

Physics of a Diode

Calculation of the Current-Voltage Relation

Patrick SCHREY

January 24, 2020

Contents

1	What is a diode?	3
2	How does a diode work?	3
2.1	Forming a p-n junction	4
2.2	Energy bands	5
2.3	Charge carrier concentration	8
2.3.1	Homogeneous Solution	9
2.3.2	Inhomogeneous Solution	10
2.3.3	Boundary Conditions	11
2.4	Current density	15
	References	16
	Legal Notice	17

1 What is a diode?

The diode is the working horse and the mother of all integrated devices. Understanding the diode is key for almost all other devices encountered in integrated electronic circuits.

Without going into much detail on the actual fabrication, a diode is formed by placing a p-doped semiconductor next to an n-doped semiconductor. The intersection of p- and n-doped semiconductor is called a p-n junction or simply diode. A cross-section of a diode can be seen in Figure 1.

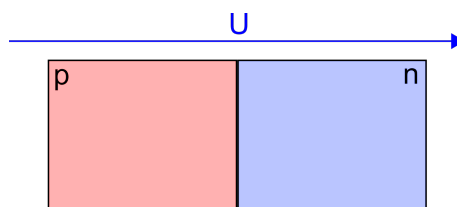


Figure 1: Simplified cross-section of a diode. On the left hand side, there is the p-doped semiconductor. On the right hand side, there is the n-doped semiconductor. At the interface between these two semiconductors, a p-n junction is formed.

A diode has the interesting property to conduct current in one direction only. It conducts current while being forward biased ($U > 0\text{ V}$) and almost behaves like an insulator while reverse biased ($U < 0\text{ V}$).

But how does a diode achieve this behaviour? Let's see how a diode works.

2 How does a diode work?

In order to understand how a diode operates, this section is divided into three parts:

First, we will take a look on what happens at the interface between p- and n-doped semiconductor when we form a p-n junction. Especially the movement of charge carriers and thermal equilibrium will be of utter importance.

Second, we will investigate how the energy band diagram along the direction of current flow in a diode changes. This gives a qualitative picture on why a diode behaves the way it does.

Third, we will grab the math hammer and derive the current-voltage relation of a diode. We will solve the continuity equation to get the charge carrier concentration. The obtained charge carrier concentration is used along with the drift-diffusion equation to get the current density \vec{J} as a function of applied voltage U . This gives a quantitative explanation of diode behaviour.

2.1 Forming a p-n junction

As we have seen in previous lectures, doping can be used to change the concentration of electrons n or holes p in a semiconductor. The concentrations of electrons and holes right after forming the p-n junction are shown in Figure 2.

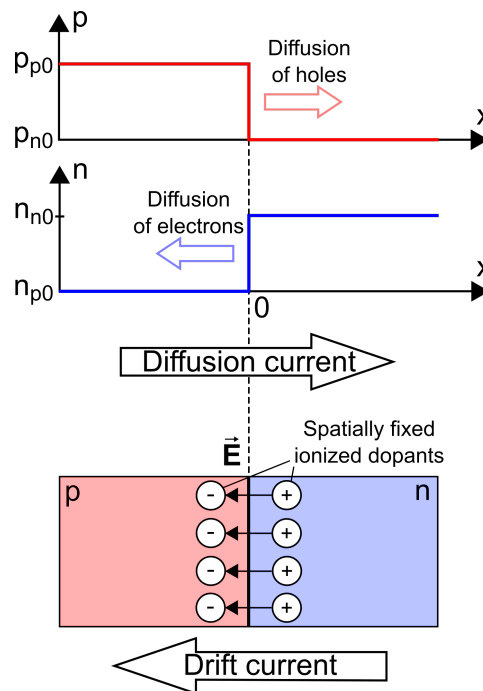


Figure 2: The charge carrier concentration for adjacent p- and n-doped semiconductors change along distance. This gradient causes a diffusion current flow. The ionized dopants create an electric field \vec{E} . The resulting drift current cancels the diffusion current.
 Top: Hole concentration p and electron concentration n over distance x .
 Bottom: Cross-section of a p-n junction.

Placing a p- and n-doped semiconductor next to each other causes a gradient in the concentrations of electrons n and holes p . A gradient in concentration leads to diffusion as governed by Fick's laws. The charge carriers start to move towards each other and eventually recombine. As the charge carrier move and recombine, the ionized dopants stay behind and create an electric field \vec{E} . This electric field causes a drift current into the opposite direction of the aforementioned diffusion current. In thermal equilibrium, the drift and diffusion currents are exactly equal but of opposite sign and perfectly cancel each other. Hence, there is no net current flow $\vec{J} = 0 \text{ A/m}^2$. The layer in between the p- and n-doped regions has two different names. As there are no mobile charge carriers present, in other words the region is depleted of mobile charge carrier, the layer is called depletion layer. But there are still the immobile ionized dopants which are fixed in space. This lead to the second name of space charge region. Both names, depletion layer and space charge region can be used synonymously.

In a next step, we will take a look at the energy bands inside a diode. Because, as the charge carrier concentration p and n have changed, surely the energy bands must be affected.

How does a diode work?

2.2 Energy bands

Previously, we noted that a diode in thermal equilibrium does not feature any net current flow $\vec{J} = 0 \text{ A/m}^2$. We could try to explain this with the drift-diffusion equation. But there is a much more simplified expression for the current density. The upcoming explanation will focus on electrons. But as holes are the dual particles of electrons, the same reasoning can be done for holes.

Think of electrons like students. Both want to go to a state of minimal effort or energy. So, the motivation for an electron to go somewhere could be a lower energy some-place else. The direction to a place of lower energy is given by the gradient. The most dominant energy in semi-conductors is of course the Fermi energy E_F . The electron current density \vec{J}_n is proportional to the change (or gradient) in Fermi energy E_F along x .

$$\vec{J}_n = \mu_n \cdot n \cdot \frac{dE_F}{dx} \quad (1)$$

And likewise for the hole current density \vec{J}_p .

$$\vec{J}_p = \mu_p \cdot p \cdot \frac{dE_F}{dx} \quad (2)$$

If we ask for zero net current flow there are three possible options:

1. $\mu_n = 0 \text{ m}^2/\text{Vs}$ and $\mu_p = 0 \text{ m}^2/\text{Vs}$
2. $n = 0 \text{ /m}^3$ and $p = 0 \text{ /m}^3$
3. $\frac{dE_F}{dx} = 0 \text{ J/m}$

From these option only the last one is reasonable, which means the Fermi energy is constant along x . For zero current flowing through the diode in thermal equilibrium, the Fermi energy E_F has to be constant throughout the diode.

We now know how the Fermi energy looks like in thermal equilibrium. But what about the valence and conduction band edges, E_V and E_C ?

In a p-doped semiconductor, the Fermi energy is close to the valence band. This can be seen on the left hand side of [Figure 3 on the following page](#). Contrary for an n-doped semiconductor, in which the Fermi energy is close to the conduction band as shown on the right hand side of [Figure 3 on the next page](#).

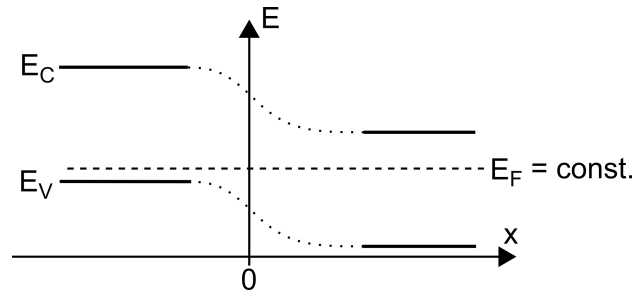


Figure 3: Energy band diagram of a diode in thermal equilibrium. The Fermi energy E_F is constant throughout the diode. The valence and conduction band edges E_V and E_C change along x . The curvature of E_V and E_C can be derived from electrostatics.

While the Fermi energy E_F is constant, the conduction band edge E_C and valence band edge E_V change along x . How E_V and E_C change can be explained by electrostatics. The argumentation will be supported by [Figure 4 on the facing page](#).

In [Section 2.1 on page 4](#), we noted that the ionized dopants create an electric field \vec{E} . Gauss law in [Equation 3](#) relates the charge density ρ to the divergence of the electric field.

$$\nabla \cdot \vec{E} = \frac{\rho}{\epsilon} \quad (3)$$

Restricting to one dimension and integrating along distance x gives us the electric field.

$$\vec{E} = \int_{-\infty}^{+\infty} \frac{d\vec{E}}{dx} dx = \int_{-\infty}^{+\infty} \frac{\rho}{\epsilon} dx \quad (4)$$

The electric potential distribution $\Phi(x)$ can be calculated by the line integral of the electric field.

$$\Phi(x) = \int_{-\infty}^x \vec{E}(x') dx' \quad (5)$$

As the electric potential is the electric potential energy E normalized by the charge, we simply de-normalize by multiplying with the elementary charge q .

$$E(x) = -q \cdot \Phi(x) \quad (6)$$

The valence band edge E_V and the conduction band edge E_C can be calculated by [Equation 6¹](#). Furthermore, the so called mid-gap energy level E_i is defined

$$E_i = \frac{E_V + E_C}{2} \quad (7)$$

The mid-gap energy level is similar to the Fermi energy of an intrinsic semiconductor.

¹ Note that the energy diagram gives the energy of electrons and their charge is negative. $-q = -1.602 \cdot 10^{-19} \text{ C}$. Hence, the energy bands bend in the opposite direction of the electric potential.

How does a diode work?

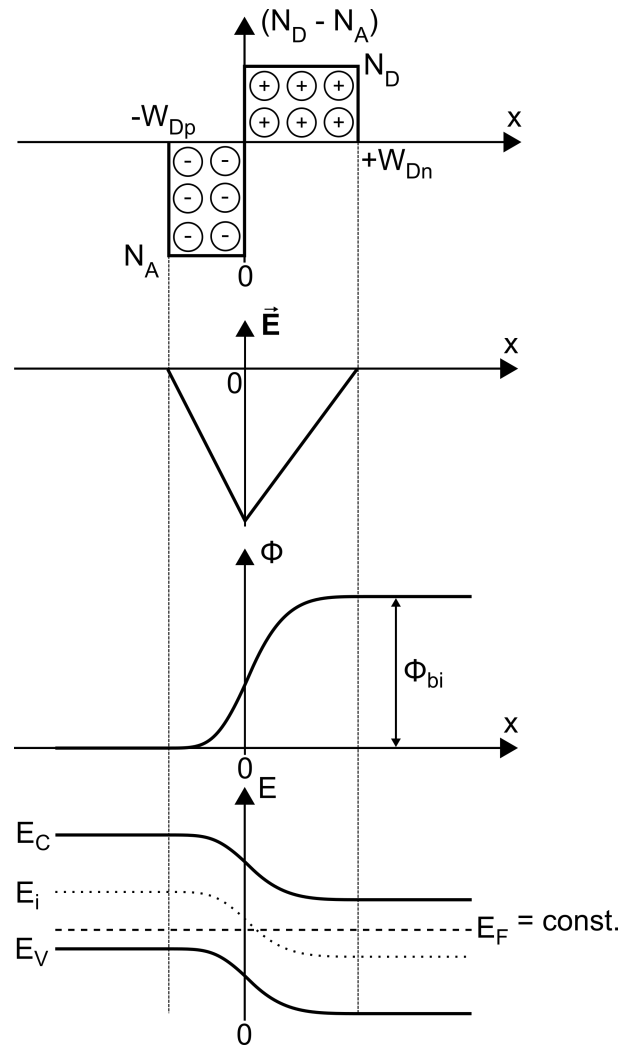


Figure 4: The change in band edges E_V and E_C is derived from electrostatics.

Top: The concentration of ionized dopants ($N_D - N_A$) is the starting point.

Above centre: The electric field \vec{E} can be obtained via Gauss law.

Below centre: The electric potential Φ results from the line integral of the electric field.

Bottom: Energy bands equals the electric potential times elementary charge q .

Observing [Figure 4](#), a difference in electric potential from p- to n-doped side can be identified.

This difference is called the built-in potential Φ_{bi} and can be approximated by

$$\Phi_{bi} \approx \frac{k \cdot T}{q} \cdot \ln \left| \frac{N_A \cdot N_D}{n_i^2} \right| \quad (8)$$

here k is the Boltzmann constant, T is the temperature, N_A and N_D are the acceptor and donor doping concentrations, and n_i is the intrinsic charge carrier concentration.

For Silicon with typical doping concentrations, the built-in potential is around 0.7 V. The built-in potential forms a barrier for current flow in reverse bias and gives a qualitative explanation for the behaviour of a diode.

In order to calculate the current-voltage relation of a diode, we have to solve the continuity equation.

2.3 Charge carrier concentration

To get a quantitative explanation for diode operation, we first think about how to tackle this problem. The current flowing through the diode is the very same current flowing through the depletion layer. If we know how many charge carriers make it to the particular opposite end of the depletion layer, we immediately are ready to solve for the current. Hence, we have to calculate how many electrons travel from the n-doped region through the depletion layer and reach the p-side. And vice versa, we have to know how many holes from the p-doped region go through the depletion layer and reach the n-side. Lets focus on the p-doped region. Instead of dealing with the holes in the p-doped semiconductor, we work with the electrons. We will call the holes in the p-doped region majority charge carriers p_p^2 , while the electrons are called minority charge carriers n_p . Dual for the n-doped region, where electrons are the majority charge carriers n_n and holes are the minority charge carriers p_n . This leads to something very interesting and can be encountered in almost every semi-conducting device. With doping we determine the majority charge carriers, but the operation is governed by the minority charge carriers.

To summarize the last paragraph: We have to calculate the minority charge carrier concentration in the p- and n-doped regions. Further on we will focus on the electrons as minority charge carriers in the p-doped region n_p .

We already derived the continuity equation in a previous lecture. Here once again restricted to one direction

$$\frac{\partial n_p}{\partial t} = D_n \cdot \frac{\partial^2 n_p}{\partial x^2} + \mu_n \cdot \vec{E} \cdot \frac{\partial n_p}{\partial x} + G - R \quad (9)$$

Yes, you are right. This is an inhomogeneous partial differential equation. If you wonder, now would be the perfect moment to start running around the room in panic and scream your terror out for the world to hear. Because we are set to solve this thing.

To tame this beast of an equation, we simplify things by making a few assumptions and omitting units.

We are mainly interested in the static current-voltage relation. Hence, we consider steady state only $\frac{\partial n_p}{\partial t} = 0$.

The entire electric field is assumed to be located in the depletion layer. The p- and n-doped regions are free of an electric field $\vec{E} = 0$.

The number of generated charge carriers in the diode are considered to be negligible, i.e. the generation rate $G = 0$.

Last but not least, the recombination rate is assumed to be proportional to the concentration of excess charge carrier. Excess charge carriers are there in addition to the ones in thermal

²The index denotes the region of interest. p_p are the holes in the p-region. p_n are the holes in the n-region.

How does a diode work?

equilibrium. The total concentration is n_p and the thermal equilibrium concentration is n_{p0} . This results in an excess charge carrier concentration of $n_p - n_{p0}$. Every charge carrier recombines after its life time τ_n has passed. The recombination rate becomes $R = \frac{n_p - n_{p0}}{\tau_n}$.

Summarizing the assumptions:

- Steady state $\frac{\partial n_p}{\partial t} = 0$
- No electric field in the p-doped region $\vec{E} = 0$
- No generation $G = 0$
- Recombination rate equals $R = \frac{n_p - n_{p0}}{\tau_n}$

The above assumptions simplify the continuity equation in [Equation 9](#) on the preceding page to

$$D_n \cdot \frac{d^2 n_p}{dx^2} - \frac{n_p - n_{p0}}{\tau_n} = 0$$

$$D_n \cdot \frac{d^2 n_p}{dx^2} - \frac{n_p}{\tau_n} = -\frac{n_{p0}}{\tau_n} \quad (10)$$

This is an inhomogeneous differential equation. As you know from your math classes, such an equation is solved in two steps. First, we deal with the homogeneous equation by setting the right hand side in [Equation 10](#) to 0. Second, we solve the inhomogeneous equation by using the solution from the homogeneous equation.

2.3.1 Homogeneous Solution

Setting the right hand side of [Equation 10](#) to 0 gives

$$D_n \cdot \frac{d^2 n_p}{dx^2} - \frac{n_p}{\tau_n} = 0$$

$$\frac{d^2 n_p}{dx^2} = \frac{1}{D_n \cdot \tau_n} \cdot n_p \quad (11)$$

Let's take a closer look at the above equation. We are searching for a function n_p which derived twice $\frac{d^2 n_p}{dx^2}$ gives exactly the same function n_p times a constant $\frac{1}{D_n \cdot \tau_n}$. Anything in mind? Yes, right! The Euler function is the way to go.

We use the ansatz

$$n_p = k \cdot e^{\lambda \cdot x} \quad (12)$$

Calculating the second derivative

$$\frac{d^2 n_p}{dx^2} = \lambda^2 \cdot k \cdot e^{\lambda \cdot x} \quad (13)$$

Substituting the ansatz and the second derivative in [Equation 11 on the preceding page](#) yields

$$\begin{aligned}\lambda^2 \cdot k \cdot e^{\lambda \cdot x} &= \frac{1}{D_n \cdot \tau_n} \cdot k \cdot e^{\lambda \cdot x} \\ \lambda^2 &= \frac{1}{D_n \cdot \tau_n}\end{aligned}\quad (14)$$

The product of the diffusion constant D_n and the life time τ_n is defined as the squared diffusion length L_n^2 .

$$\begin{aligned}L_n^2 &= D_n \cdot \tau_n \\ L_n &= \sqrt{D_n \cdot \tau_n}\end{aligned}\quad (15)$$

λ can now be written as

$$\lambda = \pm \frac{1}{L_n}\quad (16)$$

We got two different λ , $\lambda_1 = +\frac{1}{L_n}$ and $\lambda_2 = -\frac{1}{L_n}$. But we only had one k in our ansatz up to now. For the sake of generality, we re-write our ansatz and include k_1 and k_2 .

$$n_p = k_1 \cdot e^{+\frac{x}{L_n}} + k_2 \cdot e^{-\frac{x}{L_n}}\quad (17)$$

The above equation is the solution for the homogeneous equation in [Equation 11 on the previous page](#). Now we have to determine the inhomogeneous solution.

2.3.2 Inhomogeneous Solution

To solve the inhomogeneous differential equation in [Equation 10 on the preceding page](#) we again choose a suitable ansatz. The inhomogeneous solution has to fulfill the homogeneous equation of course. Thus, the homogeneous solution is part of the inhomogeneous. But we have to add something else in order to satisfy the inhomogeneous equation. Investigating the inhomogeneous equation in [Equation 10 on the previous page](#), we can see that the right hand side is a constant over x . We might also want to use a constant C for our ansatz.

$$n_p = k_1 \cdot e^{+\frac{x}{L_n}} + k_2 \cdot e^{-\frac{x}{L_n}} + C\quad (18)$$

Calculating the second derivative

$$\frac{d^2 n_p}{dx^2} = \frac{k_1}{L_n^2} \cdot k_1 \cdot e^{\lambda_1 \cdot x} + \frac{k_2}{L_n^2} \cdot k_2 \cdot e^{-\lambda_2 \cdot x}\quad (19)$$

How does a diode work?

Using the ansatz and second derivative in the inhomogeneous differential equation in [Equation 10 on page 9](#) gives

$$D_n \cdot \frac{k_1}{L_n^2} \cdot e^{+\frac{x}{L_n}} + D_n \cdot \frac{k_2}{L_n^2} \cdot e^{-\frac{x}{L_n}} - \frac{k_1}{\tau_n} \cdot e^{+\frac{x}{L_n}} - \frac{k_2}{\tau_n} \cdot e^{-\frac{x}{L_n}} - \frac{C}{\tau_n} = -\frac{n_{p0}}{\tau_n}$$

Reformulating [Equation 15 on the preceding page](#) to

$$\frac{1}{\tau_n} = \frac{D_n}{L_n^2}$$

results in

$$\frac{k_1}{\tau_n} \cdot e^{+\frac{x}{L_n}} + \frac{k_2}{\tau_n} \cdot e^{-\frac{x}{L_n}} - \frac{k_1}{\tau_n} \cdot e^{+\frac{x}{L_n}} - \frac{k_2}{\tau_n} \cdot e^{-\frac{x}{L_n}} - \frac{C}{\tau_n} = -\frac{n_{p0}}{\tau_n}$$

Hence, the constant C in our ansatz equals

$$C = n_{p0} \quad (20)$$

And our inhomogeneous solution becomes

$$n_p = k_1 \cdot e^{+\frac{x}{L_n}} + k_2 \cdot e^{-\frac{x}{L_n}} + n_{p0} \quad (21)$$

This function for n_p solves the continuity equation. But there are two missing pieces. What about k_1 and k_2 ?

To uniquely determine the charge carrier concentration, we have to use boundary conditions.

2.3.3 Boundary Conditions

Boundary conditions allow to uniquely define the charge carrier concentration. As we have still two unknown quantities k_1 and k_2 , we will need two boundary conditions.

The first and more simple one is related to the far left hand side of the p-region. There, the depletion layer and n-doped region are far away and have negligible influence on the p-region. The electron concentration n_p has to converge towards its thermal equilibrium value n_{p0} for x going to $-\infty$.

$$n_p(x \rightarrow -\infty) \stackrel{!}{=} n_{p0} \quad (22)$$

Applying the limit to [Equation 21](#) yields

$$\lim_{x \rightarrow -\infty} n_p(x) = \lim_{x \rightarrow -\infty} k_1 \cdot e^{+\frac{x}{L_n}} + k_2 \cdot e^{-\frac{x}{L_n}} + n_{p0} \stackrel{!}{=} n_{p0} \quad (23)$$

k_1 cancels as its exponential drops to 0. But the exponential with k_2 approaches ∞ . Thus,

$$k_2 = 0 \quad (24)$$

in order to satisfy the boundary condition.

To get k_1 , we need a boundary condition for the right hand side of the p-region $n_p(x = -W_{Dp})$. This is the interface to the depletion layer which will be subject to further discussions.

What we will need is the so called mass-action law. Remember the formula for the intrinsic charge carrier concentration n_i and p_i of the previous lecture.

$$p_i = n_i = \sqrt{N_V \cdot N_C} \cdot e^{-\frac{E_g}{2 \cdot k \cdot T}} \quad (25)$$

where N_V and N_C are the density of states in the valence and conduction band, E_g is the band gap energy, k is the Boltzmann constant, and T is the temperature.

Multiplying Equation 25 with n_i gives

$$p_i \cdot n_i = n_i^2 = N_V \cdot N_C \cdot e^{-\frac{E_g}{k \cdot T}} \quad (26)$$

The right hand side depends on properties of silicon N_V , N_C , and E_g , and the temperature T . Moderate doping does not change any of these properties significantly. The individual concentrations for electrons n and holes p change tremendously due to doping of course. But their product stays n_i^2 constantly. Hence, the relation

$$p \cdot n = n_i^2 \quad (27)$$

is also valid for doped semiconductors, as long as the band gap does not change due to heavy doping. The relation in Equation 27 is called the mass-action law and is the first piece of the puzzle. Note the mass-action law is valid for thermal equilibrium only, i.e. no voltage is applied. The individual concentrations for holes and electrons depend on the Fermi energy E_F . If we change the Fermi energy from its intrinsic value or the mid-gap energy level E_i , the hole and electron concentrations can be written as

$$p = n_i \cdot e^{\frac{E_i - E_F}{k \cdot T}} \quad n = n_i \cdot e^{\frac{E_F - E_i}{k \cdot T}} \quad (28)$$

Note, the product $p \cdot n$ equals n_i^2 as given by the mass-action law in Equation 27.

$$p \cdot n = n_i \cdot e^{\frac{E_i - E_F}{k \cdot T}} \cdot n_i \cdot e^{\frac{E_F - E_i}{k \cdot T}} = n_i^2 \quad (29)$$

How does a diode work?

Up to now we worked in thermal equilibrium, i.e. without any external voltage. If we apply an external voltage U , we change the Fermi energies of holes and electrons. We introduce the quasi-Fermi energies for holes E_{Fp} and electrons E_{Fn} and replace the previous Fermi energy in [Equation 28 on the facing page](#).

$$p = n_i \cdot e^{\frac{E_i - E_{Fp}}{k \cdot T}} \quad n = n_i \cdot e^{\frac{E_{Fn} - E_i}{k \cdot T}} \quad (30)$$

Calculating the product $p \cdot n$ gives

$$\begin{aligned} p \cdot n &= n_i \cdot e^{\frac{E_i - E_{Fp}}{k \cdot T}} \cdot n_i \cdot e^{\frac{E_{Fn} - E_i}{k \cdot T}} \\ p \cdot n &= n_i^2 \cdot e^{\frac{E_{Fn} - E_{Fp}}{k \cdot T}} \end{aligned} \quad (31)$$

As we can see from the equation above, the mass-action law of [Equation 27 on the preceding page](#) is no longer valid. This is expected since the mass-action law only holds for thermal equilibrium. With an applied voltage the diode is no longer in thermal equilibrium.

The difference in the quasi-Fermi energies is proportional to the applied voltage.

$$E_{Fn} - E_{Fp} = q \cdot U \quad (32)$$

Rewriting [Equation 31](#) results in

$$p \cdot n = n_i^2 \cdot e^{\frac{q \cdot U}{k \cdot T}} \quad (33)$$

And this is huge. We can now control the charge carrier concentration in the depletion layer by applying a voltage. For reverse bias, i.e. $U < 0 \text{ V}$ the charge carrier concentration becomes smaller compared to thermal equilibrium. For forward bias, i.e. $U > 0 \text{ V}$ the charge carrier concentration becomes larger compared to thermal equilibrium.

$$\begin{aligned} p \cdot n &< n_i^2 & \text{for } U < 0 \text{ V} \\ p \cdot n &> n_i^2 & \text{for } U > 0 \text{ V} \end{aligned} \quad (34)$$

This is great. But back to our initial problem with the boundary conditions.

We were looking for a suitable boundary condition at the interface p-doped region and depletion layer. What we need is $n_p(x = -W_{Dp})$. Fortunately, the charge carrier concentration is a steady function in space, i.e. it does not abruptly change along x and we can state

$$p_p(x = -W_{Dp}) \cdot n_p(x = -W_{Dp}) = p \cdot n = n_i^2 \cdot e^{\frac{q \cdot U}{k \cdot T}} \quad (35)$$

We divide Equation 33 on the previous page by the majority charge carrier concentration p_p .

$$n_p(x = -W_{Dp}) = \frac{n_i^2}{p_p(x = -W_{Dp})} \cdot e^{\frac{q \cdot U}{k \cdot T}} \quad (36)$$

The concentration of majority charge carriers $p_p(x = -W_{Dp})$ is a very large number and assumed not to change significantly with applied voltage.

$$p_p(x = -W_{Dp}) \approx p_{p0} \quad (37)$$

Combining the two equation above gives

$$n_p(x = -W_{Dp}) = \frac{n_i^2}{p_{p0}} \cdot e^{\frac{q \cdot U}{k \cdot T}} \quad (38)$$

With the mass-action law of Equation 27 on page 12 we can re-write the fraction $\frac{n_i^2}{p_{p0}}$ as n_{p0} and get

$$n_p(x = -W_{Dp}) = n_{p0} \cdot e^{\frac{q \cdot U}{k \cdot T}} \quad (39)$$

This is the missing boundary condition to obtain k_1 in Equation 21 on page 11. As $k_2 = 0$, we can write

$$n_p(x = -W_{Dp}) = k_1 \cdot e^{-\frac{W_{Dp}}{L_n}} + n_{p0} \stackrel{!}{=} n_{p0} \cdot e^{\frac{q \cdot U}{k \cdot T}} \quad (40)$$

and get k_1

$$k_1 = n_{p0} \cdot \left(e^{\frac{q \cdot U}{k \cdot T}} - 1 \right) \cdot e^{+\frac{W_{Dp}}{L_n}} \quad (41)$$

Finally, we have the electron concentration in the p-doped region.

$$n_p = n_{p0} \cdot \left(e^{\frac{q \cdot U}{k \cdot T}} - 1 \right) \cdot e^{\frac{x+W_{Dp}}{L_n}} + n_{p0} \quad (42)$$

Without repeating the entire calculations, the hole concentration in the n-doped region is given by

$$p_n = p_{n0} \cdot \left(e^{\frac{q \cdot U}{k \cdot T}} - 1 \right) \cdot e^{-\frac{x-W_{Dn}}{L_p}} + p_{n0} \quad (43)$$

Now marvel at the beauty of these two equations. We can see that the minority charge carrier concentrations are an exponential function in space. But what really matters is that we can control the sign of this exponential by applying a voltage.

$$n_p \sim \begin{cases} -n_{p0} \cdot e^{\frac{x+W_{Dp}}{L_n}} + n_{p0} & \text{for } U < 0 \text{ V} \\ n_{p0} & \text{for } U = 0 \text{ V} \\ +n_{p0} \cdot e^{\frac{x+W_{Dp}}{L_n}} + n_{p0} & \text{for } U > 0 \text{ V} \end{cases} \quad (44)$$

How does a diode work?

This control over the minority charge carrier concentrations by means of a voltage is shown in Figure 5.

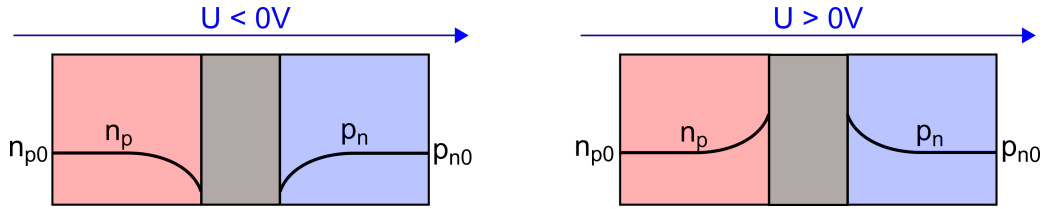


Figure 5: The minority charge carrier concentration can be controlled by applying a voltage.

Left: For reverse bias $U < 0\text{ V}$, the concentrations decrease at the depletion layer.

Right: For forward bias $U > 0\text{ V}$, the concentrations increase at the depletion layer.

The most exhausting part lies behind us. The only thing that's left is the current density.

2.4 Current density

The current can easily be calculated using the drift-diffusion equation. Again we focus on the electrons in the p-doped region.

$$\vec{J}_n = q \cdot n_p \cdot \mu_n \cdot \vec{E} + q \cdot D_n \cdot \nabla n_p \quad (45)$$

On [Section 2.3 on page 9](#), we assumed an electric field of $\vec{E} = 0$. This eliminates the drift component of the current density and leaves only the diffusion component. Also, we use one dimension for our calculations only.

$$\vec{J}_n = q \cdot D_n \cdot \frac{dn_p}{dx} \quad (46)$$

Substituting the minority charge carrier concentration n_p from [Equation 42 on the facing page](#) results in

$$\vec{J}_n = \frac{q \cdot D_n \cdot n_{p0}}{L_n} \cdot \left(e^{\frac{q \cdot U}{k \cdot T}} - 1 \right) \cdot e^{\frac{x + W_{DP}}{L_n}} \quad (47)$$

As detailed at the very beginning, the current flowing through the diode is the one through the depletion layer. The electron current density at the depletion layer interface can be obtained by setting $x = -W_{DP}$.

$$\vec{J}_n = \frac{q \cdot D_n \cdot n_{p0}}{L_n} \cdot \left(e^{\frac{q \cdot U}{k \cdot T}} - 1 \right) \quad (48)$$

A similar argumentation can be done for holes in the n-doped region p_n to get the hole current density \vec{J}_p .

$$\vec{J}_p = \frac{q \cdot D_p \cdot p_{n0}}{L_p} \cdot \left(e^{\frac{q \cdot U}{k \cdot T}} - 1 \right) \quad (49)$$

The total diode current density is the some of hole and electron current density.

$$\vec{J} = \vec{J}_p + \vec{J}_n = \left(\frac{q \cdot D_p \cdot p_{n0}}{L_p} + \frac{q \cdot D_n \cdot n_{p0}}{L_n} \right) \cdot \left(e^{\frac{q \cdot U}{k \cdot T}} - 1 \right) \quad (50)$$

Defining the saturation current density J_0

$$J_0 = \frac{q \cdot D_p \cdot p_{n0}}{L_p} + \frac{q \cdot D_n \cdot n_{p0}}{L_n} \quad (51)$$

results in the famous Shockley equation.

$$\vec{J} = J_0 \cdot \left(e^{\frac{q \cdot U}{k \cdot T}} - 1 \right) \quad (52)$$

The Shockley equation is shown in [Figure 6](#) for a very small range of voltage U .

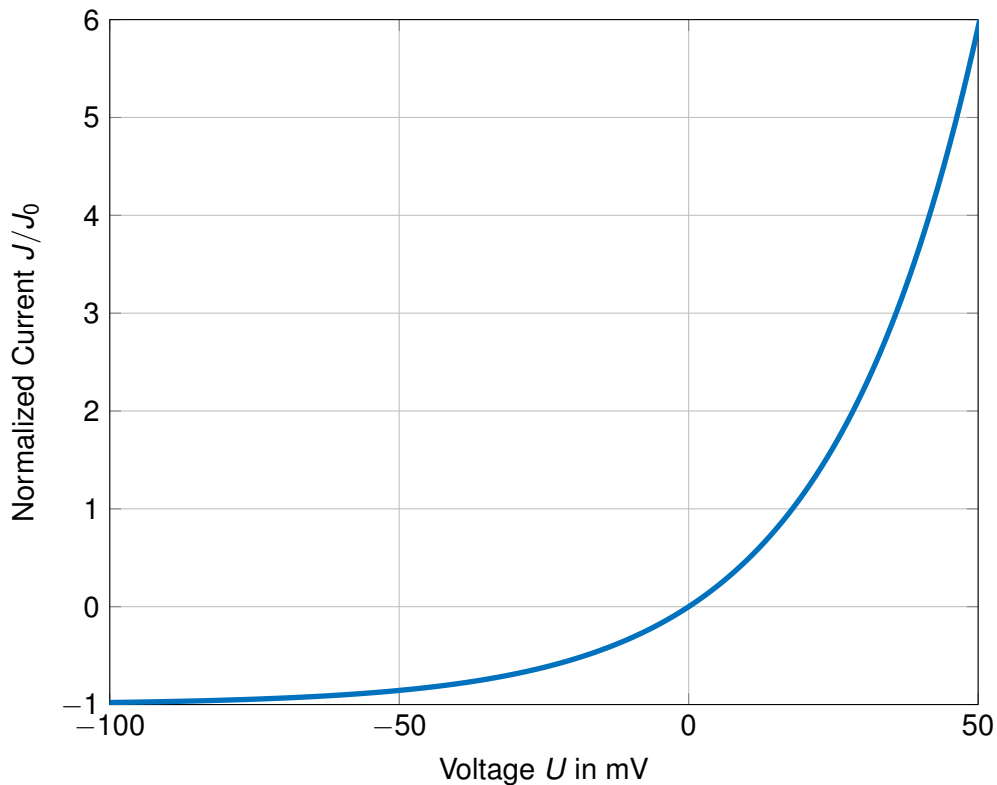


Figure 6: Normalized diode current density J/J_0 over applied voltage U according to the Shockley equation. For negative voltages, the current density approaches the saturation current density J_0 . Positive voltages cause an exponential increase in current density.

References

- [1] S. M. Sze, K. K. Ng. *Physics of Semiconductor Devices*. 3rd ed. John Wiley & Sons, 2007.

Legal Notice

Physics of a Diode by Patrick Schrey is licensed under a [Creative Commons Attribution 4.0 International License](#).

This includes all pictures with the exception of the title page and all logos.