

Supplemental Material

Light-Semiconductor-Interaction

Fundamentals

In this section the basics of light-semiconductor interaction are explained.

Band structure

The basis of the interaction between light and semiconductors is the band structure: when an electron is placed in a periodic potential, its wave function can be written as a Bloch wave with the eigenenergies $E_n(\mathbf{k})$, whereby n and $\mathbf{k}$ determine the band index and the wave vector accordingly. Plotting the possible eigenenergies against their respective wave vectors leads to the well known band structures. As the plot of a band structure needs to represent a scalar function with a three dimensional argument, most representations only show the band structure along the high symmetry k-vectors. Additionally the symmetry of the eigenenergies can be used to reduce the plot to the first Brillouin zone, which contains all the information about the band structure within the crystal.

Fig. 1 shows the band structure of silicon as well as its first Brillouin zone (green structure). As mentioned above, every point on a line is equivalent to the energy value of a Bloch wave for the corresponding wave vector. It is noticeable that certain energy regions do not contain such lines (grey area). These energy regions, which can not be filled by electrons in the solid state material, are called band gaps. If the Fermi-level for a temperature of 0 K (equivalent to the Fermi-energy), which shows the energy of the lowest state, filled with electrons, lies in such a band gap and if the band gap is not larger than a few eV, one can speak of a semiconductor. The Fermi-level is typically placed in the middle of the band gap, even though no states are present there. The reason for that is discussed in the subchapter about PN-junctions.

To analyze the effects of the band structure on the electrical conductivity, the energy lines are often combined into energy bands, which cover the energy regions that can be filled with electrons. These energy bands are of central importance to the conductivity of a material. First of all it can be noted that bands, which are fully filled with electrons are not conductive. This can be shown rather easily, by integrating the overall current density within the first Brillouin zone:

$$\mathbf{J} = \sum_n \frac{-e}{4\pi^3 \hbar} \int_{1.BZ} \nabla_{\mathbf{k}} E_n(\mathbf{k}) d^3\mathbf{k},$$

whereby the integration has to contain all $E_n(\mathbf{k})$ within the band. As the energy is periodic under the addition or subtraction of a reciprocal lattice vector and since the first Brillouin zone just contains one period of the k-space, the upper integral has to be zero for a fully filled band. For a non-fully filled band a non-zero conductivity is possible since not all the $E_n(\mathbf{k})$ have to be used in the integral, but only the once filled by electrons. By creating an asymmetric distribution of electrons in the k-space, which can be realized in practice by the use of an electric field, an electric current can be created.

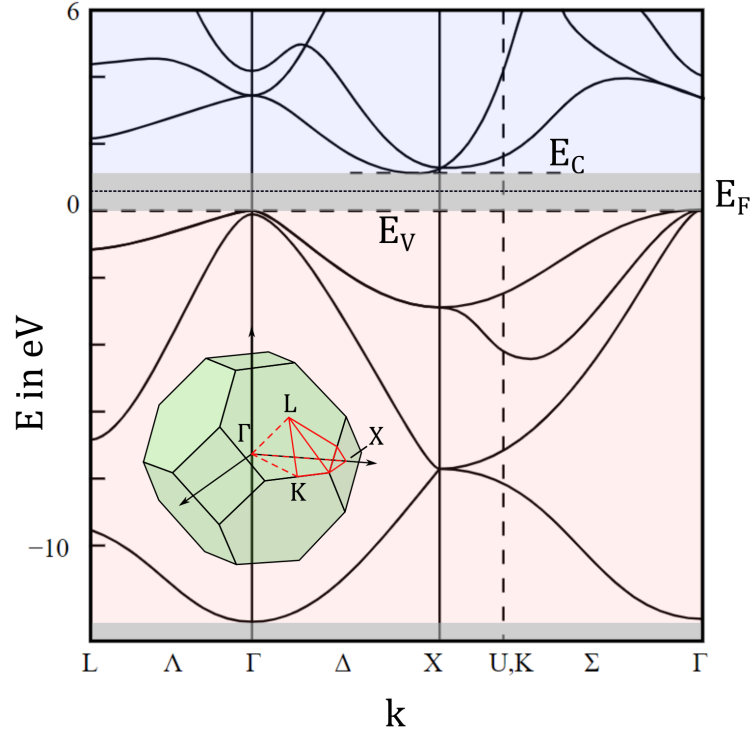


Fig. 1: Band structure of silicon along the high symmetry k -vectors. The different colours mark the valence- and the conduction band (red and blue respectively) as well as the band gaps (grey). The energy axis labels are aligned in respect to the upper edge of the valence band ($E_V = 0$); E_C and E_F determine the lower edge of the conduction band and the Fermi-energy. Additionally the image contains the first Brillouin zone of silicon, in which some of the high symmetry points, which are of special importance for the band structure, are highlighted (green sketch). Band structure taken from J. R. Chelikowsky, M. L. Cohen: *Electronic structure of silicon*, Phys. Rev. B **10**, 12 (1974) (edited).

Due to the absent conductivity of a fully filled band, the energetically highest, fully filled band is called the valence band, while the energetically lowest, non-fully filled band is called the conduction band. To produce a non-zero conductivity within a semiconductor, an electron has to be lifted from the valence band into the conduction band. This can be achieved either via thermal excitation or via electron-photon interaction, which is discussed in the following chapter. In this scenario, it is important to note that such an excitation leads to a finite conductivity in both bands, as the removal of some electrons from the valence band means that this band is no longer fully filled. The resulting conductivity in the valence band is often called hole conductivity, as the missing electrons can be regarded as quasi particles in a simplified picture.

Typically one distinguishes between two types of semiconductors: the direct and the indirect ones. For the direct semiconductors the minimum of the conduction band and the maximum of the valence band are present at the same value of k , while for an indirect semiconductor they are found to be at different k -values. This distinction is useful, as direct and indirect semiconductors have different characteristics in their interaction with photons, especially in the case of absorption (see chapter about electron-photon-interaction).

Electron-Photon-Interaction

Due to the wave-particle duality, the interaction between an electron and a photon can be regarded as a scattering process of two particles. As is the case for every scattering process, the energies of the particles as well as the impulses have to be conserved after the scattering process. Since the photon only has a marginal impulse compared to the electrons in the material, the change of the wave vectors of the electron and the photon can be neglected. Accordingly the photon (approximately) only transfers energy to the electrons. However this transfer of energy is only possible if a free state for the interacting electron exists at the according energy in one of the energy bands. This is the case due to the fermionic nature of the electrons, which prohibits that two electrons share the exact same state. In fig. 2 typical electron-photon interactions are presented. The most common of these processes are the absorption and emission of a photon.

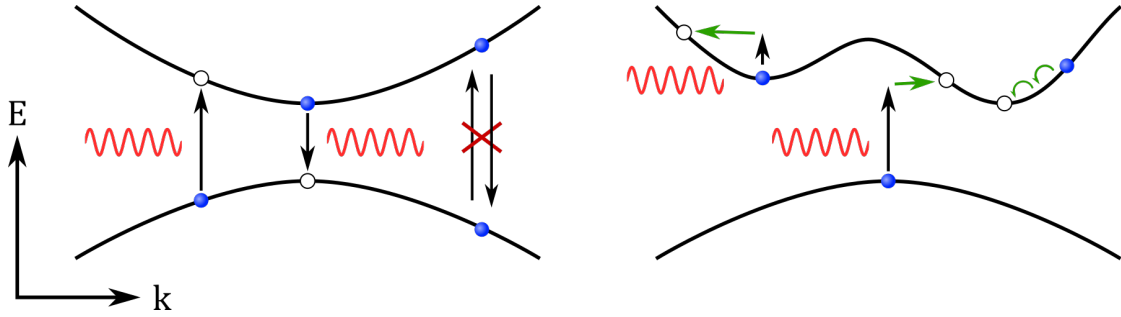


Fig. 2: Typical electron-photon processes (left) as well as typical electron-photon-phonon processes (right). In the left sketch the leftmost process shows the absorption of a photon by an electron, while the process in the center shows the emission of a photon by an electron. The right part of the sketch indicates that a transition is not possible if the according energy levels are already occupied. In the right sketch on the left, an intraband transition is presented, which necessitates the absorption of a photon and a phonon (green). The interband transition presented in the center works completely analogous. The rightmost sketch shows the relaxation of an excited electron into the local band minimum through the emission of phonons.

The energy of the transition for an absorption is given by the energy or the wavelength of the absorbed photon respectively. This means that electrons, which absorb a photon with a large wavelength can overcome only small energy gaps, while the ones that absorbed a photon with a small wavelength are able to transition through larger energy gaps. The energies for which absorption is possible are given by the band structure, more specifically by the so called jDOS (joint density of states) of a material. In this way certain information about a material can be collected by looking at its absorption spectrum.

The emission of a photon works completely analogous to the absorption of a photon, the only difference being that the electron is transitioning from a higher to a lower energy state. As we will see later, the emission of photons is mostly occurring at certain energies, these being the ones which describe the transition from a local band minimum of the conduction band to the local band maximum of the valence band.

Next to the aforementioned processes, several more electron-photon processes exist, which become relevant in certain situations. A few examples being the stimulated emission, which is relevant for the physical processes of a laser, and the 2-photon absorption, which only becomes relevant for very high light intensities.

Up to this point only pure electron-photon interactions have been discussed. However the overall interaction between light and semiconductors cannot be explained by only looking at these interactions. The reason for that is that an important factor has been disregarded up to this point: the atomic lattice, which can transfer energy and impulse to the electrons in the form of lattice vibrations, which are usually known as phonons. While the energy of phonons for room temperature is in the range of 26 meV, which is small compared that of photons in the visible range (few eV), the impulse that the phonons can transfer is gigantic compared to the photons'. Due to this, the wave vector of an electron can be changed by an electron-photon-phonon process. This enables the processes, which are presented in fig. 2 (right): in the interband absorption an electron and a phonon are absorbed by the electron. Thereby the electron can transition into a band even if the energy of the absorbed photon would not be large enough to facilitate a direct electron-photon transition, namely when an empty state exists for the given energy at a different wave vector. This effect is especially important in indirect semiconductors, since it enables the absorption of photons and the according transition of electrons at the energy of the band gap, which would not be possible without the influence of the phonons.

The interband absorption, i.e. the transition of an electron into a higher (or lower) energy level in the same band, can be described analogous via the electron-photon-phonon interaction.

One could think that the inverted processes, namely the emission of a photon and a phonon are occurring in semiconductors. While this is technically correct, the processes have a much lower chance of occurring, due to the necessary presence of another particle, compared to the electron-photon interaction. Accordingly the effects of electron-photon-phonon processes are mostly observed, when other processes are suppressed or when certain processes can only be triggered by electron-photon-phonon processes (for example the intraband transition).

The pure electron-phonon interaction is also important for the understanding of electronic behaviour in a semiconductor. It allows for the relaxation of electrons at high energies (often called thermalization) into empty lower states by the emission of photons. This process is mostly happening on a picosecond timescale, which is usually much faster than the spontaneous emission. This explains the aforementioned facts about the emission of photons: the electrons are lifted into a higher band, thermalize into the local energy minimum and, from there, transition into the maximum of the valence band, in which a similar process of thermalization has occurred. Due to this, the transitions are happening primarily between the minimum of the conduction band and the maximum of the valence band. Since the spectrum of emission is defined by the band structure of the material, a measurement of the emission spectrum allows for conclusions about the material in question.

p-n junction

For the measurement of photons so called photodiodes are used most of the time. The fundamental basis of photodiodes is given by the p-n junction, which describes a physical transition between the band structures of two semiconductors with different doping.

To understand the behaviour at the physical transition of two semiconductors as well as the effects of doping, another important parameter of solid state physics has to be explained: the Fermi level. According to the Fermi-Dirac distribution, electrons with a finite temperature have a certain probability to occupy higher energy levels. For every temperature an energy value exists, for which the probability of occupation is exactly 50%. This energy value is

defined as the Fermi level and it describes the highest occupied state under consideration of the thermodynamic distribution of the carriers. It has to be kept in mind however that there is not necessarily a state present at the Fermi level. For example a semiconductor is defined by the fact that the Fermi level is within the energy gap, which is why in most cases there is no energy at the Fermi level for semiconductors.

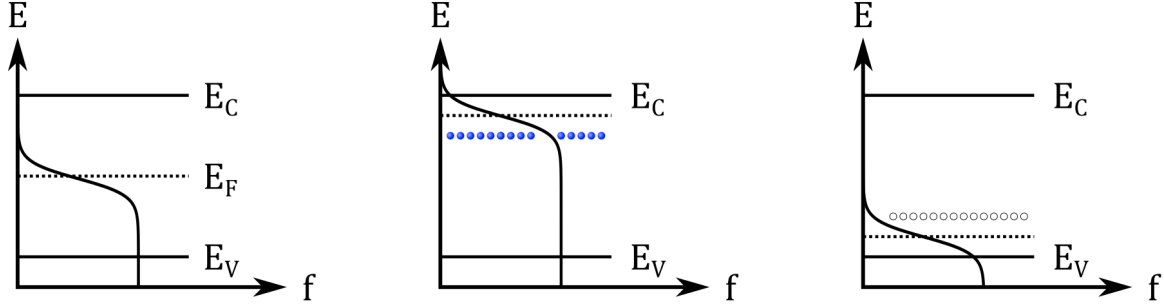


Fig. 3: Probability of occupation according to the Fermi-Dirac distribution for three different dopings (left: undoped, center: n-doped, right: p-doped). Additionally the energies of the lower edge of the conduction band and the upper edge of the valence band as well as the Fermi level (dotted line) are shown. For the doped semiconductors the additional states are presented as well.

By doping a semiconductor, the Fermi level can be shifted in both energetic directions. To accomplish this, foreign atoms are placed inside the crystal lattice of the pure semiconductor. One can distinguish between an n-doping, for which atoms are inserted, which contain more electrons than the other atoms in the semiconductor, or a p-doping, for which electrons with less electrons are inserted. In the case of n-doping, for each foreign atom, there is a single electron, which is bound less strongly to the lattice, since it cannot create a covalent bond with another atom in the semiconductor. This leads to additional (filled) states being present in the band structure, which are placed close to the conduction band edge. For a p-doped system the situation is exactly inverted: here additional (empty) states are added in close proximity to the valence band. As can be seen in fig. 3, this leads to an increase or decrease of the Fermi level.

If two semiconductors with different dopings are brought into physical contact, the semiconductors will have different dopings when first touching each other. This means that from the viewpoint of the electrons in the n-doped material energetically more favorable states exist in the other material, which is why the electrons of the n-doped region transition into the free states of the p-doped semiconductor. Due to this the two materials are no longer electrically neutral, which leads to an electric field that points in the opposite direction compared to the direction of the electron migration. This electron migration will take place until the Fermi levels of both materials are at the same value, implying that there are no more energetically favorable states in the other material. The formation of a p-n junction is graphically presented in fig. 4. In the center of the right sketch, the depletion region is visible. Only a small number of carriers exist in this region, because the previously existing carriers have either recombined during the moment of contact of the two materials or have been pushed out of the region due to the electric field. This depletion region is of fundamental importance for the functioning of a photodiode.

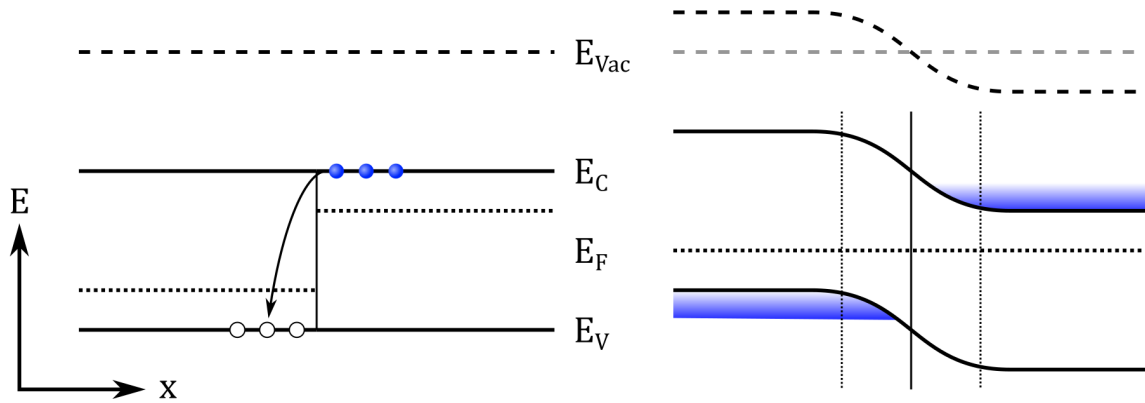


Fig. 4: Formation of a p-n junction (left the situation directly after the moment of contact, right the situation after the alignment of the Fermi levels). The position of the energy levels is given by their relative distance to the vacuum level. In the right sketch the vacuum energy is no longer constant, which is caused by the electric field in the semiconductor. The depletion region is limited by the two dotted vertical lines.

In the following part the light-semiconductor interaction of a photodiode as well as its measuring schemes are elaborated. It should be noted that the p-n junction can show further optoelectronic behaviours, which are of special importance for its usages as a (photo-) diode or as a solar cell. Both of these applications are of central relevance for our modern society, however they are not the topic of this text.

To understand the usage of a photodiode for the measurement of photons, it should be understood first how the photons interact with electrons in the material. In this chapter only the interactions in the depletion region are discussed, since the behaviour of electrons in a pure semiconductor should be known already from the former chapters: in this case the photons that have enough energy will excite electrons, which will recombine with the holes after a given time, while emitting a photon.

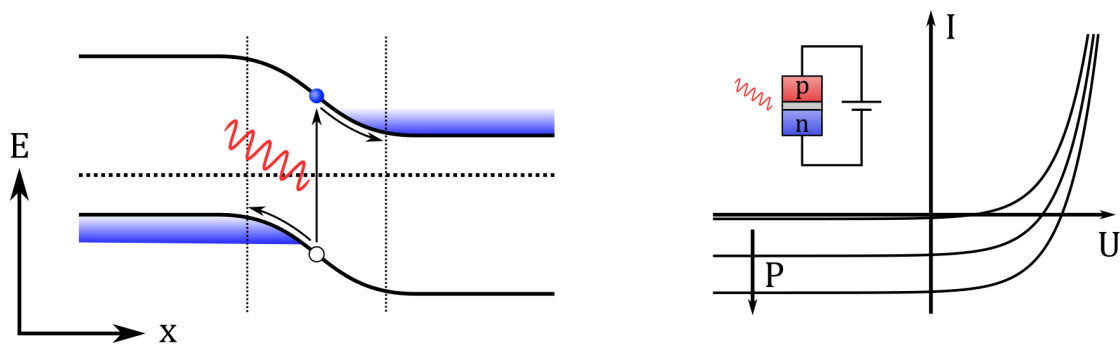


Fig. 5: Left: absorption of a photon within the depletion region. First an electron-hole pair is created, which drifts apart due to the internal electric field. Right: I-V curves of a p-n junction. The small sketch shows the circuit for a positive voltage $U > 0$ V. The lower curves are valid, when light with a higher intensity is interacting with the p-n junction.

In fig. 5 the absorption of a photon within the depletion region is shown. First an electron-hole pair is created through the absorption of a photon. This electron-hole pair is being

influenced by the electric field within the depletion region, which immediately separates the charge carriers from one another and thus suppresses their recombination. The acceleration of the charge carriers leads to a measurable current through the photodiode.

To understand the different measuring schemes of a photodiode, one can take a look at the I-V curves (fig. 5). Without photons and under the influence of a reverse bias ($U < 0\text{ V}$), only a marginal current flows through the device. For a certain voltage in the forward bias ($U > 0\text{ V}$) the current through the device is exponentially increasing. This current is created, because the internal electric field within the depletion region is being attenuated so much by the external field that the electrons and holes can recombine in a large number.

If photons are hitting the depletion region, the I-V curves are being shifted downwards on the I-axis. This means that a current flows through the device, without the need for an electric field. The reason for that has been explained in the previous parts of this text.

For photodiodes one can distinguish between three measuring schemes: the first one being the measurement in a short circuit. In this case the situation described before is being used and thus current flows through the device without a voltage. By measuring this current, information about the intensity of the incoming light can be gained. In the I-V curve this measurement scheme is equivalent to measuring the intersection of the curves with the current axis. For the second measuring scheme the same principle is being used. Additionally the depletion region and the electric field in it is increased by an additional voltage in reverse bias. Due to the increased electric field a higher efficiency and a faster measurement is possible. The third measuring scheme does not use the current for its measurement, but instead the voltage. For this, a voltagemeter can be connected to the photodiode, which sets the current to zero. This measuring scheme is equivalent to measuring the intersections of the I-V curves with the voltage axis. It should be noted that the measurement is strongly non-linear, which is why the other methods are usually preferred.

Setup

The setup makes use of a commercial supercontinuum laser manufactured by the company NKT-Photonics as well as two acousto optic modulators (from the same company) for the filtering of the wavelengths. Furthermore mirrors and other optomechanical components from the company Thorlabs are being used.

Supercontinuum laser

The fundamental principle of the supercontinuum laser is based on the self-phase modulation, a second order non-linear effect, for which the frequency spectrum of a pulse widens, while passing a material. The basis for this behaviour is the Kerr effect, which adds an intensity dependant component to the refractive index:

$$n = n_0 + n_2 I(z, t)$$

Typically this effect is negligible due to the small size of n_2 (typically in the range of $10^{-14}\text{ m}^2/\text{W}$), however the effect can be enhanced significantly by selecting a proper material.

To understand the effects of a frequency dependant refractive index on the wave pulse, the

numerical basics of the self-phase modulation are explained.

For a one-dimensional electromagnetic wave with a slowly varying amplitude $A(z, t)$:

$$E(z, t) = A(z, t) e^{-i(\omega_0 t - k_0 z)} + \text{c.c.},$$

the form of the field, while passing through the medium, is approximately given by

$$\frac{\partial A_s(z, \tau)}{\partial z} - \frac{1}{2} i k_2 \frac{\partial^2 A_s(z, \tau)}{\partial \tau^2} = -i \Delta k_{\text{NL}} A_s(z, \tau),$$

where A_s is connected to A by the equation $A_s(z, \tau) = A_s(z, t - z/v_G) = A(z, t)$ (with v_G as the group velocity). k_2 describes the change of the group velocity according to $k_2 = (d^2 k / d\omega^2)_{\omega=\omega_0}$, while $\Delta k_{\text{NL}} = n_{\text{NL}} \omega_0 / c$ describes the non-linear dispersion. For a medium, in which only the self-phase modulation should occur and all other dispersive effects are neglected, the equation can be simplified to

$$\frac{\partial A_s(z, \tau)}{\partial z} = -i \gamma |A_s(z, \tau)|^2 A_s(z, \tau).$$

Here the assumption that no frequency dependent dispersive effects are present (equivalent to $k_2 = 0$) and the assumption that the Kerr effect should occur have been used, which means that Δk_{NL} can be written as:

$$\Delta k_{\text{NL}} = n_2 I(z, t) \frac{\omega_0}{c} = \frac{1}{2} n_2 n_0 \epsilon_0 \omega_0 |A_s(z, \tau)|^2 = \gamma |A_s(z, \tau)|^2.$$

If $A_s(z, \tau) = |A_s(z, \tau)| e^{i\varphi(z, \tau)}$ is being substituted into the equation one obtains:

$$\frac{\partial |A_s|}{\partial z} e^{i\varphi} + i |A_s| e^{i\varphi} \frac{\partial \varphi}{\partial z} = -i \gamma |A_s|^2 |A_s| e^{i\varphi}.$$

After dividing the former equation by $e^{i\varphi}$, it can be noted that the first term on the left side is real, while all other parts are imaginary. To satisfy the equation

$$\frac{\partial |A_s(z, \tau)|}{\partial z} = 0$$

has to be valid. This indicates that the form of the envelope does not change during the transition through the material. The rest of the equation is leading to:

$$\frac{\partial \varphi(z, \tau)}{\partial z} = -\gamma |A_s(0, \tau)|^2.$$

Thus the self-phase modulation is leading to an additional time dependent phase factor, according to

$$\varphi(z, \tau) = \varphi(0, \tau) - \gamma |A_s(0, \tau)|^2 z.$$

The influence of such a phase factor is being presented in fig. 6. As mentioned before, the form of the enveloping function does not change during the transition through the material. Only the influence of the time- and space-dependent phase factor is leading to a change in the form of the wave.

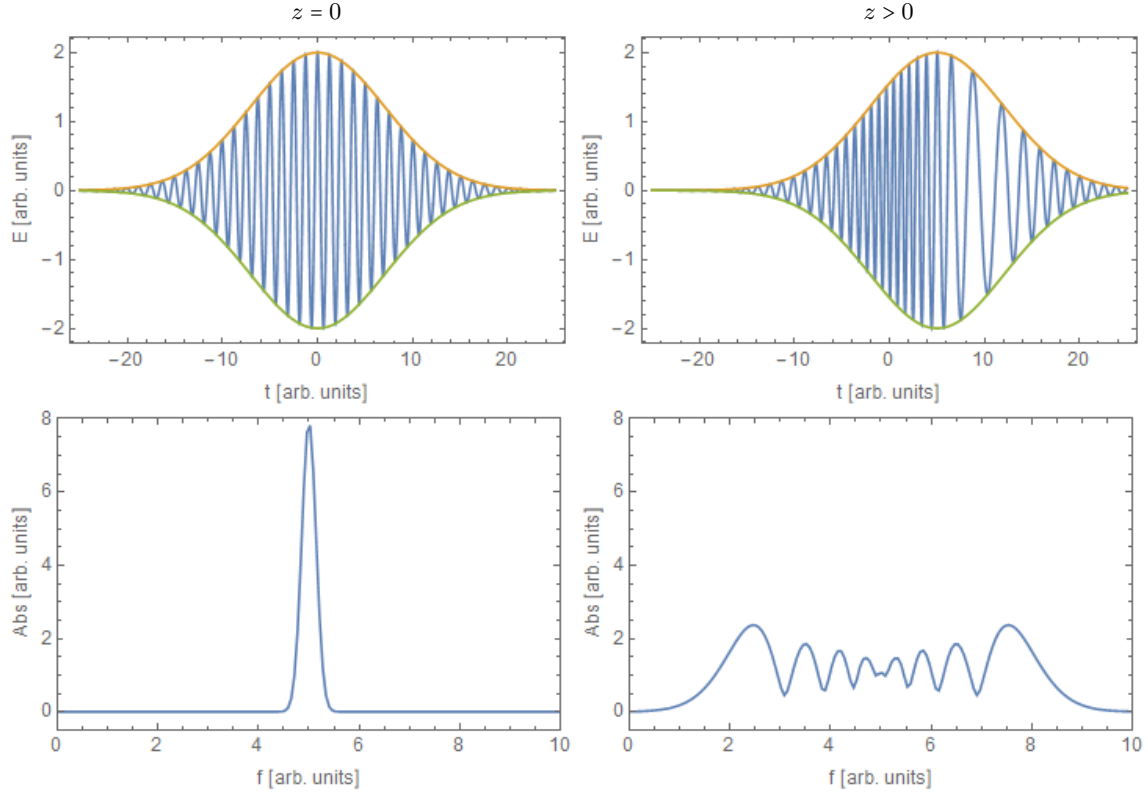


Fig. 6: Presentation of a wave packet and its frequency spectrum (absolute values) before (left) and after passing a non-linear medium (right). As described in the equations above, all dispersive effects except for the self-phase modulation have been neglected. The shift of the wave packet for higher values of z is due to the use of the retarded time (i.e. the finite speed of light) and is not an effect of the self-phase modulation.

Compared to that, a distinct change in the frequency spectrum can be observed. While the frequency distribution could be described by a Gaussian pulse for the starting condition, the wave packet for higher values of z contains components of further distant frequencies. These frequencies are symmetrical around the main carrier frequency, which is a direct consequence of the symmetry of the starting pulse.

By understanding the former calculations, the working mechanisms of the supercontinuum laser can be understood. At first a pulsed Nd:YAG-laser emits an electromagnetic wave with a carrier frequency of 1064 nm, which in turn gets focussed into an optical fiber. This optical fiber consists of an arrangement of different materials, in which the light can be guided via the means of total reflection (for a certain wavelength range). The most simple realization of such a fiber uses a cylindrical material core, which is surrounded by a material with a lower refractive index. The special feature of the fiber used in this experiment is its optical nonlinearity. Due to this, the aforementioned effects of self-phase modulation occur, which drastically widens the spectrum of the previously strongly localized wave pulse. At the exit of the fiber, broadbanded light gets emitted, which can be filtered in further parts of the setup. The described parts of the supercontinuum laser are presented again in fig. 7.

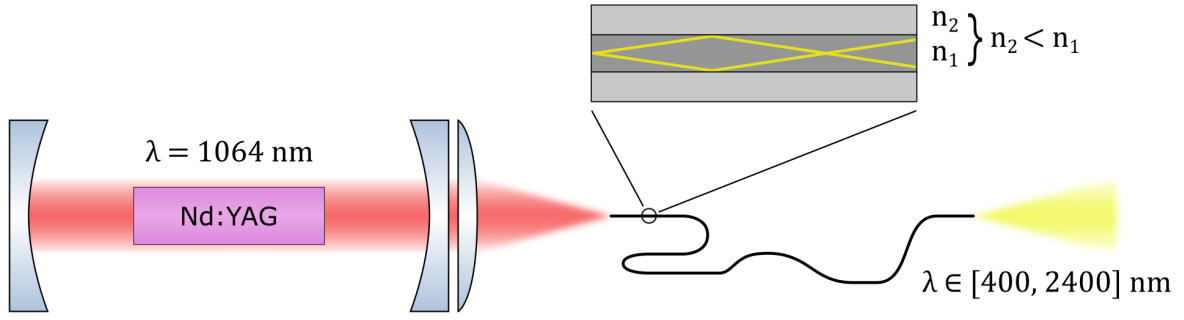


Fig. 7: Sketch of the parts of the supercontinuum laser. Beginning on the left side, a pulsed Nd:YAG-laser emits light with a frequency of 1064 nm. This light is being coupled (internally) into an optical fiber, in which the light is directed via total reflection. The optical fiber consists of optically non-linear materials, which lead to self-phase modulation. Because of that, the spectrum of the light leaving the fiber is noticeably widened.

In fig. 8 a typical frequency spectrum of the laser model used in this experiment is being shown. In contrast to the idealized self-phase modulation described above, other dispersive effects of the optical fiber are noticeable in the real spectrum. These dispersive elements lead to a spatially dependent change in the envelope, which in turn has effects on the self-phase modulation, as the form of the envelope is directly influencing the non-linear part of the refractive index.

This dependency of the self-phase modulation is also the reason why a change in intensity of the laser system cannot be achieved by changing the strength of the electric field emitted by the Nd:YAG-laser, as this would influence the strength of the second order non-linear effects in the fiber, which are responsible for the self-phase modulation. To allow for a comparability of the frequency spectrum of different output intensities, only the repetition rate (i.e. the number of pulses per timeframe) gets changed. The form of the single pulses as well as their intensity remain unchanged.

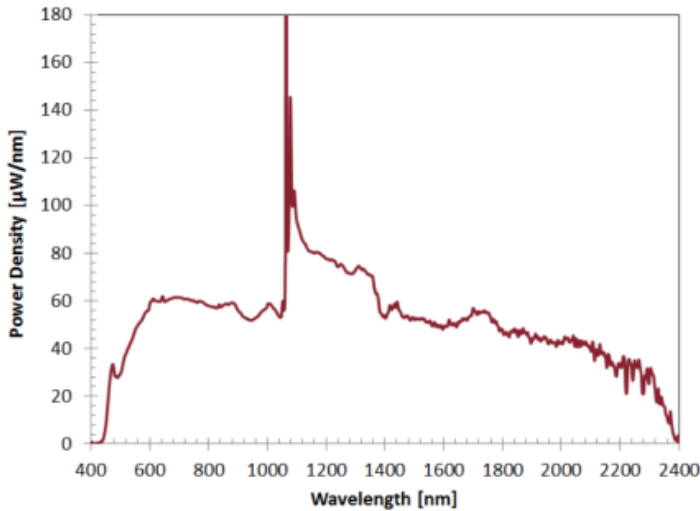


Fig. 8: Typical frequency spectrum for a laser of the model used in this experiment. The strongly pronounced part of the frequency spectrum at 1064 nm is a direct consequence of the Nd:YAG-lasers output at 1064 nm, while the other changes compared to the idealized self-phase modulation are the result of other dispersive effects in the fiber.

Acousto optic modulators

Since narrowbanded wave pulses should be used to analyze the properties of the semiconductor materials used in this experiment, the broadband spectrum of the laser has to be filtered. As the filtered wavelength should be adjustable during the experiment, no usual bandpass filters can be used. Instead several so called acousto optic modulators (AOMs) will be used. These operate according to the principle of Bragg reflection and can change their transmitted wavelength dynamically for a given wavelength range. The principle of an acousto optic modulator is being presented in fig. 9.

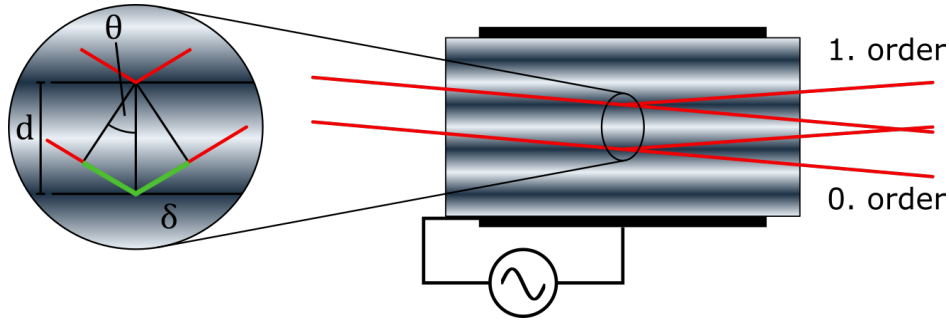


Fig. 9: Principle of an acousto optic modulator. An ultrasonic wave generator (AC-symbol) creates a density variation in the crystal, which acts on the incoming light like a lattice. While the Bragg condition is fulfilled (see left side and text), refractive effects occur, from which the zeroth and first order is being presented in the sketch.

The fundamental idea behind the acousto optic modulators is based on Braggs law: if a beam of light hits a lattice according to fig. 9 (left side), an optical path difference (green) occurs between the reflection at the upper part of the lattice and the lower one. If the path difference is a multiple of the wavelength, constructive interference arises, and the beam gets diffracted. The condition for constructive interference, sometimes called Bragg condition, is:

$$2d \sin(\theta) = \delta = m \frac{\lambda}{n},$$

where d represents the lattice constant, θ the angle of incidence, δ the difference in path lengths and λ/n the wavelength in the material.

If light with a broad frequency spectrum falls onto the lattice, the Bragg condition for a given angle is fulfilled only for one wavelength. Accordingly, for a fixed angle, light of a specific wavelength can be filtered.

In the material of the acousto optic modulators, which is a quartz crystal most of the time, the lattice gets created by inducing ultrasonic waves into it, which are created by a piezo electric transducer in combination with an AC-source. The piezo electric transducer is based on a piezo electric crystal, i.e. a crystal which changes its size, while being under the influence of an electric field. By applying an AC-source to this crystal an oscillation can be induced into the quartz. The induced ultrasonic waves lead to a density fluctuation in the material, which in turn leads to the desired change in the refractive indices.

These variations of the refractive indices act on the incoming light similar to a lattice and can lead to the above mentioned Bragg diffraction. A special feature of the acousto optic modulators in comparison to simple bandpass filters is the fact that they can change the

parameters of their Bragg condition to a certain degree by changing the frequency of the ultrasonic waves and thus changing the lattice constant d . This way the filtered wavelength can be changed.

In the real case this filtering has its experimental limits, which is why not only one acousto optic modulator is being used, but instead three. These AOMs cover different wavelength ranges and account for a range from 450 nm to 2000 nm together.

Beam path

To be able to focus the narrowbanded beam leaving the acousto optic modulators onto the sample, the beam has to be guided by optical components. Since the supercontinuum laser can be directly connected to the acousto optic modulators by its optical fiber, this part does not have to be build manually.

For the beam path from the outputs of the AOMs up to the sample, plane silver mirrors are used, which can be finely adjusted through the use of optomechanical holders from the company Thorlabs. In the experiment an arbitrary number of mirrors can be used, however you should try to use as few as possible, due to the limited size of the breadboard.

In fig. 10 a possible setup for a transmission measurement is shown. Here the three possible beam lines are parallelized first and then focussed onto the detector by a parabolic mirror. Two of the mirrors are so called flip mirrors, which contain a spring mechanism, allowing them to be flipped to the side, to make room for another beam path.

The intensity of the transmitted beam is then measured by a photodetector and being read by a sourcemeter. The detector is either a Si-detector or an InGaAs-detector, depending on the sample.

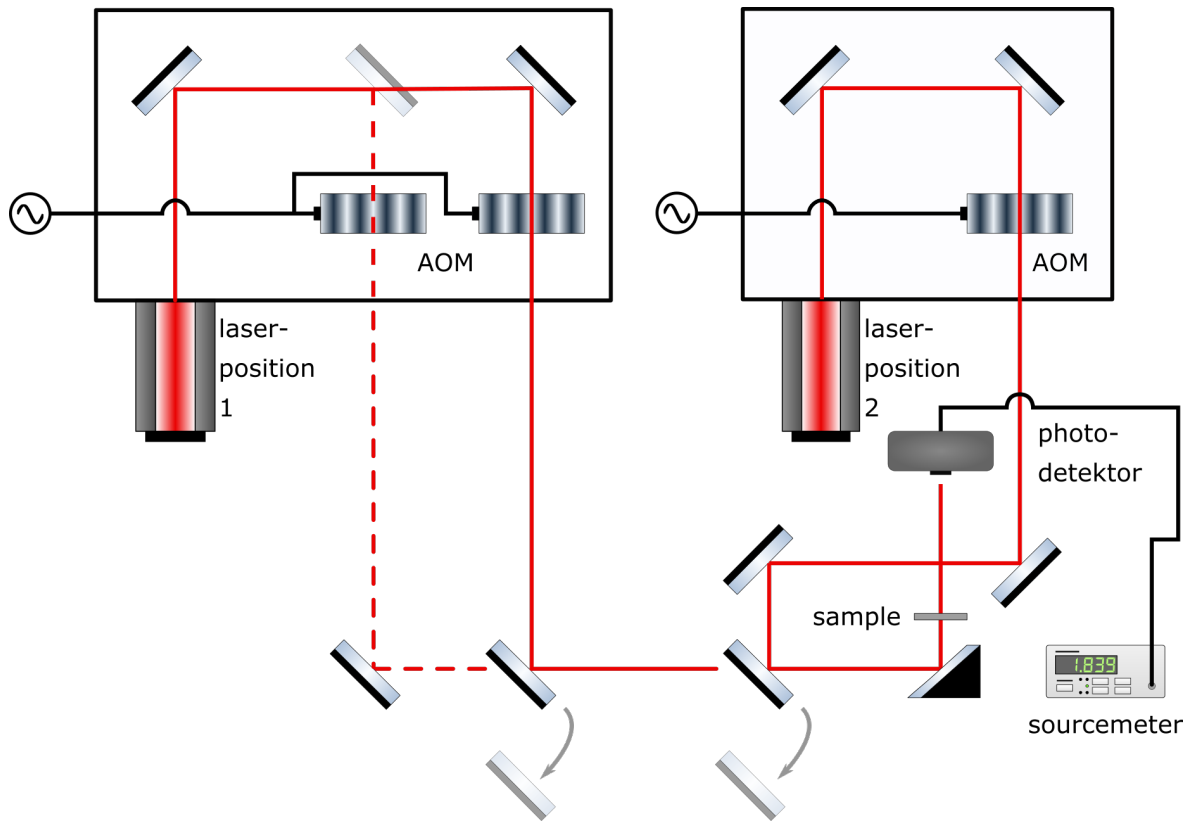


Fig. 10: Sketch of a setup for a transmission measurement containing one supercontinuum laser and several AOMs. The supercontinuum laser can be connected to one of the two AOM-housings, to activate one of the highlighted beam paths. The two rightmost mirrors in the lowest part of the beampaths are flip mirrors, which can be flipped to the side to make room for the beam paths behind them. The last mirror in front of the sample is a parabolic mirror, which focusses the beam onto the photodetector. Afterwards the transmitted light intensity is detected by a combination of a photodetector and a sourcemeter.