^2
C_j	Junction capacitance per unit area	F/m^2
C_{ox}	Oxide capacitance per unit area	F/m^2
D_n	Electron diffusion constant	m^2/s
D_p	Hole diffusion constant	m^2/s
E	Energy	Joule
\mathcal{E}	Electric field	V/m
E_a	Acceptor energy	Joule
E_c	Conduction band energy of a semiconductor	Joule
E_d	Donor energy	Joule
E_F	Fermi energy (thermal equilibrium)	Joule
E_g	Energy bandgap of a semiconductor	Joule
E_i	Intrinsic Fermi energy	Joule
E_v	Valence band energy of a semiconductor	Joule
E_{vacuum}	Electron energy in vacuum	Joule
$f(E)$	Distribution function (probability density function)	
F_n	Quasi-Fermi energy of electrons	Joule
F_p	Quasi-Fermi energy of holes	Joule
$g_c(E)$	Density of states in the conduction band per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
$g_v(E)$	Density of states in the valence band per unit energy and per unit volume	$\text{m}^{-3}\text{J}^{-1}$
G_n	Electron generation rate	$\text{m}^{-3}\text{s}^{-1}$
G_p	Hole generation rate	$\text{m}^{-3}\text{s}^{-1}$
h	Plank's constant	Js
\hbar	Reduced Plank's ($= h /2\pi$)	Js
I	Current	A

J	Current density	A/m^2
J_n	Electron current density	A/m^2
J_p	Hole current density	A/m^2
k	Boltzmann's constant	J/K
l	Mean free path	m
L	Length	m
L_n	Electron diffusion length	m
L_p	Hole diffusion length	m
m	Mass	kg
m_0	Free electron mass	kg
m_e^*	Effective mass of electrons	kg
m_h^*	Effective mass of holes	kg
n	Electron density	m^{-3}
n_i	Intrinsic carrier density	m^{-3}
$n(E)$	Electron density per unit energy and per unit volume	m^{-3}
n_0	Electron density in thermal equilibrium	m^{-3}
n_i	Intrinsic carrier density	m^{-3}
N	Doping density	
N_a	Acceptor doping density	m^{-3}
N_a^-	Ionized acceptor density	m^{-3}
N_B	Base doping density	m^{-3}
N_c	Effective density of states in the conduction band	m^{-3}
N_C	Collector doping density	m^{-3}
N_d	Donor doping density	m^{-3}
N_d^+	Ionized donor density	m^{-3}
N_E	Emitter doping density	m^{-3}
N_v	Effective density of states in the valence band	m^{-3}
p	Hole density	m^{-3}
$p(E)$	Hole density per unit energy	m^{-3}
p_0	Hole density in thermal equilibrium	m^{-3}
p_n	Hole density in an n-type semiconductor	m^{-3}
q	electronic charge	C

Q	Charge	C
Q_d	Charge density per unit area in the depletion layer of an MOS structure	C/m^2
$Q_{d,T}$	Charge density per unit area at threshold in the depletion layer of an MOS structure	C/m^2
Q_i	Interface charge density per unit area	C/m^2
R	Resistance	Ohm
R_n	Electron recombination rate	$\text{m}^{-3}\text{s}^{-1}$
R_p	Hole recombination rate	$\text{m}^{-3}\text{s}^{-1}$
t	Thickness	m
t_{ox}	Oxide thickness	m
T	Temperature	Kelvin
U_n	Net recombination rate of electrons	$\text{m}^{-3}\text{s}^{-1}$
U_p	Net recombination rate of holes	$\text{m}^{-3}\text{s}^{-1}$
v	Velocity	m/s
v_{th}	Thermal velocity	m/s
V_a	Applied voltage	V
V_B	Base voltage	V
V_C	Collector voltage	V
V_D	Drain voltage	V
V_E	Emitter voltage	V
V_{FB}	Flatband voltage	V
V_G	Gate voltage	V
V_t	Thermal voltage	V
V_T	Threshold voltage of an MOS structure	V
w	Depletion layer width	m
w_B	Base width	m
w_C	Collector width	m
w_E	Emitter width	m
w_n	Width of an n-type region	m
w_p	Width of a p-type region	m
x	Position	m
x_d	Depletion layer width in an MOS structure	m
$x_{d,T}$	Depletion layer width in an MOS structure at threshold	m

x_j	Junction depth	m
x_n	Depletion layer width in an n-type semiconductor	m
x_p	Depletion layer width in a p-type semiconductor	m
α	Transport factor	
β	Current gain	
γ	Body effect parameter	V ^{1/2}
γ_E	Emitter efficiency	
δ_n	Excess electron density	m ⁻³
δ_p	Excess hole density	m ⁻³
$\Delta Q_{n,B}$	Excess electron charge density in the base	C/m ²
ϵ_{ox}	Dielectric constant of the oxide	F/m
ϵ_s	Dielectric constant of the semiconductor	F/m
μ_n	Electron mobility	m ² /V-s
μ_p	Hole mobility	m ² /V-s
	Charge density per unit volume	C/m ³
ρ	Resistivity	Ω m
ρ_{ox}	Charge density per unit volume in the oxide	C/m ³
σ	Conductivity	$\Omega^{-1} m^{-1}$
τ_n	Electron lifetime	s
τ_p	Hole lifetime	s
ϕ	Potential	V
ϕ_B	Barrier height	V
ϕ_F	Bulk potential	V
ϕ_i	Built-in potential of a p-n diode or Schottky diode	V
ϕ_s	Potential at the semiconductor surface	V
Φ_M	Workfunction of the metal	V
Φ_{MS}	Workfunction difference between the metal and the semiconductor	V
Φ_S	Workfunction of the semiconductor	V
χ	Electron affinity of the semiconductor	V

Chapter 7: MOS Field Effect Transistors

7.3. MOSFET analysis

$$I_D = -\frac{Q_{inv}WL}{t_r} \quad (7.3.1)$$

$$t_r = \frac{L}{v} \quad (7.3.2)$$

$$v = mE = m \frac{V_{DS}}{L} \quad (7.3.3)$$

$$I_D = -m \cdot Q_{inv} \cdot \frac{W}{L} \cdot V_{DS} \quad (7.3.4)$$

$$Q_{inv} = -C_{ox}(V_{GS} - V_T), \text{ for } V_{GS} > V_T \quad (7.3.5)$$

$$Q_{inv} = 0, \text{ for } V_{GS} \leq V_T$$

$$I_D = mC_{ox} \frac{W}{L} (V_{GS} - V_T) V_{DS}, \text{ for } |V_{DS}| \ll (V_{GS} - V_T) \quad (7.3.6)$$

$$I_D = mC_{ox} \frac{W}{dy} (V_G - V_S - V_C - V_T) dV_C \quad (7.3.7)$$

$$\int_0^L I_D dy = mC_{ox} W \int_0^{V_{DS}} (V_G - V_S - V_C - V_T) dV_C \quad (7.3.8)$$

$$I_D = mC_{ox} \frac{W}{L} [(V_{GS} - V_T) V_{DS} - \frac{V_{DS}^2}{2}], \quad \text{for } V_{DS} < V_{GS} - V_T \quad (7.3.9)$$

$$I_{D,sat} = mC_{ox} \frac{W}{L} \frac{(V_{GS} - V_T)^2}{2}, \quad \text{for } V_{DS} > V_{GS} - V_T \quad (7.3.10)$$

$$I_D = 0, \quad \text{for } V_{GS} < V_T \quad (7.3.11)$$

$$g_m = \frac{\Delta I_D}{\Delta V_{GS}} \Bigg|_{V_{DS}} \quad (7.3.12)$$

$$g_{m,quad} = mC_{ox} \frac{W}{L} V_{DS} \quad (7.3.13)$$

$$g_{m,sat} = mC_{ox} \frac{W}{L} (V_{GS} - V_T) \quad (7.3.14)$$

$$g_d = \frac{\Delta}{\Delta} \left. \frac{\mathbf{I}_D}{\mathbf{V}_{DS}} \right|_{V_{GS}} \quad (7.3.15)$$

$$g_{d,quad} = \mathbf{m} \mathbf{C}_{ox} \frac{W}{L} (V_{GS} - V_T - V_{DS}) \quad (7.3.16)$$

$$g_{d,sat} = 0 \quad (7.3.17)$$

$$I_{D,sat} = \mathbf{m} \mathbf{C}_{ox} \frac{W}{L} \frac{(V_{GS} - V_T)^2}{2} (1 + I_{V_{DS}}), \quad \text{for } V_{DS} > V_{GS} - V_T \quad (7.3.18)$$

$$Q_{inv} = -C_{ox} (V_{GS} - V_T), \quad \text{for } V_{GS} > V_T \quad (7.3.19)$$

$$V_T = V_{FB} + V_C + 2\mathbf{f}_F + \frac{\sqrt{2\mathbf{e}_s q N_a (2\mathbf{f}_F + V_{SB} + V_C)}}{C_{ox}} \quad (7.3.20)$$

$$I_D = \mathbf{m}_n \mathbf{C}_{ox} \frac{W}{dy} (V_{GS} - V_{FB} - 2\mathbf{f}_F - V_C - \frac{\sqrt{2\mathbf{e}_s q N_a (2\mathbf{f}_F + V_{SB} + V_C)}}{C_{ox}}) dV_C \quad (7.3.21)$$

$$\int_0^L I_D dy = \mathbf{m}_n \mathbf{C}_{ox} W \int_0^{V_{DS}} (V_{GS} - V_{FB} - 2\mathbf{f}_F - V_C) dV_C - \mathbf{m}_n W \int_0^{V_{DS}} \sqrt{2\mathbf{e}_s q N_a (2\mathbf{f}_F + V_{SB} + V_C)} dV_C \quad (7.3.22)$$

$$I_D = \frac{\mathbf{m}_n \mathbf{C}_{ox} W}{L} (V_{GS} - V_{FB} - 2\mathbf{f}_F - \frac{V_{DS}}{2}) V_{DS} - \frac{2}{3} \mathbf{m}_n \frac{W}{L} \sqrt{2\mathbf{e}_s q N_a} ((2\mathbf{f}_F + V_{DB})^{3/2} - (2\mathbf{f}_F + V_{SB})^{3/2}) \quad (7.3.23)$$

$$V_{DS,sat} = V_{GS} - V_{FB} - 2\mathbf{f}_F - \frac{q N_a \mathbf{e}_s}{C_{ox}^2} \left\{ \sqrt{1 + 2 \frac{C_{ox}^2}{q N_a \mathbf{e}_s} (V_{GB} - V_{FB})} - 1 \right\} \quad (7.3.24)$$

$$g_{m,sat} = \mathbf{m}_n \mathbf{C}_{ox} \frac{W}{L} [V_{GS} - V_{FB} - 2\mathbf{f}_F - \frac{q N_a \mathbf{e}_s}{C_{ox}^2} \left\{ \sqrt{1 + 2 \frac{C_{ox}^2}{q N_a \mathbf{e}_s} (V_{GB} - V_{FB})} - 1 \right\}] \quad (7.3.25)$$

$$g_{m,sat} = \mathbf{m}_n^* \mathbf{C}_{ox} \frac{W}{L} (V_{GS} - V_T) \quad (7.3.26)$$

$$\mathbf{m}_n^* = \mathbf{m}_n \left(1 - \frac{1}{\sqrt{1 + \frac{2(2\mathbf{f}_F + V_{SB})C_{ox}^2}{qN_a \mathbf{e}_s}}} \right) \quad (7.3.27)$$

7.4. Threshold voltage

$$V_T = V_{FB} + 2\mathbf{f}_F + \frac{\sqrt{2\mathbf{e}_s q N_a (2\mathbf{f}_F + V_{SB})}}{C_{ox}} \quad (7.4.1)$$

$$V_{FB} = \Phi_{MS} - \frac{Q_f}{C_{ox}} - \frac{1}{C_{ox}} \int_0^{t_{ox}} \frac{x}{x_{ox}} \mathbf{r}_{ox}(x) dx \quad (7.4.2)$$

$$\Phi_{MS} = \Phi_M - \Phi_S = \Phi_M - (\mathbf{c} + \frac{E_g}{2q} + \mathbf{f}_F) \quad (7.4.3)$$

$$\mathbf{f}_F = V_t \ln \frac{N_a}{n_i}, \text{ p-substrate} \quad (7.4.4)$$

$$V_T = V_{FB} - |2\mathbf{f}_F| - \frac{\sqrt{2\mathbf{e}_s q N_d (|2\mathbf{f}_F| - V_{SB})}}{C_{ox}} \quad (7.4.5)$$

$$V_{FB} = \Phi_{MS} - \frac{Q_f}{C_{ox}} - \frac{1}{C_{ox}} \int_0^{t_{ox}} \frac{x}{x_{ox}} \mathbf{r}_{ox}(x) dx \quad (7.4.6)$$

$$\Phi_{MS} = \Phi_M - \Phi_S = \Phi_M - (\mathbf{c} + \frac{E_g}{2q} - |\mathbf{f}_F|) \quad (7.4.7)$$

$$|\mathbf{f}_F| = V_t \ln \frac{N_d}{n_i}, \text{ n-substrate} \quad (7.4.8)$$

$$V_T = V_{FB} + 2\mathbf{f}_F + \frac{\sqrt{2\mathbf{e}_s q N_a (2\mathbf{f}_F + V_{SB})}}{C_{ox}} \quad (7.4.9)$$

$$\Delta V_T = \mathbf{g} (\sqrt{(2\mathbf{f}_F + V_{SB})} - \sqrt{2\mathbf{f}_F}) \quad (7.4.10)$$

$$\mathbf{g} = \frac{\sqrt{2\mathbf{e}_s q N_a}}{C_{ox}} \quad (7.4.11)$$

7.5. MOSFET SPICE MODEL

7.6. MOSFET Circuits and Technology

7.7. Advanced MOSFET issues

$$I_D \propto \exp\left(\frac{V_G - V_T}{nV_t}\right) \quad (7.7.1)$$

$$n = 1 + \frac{1}{2C_{ox}} \sqrt{\frac{q\mathbf{e}_s N_a}{\mathbf{f}_F}} \quad (7.7.2)$$

$$Q_d \propto \exp\left(\frac{\mathbf{f}_s}{V_t}\right) \quad (7.7.3)$$

$$V_G = V_{FB} + \mathbf{f}_s + V_{ox} = V_{FB} + \mathbf{f}_s + \frac{\sqrt{2q\mathbf{e}_s \mathbf{f}_s}}{C_{ox}} \quad (7.7.4)$$

$$\frac{dV_G}{d\mathbf{f}_s} = 1 + \frac{1}{2C_{ox}} \sqrt{\frac{2q\mathbf{e}_s}{\mathbf{f}_s}} \cong 1 + \frac{1}{2C_{ox}} \sqrt{\frac{q\mathbf{e}_s}{2\mathbf{f}_F}} = n \quad (7.7.5)$$

$$I_D \propto Q_d \propto \exp\left(\frac{\mathbf{f}_s}{V_t}\right) \propto \exp\left(\frac{V_G}{nV_t}\right) \quad (7.7.6)$$

$$\mathbf{m}_{surface} \propto E^{-1/3} \quad (7.7.7)$$

Acceptor An atom which is likely to take on one or more electrons when placed in a crystal

Bandgap The range of energies between existing energy bands where no energy levels exist

Compensation The process of adding donors and acceptor to a crystal

Conduction band Lowest empty or partially filled band in a semiconductor

Conductivity The ratio of the current density to the applied electric field

Continuity equation Equation which states that the rate of change of a density of particles equals the net flux of particles coming in.

Crystal A solid which consists of atoms placed in a periodic arrangement

Crystalline Made of one or multiple crystals

Density of states The density of electronic states per unit energy and per unit volume

Diffusion Motion of carriers caused by thermal energy

Donor An atom which is likely to give off one or more electrons when placed in a crystal

Drift Motion of carriers caused by an electric field

Energy band A collection of closely spaced energy levels

Epitaxial layer Thin layer of a single crystalline semiconductor grown on a substrate

Generation Process by which electron-hole pairs are generated

Hole Particle associated with an empty electron level in an almost filled band

Impurity A foreign atom in a crystal

Intrinsic carrier density The density of electrons and holes in an intrinsic semiconductor

Intrinsic semiconductor A semiconductors free of defects or impurities

Ionization The process of adding or removing an electron to/from an atom thereby creating a charged atom (ion)

Majority Carrier Density The larger density of the two carrier types (electrons and holes). The majority carrier density is frequently - but not always - equal to the doping density.

Mass action law The law which describes the relation between the densities of species involved in a chemical reaction

Minority Carrier Density The lower density of the two carrier types (electrons and holes). The minority carrier density is typically orders of magnitude lower than the majority carrier density, yet plays an important role in p-n diodes and bipolar transistors.

Mobility The ratio of the carrier velocity to the applied electric field

n- semiconductor n-type semiconductor with low donor density (< 10^{16} cm $^{-3}$)

n+ semiconductor n-type semiconductor with high donor density (< 10^{18} cm $^{-3}$)

p- semiconductor p-type semiconductor with low donor density (< 10^{16} cm $^{-3}$)

p+ semiconductor p-type semiconductor with high donor density (< 10^{18} cm $^{-3}$)

Poly-silicon Poly-crystalline silicon. Sometimes referred to as poly.

Recombination Process by which electron-hole pairs are removed

Resistivity The ratio of the applied voltage to the current

Saturation Velocity Maximum velocity which can be obtained in a specific semiconductor

Valence band Highest filled or almost filled band in a semiconductor

Valence electrons Electrons in the outer shell of an atom

Chapter 2: Semiconductor Fundamentals

2.2. Crystals and crystal structures

$$\vec{r} = k \vec{a}_1 + l \vec{a}_2 + m \vec{a}_3 \quad (2.2.1)$$

2.3. Energy bands

$$E_g(T) = E_g(0) - \frac{aT^2}{T + b} \quad (2.3.1)$$

$$J_{vb} = \frac{1}{V} \sum_{\substack{\text{filled} \\ \text{states}}} (-q)v_i \quad (2.3.2)$$

$$J_{vb} = \frac{1}{V} \left(\sum_{\substack{\text{all} \\ \text{states}}} (-q)v_i - \sum_{\substack{\text{empty} \\ \text{states}}} (-q)v_i \right) \quad (2.3.3)$$

$$J_{vb} = \frac{1}{V} \sum_{\substack{\text{empty} \\ \text{states}}} (+q)v_i \quad (2.3.4)$$

2.4. Density of states

$$\Psi = A \sin(k_x x) + B \cos(k_x x) \quad (2.4.1)$$

$$k_x = \frac{n\mathbf{p}}{L}, n = 1, 2, 3, \dots \quad (2.4.2)$$

$$N = 2 \times \frac{1}{8} \times \left(\frac{L}{\mathbf{p}}\right)^3 \times \frac{4}{3} \times \mathbf{p} \times k^3 \quad (2.4.3)$$

$$\frac{dN}{dE} = \frac{dN}{dk} \frac{dk}{dE} = \left(\frac{L}{\mathbf{p}}\right)^3 \mathbf{p} k^2 \frac{dk}{dE} \quad (2.4.4)$$

$$E(k) = \frac{\hbar^2 k^2}{2m^*} \quad (2.4.5)$$

$$g(E) = \frac{1}{L^3} \frac{dN}{dE} = \frac{8\mathbf{p}\sqrt{2}}{h^3} m^*{}^{3/2} \sqrt{E}, \text{ for } E \geq 0 \quad (2.4.6)$$

$$g_c(E) = \frac{1}{L^3} \frac{dN}{dE} = \frac{8\mathbf{p}\sqrt{2}}{h^3} m^*{}^{3/2} \sqrt{E - E_c}, \text{ for } E \geq E_c \quad (2.4.7)$$

$$g_c(E) = 0, \text{ for } E < E_c$$

2.5. Carrier distribution functions

$$f(E) = \frac{1}{1 + e^{(E-E_F)/kT}} \quad (2.5.1)$$

$$f_{donor}(E_d) = \frac{1}{1 + \frac{1}{2}e^{(E_d-E_F)/kT}} \quad (2.5.2)$$

$$f_{acceptor}(E_a) = \frac{1}{1 + 4e^{(E_a-E_F)/kT}} \quad (2.5.3)$$

$$f_{BE}(E) = \frac{1}{e^{(E-E_F)/kT} - 1} \quad (2.5.4)$$

$$f_{MB}(E) = \frac{1}{e^{(E-E_F)/kT}} \quad (2.5.5)$$

2.6. Carrier densities

$$n(E) = g_c(E)f(E) \quad (2.6.1)$$

$$p(E) = g_v(E)[1 - f(E)] \quad (2.6.2)$$

$$g_c(E) = \frac{8p\sqrt{2}}{h^3} m_e^{*3/2} \sqrt{E - E_c}, \text{ for } E \geq E_c \quad (2.6.3)$$

$$g_v(E) = \frac{8p\sqrt{2}}{h^3} m_e^{*3/2} \sqrt{E_v - E}, \text{ for } E \leq E_v \quad (2.6.4)$$

$$n = \frac{\text{top of the conduction band}}{\int_{E_c}^{\infty} n(E)dE} = \frac{\text{top of the conduction band}}{\int_{E_c}^{\infty} g_c(E)f(E)dE} \quad (2.6.5)$$

$$n_o = \int_{E_c}^{\infty} g_c(E)f(E)dE \quad (2.6.6)$$

$$n_o = \int_{E_c}^{\infty} \frac{8p\sqrt{2}}{h^3} m_e^{*3/2} \sqrt{E - E_c} \frac{1}{1 + e^{\frac{E-E_F}{kT}}} dE \quad (2.6.7)$$

$$p_o = \int_{-\infty}^{E_v} g_v(E)[1 - f(E)]dE \quad (2.6.8)$$

$$p_o = \int_{-\infty}^{E_v} \frac{8p\sqrt{2}}{h^3} m_h^{*3/2} \sqrt{E_v - E} \frac{1}{1 + e^{\frac{E_F-E}{kT}}} dE \quad (2.6.9)$$

$$n_o = \int_{E_c}^{E_F} g_c(E) dE \text{ at } T = 0 \text{ K} \quad (2.6.10)$$

$$n_o = \frac{2}{3} \frac{\sqrt{2}}{p^2} \left(\frac{qm^*}{\hbar^2} \right)^{3/2} (E_F - E_c)^{3/2}, \text{ for } E_F \geq E_c \quad (2.6.11)$$

$$n_o \equiv \int_{E_c}^{\infty} \frac{8p\sqrt{2}}{h^3} m_e^{*3/2} \sqrt{E - E_c} e^{\frac{E_F - E}{kT}} dE = N_c e^{\frac{E_F - E_c}{kT}} \quad (2.6.12)$$

$$N_c = 2 \left[\frac{2p m_e^* kT}{h^2} \right]^{3/2} \quad (2.6.13)$$

$$p_o \equiv \int_{-\infty}^{E_v} \frac{8p\sqrt{2}}{h^3} m_h^{*3/2} \sqrt{E_v - E} e^{\frac{E - E_F}{kT}} dE = N_v e^{\frac{E_v - E_F}{kT}} \quad (2.6.14)$$

$$N_v = 2 \left[\frac{2p m_h^* kT}{h^2} \right]^{3/2} \quad (2.6.15)$$

$$\frac{E_F - E_c}{kT} \equiv \ln \frac{n_o}{N_c} + \frac{1}{\sqrt{8}} \frac{n_o}{N_c} - \left(\frac{3}{16} - \frac{\sqrt{3}}{9} \right) \left(\frac{n_o}{N_c} \right)^2 + \dots \quad (2.6.16)$$

$$\frac{E_v - E_F}{kT} \equiv \ln \frac{p_o}{N_v} + \frac{1}{\sqrt{8}} \frac{p_o}{N_v} - \left(\frac{3}{16} - \frac{\sqrt{3}}{9} \right) \left(\frac{p_o}{N_v} \right)^2 + \dots \quad (2.6.17)$$

$$n_i = n_o \Big|_{(E_F = E_i)} = N_c e^{(E_i - E_c)/kT} \quad (2.6.18)$$

$$n_i = p_o \Big|_{(E_F = E_i)} = N_v e^{(E_v - E_i)/kT}$$

$$n_i = \sqrt{N_c N_v} e^{-E_g/2kT} \quad (2.6.19)$$

$$n_o \cdot p_o = N_c N_v e^{(E_v - E_c)/kT} = n_i^2 \quad (2.6.20)$$

$$E_i = \frac{E_c + E_v}{2} + \frac{1}{2} kT \ln \left(\frac{N_v}{N_c} \right) \quad (2.6.21)$$

$$E_i = \frac{E_c + E_v}{2} + \frac{3}{4} kT \ln \left(\frac{m_h^*}{m_e^*} \right) \quad (2.6.22)$$

$$n_o = n_i e^{(E_F - E_i)/kT} \quad (2.6.23)$$

$$p_o = n_i e^{(E_i - E_F)/kT} \quad (2.6.24)$$

$$E_F = E_i + kT \ln \frac{n_o}{n_i} \quad (2.6.25)$$

$$E_F = E_i - kT \ln \frac{p_o}{n_i} \quad (2.6.26)$$

$$N_d^+ \cong N_d \quad (2.6.27)$$

$$N_a^- \cong N_a \quad (2.6.28)$$

$$n_o \cong N_d^+ - N_a^-, \quad \text{if } N_d^+ - N_a^- \gg n_i \quad (2.6.29)$$

$$p_o \cong N_a^- - N_d^+, \quad \text{if } N_a^- - N_d^+ \gg n_i \quad (2.6.30)$$

$$E_c - E_d = 13.6 \frac{m_{cond}^*}{m_0 \epsilon_r^2} \text{ eV} \quad (2.6.31)$$

$$\mathbf{r} = q(p_o - n_o + N_d^+ - N_a^-) = 0 \quad (2.6.32)$$

$$n_o = \frac{n_i^2}{n} + N_d^+ - N_a^- \quad (2.6.33)$$

$$n_o = \frac{N_d^+ - N_a^-}{2} + \sqrt{\left(\frac{N_d^+ - N_a^-}{2}\right)^2 + n_i^2} \quad (2.6.34)$$

$$p_o = \frac{N_a^- - N_d^+}{2} + \sqrt{\left(\frac{N_a^- - N_d^+}{2}\right)^2 + n_i^2} \quad (2.6.35)$$

$$p_o + N_d^+ = n_o + N_a^- \quad (2.6.36)$$

$$n = n_o + \mathbf{d} \quad n = n_i \exp\left(\frac{F_n - E_i}{kT}\right) \quad (2.6.37)$$

$$p = p_o + \mathbf{d} \quad p = n_i \exp\left(\frac{E_i - F_p}{kT}\right) \quad (2.6.38)$$

2.7. Carrier Transport

$$I = \frac{Q}{t_r} = \frac{Q}{L/v} \quad (2.7.1)$$

$$\vec{J} = \frac{Q}{AL} \vec{v} = \mathbf{r} \vec{v} = qn \vec{v} \quad (2.7.2)$$

$$\vec{F} = m \vec{a} = m \frac{d < \vec{v} >}{dt} \quad (2.7.3)$$

$$\vec{F} = q \vec{E} - \frac{m < \vec{v} >}{t} \quad (2.7.4)$$

$$q \vec{E} = m \frac{d < \vec{v} >}{dt} + \frac{m < \vec{v} >}{t} \quad (2.7.5)$$

$$\mathbf{m} = \frac{\Delta |\vec{v}|}{|\vec{E}|} = \frac{q \mathbf{t}}{m} \quad (2.7.6)$$

$$\vec{J} = qn \mathbf{m}_n \vec{E} \quad (2.7.7)$$

$$\mathbf{m} = \frac{q \mathbf{t}}{m^*} \quad (2.7.8)$$

$$\mathbf{m} = \mathbf{m}_{\min} + \frac{\mathbf{m}_{\max} - \mathbf{m}_{\min}}{1 + \left(\frac{N}{N_r} \right)^a} \quad (2.7.9)$$

$$J = qnv_e + qp v_h = q(n \mathbf{m}_n + p \mathbf{m}_p) E \quad (2.7.10)$$

$$\frac{\Delta J}{E} = q(n \mathbf{m}_n + p \mathbf{m}_p) \quad (2.7.11)$$

$$\mathbf{r} = \frac{1}{s} = \frac{1}{q(\mathbf{m}_n n + \mathbf{m}_p p)} \quad (2.7.12)$$

$$R_s = \frac{\mathbf{r}}{t} \quad (2.7.13)$$

$$R = R_s \frac{L}{W} \quad (2.7.14)$$

$$v(E) = \frac{\mathbf{m} E}{1 + \frac{\mathbf{m} E}{v_{sat}}} \quad (2.7.15)$$

$$v_{th} = \frac{l}{t} \quad (2.7.16)$$

$$\Phi_{n, left \rightarrow right} = \frac{1}{2} v_{th} n(x = -l) \quad (2.7.17)$$

$$\Phi_{n,right \rightarrow left} = \frac{1}{2}v_{th}n(x=l) \quad (2.7.18)$$

$$\Phi_n = \Phi_{n,left \rightarrow right} - \Phi_{n,right \rightarrow left} = \frac{1}{2}v_{th}[n(x=-l) - n(x=l)] \quad (2.7.19)$$

$$\Phi_n = -lv_{th} \frac{n(x=l) - n(x=-l)}{2l} = -lv_{th} \frac{dn}{dx} \quad (2.7.20)$$

$$J_n = -q\Phi_n = qlv_{th} \frac{dn}{dx} \quad (2.7.21)$$

$$J_n = qD_n \frac{dn}{dx} \quad (2.7.22)$$

$$J_p = -qD_p \frac{dp}{dx} \quad (2.7.23)$$

$$v_{th} = \frac{l}{t} \quad (2.7.24)$$

$$\frac{kT}{2} = \frac{m^* v_{th}^2}{2} \quad (2.7.25)$$

$$lv_{th} = \frac{m^* v_{th}^2}{q} \frac{q}{m^*} = \frac{kT}{q} m \quad (2.7.26)$$

$$D_n = m_n \frac{kT}{q} = m_n V_t \quad (2.7.27)$$

$$D_p = m_p \frac{kT}{q} = m_p V_t \quad (2.7.28)$$

$$J_n = qn m_n E + qD_n \frac{dn}{dx} \quad (2.7.29)$$

$$J_p = qp m_p E - qD_p \frac{dp}{dx} \quad (2.7.30)$$

$$I_{total} = A(J_n + J_p) \quad (2.7.31)$$

2.8. Carrier recombination and generation

$$U_n = R_n - G_n = \frac{n_p - n_{p0}}{t_n} \quad (2.8.1)$$

$$U_p = R_p - G_p = \frac{p_n - p_{n0}}{t_p} \quad (2.8.2)$$

$$U_{b-b} = b(np - n_i^2) \quad (2.8.3)$$

$$U_{SHR} = \frac{pn - n_i^2}{p + n + 2n_i \cosh(\frac{E_i - E_t}{kT})} N_t v_{th} \mathbf{s} \quad (2.8.4)$$

$$U_n = R_n - G_n = \frac{n_p - n_{p0}}{\mathbf{t}_n} \quad (2.8.5)$$

$$U_p = R_p - G_p = \frac{p_n - p_{n0}}{\mathbf{t}_p} \quad (2.8.6)$$

$$\mathbf{t}_n = \mathbf{t}_p = \frac{1}{N_t v_{th} \mathbf{s}} \quad (2.8.7)$$

$$U_{s,SHR} = \frac{pn - n_i^2}{p + n + 2n_i \cosh(\frac{E_i - E_{st}}{kT})} N_{st} v_{th} \mathbf{s}_s \quad (2.8.8)$$

$$U_{s,n} = R_{s,n} - G_{s,n} = v_s (n_p - n_{p0}) \quad (2.8.9)$$

$$v_s = N_{st} v_{th} \mathbf{s}_s \quad (2.8.10)$$

$$U_{Auger} = \Gamma_n n(np - n_i^2) + \Gamma_p p(np - n_i^2) \quad (2.8.11)$$

$$G_{p,light} = G_{n,light} = \mathbf{a} \frac{P_{opt}(x)}{E_{ph} A} \quad (2.8.12)$$

$$\frac{dP_{opt}(x)}{dx} = -\mathbf{a} P_{opt}(x) \quad (2.8.13)$$

2.9. Continuity equation

$$\frac{\frac{d}{dt} n(x,t)}{q} A dx = \left(\frac{J_n(x)}{-q} - \frac{J_n(x+dx)}{-q} \right) A + (G_n(x,t) - R_n(x,t)) A dx \quad (2.9.1)$$

$$J_n(x+dx) = J_n(x) + \frac{d J_n(x)}{dx} dx \quad (2.9.2)$$

$$\frac{\frac{d}{dt} p(x,t)}{q} = \frac{1}{q} \frac{\partial J_n(x,t)}{\partial x} + G_p(x,t) - R_p(x,t) \quad (2.9.3)$$

$$\frac{\frac{d}{dt} p(x,t)}{q} = -\frac{1}{q} \frac{\partial J_p(x,t)}{\partial x} + G_p(x,t) - R_p(x,t) \quad (2.9.4)$$

$$\frac{\frac{d}{dt}n(x,t)}{t} = \quad (2.9.5)$$

$$\mathbf{m}_n n \frac{\partial E(x,t)}{\partial x} + \mathbf{m}_n E \frac{\partial n(x,t)}{\partial x} + D_n \frac{\partial^2 n(x,t)}{\partial x^2} + G_n(x,t) - R_n(x,t)$$

$$\frac{\frac{d}{dt}p(x,t)}{t} = \quad (2.9.6)$$

$$- \mathbf{m}_p p \frac{\partial E(x,t)}{\partial x} - \mathbf{m}_p E \frac{\partial p(x,t)}{\partial x} + D_p \frac{\partial^2 p(x,t)}{\partial x^2} + G_p(x,t) - R_p(x,t)$$

$$\frac{\frac{d}{dt}n(x,y,z,t)}{t} = \frac{1}{q} \vec{\nabla} \vec{J}_n(x,y,z,t) + G_n(x,y,z,t) - R_n(x,y,z,t) \quad (2.9.7)$$

$$\frac{\frac{d}{dt}p(x,y,z,t)}{t} = - \frac{1}{q} \vec{\nabla} \vec{J}_p(x,y,z,t) + G_p(x,y,z,t) - R_p(x,y,z,t) \quad (2.9.8)$$

$$\frac{\frac{d}{dt}n(x,t)}{t} = D_n \frac{\frac{d^2 n_p(x,t)}{dx^2} - \frac{n_p(x,t) - n_{p0}}{t_n}}{t_n} \quad (2.9.9)$$

$$\frac{\frac{d}{dt}p(x,t)}{t} = D_p \frac{\frac{d^2 p_n(x,t)}{dx^2} - \frac{p_n(x,t) - p_{n0}}{t_p}}{t_p} \quad (2.9.10)$$

$$0 = D_n \frac{d^2 n_p(x)}{dx^2} - \frac{n_p(x) - n_{p0}}{t_n} \quad (2.9.11)$$

$$0 = D_p \frac{d^2 p_n(x)}{dx^2} - \frac{p_n(x) - p_{n0}}{t_p} \quad (2.9.12)$$

$$p_n(x \geq x_n) = p_{n0} + A e^{-(x-x_n)/L_p} + B e^{(x-x_n)/L_p} \quad (2.9.13)$$

$$n_p(x \leq -x_p) = n_{p0} + C e^{-(x+x_p)/L_p} + D e^{(x+x_p)/L_p} \quad (2.9.14)$$

$$L_n = \sqrt{D_n t_n} \quad (2.9.15)$$

$$L_p = \sqrt{D_p t_p} \quad (2.9.16)$$

$$n = n_0 + \mathbf{d}_n \quad (2.9.17)$$

$$p = p_0 + \mathbf{d}_p \quad (2.9.18)$$

$$0 = \frac{d^2(\mathbf{d}n_p)}{dx^2} - \frac{\mathbf{d}n_p}{L_n^2} \quad (2.9.19)$$

$$0 = \frac{d^2(\mathbf{d}p_n)}{dx^2} - \frac{\mathbf{d}p_n}{L_p^2} \quad (2.9.20)$$

2.10. The drift-diffusion model

$$\mathbf{r} = q(p - n + N_d^+ - N_a^-) \quad (2.10.1)$$

$$\frac{dE}{dx} = \frac{\mathbf{r}}{\mathbf{e}} \quad (2.10.2)$$

$$\frac{d\mathbf{f}}{dx} = -E \quad (2.10.3)$$

$$\frac{dE_i}{dx} = qE \quad (2.10.4)$$

$$n = n_i e^{(F_n - E_i)/kT} \quad (2.10.5)$$

$$p = n_i e^{(E_i - F_p)/kT} \quad (2.10.6)$$

$$J_n = qn\mathbf{m}_n E + qD_n \frac{dn}{dx} \quad (2.10.7)$$

$$J_p = qp\mathbf{m}_p E - qD_p \frac{dp}{dx} \quad (2.10.8)$$

$$0 = \frac{1}{q} \frac{\partial J_n}{\partial x} - \frac{np - n_i^2}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} \frac{1}{t} \quad (2.10.9)$$

$$0 = -\frac{1}{q} \frac{\partial J_p}{\partial x} - \frac{np - n_i^2}{n + p + 2n_i \cosh(\frac{E_t - E_i}{kT})} \frac{1}{t} \quad (2.10.10)$$

Principles of Semiconductor Devices



- [Table of Contents](#)
- [Slideshow](#)
- [Movie](#)

Principles of Semiconductor Devices



Principles of Semiconductor Devices

by
Bart Van Zeghbroeck

- [Table of Contents](#)
- [Slideshow](#)
- [Movie](#)

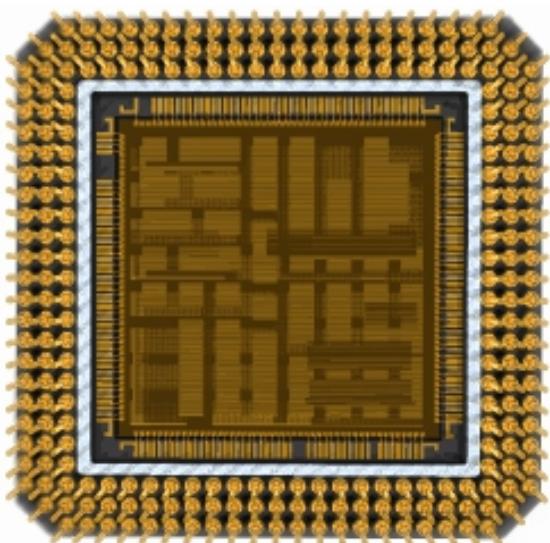
Principles of Semiconductor Devices



The Physics and Devices behind Modern-day Electronics

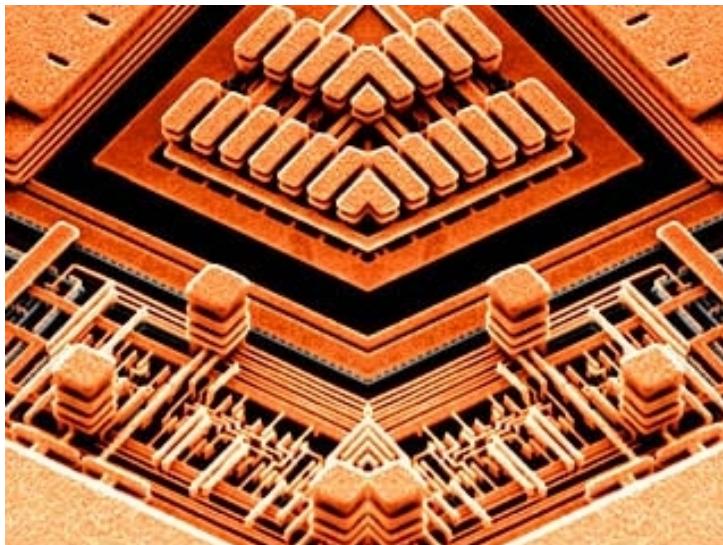
- [Table of Contents](#)
- [Slideshow](#)
- [Movie](#)

Principles of Semiconductor Devices



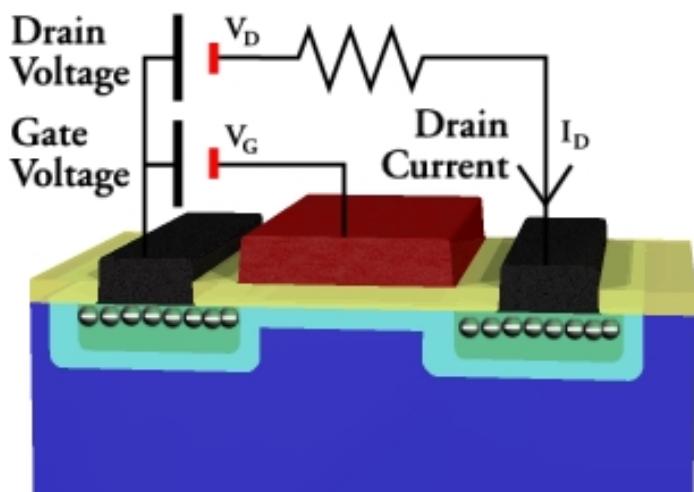
- [Table of Contents](#)
- [Slideshow](#)
- [Movie](#)

Principles of Semiconductor Devices



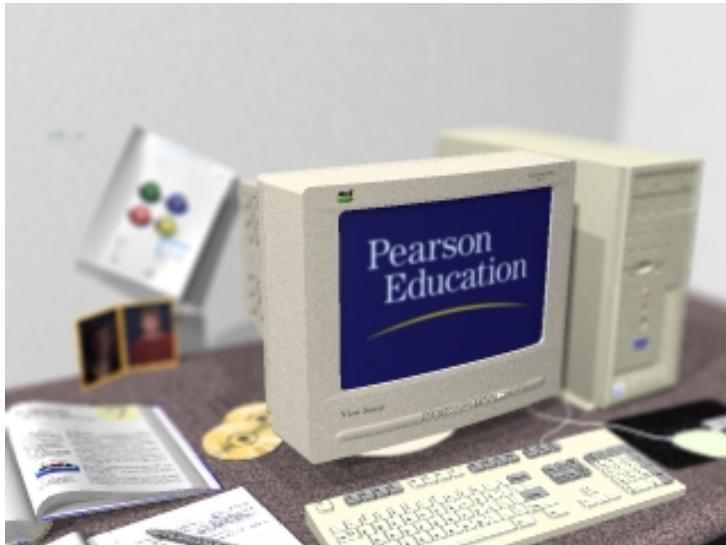
- [Table of Contents](#)
- [Slideshow](#)
- [Movie](#)

Principles of Semiconductor Devices



- [Table of Contents](#)
- [Slideshow](#)
- [Movie](#)

Principles of Semiconductor Devices



- Table of Contents
- Slideshow
- Movie

Principles of Semiconductor Devices



- [Table of Contents](#)
- [Slideshow](#)
- [Movie](#)

Principles of Semiconductor Devices

