Block 11 — Semiconductor Fundamentals and Diodes

Student Group

First Name	Surname	Matrikel Nr.

Table of Contents

Block 11 — Semiconductor Fundamentals and Diodes

Learning objectives

After this 90-minute block, you can

- distinguish conductors, semiconductors, and insulators using the band model.
- explain intrinsic conduction, electron conduction, and hole conduction.
- explain how n-doping and p-doping change the number of mobile charge carriers.
- describe the formation of a pn junction and the depletion region.
- decide whether a diode is forward-biased or reverse-biased from $u_{\rm AK}$.
- compare the ideal diode model, the constant-voltage model, and the piecewise-linear diode model.
- use the diode equation

$$\begin{align*} i_{\rm D}=I_{\rm S}(T)\left({\rm e}^{\frac{u_{\rm AK}}{mU_{\rm T}}}-1\right) \end{align*}$$ at a qualitative level.

90-minute plan

- **Warm-up (10 min):**
  - Why does a diode conduct in one direction but not in the other?
  - Recall from EEE1: voltage, current direction, power, and resistors.
  - Recall from EEE2: transient overvoltages at inductive loads will later need diode protection.
- **Core concepts (55 min):**
  - Conductors, semiconductors, insulators, and the band gap.
  - Intrinsic conduction, electron conduction, hole conduction.
  - Doping: n-type and p-type material.
  - pn junction, depletion region, diffusion voltage.
  - Diode operation in forward and reverse direction.
  - Ideal and real diode characteristics.
  - Practical diode models for circuit calculations.
- **Practice (20 min):**
  - Determine diode polarity and conduction state.
  - Calculate current with a constant-voltage diode model.
  - Estimate differential resistance at a given operating point.
  - Compare ideal and real diode assumptions.
- **Wrap-up (5 min):**
  - Key messages: pn junction, forward/reverse bias, current limiting, diode models.
  - Preview: rectifiers, smoothing, protection circuits, LEDs, and Z-diode stabilizers in Block 12.

Conceptual overview

- A semiconductor is neither a good conductor nor a perfect insulator. Its conductivity can be controlled by material, temperature, light, and doping.
- A diode is a pn junction with two terminals:
  - **anode A** on the p-side,
  - **cathode K** on the n-side.
- In forward direction, the external voltage reduces the depletion region and current can flow.
- In reverse direction, the depletion region becomes wider and only a very small leakage current flows, until breakdown occurs.
- A diode is nonlinear. It is not a resistor.
- In circuits, diode current must usually be limited by another component, often a resistor.

Scope of this block

This block explains **why** diodes behave as they do and how we model them.

Diode applications such as

- rectifiers,
- smoothing capacitors,
- freewheeling diodes,
- input protection circuits,
- LED circuits,
- Z-diode voltage stabilizers

are continued in Block 12.

Core content

Introductory Example

Microcontrollers often have many pins that evaluate signals between $0...5~\rm V$ as a digital signal. However, the input signal can be disturbed during transmission by small coupled pulses, e.g. from HF sources like mobile phones. This interference can cause the signal to leave the permitted voltage range of approx. $-0.5...5.5~\rm V$ and thus destroy the logical unit.

To prevent such destruction, an over-voltage protection circuit consisting of diodes is installed (see e.g. ATmega 328). In case of an over-/under-voltage one of the diodes becomes conductive and lowers the input voltage by the resulting current. In the simulation, it can be seen that the interference on the input side can be reduced to an acceptable, low level by the protection circuit.

This chapter explains why a diode becomes conductive at a certain voltage, what has to be considered when using diodes, and which different types of diodes are available.

For the protection of digital interfaces that leave the device housing (e.g. USB), additional separate ICs are used that support this protection of the data processing chips. These protection diode ICs suppress the short-time voltages and are called Transient Voltage Suppressor or TVS diodes. Typical TVS ICs are NUP2301 or for USB NUP4201.

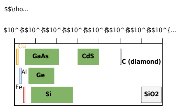
Further reading

- An introductory is available at electronics-tutorials

A short quantum view: why energy bands matter

Materials differ strongly in their specific resistance \(\rho\).

Fig. 1: specific resistance for selected conductors, semiconductors, and insulators.



Why we need a band model

The simple circuit view says: conductors conduct, insulators block, semiconductors are somewhere in between. To understand **why** semiconductors can be controlled so well, we need a short look at the energy of electrons in a solid.

In a single atom, electrons can only have certain discrete energies (figure 2 1a and 1b). This is one result of quantum physics. A simple picture is the Bohr model: electrons are not allowed to move on arbitrary paths, but only on certain allowed energy levels.

In a solid, many atoms are very close together (figure 2 2a and 2b). Their individual energy levels interact and broaden into **energy bands**.

Fig. 2: From discrete energy levels in an atom to energy bands in a solid.

Bohr atomic model

The two most important bands are:

- the **valence band**: the highest occupied band. Electrons here are still bound in the crystal.
- the **conduction band**: the next higher band. Electrons here can move through the crystal and contribute to current.

The energetic distance between them is the **band gap** $E_{\rm g}$.

$$\begin{align*} E_{\rm g} = W_{\rm conduction~band} - W_{\rm valence~band} \end{align*}$$

For semiconductors, the band gap is small enough that some electrons can be lifted from the valence band into the conduction band.

Material type	Band model	Electrical behavior
conductor	conduction band available or overlapping	many mobile charge carriers
semiconductor	small band gap, typically a few ${\rm eV}$	conductivity can be controlled
insulator	large band gap	almost no mobile charge carriers

> An electron can become mobile if it receives enough energy to cross the band gap. This energy can come, for example, from
>
> - light, i.e. photons,
> - lattice vibrations, i.e. thermal energy or phonons.

When an electron reaches the conduction band, it leaves behind a missing electron in the valence band. This missing electron behaves like a positive mobile charge carrier and is called a **hole**.

$$\begin{align*} \text{energy input} \quad\Rightarrow\quad \text{electron-hole pair} \end{align*}$$

The opposite process is called **recombination**:

$$\begin{align*} \text{electron}+\text{hole} \quad\Rightarrow\quad \text{released energy}. \end{align*}$$

Analogy: soccer stadium with two tribunes

Imagine a soccer stadium with two tribunes.

- The **lower tribune** is close to the field and very popular. It is normally fully booked. This is the **valence band**.
- The **upper tribune** is farther away and less attractive. To get there, a person needs an extra “energy ticket”. This is the **conduction band**.
- The required energy ticket is the **band gap** $E_{\rm g}$.

If one person receives enough energy, they move from the full lower tribune to the upper tribune. Now the person upstairs can move around more freely, like an electron in the conduction band.

At the same time, an empty seat remains in the lower tribune. When neighboring people move into that empty seat, the empty seat itself seems to move. This moving empty seat is the analogy for a **hole**.

Intrinsic conduction, electrons, and holes

In a pure semiconductor, some electrons can gain enough energy to leave their bonds. Then

1. The electron becomes mobile in the conduction band.
2. A 'missing electron' (technically: a **defect electron** or a **hole**) remains in the valence band.
3. This hole behaves like a positive mobile charge carrier.

> There are two types of mobile charge carriers in semiconductors:
>
> - **electrons** with negative charge,
> - **holes** with positive effective charge.

At room temperature, only a very small fraction of thermal vibrations has enough energy to generate such electron-hole pairs in pure silicon. Nevertheless, this already creates measurable **intrinsic conduction**.

Doping: n-type and p-type semiconductors

Doping increases the number of mobile charge carriers much more effectively (see figure 3):

- n-doping adds donor atoms and therefore additional mobile electrons,
- p-doping adds acceptor atoms and therefore additional mobile holes.

Doping only works predictably when the semiconductor crystal is very pure.
The desired dopant atoms should dominate over unwanted impurities.

Doping means adding a very small amount of foreign atoms to the semiconductor crystal.

Fig. 3: Doping: donor and acceptor atoms change the number of mobile charge carriers.

Doping type	Typical dopant atoms	Main mobile charge carriers	Name of dopant
n-type	phosphorus, arsenic, antimony	electrons	donors
p-type	boron, aluminium, indium	holes	acceptors

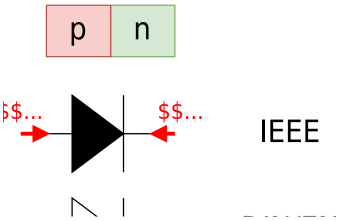
Tab. 1: Doping of silicon

> Doping does **not** mean that the semiconductor becomes strongly charged as a whole.
> The crystal is still approximately electrically neutral.
> Doping mainly changes how many mobile charge carriers are available.

The pn junction

A diode is formed when p-doped and n-doped regions meet.

Fig. 4: Diode symbol and pn junction with anode A and cathode K.



At the junction:

- electrons diffuse from the n-side into the p-side,
- holes diffuse from the p-side into the n-side,
- electrons and holes recombine,
- a region with almost no mobile charge carriers forms.

> The region on the junction has virtually no mobile charge carriers.
> This is called the **depletion region** or **space-charge region**.

Fig. 5: Formation of the depletion region at a pn junction.

The depletion region behaves like an internal barrier.
Without an external voltage, it prevents a large current.

Mnemonic: PANIC!

\[\begin{align*} \text{Positive Anode, Negative Is Cathode} \end{align*} \]

This helps to remember the forward direction of a diode.

Forward and reverse operation

We define the diode voltage

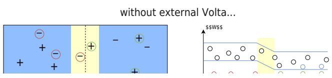
\[\begin{align*} u_{\rm AK}=u_{\rm A}-u_{\rm K}. \end{align*} \]

- \(u_{\rm AK}>0\): anode is more positive than cathode.
- \(u_{\rm AK}<0\): anode is more negative than cathode.

Condition	Name	Effect on depletion region	Current
$u_{\rm AK}>0$	forward bias forward voltage is $U_{\rm F} = u_{\rm AK}$	depletion region becomes smaller	large current possible
$u_{\rm AK}<0$	reverse bias reverse voltage is $U_{\rm R} = -u_{\rm AK}$	depletion region becomes larger	only small leakage current, until breakdown

Tab. 2: Diode operation depending on $u_{\rm AK}$

Fig. 6: PN-Junction for a Forward voltage and Blocking voltage.



Analogy: two tribunes and an empty border zone

Imagine two neighboring tribunes in a stadium (e.g. fan section and main tribune).

- On the **n-side**, there are many extra people. They represent mobile **electrons**.
- On the **p-side**, there are many empty seats. They represent mobile **holes**.

At first, people near the border can move into empty seats on the other side.
After this happens, there are fewer mobile people and fewer mobile empty seats close to the border.
A locally empty border zone appears. This represents the **depletion region**.

The depletion region is therefore not an extra part inserted between the two sides.
It forms automatically because electrons and holes recombine near the pn junction.

In **forward bias**, the external voltage pushes people and empty seats toward the border.
The empty border zone becomes narrower, and new people and empty seats are continuously supplied from the outside. A current can flow.

In **reverse bias**, the external voltage pulls people and empty seats away from the border.
The empty border zone becomes wider, so crossing becomes very unlikely. Only a tiny leakage current remains.

Ideal diode model

The simplest model is the ideal diode.

$$\begin{align*} \text{forward direction: } u_{\rm AK}=0,\quad i_{\rm D}>0 \end{align*}$$

\[\begin{align*} \text{reverse direction: } i_{\rm D}=0,\quad u_{\rm AK}<0 \end{align*} \]

Engineering meaning

The ideal diode is useful for a first decision:

- Is the diode conducting?
- Is the diode blocking?
- Which path can current take?

It is too simple for accurate voltage and current calculations.

Real diode characteristic

A real diode has an exponential current-voltage characteristic.

\[\begin{align*} \boxed{ i_{\rm D} = {\color{red}{I_{\rm S}(T)}} \left({\rm e}^{\frac{{\color{blue}{u_{\rm AK}}}}{{\color{green}{mU_{\rm T}}}}} -1 \right) } \end{align*} \]

with

\[\begin{align*} U_{\rm T}=\frac{kT}{e}. \end{align*} \]

Symbol	Meaning
\(I_{\rm S}(T)\)	reverse saturation current, strongly temperature-dependent
\(m\)	emission coefficient, typically \(1\ldots 2\), material constant
\(U_{\rm T}\)	thermal voltage ($U_{\rm T}\approx 26~{\rm mV}$ at room temperature)
\(k\)	Boltzmann constant
\(e\)	elementary charge
\(T\)	absolute temperature in \({\rm K}\)

Tab. 3: Symbols in the diode equation

Often a **turn-on voltage** $U_{\rm TO}$ for typical currents (some $\rm mA$) at \(25^\circ{\rm C}\) are used.

Diode material	Approximate threshold voltage \(U_{\rm TO}\)	Reverse saturation current \(I_{\rm S}\)
silicon	\(\approx 0.7~{\rm V}\)	some \({\rm pA}\)
germanium	\(\approx 0.3~{\rm V}\)	some \(\mu{\rm A}\)

Tab. 4: Typical diode values

- the turn-on voltage has also some alternative labeling: knee voltage, threshold voltage, diode voltage $U_{\rm D}$, forward voltage $U_{\rm F}$
- The value \(U_{\rm TO} = 0.7~{\rm V}\) for a silicon diode is not a physical constant.

- It is a useful approximation for typical currents in small signal and basic power circuits.

Practical diode models for circuit calculation

For hand calculations we usually do not use the full exponential equation, because it is often too complex for a quick solution.
Instead the following is often used:

Model	**Forward direction**	**Reverse direction**	**Use**	**Example**
ideal diode	$u_{\rm AK}=0$	$i_{\rm D}=0$	switching logic, first estimate	Is the rectifier path conducting?
constant-voltage model	$u_{\rm AK}\approx U_{\rm TO}$	$i_{\rm D}\approx 0$	quick current calculations	Which current flows through an LED and its series resistor?
piecewise-linear model	$u_{\rm AK}\approx U_{\rm TO}+r_{\rm F}\cdot i_{\rm D}$	$i_{\rm D}\approx 0$	more accurate operating point	How does the diode voltage change when the current changes?

Tab. 5: Diode models for circuit calculations

The differential forward resistance is

$$\begin{align*} r_{\rm F} = \frac{\Delta U_{\rm F}}{\Delta I_{\rm F}}. \end{align*}$$

For large forward voltages compared with $U_{\rm T}$, the diode equation leads approximately to

$$\begin{align*} r_{\rm D} = \frac{{\rm d}u_{\rm D}}{{\rm d}i_{\rm D}} \approx \frac{mU_{\rm T}}{I_{\rm D}}. \end{align*}$$

Unit check

$$\begin{align*} [r_{\rm D}] = \frac{[U_{\rm T}]}{[I_{\rm D}]} = \frac{{\rm V}}{{\rm A}} = \Omega. \end{align*}$$

Exercises

Exercise E1.1 Quick check: doping and charge carriers

Complete the table.

Doping type	Typical dopant atom	Main mobile charge carrier	Dopant name
n-type	?	?	?
p-type	?	?	?

Result

Doping type	Typical dopant atom	Main mobile charge carrier	Dopant name
n-type	phosphorus, arsenic, or antimony	electrons	donor
p-type	boron, aluminium, or indium	holes	acceptor

N-type material has additional mobile electrons. P-type material has additional mobile holes.

The semiconductor as a whole remains approximately electrically neutral.

Exercise E2.1 Quick check: diode polarity

A diode has the anode voltage

$$\begin{align*} U_{\rm A}=4.8~{\rm V} \end{align*}$$

and the cathode voltage

$$\begin{align*} U_{\rm K}=4.1~{\rm V}. \end{align*}$$

- Calculate $u_{\rm AK}$.
- Is the diode forward-biased or reverse-biased?
- For a silicon diode, is a noticeable current likely?

Result

$$\begin{align*} u_{\rm AK} = U_{\rm A}-U_{\rm K} = 4.8~{\rm V}-4.1~{\rm V} = 0.7~{\rm V}. \end{align*}$$

Since

$$\begin{align*} u_{\rm AK}>0, \end{align*}$$

the diode is forward-biased.

For a silicon diode, $0.7~{\rm V}$ is a typical forward voltage in the mA range. Therefore a noticeable current is likely.

Exercise E3.1 Quick check: current with the constant-voltage model

A silicon diode is connected in series with a resistor.

$$\begin{align*} U_{\rm I}=5.0~{\rm V}, \qquad R=1.0~{\rm k}\Omega. \end{align*}$$

Use the constant-voltage model

$$\begin{align*} U_{\rm D}\approx 0.7~{\rm V}. \end{align*}$$

Calculate the diode current $I_{\rm D}$.

Result

The voltage across the resistor is

$$\begin{align*} U_R = U_{\rm I}-U_{\rm D} = 5.0~{\rm V}-0.7~{\rm V} = 4.3~{\rm V}. \end{align*}$$

Therefore

$$\begin{align*} I_{\rm D} = \frac{U_R}{R} = \frac{4.3~{\rm V}}{1.0~{\rm k}\Omega} = 4.3~{\rm mA}. \end{align*}$$

Exercise E4.1 Quick check: differential diode resistance

A diode operates at

$$\begin{align*} I_{\rm D}=10~{\rm mA}. \end{align*}$$

Assume

$$\begin{align*} m=1, \qquad U_{\rm T}=26~{\rm mV}. \end{align*}$$

Estimate the differential diode resistance

$$\begin{align*} r_{\rm D}\approx \frac{mU_{\rm T}}{I_{\rm D}}. \end{align*}$$

Result

> $$\begin{align*} r_{\rm D} &\approx \frac{mU_{\rm T}}{I_{\rm D}} \\ &= \frac{1\cdot 26~{\rm mV}}{10~{\rm mA}} \\ &= 2.6~\Omega. \end{align*}$$
>
> This is a small-signal resistance around the operating point. It is not the same as the large-signal ratio $\frac{U_{\rm D}}{I_{\rm D}}$.

Exercise E5.1 Longer exercise: operating point with a piecewise-linear diode

A diode is connected in series with a resistor.

$$\begin{align*} U_{\rm I}=12~{\rm V}, \qquad R=560~\Omega. \end{align*}$$

For the diode, use the piecewise-linear forward model

$$\begin{align*} U_{\rm D} = U_{\rm TO}+r_{\rm F}I_{\rm D} \end{align*}$$

with

$$\begin{align*} U_{\rm TO}=0.65~{\rm V}, \qquad r_{\rm F}=5.0~\Omega. \end{align*}$$

- Draw the loop equation.
- Calculate $I_{\rm D}$.
- Calculate $U_{\rm D}$.
- Calculate the diode power $P_{\rm D}$.
- Compare briefly with the constant-voltage model $U_{\rm D}=0.65~{\rm V}$.

Result

> The loop equation is
>
> $$\begin{align*} U_{\rm I} = RI_{\rm D} + U_{\rm D}. \end{align*}$$

Insert the piecewise-linear diode model:

$$\begin{align*} U_{\rm I} = RI_{\rm D} + U_{\rm TO} + r_{\rm F}I_{\rm D}. \end{align*}$$

Thus

$$\begin{align*} I_{\rm D} = \frac{U_{\rm I}-U_{\rm TO}}{R+r_{\rm F}}. \end{align*}$$

Insert the values:

$$\begin{align*} I_{\rm D} &= \frac{12~{\rm V}-0.65~{\rm V}}{560~\Omega+5.0~\Omega} \\ &= \frac{11.35~{\rm V}}{565~\Omega} \\ &= 20.1~{\rm mA}. \end{align*}$$

The diode voltage is

$$\begin{align*} U_{\rm D} &= U_{\rm TO}+r_{\rm F}I_{\rm D} \\ &= 0.65~{\rm V} + 5.0~\Omega\cdot 20.1~{\rm mA} \\ &= 0.65~{\rm V}+0.101~{\rm V} \\ &= 0.751~{\rm V}. \end{align*}$$

The diode power is

$$\begin{align*} P_{\rm D} = U_{\rm D}I_{\rm D} = 0.751~{\rm V}\cdot 20.1~{\rm mA} = 15.1~{\rm mW}. \end{align*}$$

With the constant-voltage model,

$$\begin{align*} I_{\rm D} = \frac{12~{\rm V}-0.65~{\rm V}}{560~\Omega} = 20.3~{\rm mA}. \end{align*}$$

The difference is small here because $r_{\rm F}\ll R$.

Common pitfalls

- **Thinking a diode is just a resistor:** A diode is nonlinear. The ratio U/I is not constant.
- **Forgetting current limitation:** A forward-biased diode needs a current-limiting component.
- **Treating $0.7~{\rm V}$ as exact:** The forward voltage depends on current, temperature, and semiconductor material.
- **Mixing anode and cathode:** Current flows easily from anode to cathode when the diode is forward-biased.
- **Ignoring reverse limits:** Real diodes have maximum reverse voltage. LEDs often tolerate only small reverse voltages.
- **Confusing hole movement with electron movement:** Holes are missing electrons, but they behave like positive mobile charge carriers.
- **Using the exponential diode equation without unit care:** $U_{\rm T}$ must be in volts and T in kelvin.

Embedded resources