Handbook
for
Generic Photonic IC Design

Editors: Meint Smit and Xaveer Leijtens

4-4-2026

Chapter 14

Semiconductor Optical Amplifiers

Erwin Bente

The semiconductor optical amplifier (SOA) is a building block that can be used in many ways in photonic integrated circuits. First of all as an amplifier for all sorts of optical signals: continuous single frequency light, pulses and/or phase modulated light, and light signals with many wavelengths in parallel. It is a key component in semiconductor lasers. Depending on the application the SOA might need to produce high or low average output power, or high or low peak power levels. An SOA can however also change the shape of an optical pulse and the spectral content of the signal. Such distortions are mostly unwanted for pure amplification of signals. However the nonlinear response and phase effects of the SOA can also be utilized for the manipulation of signals. The SOA can be used as a component that controls the phase of the optical signals, as an optical gate, or a wavelength converter even. In this chapter the quantitative aspects of amplification and other effects on the optical signals of the SOA are discussed.

semiconductor optical amplifier

14.1 Introduction

This chapter starts with a discussion of the amplification of continuous wave light in a steady state situation, first with low intensity where amplification is linear, followed by discussion of high intensity where saturation effects kick in. Then the amplification of fast optical signals will be discussed. Fast signals means signals that vary in intensity in a way that the behaviour of the amplifier cannot be described as a steady state. The relevant time scale for a number of different time dependent signals will be discussed.

Amplifiers fundamentally add noise to the signal they are amplifying. This is due to the light that is generated by the spontaneous emission (SE) process which always takes place if an SOA is forward biased. The optical gain and spontaneous emission processes are linked. The spontaneous emission generated light will also be amplified in case the SOA has optical gain. This is named amplified spontaneous emission (ASE) and again this affects, i.e. perturbs the amplified signal quality, however the generated ASE can also be used in some applications. The effects of ASE will be discussed at each point where it is considered relevant.

spontaneous emission

amplified spontaneous emission

laser An important utilization of an amplifier is inside a laser cavity. In integrated optical circuits this will often be an extended cavity laser system in which the amplifier, passive waveguides, reflectors and other components are combined. Characteristic for SOAs in laser cavities is the high sensitivity of the behaviour of the laser to the properties of the SOA. The use of the SOA in lasers will be discussed in the chapter on lasers (Chapter 29). In the present chapter the focus is on the amplification of signals.

The properties of an SOA are temperature dependent. Since energy dissipation in photonic integrated circuits originates mainly from the SOAs in it, the operating conditions of the SOAs in the circuit in combination with the environment, determine its temperature. These effects need to be taken into account, or at least need to have been considered in the design of the PIC. The temperature effects will be discussed at a number of relevant points in this chapter.

multi-quantum well In this chapter a number of assumptions about the SOA are made. The first is that the SOA is assumed to be a multi-quantum well (MQW) based amplifier. Such an SOA amplifies mainly light in the TE polarized modes, i.e. the polarization is in the plane of the quantum wells. The gain for TM-polarized light is lower (for unstrained quantum wells) [187], [188]. The second assumption is that the SOAs discussed here amplify only one single transverse mode. The third assumption is that the fastest pulse signals are about 1 ps.

The examples shown in this chapter to illustrate different aspects are based on parameters for an amplifier on InP with four InGaAs/InGaAsP quantum wells.

The semiconductor optical amplifier is a rich and much studied subject. Many different effects can play an important role in some applications and may be completely irrelevant in others. There are many text books in this field that contain a lot more detailed information than can be described in here. We will refer to those texts for more detailed information. This chapter tries to present an overview of the most relevant aspects of SOAs in photonic integrated circuits and simplified guidelines to start your design using phenomenological, more compact models. But it should be stressed that particular choices for a description have been made and that other, more complete modelling techniques exist. Those techniques might be required for the particular application the reader is interested in.

14.2 Amplifier structure

In Figure 14.1 the structure of a ridge waveguide optical amplifier as used in the generic integration process on an n-doped InP substrate is presented.

quantum well Starting from the lower end in Figure 14.1a there is the bottom metal contact to the highly n-doped substrate (dark blue). Then follows the InP growth layer that is lower n-doped (light blue). On top of that is the quaternary waveguiding layer (green) with a higher index. In the middle of the quaternary waveguide stack is a set of quantum wells separated by barrier layers (indicated as one red layer). The quantum well material has the lowest bandgap in the stack and the interaction with the light takes place in this material. In the SOA design described in this chapter, there are four quantum wells in this region, each 7 nm thick. On top of the waveguiding layer there are p-doped InP layers (blue) with increasing p-doping concentration up to the contacting layer (red) with the metal contact on top. In Figure 14.1b a SEM picture is presented of a realization of such an SOA. The waveguiding layer, the contacting layer and metallization on top clearly stand out. In the middle of the waveguiding layer you can just make out the area with the quantum layers.

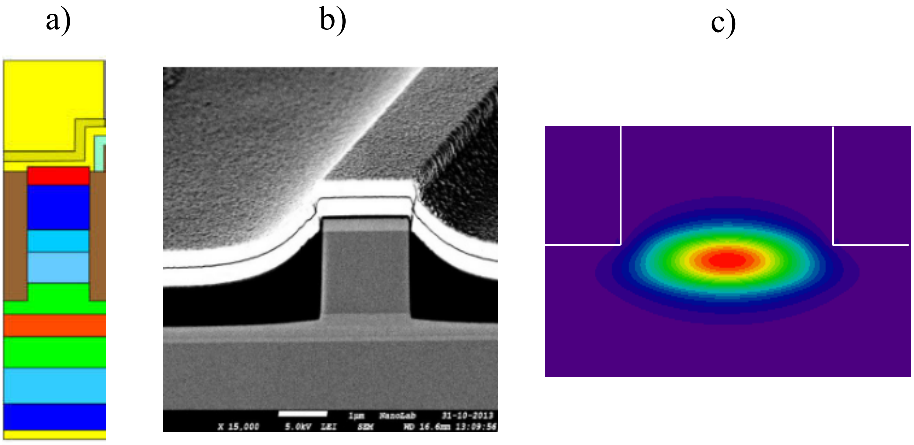


Figure 14.1: a) Schematic cross section of the SOA on an n-doped substrate; b) SEM picture of an SOA cross-section; c) a calculated intensity distribution of the fundamental optical mode in 2 µm wide SOA ridge waveguide using a linear color coded scale. The edges of the 2 µm wide waveguide ridge are indicated in white and show the scale of the plot.

For operation of the SOA the p-side (top) is connected to the plus of a current source and the n-side (bottom) to the minus, and thus the SOA PIN structure is forward biased. Excess electrons are injected from the n-doped layer into the waveguide layers and are trapped in the quantum wells. Similarly excess holes are injected from the p-side into the waveguiding layer and also trapped in the quantum wells. The excess electrons and holes thus collect in the quantum wells and can recombine due to stimulated emission to provide optical gain, or due to spontaneous recombination to provide spontaneous emission of light, or recombine through other processes which convert the energy of an electron-hole pair into heat.

excess electrons
excess holes
stimulated emission
spontaneous emission

As described in chapter 2, the double heterostructure also provides confinement of the light in the vertical direction and together with the ridge structure forms a waveguide. The intensity distribution of the optical mode in the ridge waveguide of the SOA is drawn in Figure 14.1c). You can see from this that the overlap of the optical mode with the quantum wells is relatively small, for this specific amplifier the confinement factor is approximately 5.5%. As mentioned above, the light is amplified by the material in the quantum wells. In these wells the injected electrons and holes are collected and the concentration of these excess carriers can be sufficiently high to provide the optical gain. Note that the quantum well as well as the barrier and waveguiding layers are typically undoped. The undoped materials are sometimes described as non-intentionally doped due to technical limitations in the production of the materials. Doping, in particular p-doping, gives rise to significant optical losses. E.g. for p-doping in InGaAsP quaternary material a doping level of 10^{18} cm^{-3} gives rise to optical losses in the order of approximately 23 cm^{-1} (i.e. about 100 dB/cm) at 1.55 µm [189]. The quantum well is therefore composed of intrinsic material and this means that the thermal equilibrium carrier concentration in the wells is negligible with respect to the injected excess carriers. The carrier concentration in the gain material can thus be described by purely the non-thermal excess carrier concentration. The material is considered to be electrically neutral. This means that the concentration of excess holes δp equals the concentration of excess electrons δn. Therefore when describing the SOA one typically only talks about the carrier concentration $N = \delta p = \delta n$ in the gain material.

double heterostructure
confinement factor
optical gain
non-intentionally doped
optical losses
excess carrier concentration

Problem 14.1: Contact Resistance.

Problem: The contact resistance for a Ti-Pt-Au contact on Zn-doped ($7\times10^{18}\text{cm}^{-3}$) InGaAs is reported to be $2\times10^{--5}\Omega\text{cm}^2$ [190]. For comparison, in the layer stack in the generic process flow (see Fig. 4.11a the Zn doping level of the p-doped InGaAs contacting layer is $1.5\times10^{19}\text{cm}^{-3}$. Calculate a) the resistance of a 500-µm-long SOA and b) of a 50-µm-long SOA for a specific contact resistance value of 2×10^{-5} Ωcm^2 . c) Calculate the dissipated power per unit length for a current density $J = 4$ kA/cm^2.

Solution: a) The surface area of the contact $A_\text{c} = w_\text{ridge} \cdot L_\text{SOA} = 2 \cdot 500 = 1000\,\mu\text{m}^2$. So the resistance $R_{500} = 2\times10^{-5}/1000\cdot10^{-8} = 2\Omega$, where the factor 10^{-8} in the denominator accounts for the conversion from µm to cm.
b) For the 50 µm long SOA $R_{50} = 2\times10^{-5}/100\cdot10^{-8} = 20\Omega$.
c) The dissipated power per mm $P = I^2 \cdot R_{1000} = (40\cdot0.002\cdot1)^2\cdot1 = 6.4\,\text{mW/mm}$ in which I is the current through a 1-mm-long SOA and the current density and the contact dimensions are expressed in kA/mm^2 and mm, respectively. When the same current density is used for both the short and long SOA the heating effect will be similar since the dissipated power per unit length is the same.

contact resistance

Let us look at the flow of the excess carriers in forward bias through the structure in a bit more detail. If we start at the top the current first flows through the metal/semiconductor contacts. These have resistance that is characterized by the specific contact resistance R_c which is typically given in units of Ωcm^2. The total resistance of the contact is then given by:

$$R_\text{contact} = \frac{R_\text{c}}{A_\text{c}},$$

where A_c is the total surface area of the contact on the semiconductor, i.e. the length of the SOA times the ridge width. The heat dissipated in this resistance can be an important part of the heating of the SOA, especially for very short SOA sections, because short SOAs tends to be operated at higher current densities to achieve the required gain. This increases the power dissipation quadratically.

current spreading

With a highly n-doped substrate the contact resistance on the bottom InP-contact is less of an issue for two reasons. The first is that typically a somewhat lower contact resistance value is achieved [191]. The second reason is that there is current spreading in the 100 µm or more thick substrate layer. The surface area through which the current will go on the substrate side is effectively at least ten times larger because of this. So the total resistance from the n contact is significantly lower than that of the p-contact.

semi-insulating substrate

When a semi-insulating substrate is used there will be a side contact on a relatively thin highly n-doped layer on top of the insulating substrate. In that case the resistance of the n-doped layer and the contact resistance on that n-doped layer may be an issue again. However the n-contact area can be made several tens of microns wide so resistive heating will be less and also the heat is generated at some distance from the SOA ridge. The n-contacts will typically be part of the SOA building block and will have been designed with the electrical resistance in mind.

The carriers that pass the doped and undoped InP and InGaAsP layers will get caught in the quantum wells where a high concentration can build up. Not all carriers will end up in the well. A fraction will overshoot the well and will recombine elsewhere not contributing to the gain. Electrons that overshoot the quantum wells will recombine in the p-doped layers and holes that overshoot will recombine in the n-doped layers. The higher the concentration in the quantum well, the more overshoot will occur. This is due to the fact that at high excess concentrations in the wells there are less empty states in the quantum well available.

overshoot

Since in the shallowly etched ridge waveguide amplifier the carriers can diffuse sideways outside the ridge driven by the concentration difference, there is a loss of excess carriers available for interaction with the optical mode, i.e. gain. This means a reduction of the maximum achievable efficiency with this type of amplifier. This can be improved by using a deeply etched ridge through the quantum well layers. The technical issue is then the passivation of the sidewalls to prevent or reduce carrier recombination at those points.

passivation

The sidewall recombination issue can be solved using a buried heterostructure [187] where the sides of the ridges are filled in with semiconductor material with a doping profile that creates a reverse biased pn-structure outside the ridge to block current not going into the original ridge. The two main advantages of the buried heterostructure are firstly the reduced loss of carriers and secondly the fact that heat can flow from the contact and active region in all directions. The main disadvantage is the reduced index contrast in the horizontal direction which tends to increase the mode size and a larger overlap with doped semiconductor material. Despite these it is possible to use buried heterostructure SOAs in a monolithic integration scheme and benefit from the advantages, see e.g. [192].

buried heterostructure

In the following sections we first present a description of the SOA and look at the steady state.

14.3 SOA description

Since we consider the light field to be a single transverse mode of the waveguide, we can describe the power in the mode by a single average intensity and a single value for the optical phase. It is most common in the description of the light in the amplifier to use an approximation where the intensity of the light and the carrier density are considered to be uniform over an effective mode size. So the distribution of the mode is described as a rectangle of uniform intensity. It is important to look at this in a bit more detail. In literature one can find a commonly used definition of the effective mode area S_{eff} [193].

However it is practical to use the same value as the waveguide ridge width w_{r} as the width of the mode. This is the choice made in this chapter. It is important to use the value of the width consistently. The surface area of the active gain material is the known total thickness of the quantum wells times the width which is assumed to be w_{r} . The effective surface area of the mode can then be calculated using the confinement factor for the quantum wells, which in turn can be calculated given the waveguide design. Therefore, $S_{\text{mode}} = (w_{\text{r}} h_{\text{qw}})/(\Gamma)$, in which h_{qw} is the total height of the quantum wells and Γ is the confinement factor of the transverse mode for the active material.

These one-dimensional approximations have limitations but appear to describe the behaviour of an SOA sufficiently accurate in many cases.

The propagation of light at wavelength λ with intensity $I(z)$ through the amplifier in the z direction along the waveguide in a steady state, can be described by the following differential equation:

$$\frac{\partial I(z)}{\partial z} = \frac{1}{1+\varepsilon_1 I(z)} g_{\text{mod}}(N(z),\lambda) \cdot I(z) + I_{\text{se}}(N(z),\lambda) \tag{14.1}$$

modal gain

spectral hole burning

In this equation $g_{\text{mod}}(N(z),\lambda)$ is the modal gain of the amplifier in units of 1/length (typically 1/cm is used). $I(z)$ is the intensity at position z in units of power over surface area, e.g. W/cm^2. Note that this is an average intensity over the cross-section of the transverse intensity distribution of the mode and thus an approximation in the description. The modal gain, that is the gain factor for the optical mode propagating in the SOA is dependent on the average local carrier density in the quantum wells $N(z)$, and on the wavelength λ. The modal gain is reduced by the term $1/(1+\varepsilon_1 I(z))$. This term describes the effect of spectral hole burning: a reduction in the gain due to the depletion of carriers by the light in the gain material at the required energy levels within the bands. The parameter ε_1 has a positive sign and describes the relation between the light intensity and the gain reduction. In this section of the chapter we can neglect this term for low intensities. The last term $I_{\text{se}}(N(z),\lambda)$ describes the average local contribution to the intensity of the spontaneous emission in the wavelength region of the signal. Please note that the linewidth of the input signal is assumed to be small (much smaller than the gain bandwidth) so it does not contain signals faster than 1 ps. This will be discussed in more detail in the time dependent signal amplification.

carrier density

The local carrier density $N(z)$ in equilibrium can be calculated using an equation derived from the time dependent differential equation for the carrier density at position z:

$$\frac{\partial N}{\partial t} = \frac{J\cdot\eta_{\text{i}}}{e\cdot h_{\text{qw}}} - \frac{1}{1+\varepsilon_1 I} g_{\text{mat}}(N,\lambda)\cdot I \cdot \frac{1}{E_{\text{ph}}\cdot v_{\text{g}}} - A\cdot N - B\cdot N^2 - C\cdot N^3 - D\cdot N^{5.5} \tag{14.2}$$

injection efficiency

current density

In this equation the first term $(J\cdot\eta_{\text{i}})/(e\cdot h_{\text{qw}})$ represents the injection of carriers at current density J, e is the elementary charge, η_{i} is the injection efficiency, h_{qw} is the total thickness of the quantum wells, $E_{\text{ph}} = h\nu = hc/\lambda$ is the photon energy and $v_{\text{g}} = c/n_{\text{g}}$ is the group velocity. Note that $I/(E_{\text{ph}}\cdot v_{\text{g}})$ is the photon density. The injection efficiency η_i takes into account that not all injected carriers end up in the volume of the optical mode in the active material. The value of η_i can be difficult to determine accurately for real devices. The current density will typically be uniform over the surface of the contact on top of the waveguide, so

$$J = \frac{I_{\text{inj}}}{w_{\text{r}}\cdot L_{\text{SOA}}} \tag{14.3}$$

modal gain

where I_{inj} is the total injection current, w_{r} is the width of the ridge and L_{SOA} is the length of the SOA. The second term in Eq. 14.2 is the reduction of carrier density due to optical gain of the mode. Since the carrier density is only present in the active gain material, the material gain is $g_{\text{mat}}(N,\lambda)$ in this term. The optical material gain inside a uniform active material is linked to the gain experienced by the optical mode, i.e. the modal gain $g_{\text{mod}}(N,\lambda)$ by:

$$g_{\text{mod}}(N,\lambda) = \Gamma\cdot g_{\text{mat}}(N,\lambda) - \alpha_s(N) \tag{14.4}$$

confinement factor

Here Γ is the confinement factor of the optical mode for all the active material, i.e. the quantum well layers. $\alpha_s(N)$ is the loss of the optical mode. This loss is due to scattering at the waveguide edges, overlap of the mode with other layers where free carriers can absorb light, and absorption due to the injected excess carriers in all layers. We will see that the latter is dominant and hence the dependency on N is included. The losses are in principle also wavelength dependent but we will assume them to be constant over the wavelength range with optical gain in the SOA.

The second term in Eq. 14.2 is followed by four more terms that represent other effects that reduce the excess carrier concentration in the quantum wells. Three of these terms represent losses due to non-radiative recombination processes. The energy that comes free in these recombination processes becomes heat and is thus lost. The term $A \cdot N$ represents the loss due to material defects and surfaces, the term $C \cdot N^3$ represents the three-particle Auger recombination processes. The third term $D \cdot N^{5.5}$ is less commonly used but it turns out to be important for describing realistic amplifiers. The term represents a carrier leakage process by carrier drift, i.e. carriers that do not end in the quantum well but overshoot the active region and recombine in the doped region and metal contacts [188] [194]. The term $B \cdot N^2$ describes the radiative recombination process of spontaneous emission. Since spontaneous emission is a process that involves two carriers recombining it is a quadratic term in N. The spontaneous emission term $I_{\mathrm{se}}(N,\lambda)$ in Eq. 14.1 and the $B \cdot N^2$ term in Eq. 14.2 are linked:

non-radiative recombination

Auger recombination

carrier leakage

overshoot

radiative recombination

spontaneous emission

$$I_{\mathrm{se}}(N,\lambda) = E_{\mathrm{ph}} \cdot v_{\mathrm{g}} \cdot \beta \cdot \Gamma \cdot B \cdot N^2 \tag{14.5}$$

Here Γ is the confinement factor of the optical mode for all the active material and β is the fraction of all spontaneously emitted photons that couple to the optical mode in the amplifier. It is often called the Petermann factor [195]. Typical values for β are 10^{-4} to 10^{-5}. Please note that at this point we do not consider what happens with light from spontaneous emission at other wavelengths than the light sent into the SOA. We will describe those effects in more detail in the time dependent description sections. There are several methods of describing the spontaneous emission and amplified spontaneous emission (ASE) [188], [196], [197].

confinement factor

Petermann factor

14.4 SOA in steady state

Now let us look at Eq. 14.2 for the steady state situation. The slowest process in the equation is the $A \cdot N$ term. This means that a steady state situation is reached after three to four times the carrier lifetime of $1/A$ s. The lifetime is typically in the order of a few hundred picoseconds to one nanosecond, so equilibrium tends to be achieved within a few nanoseconds and $\partial N/\partial t = 0$. Given that values A, B, C and D, the optical modal gain $g_{\mathrm{mod}}(N,\lambda)$ function and the injection current are known, one can set the equation 14.2 to zero. One then obtains a relation between the local light intensity and carrier concentration. So given the constant intensity $I(z)$ at position z and the current, one can calculate (numerically) the steady state carrier density $N_{\mathrm{eq}}(z, I(z))$. The intensity distribution $I(z)$ over the length of the amplifier, considering light in one (the $+z$) direction, can be calculated by integration over the length. If we then also neglect the spontaneous emission we get:

steady state

carrier lifetime

equilibrium

intensity distribution

$$I(z) = \int_0^z g_{\mathrm{mod}}(N_{\mathrm{eq}}(z', I(z')), \lambda) \cdot I(z') \mathrm{d}z' \tag{14.6}$$

When the intensity of the light is sufficiently low, the concentration of carriers is not affected much by the optical signals. The carrier concentration N_{eq} then becomes independent of the signal intensity at position z, assuming a uniform current injection.

The steady state carrier concentration under these conditions can be calculated from Eq. 14.2 by setting $\partial N/\partial t = 0$ and neglecting the carrier loss rate term due to the amplification. The resulting equation gives a direct relation between the injected current density and N_{eq}. Therefore, in this situation there is a direct link between the injected current density and the modal gain g_{mod}. This is the reason why the unsaturated modal gain of an SOA for continuous wave or slowly varying optical signals can be presented as a function of current density and wavelength.

The solution to equation 14.6 with g_{mod} being independent from $I(z)$ due to the low intensity and, consequently, N_{eq} being independent from $I(z')$ and z', is straightforward:

$$I(z) = I(0) \cdot \exp(g_{\mathrm{mod}}(N_{\mathrm{eq}}, \lambda) \cdot z) \tag{14.7}$$

The intensity of the signal grows exponentially if the gain is positive. Such a solution can be used if the intensity of the signal everywhere inside the SOA is sufficiently low such that the term $-1/(1+\varepsilon_1 I) g_{\mathrm{mat}}(N, \lambda) \cdot I$ in Eq. 14.2 stays small with respect to the other terms. We will discuss examples of the onset of saturation in section 14.8.

The value of the modal gain g_{mod} is typically given as a function of injection current density in documentation on SOAs e.g. in the foundry design manual. This can be in the form of a graph, but it can also be presented in the form of analytical expressions that describe the gain as a function of wavelength for a range of current densities. An example of such a set of expressions is discussed in Sec. 14.8. These modal gain values can be used in Eq. 14.7 for a useful estimate of the maximum gain possible from an amplifier. E.g. one can estimate the minimum SOA length required to achieve the threshold of lasing if the losses in the cavity are known.

There are a number of clear limitations of such an estimate using Eq. 14.7. The first is that saturation of the SOA is not taken into account. The estimate thus presents an upper limit to the amplification that can be achieved. The second is that spontaneous emission is not taken into account. Spontaneous emission light is generated and amplified in the SOA which gives rise to amplified spontaneous emission (ASE). At high currents and/or long SOAs the ASE can contain so much power that it starts saturating the amplifier itself. An issue related to the ASE is the fact that the input signal needs to have sufficient power with respect to the spontaneous emission. The input will be amplified together with the ASE generated in the SOA and the signal will not be detectable in the output. To improve this one can filter out spectrally much of the ASE.

spontaneous emission

Another issue that is often encountered particularly in integrated optics, is that neighboring components do reflect light. Reflections that are unintentional from e.g. MMIs or manufacturing imperfections are typically small (< −30 dB). However the circuit in which the SOA is used can also contain e.g. a reflector with a reflectivity that is much higher. E.g. one can have an SOA after a linear laser cavity, which means there will be a large reflection from the output coupler of the laser next to the SOA. If there is then another small reflection on the other side of the SOA, issues come up (also for the laser). Typically such reflections can lead to etalon effects since the gain in between the reflections amplifies the reflected power and it makes the transmission of the system wavelength dependent. Reflections around the SOA can also, and in practice often do, lead to lasing.

reflection

etalon effects

lasing

Please note that the mode of operation of the SOA well below saturation is needed (i.e. sufficiently low input power levels) for a linear amplification of the signal, and this will always lead to a relatively poor energy efficiency. The poor efficiency is fundamental, since the other loss terms in Eq. 14.2 describing the non-radiative losses will be large compared to the carrier loss term related to optical gain. In the unsaturated gain mode the light should not noticeably affect the carrier concentration.

14.5 ASE from an SOA in the steady state

Next we have a look at the ASE from the SOA if no signal is sent into the SOA. The ASE is at this point described as a wavelength dependent CW signal $I_{\mathrm{ASE}}(\lambda)$. One can integrate a simplified version of equation 14.1 given below (14.8) over the length L of the SOA assuming a constant modal gain (i.e. no saturation) and with the carrier density at its equilibrium value. The input intensity at the entrance of the SOA is set to 0.

$$\frac{\partial I(z, N_{\mathrm{eq}})}{\partial z} = g_{\mathrm{mod}}(N_{\mathrm{eq}}, \lambda) \cdot I(z) + I_{\mathrm{se}}(N_{\mathrm{eq}}, \lambda) \tag{14.8}$$

The resulting expression for the ASE intensity at the output of the SOA is [198]:

$$I_{\mathrm{ASE}}(\lambda, L) = \frac{I_{\mathrm{se}}(N_{\mathrm{eq}}, \lambda}{g_{\mathrm{mod}}(N_{\mathrm{eq}}, \lambda)} \left(e^{g_{\mathrm{mod}}(N_{\mathrm{eq}}, \lambda) L} - 1 \right) \tag{14.9}$$

ASE spectrum

This relation enables the measurement of the modal gain and the spontaneous emission intensity as a function of wavelength and injection current density. One can measure the ASE spectrum from a series of amplifiers at a specific injection current density at different lengths and fit the modal gain $g_{\mathrm{mod}}(N_{\mathrm{eq}}, \lambda)$ and spontaneous emission intensity $I_{\mathrm{se}}(N_{\mathrm{eq}}, \lambda)$ for each wavelength. The remaining issue is then the link between the practical parameter of the injection current density J and the carrier density $N_{\mathrm{eq}}(J)$. However one can just use the modal gain $g_{\mathrm{mod}}(J, \lambda)$ as a function of current density instead, as is commonly done since this is more practical.

The results of a series of measurements [199] of the modal gain $g_{\mathrm{mod}}(J, \lambda)$ as a function of current density J and wavelength λ of four quantum wells in 2 µm wide ridge waveguide amplifiers on InP are presented in Figure 14.2.

14.6 SOA in steady state – modal gain spectrum analysis and compact model

14.6.1 Description for all wavelengths

unsaturated modal gain

The unsaturated modal gain of the SOA in equilibrium as a function of injection current and wavelength data can be described using parametrized analytical expressions derived from theoretical modelling. This means one can have a simple expression for g_{mod} as a function of carrier density that can be used in the modelling of amplifiers and lasers. But also, the parameters in those expressions give insight into the performance of the amplifier and allow for further analysis to establish a link between current density and carrier density.

material gain

parabolic bands

Here we will utilize the parametrization of the description of the material gain $g_{\mathrm{mat}}(N, f)$ as a function of excess carrier density N and optical frequency f from [200] which is derived for quantum well material, assuming parabolic bands and for a temperature of 0 Kelvin (essentially ignoring temperature effects):

$$g_{\mathrm{mat}}(N, f) = \chi \left[\tan^{-1}\left(\frac{f - f_0}{\gamma} \right) - 2\tan^{-1}\left(\frac{f - f_0}{\gamma}) - \frac{N}{N_{\mathrm{t}}} \right) - \frac{\pi}{2} \right] \tag{14.10}$$

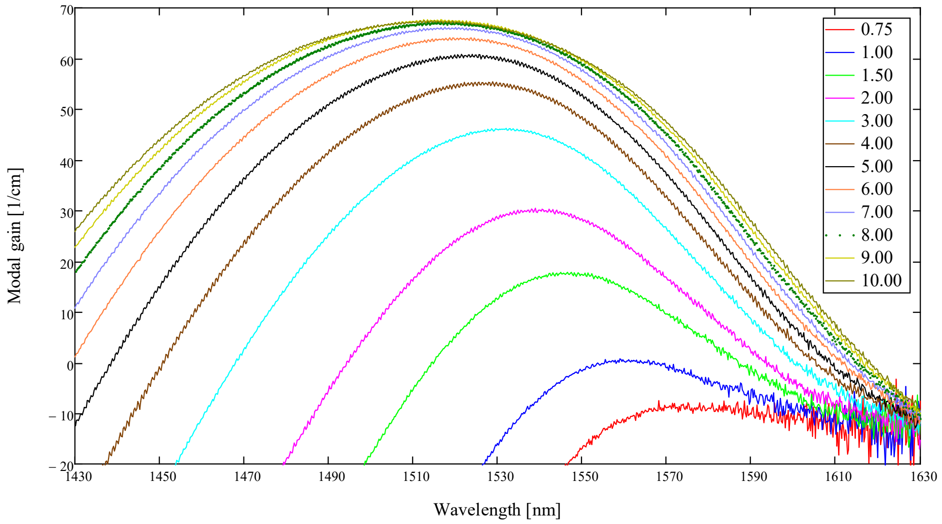


Figure 14.2: Measured unsaturated modal gain of a four quantum well SOA in units of cm^{-1} as a function of wavelength and for a range of injection current densities. The temperature of the bottom of the SOA substrate was 18 °C. The current density values are indicated in the legend in units of kA/cm^2. Measurement data made available by D. Pustakhod [199].

homogeneous linewidth
transparency carrier density
scaled carrier density

Here, χ is a scaling factor, f_0 is derived from the energy E_g of the bandgap: $f_0 = E_g/h$, γ is the homogeneous linewidth (in frequency units) of the states in the conduction band, and N_t is the transparency carrier density at the bandgap optical frequency. One can use Eq. 14.10 in combination with Eq. 14.4 to fit the function $g_{mod}(N,\lambda)$ to the measured data $g_{mod}(J,\lambda)$ using as fit parameters the scaling factor χ, the bandgap frequency f_0, the linewidth γ, the scaled carrier density N/N_t , and the scattering loss α_s. To illustrate what such a parametrization can do, we discuss an example using the gain spectra in Fig. 14.2. In Fig. 14.3 below an example of a fit result is shown in which the $g_{mod}(N,\lambda)$ function is fitted to the modal gain curve from Fig. 14.2 at $J = 6\,\text{kA/cm}^2$. One can see that the approximation works well for a large part of the gain spectrum, but the gain at shorter wavelengths is underestimated. That is why the fit was always done for only the longer wavelength part of the spectrum. This is indicated by the red dashed line. The discrepancy at the short wavelength side of the spectrum is due to the gain originating from the light holes which are in a band with a larger bandgap.

It is possible to fit the modal gain data in Fig. 14.2 for all values of the current density in a similar way using only a single value for the scaling factor $\chi = 950\,\text{cm}^{-1}$ and the homogeneous linewidth fixed at $\gamma = 5\cdot10^{12}\,\text{Hz}$. The other parameters f_0, α_s, and N/N_t had to be fitted separately for each current density. The results for these parameters are presented in Fig. 14.4. In this way, as can be seen in Fig. 14.4a, one can obtain the relation between the scaled carrier density N/N_t and the current density J. It shows that at the higher current densities the carrier density does not increase as much.

bandgap shrinkage
modal loss
free carrier absorption

One can then use this relationship between current density and carrier concentration to plot the bandgap frequency (Fig. 14.4b) and the loss (Fig. 14.4c) as a function of the scaled carrier density. This shows that the bandgap shrinkage and the modal loss due to carrier injection are close to linearly dependent on the carrier density. Notice that the main optical loss mechanism in the SOA is absorption by the free carriers (which will be mainly the holes) that are injected into the SOA. This loss is significant.

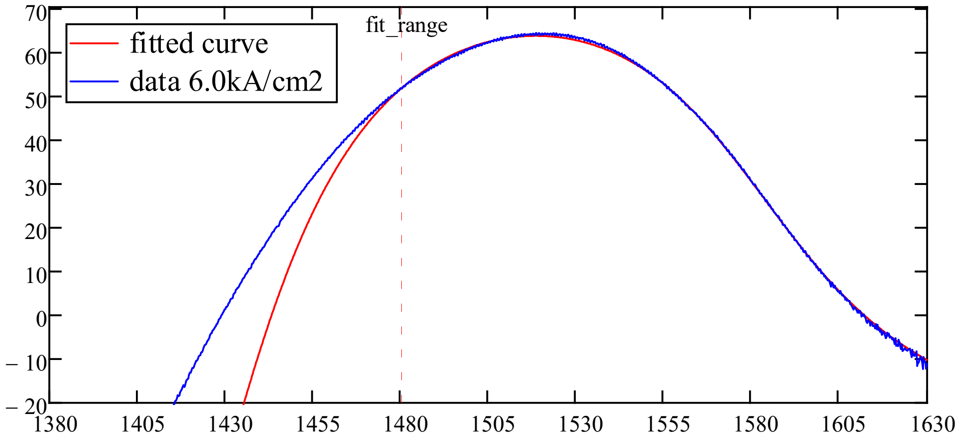


Figure 14.3: Example of fit of Eq. 14.10 (red) to the measured modal gain data (blue) of the four quantum well SOA at 6 kA/cm^2 current injection.

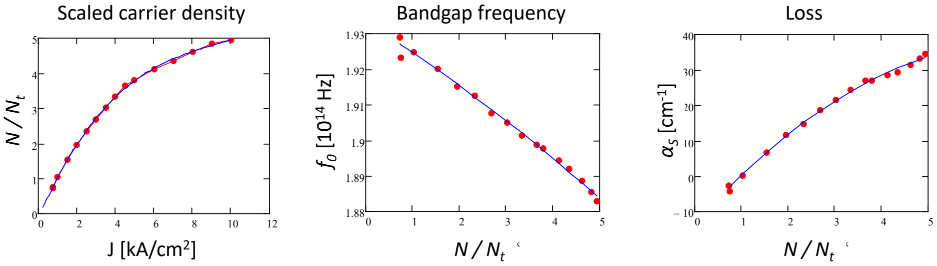


Figure 14.4: Fitted parameter results of Eq. 14.10 to the measured unsaturated gain values. The scaled carrier density N/N_{t} as a function of injection current density J, the bandgap frequency f_0 in units of 10^{14} Hz as a function of scaled carrier density, the loss α_s in units of cm^{-1} as a function scaled carrier density. The dots are the fit results for the parameters, the blue line is a polynomial fit.

Problem 14.2: Unsaturated gain.

Problem: Consider an SOA in generic technology that is 800 µm long, and that is operated at three different current levels: 33 mA, 49mA and 145mA. The injection efficiency is $\eta_i = 0.63$ and the wavelength of the input signal $\lambda = 1530\,\text{nm}$.

a) What is the unsaturated amplification of the amplifier at the three current levels (you can use measured values of the modal gain from Fig. 14.2, or the expressions given in Table 14.1).

b) What is the the expected output power of the amplifier with 0.5mW input at 1530nm?

Solution: $L_{\text{SOA}} = 800\,\mu\text{m}$, $w_{\text{ridge}} = 2\,\mu\text{m}$, $I_1 = 33\,\text{mA}$, $I_2 = 49\,\text{mA}$, $I_3 = 145\,\text{mA}$,

a) $J_1 = I_1/(L_{\text{SOA}} \cdot w_{\text{ridge}}) = 2.1\ \text{kA/cm}^2$, $J_2 = 3.1\ \text{kA/cm}^2$, $J_3 = 9.1\ \text{kA/cm}^2$.

From Fig. 14.2: $g_1 = 30\,\text{cm}^{-1}$, $g_2 = 46\,\text{cm}^{-1}$, $g_3 = 67\,\text{cm}^{-1}$.

The amplification is then:

$A_1 = \exp(g_1 \cdot L_{\text{SOA}}) = 11.0\ \text{cm}^{\text{-1}}$, $A_2 = 39.6\ \text{cm}^{\text{-1}}$, $A_3 = 212\ \text{cm}^{\text{-1}}$.

b) P_{in}=0.5 mW. For the unsaturated output power we find $P_{1,\text{out}} = A_1 \cdot P_{\text{in}} = 5.51$ mW, $P_{2,\text{out}} = A_2 \cdot P_{\text{in}} = 19.8$ mW, $P_{3,\text{out}} = A_3 \cdot P_{\text{in}} = 106$ mW.

The photon energy $E_{\text{ph}} = hc/\lambda = 1.298 \cdot 10^{-19}\,\text{J}$. So for the input current $I_3 = 145\,\text{mA}$ the maximum output power $P_{3,\text{max}} = \frac{I_3}{e} \cdot \eta_{\text{i}} \cdot E_{\text{ph}} = 74$ mW. So the the amplifier will be heavily saturated and the expected output power for I_3 will be well below 74 mW. For I_2 there will be some degree of saturation.

The blue curves in the graphs in Fig. 14.4 are polynomial fits to the data points (red dots). The values of the parameters and the polynomial fit results are listed in Table 14.1. These polynomial fits for the parameters can be used for fast calculations in which the value of $g_{\text{mod}}\,(J,\lambda)$ needs to be evaluated frequently, e.g. in laser modelling.

The relation between the scaled carrier density and the current density can be analyzed further by calculating the equilibrium carrier concentration using Eq. 14.2. Using this with $\partial N/\partial t = 0$ and assuming the light intensity is sufficiently low to not noticeably affect the carrier concentration, one can set the photon density $I = 0$. One then ends up with a relation between the current density and carrier density:

$$0 = \frac{J \cdot \eta_{\text{i}}}{e \cdot h_{\text{qw}}} - A \cdot N - B \cdot N^2 - C \cdot N^3 - D \cdot N^{5.5} \tag{14.11}$$

This equation can be solved numerically for N for a given value of J and the values for the parameters h_{qw}, A, B, C, D, and η_{i}. The values of h_{qw}, A, B, and C can be known from the SOA design and material properties from literature. By comparing and fitting calculated values of N using Eq. 14.1 with the measured curve in Fig. 14.4a one can determine values for N_{t}, D, and η_{i}. An example of the result of this fit is shown in Fig. 14.5. The values of the parameters are the ones in Table 14.1. Thus a full set of parameters for a compact SOA gain model and for more detailed differential equations is obtained. In the data presented here this is valid for current densities from 1 to 10 kA/cm^2. It has to be noted that the value of h_{qw} and η_{i} can be changed by the same factor without affecting the result of the fit, since the ratio of the two numbers is in equation (14.11). There is some uncertainty in the value that needs to be used for h_{qw} which is the sum of the effective heights of the quantum wells. Measurements on gain saturation can give more information to determine a more precise value of the injection efficiency.

Table 14.1: Fitting parameters of the modal gain on data from a four quantum well SOA for 1550 nm at 18 °C, valid for current densities from 1.0 kA/cm^2 to 10 kA/cm^2 [199] and parameter values used in Eq. 14.11 for the analysis of the carrier density – current density relation. Parameters marked * are from the Apollo Photonics Laser simulator [201].

Parameter	Value
Scaling factor χ	950 cm^{-1}
Homogeneous linewidth γ	$5\cdot 10^{12} Hz$
Relation $\frac{N}{N_t}$ $and J$ [$\frac{kA}{cm^2}$]	$\frac{N}{N_t} = -0.102 + J\cdot 1.262 - J^2\cdot 0.11629 + J^3\cdot 4.091\cdot 10^{-3}$
Relation f_0 [$10^{14} Hz$] and $\frac{N}{N_t}$	$f_0 = 1.933286 - 8.289\cdot 10^{-3}\frac{N}{N_t} - 3.25851\cdot 10^{-4}\left(\frac{N}{N_t}\right)^2$
Relation loss α_s $[cm^{-1}]$ $and \frac{N}{N_t}$	$\alpha_s = -13.44 + 14.72\frac{N}{N_t} - 1.057\left(\frac{N}{N_t}\right)^2$
Confinement factor Γ	0.053
Injection efficiency η_i	0.63
Mode h_{qw} [m]	$0.0265\cdot 10^{-6}$
Tr. carrier density N_t $[m^{-3}]$	$6.5\cdot 10^{23}$
A [s^{-1}] $*$	$1.672\cdot 10^{9}$
B [$s^{-1}\cdot m^3$] $*$	$2.6202\cdot 10^{-16}$
C [$s^{-1}\cdot m^6$] $*$	$5.269\cdot 10^{-41}$
D [$s^{-1}\cdot m^{13.5}$]	$5.7\cdot 10^{-102}$

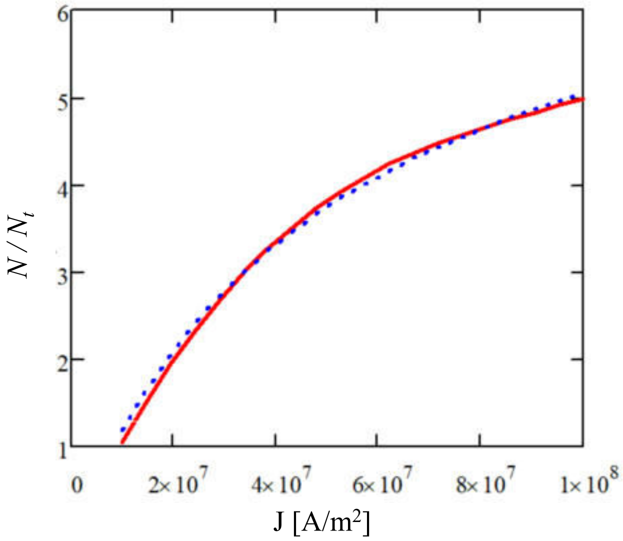


Figure 14.5: Scaled carrier density as a function of current density in the SOA, calculated using Eq. 14.11 (blue dotted curve) and the data following from the gain analysis (red).

Problem 14.3: Recombination mechanisms.

Problem: Consider a C band SOA in the generic technology. The excess carrier density in the amplifier gain medium is $N_c = 3.0 \cdot 10^{18} \text{cm}^{-3}$. Calculate the relative recombination due to defects, bimolecular recombination, Auger recombination, and carrier leakage. Use the parameters in Appendix 14.A. There is no input signal to the SOA.

Solution: The values for the recombination coefficients are $A = 1.672 \cdot 10^{9}\ \text{s}^{-1}$, $B = 2.62 \cdot 10^{-16}\ \text{m}^{3}\text{s}^{-1}$, $C = 5.269 \cdot 10^{-41}\ \text{m}^{6}\text{s}^{-1}$, $D = 5.07 \cdot 10^{-102}\ \text{m}^{13.5}\text{s}^{-1}$. The corresponding recombination rates are:
$A \cdot N_c = 5.017 \cdot 10^{27}\ \text{cm}^{-3}\text{s}^{-1}$, relative to $A \cdot N_c : 1$
$B \cdot N_c^2 = 2.358 \cdot 10^{27}\ \text{cm}^{-3}\text{s}^{-1}$, relative to $A \cdot N_c : 0.47$
$C \cdot N_c^3 = 1.423 \cdot 10^{27}\ \text{cm}^{-3}\text{s}^{-1}$, relative to $A \cdot N_c : 0.28$
$D \cdot N_c^4 = 2.134 \cdot 10^{27}\ \text{cm}^{-3}\text{s}^{-1}$, relative to $A \cdot N_c : 0.42$

The value for the linear dependency factor of the loss on the scaled carrier density N/N_t was 14.72 cm^{-1}. Since we now have determined a value for N_t of $6.5 \cdot 10^{17}$ cm^{-3}, the loss per carrier per unit volume can be calculated to be $2.26 \cdot 10^{-17}$ cm^{2}. This value compares well with a value from literature [189], [202] in which an expression is given for the loss per unit length due to holes as a function of hole concentration and wavelength. This expression gives a value of $2.28 \cdot 10^{-17}$ cm^{2} at 1550 nm for a hole density of $6.5 \cdot 10^{17}$ cm^{-3}. This value of the absorption per carrier per cm^{-3} is in good agreement with the value determined here.

thermal resistance

Please note that the parameters in Table 14.1 are valid for a fixed substrate temperature of 18 °C at the bottom of the chip with the SOA. The temperature of the SOA waveguide will be higher than that due to the thermal resistance between the SOA waveguide and the bottom of the substrate. That means the channel temperature does vary with current density, i.e. the heat load. The current density dependent parameters as presented in Table 14.1 and Fig. 14.4 do include this temperature dependence. The rise in temperature of the waveguide with respect to the bottom of the substrate is in the order of 5 – 7 K over the current density range. The effects of the substrate temperature will be discussed in more detail in Sec. 14.8.

14.6.2 Material gain description for a single wavelength

For describing the gain of the SOA at one specific wavelength λ, one can use a more simple parametric description for the material gain using different parameters as presented in the equation below ([203], Sec. 4.6.3, Table 4.4):

$$g_{\text{mat}}(N, \lambda) = \sigma(\lambda) N_0(\lambda) \ln\left(\frac{N}{N_0(\lambda)}\right) \tag{14.12}$$

gain cross-section

Here $\sigma(\lambda)$ is the gain cross-section at wavelength λ. It is the differential material gain $\frac{\delta g_{\text{mat}}(N,\lambda)}{\delta N}$ per excess carrier per unit volume at the transparency carrier density $N_0(\lambda)$ and has the units of a surface area. So at each wavelength there are only two parameters. Eq. 14.12 can be extended e.g. with an additional term to avoid a problem with the natural log function at very low values for N, as described in [203], for a more accurate description. The relation with the modal gain is then again as in Eq. 14.4. Note that

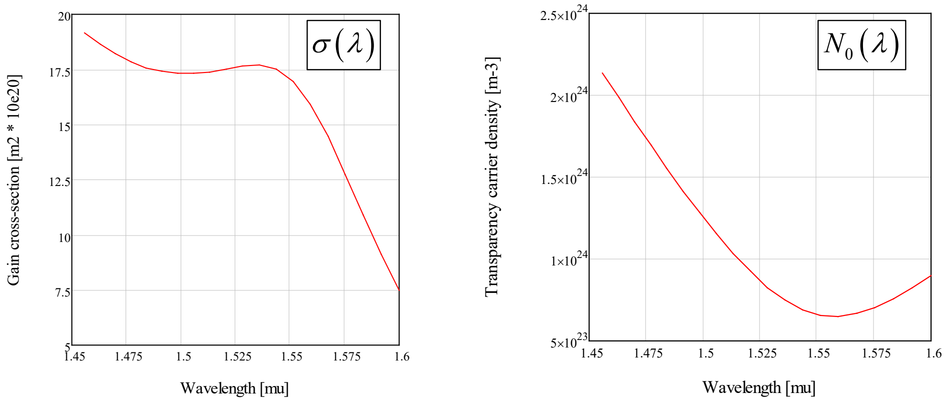


Figure 14.6: Values of cross-section $\sigma(\lambda)$ and $N_0(\lambda)$ for the 4-quantum well ridge waveguide SOA.

the parameter $N_0(\lambda)$ is also a transparency carrier density, but it is different from the parameter N_t in Eq. 14.10. It is specific for a wavelength. This parametrization will be used here in the time dependent description of SOAs and lasers. In Figure 14.6 examples of values of the cross-section $\sigma(\lambda)$ of the same 4-quantum well SOA as discussed in Sec. 14.6.1. are presented that approximate the measured modal gain values.

For situations where one specific wavelength is used and the carrier density is not changing much due to the varying intensity of the input light or injection current level, one can use a first order approximation for the description of the carrier dependency of the material gain:

$$g_{\text{mat}}(N,\lambda) = a_g(\lambda)\,(N - N_0(\lambda)) \tag{14.13}$$

Where $a_g(\lambda) = \left.\frac{\delta g_{\text{mat}}(N,\lambda)}{\delta N}\right|_{N_0(\lambda)}$ is the differential gain at $N_0(\lambda)$. It is equal to the value of $\sigma(\lambda)\,N_0(\lambda)$ in Eq. 14.12. In lasers that operate above threshold, the carrier density in the SOA is clamped at the lasing threshold value. For the description of the SOA in a laser, often this linear approximation is utilized to describe the gain deviation from the threshold gain. One then uses the differential gain at the threshold carrier density gain and the threshold density in Eq. 14.6. Note the threshold density will be significantly higher than the transparency density. For continuous lasers operating at a specific wavelength this can be a very good approximation.

differential gain

14.7 SOA effects on the optical phase – Amplitude phase coupling

If one changes the amplification in an SOA by changing the injection current level, not only the intensity of the light going through is affected, also the optical phase of the light at the end of the SOA changes. This is due to a number of physical processes. The first is a change in temperature in the waveguide. The temperature changes due to the change in current and the associated Joule heating as well as the change in heat load from non-radiative recombination of carriers. The change in temperature leads to: a) a change in the physical length due to thermal expansion, b) a change in refractive index which also depends on temperature and c) possibly a change in mechanical strain on the waveguide. The second effect is that the change in optical gain itself induces a change in refractive index.

The thermal effects on the phase are fairly large but tend to be relatively slow compared to carrier dynamics. The temperature settles to a new distribution after a change in current in a typical time of tens of milliseconds. The precise thermal settling time depends on the design of the chip and the way it is mounted/packaged and the thermal control system used. Since the changes in the optical phase due to temperature are relatively slow, the time derivative of the phase of light transmitted through the SOA is negligible for many applications. For amplification of narrow linewidth signals (tens of kHz and lower) these aspects will have to be taken into account together with other effects such as mechanical vibrations. Since most of the heat is generated by the injection current, by keeping this constant the thermal effects can be kept minimal.

phase modulation

The effect of the carrier density dependent refractive index is not as large, however the change in refractive index is as fast as the change in carrier density, and the carrier density can change very fast (faster than the picosecond time scale) due to an optical input signal to the SOA. This effect can therefore lead to fast phase modulation and the time derivative of the transmitted phase of the optical signal can significantly change the spectral content of the signal.

So for pure steady state situations, constant current and optical signal, the phase effects on the transmitted optical signal are often not considered when looking at signal amplification. When the optical signal becomes time dependent one has to consider the effect of the changing carrier density on the phase, even if just to make an estimate to conclude whether or not it is relevant for the application at hand. The local refractive index as a function of carrier density can be calculated if one knows the complete gain spectrum of the SOA as a function of carrier density, e.g. using the expression for gain presented in Eq. 14.10. This can be done, but the calculation takes more time and one does not always have the required data on the gain spectrum available. A commonly used and effective parametrization of the change in the gain dependent refractive index of the mode $n_{\mathrm{mod}}(N)$ as a function of the change in carrier density is:

$$\frac{\mathrm{d}n_{\mathrm{mod}}}{\mathrm{d}N} = -\frac{\mathrm{d}g_{\mathrm{mod}}}{\mathrm{d}N}\frac{\lambda}{4\pi}\alpha \tag{14.14}$$

linewidth enhancement factor

where α is the so called linewidth enhancement factor and λ is the wavelength of the amplified light. Eq. 14.14 links the differential gain to the change in index via parameter α. It typically has a value of 3 to 5 for a quantum well based SOA at 1550 nm in InP. The value of α varies with wavelength and operating conditions in that range, it is a useful parameter in practice for an SOA. As you can see from Fig. 14.2 and Eq. 14.10 the differential gain will vary with current density and Eq. 14.14 just says the change in index is proportional to that variation. There are semiconductor gain materials, particularly quantum dot material, that can have a much lower α parameter, even down to zero [204].

The differential equation for the phase change due to carrier density changes in φ of the complex amplitude of the envelope of the electric field is [187]:

$$\frac{\mathrm{d}\varphi}{\mathrm{d}z} = -\frac{\alpha}{2}g_{\mathrm{mod}}(N) \tag{14.15}$$

The phase delay in an amplifier can then be given by

$$\varphi(z) = \int_0^z \left\{-\frac{\alpha}{2}g_{\mathrm{mod}}\left[N(z')\right]\right\}\mathrm{d}z' \tag{14.16}$$

In steady state, the equilibrium the carrier density is directly linked to the light intensity and one can write

$$\varphi(z) = \int_0^z \left\{-\frac{\alpha}{2} g_{\text{mod}} N_{\text{eq}}\left[z', I(z')\right], \lambda\right\} \mathrm{d}z' \tag{14.17}$$

More advanced descriptions of the phase modulation due to carrier density changes exist which include effects such as gain saturation and carrier heating effects, as in e.g. [205]. A number of these effects will be introduced in the description of the time dependent behavior of the SOA in Sec. 14.11.

For time dependent signals that are much slower that the carrier concentration recovery time, typically 2 to 3 ns, one can use the equilibrium solutions for the amplifier output intensity and phase. Such a description will predict the phase variations due to the time dependent signal correctly, but of course these effects will be strongest when the changes in carrier concentration are large, i.e. saturation of the amplifier comes into play.

14.8 SOA in steady state – high intensity, saturation

Next the behavior of the SOA in steady state is discussed in the case where the input power to the SOA is at levels where a significant fraction of the injected carriers is used for amplification of the light. Ignoring ASE, the intensity and phase changes for light passing through the SOA can then be described by Eq. 14.6 and Eq. 14.15. The local value for N_{eq} in the integral is calculated from setting Eq. 14.2 to zero. This equation is solved for the steady state carrier density using the known local photon density. The result of the output power of the SOA as a function of input power can be described more simply using an approximate solution of these equations for the amplification G of a single wavelength [206]:

amplification

$$G = G_0 \frac{1 + P_{\text{in}}/P_{\text{S}}}{1 + G_0 P_{\text{in}}/P_{\text{S}}} \tag{14.18}$$

where G_0 is the unsaturated amplification factor, P_{in} is the input power, and P_{S} is the material gain saturation power. P_{S} is given by:

gain saturation

$$P_{\text{S}} = S_{\text{mode}} \cdot \frac{h\nu}{a_{\text{g}}\tau} \tag{14.19}$$

where a_{g} is the differential modal gain and τ is the carrier lifetime (see also Eq. 14.13 and below). The input power P_{in3dB} at which the gain is a factor of 2 (3dB) less is:

carrier lifetime

$$P_{\text{in3dB}} = \frac{P_{\text{S}}}{G_0 - 2} \tag{14.20}$$

Note that the carrier lifetime is dependent on the carrier density which will in practice not be uniform throughout the SOA in saturation. However the approximation works quite well in practice. This is depicted in the example presented in Figure 14.7.

Let us now look in a bit more detail what happens in an amplifier of specific length with a fixed current level, but with a variable input power in steady state. This is depicted in Fig. 14.8 where simulation results are presented for an SOA of length 810 μm (this is relatively long for an SOA) and with 50 mA current (current density 3.1 kA/cm^2) as an example. The calculation essentially involves the integration of Eq. 14.6.

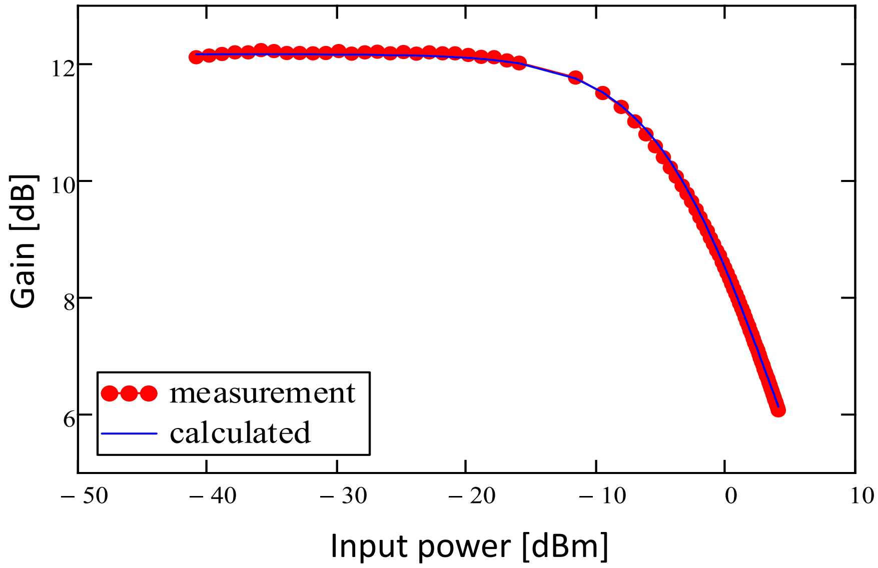


Figure 14.7: Example of the measured saturation curve of an SOA (500 µm length, 50 mA current). The gain is plotted in dB as a function of input power in dBm (red bullets). The blue curve is the curve presented in Eq. 14.18 with $G_0 = 16.5$ and $P_S =$ 10.6 mW, $P_{in3dB} = 0.731$mW.

Presented are simulation results for: a) the optical power in units of dBm and b) in Watt along the length of the SOA, and c) the carrier density distribution over the length of the SOA. The different coloured curves indicate different input power levels. In the simulation the spontaneous emission was switched off to get cleaner lines illustrate the main issues here. At low input power (red lines, 10 µW) one can see clearly in Fig. 14.8a that this is the near unsaturated gain situation. The line is near linear in the dBm scale which means an exponential increase in power. In Fig. 14.8c one can see from the red line that the carrier density is nearly constant over the SOA, and starts to deviate from its maximum around 500 µm.

maximum power
amplification factor
efficiency

The opposite situation is represented by the yellow lines with an input power of 20 mW (13 dBm). This input power level saturates the amplifier right from the entry point, which means that the carriers are all used up for stimulated emission / amplification down to a value close to the transparency carrier density. The light takes out all the carriers and therefore in this situation the light takes up the maximum amount of power it can get from the SOA. The amplification factor is now of course low, but the efficiency of this operating point is the highest. As can be seen clearly in Fig. 14.8b the power increases linearly with distance. Per unit length SOA there is an increase in power by a fixed amount that depends on the current. Increasing current in this situation does not increase the carrier density much. There is only some increase in temperature mainly due to Joule heating in the contact.

The other curves show the simulation results for input power levels that lie in between the unsaturated and fully saturated regimes. The power levels belonging to the different colours are: blue 30 µW, green 100 µW, cyan 200 µW, magenta 1 mW and olive green 4 mW. One can see at the lower input values the power grows exponentially until the carrier density starts to decrease and the power level goes towards the saturation regime. The difference in carrier density over the length of the SOA will drive some carrier diffusion, but the diffusion length is in the order of 5 µm and this will not signif-

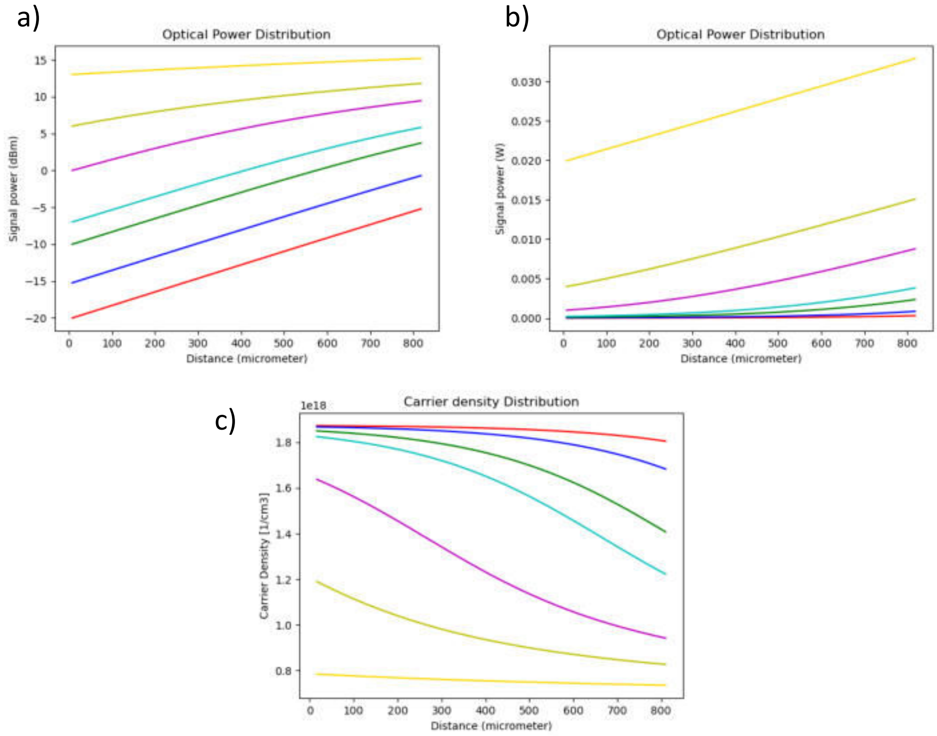


Figure 14.8: The calculated optical power in dBm (a) and in Watt (b) as well as the carrier density per cm^3 (c) as a function of the relative position in the SOA at 7 different input power levels. The SOA is 810 μm long, the injection current is 50 mA. The power level is indicated by the colour: red 10 μW, blue 30 μW, green 100 μW, cyan 200 μW, magenta 1 mW, olive green 4 mW and yellow 20 mW. Calculated using PHIsim [207] with spontaneous emission switched off. The input wavelength is 1552 nm.

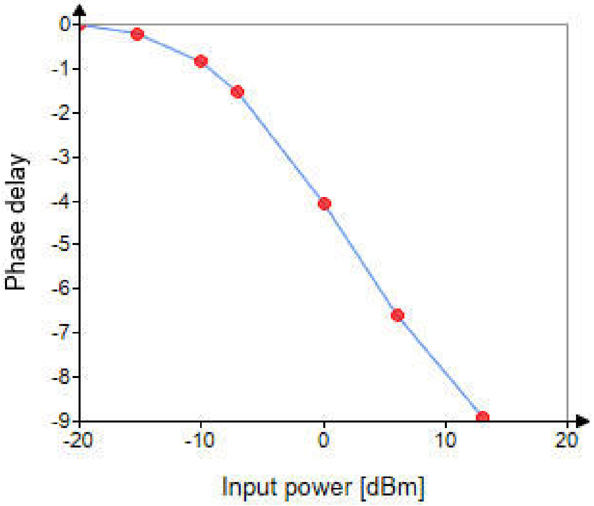


Figure 14.9: The phase delay (in rad) of the 810 μm SOA at 50 mA injection current as a function of input power [dBm]. The dots indicate the same input power levels as in Fig 14.8.

carrier diffusion

icantly affect the distribution along the length of the SOA. Note that carrier diffusion was not included in the simulation.

transparency density

The different carrier densities over the SOA mean that the gain spectrum in the amplifier changes. In the fully saturated case the carrier density is close to the transparency density at the input wavelength. If you look at Fig. 14.2 you can see that in the gain spectrum, the spectral region for positive gain will be to the longer wavelength side. So amplified spontaneous emission (ASE) will be much reduced in saturation when compared to the unsaturated SOA and will be mainly at wavelengths longer than the input wavelength. In a situation where there is no gain saturation, the ASE spectra will resemble those of the signals from the amplifier without any input.

refractive index

The optical path length of the SOA will change with changing input power since the carrier density in the SOA changes and this will change the refractive index in the SOA. The change in phase can be calculated using Eq. 14.17. The results of such a calculation for the same SOA and input power levels are presented in Fig. 14.9. The change in phase delay of the SOA as a function of input power is depicted. You can see that a total change in phase of almost 3π is achieved. And a change of π can be achieved going from 10 μW (−20 dBm) to 0.6 mW (−2 dBm) approximately. Such a change can e.g. be used in an interferometric optical switch. However one must realize that thermal effects can negate such a phase shift, but this depends on the details and in particular thermal effects do not play a role on a short timescale.

The limit of increasing current in the SOA originates from a number of factors. If the SOA is not saturated the first limit is that the carrier density does not increase much with current density above a certain value of the current density (see Fig. 14.2). This is mainly due to the carrier loss related to drift and Auger recombination and due to the temperature rise. For longer SOAs (a few hundred μm), the temperature effect is secondary. This provides effectively an upper limit on the achievable unsaturated gain. If the SOA is long and for high current densities, the ASE intensity can get to a level where it starts saturating the SOA and limit the gain for a signal. Another practical limit is caused by the circuit in which the amplifier is placed. Reflections from interfaces at the end of the SOA or from other components in the chip can make the SOA go lasing. In practice this tends to happen before issues occur with ASE. However, if one injects a signal with sufficient input power to start to saturate the SOA, the carrier density is reduced and one can stop the SOA from lasing, and use the device as an amplifier for a CW signal.

When the SOA is saturated by the input signal, one can use higher injection current densities since the carrier density is then kept low, i.e. just above the transparency density, by the input light. The output power limits are then formed by the heating in the SOA, and the intensity level of the light which can lead to non-linear processes and damage. The increase in current density also can lead to an increase in absorption by free carriers and reduction of efficiency.

More complete modelling of the SOA in the steady state is possible. E.g. one can model the ASE as well as the signal in more detail. The ASE intensity can be modeled using another differential equation or by dividing the ASE spectrum in narrow spectral segments and describing the ASE in each segment with its own equation with different parameters such as described in [188].

14.9 SOA in steady state - Temperature effects

The effect of temperature increase on the SOA, is that the excess carrier density will decrease for the same injection current density. The reason for this is that the Fermi distribution of the carriers over the energy levels in the conduction and valence bands spreads out more when the temperature is increased, and that more excess carriers can therefore escape the quantum wells. To maintain the same gain level one typically then increases the current, but this increases the temperature as well due to an increase in Joule heating. So there is some positive feedback when increasing the current. At higher current levels this can lead to a reduced output from the SOA when increasing the injection current.

Heat is generated in the SOA in a) the contact due to the contact resistance, b) the resistive heating in the different doped and undoped semiconductor layers, c) heat from non-radiative recombination of excess carriers. The latter process is in competition with the optical processes so that thermal load will depend on the optical input to the SOA.

temperature distribution

For the InP based quantum well ridge waveguide SOA a reasonable estimate of the heat load on the SOA is to assume that 80% of the electrical power injected into the SOA will be converted into heat. Using a thermal finite element simulator one can make an estimate of the temperature distribution. To illustrate this, a simplified simulation using QuickField Student Edition [208] is discussed. In this simulation it is assumed that all heat is generated in the 500 nm thick waveguide layer in the 2 μm wide shallowly etched ridge. The simulation is over a 20 μm wide cross-section of the chip. If we assume 80% of the electrical input power is dissipated, the current density is 10 $\mathrm{kA/cm^2}$, and the voltage over the junction is 1.2 V. The thickness of the substrate in this simulation is only 10 μm and the bottom is held at a fixed temperature. We get a result presented in Fig. 14.10. This is all very much a simplification, e.g. we do not take the effect of the metal contact on top into account. But the result shows qualitatively what happens, a two-dimensional heat flow pattern and temperature distribution builds up over much of the length of the SOA. You can see that the increase in temperature in the waveguide is approximately 6 K. The quantum wells can actually get warmer since much of the energy is dissipated in these small volumes and due to the heat flow in a

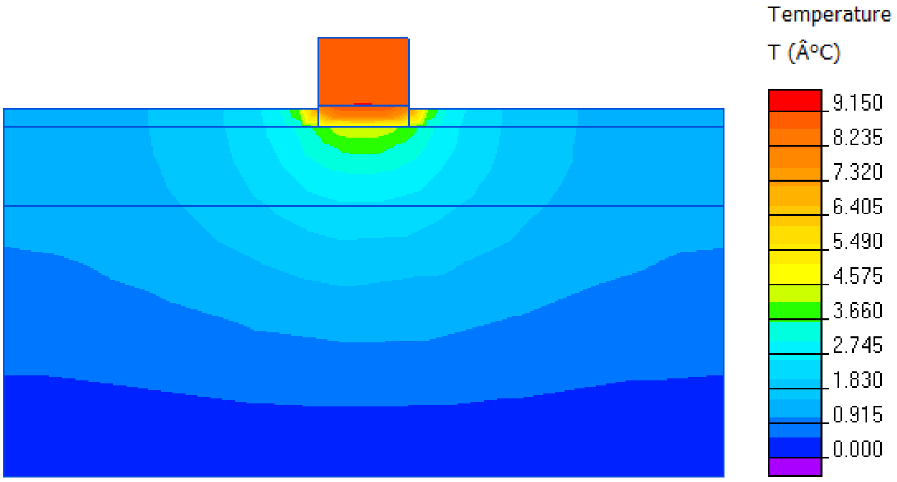


Figure 14.10: Simulated temperature distribution of a simplified SOA structure on an InP substrate operating at 10kA/cm^2. Indicated is the temperature increase (in °C) from the heatsink temperature at the bottom (Quickfield).

thicker substrate. This will also introduce additional strain. An example of calculations of heat loads can be found in [209].

thermal roll-over

The effect of temperature increase on gain can be described by a decrease of the carrier density at the same current density. The consequence is that in order to maintain the same gain level, one must increase the current, which leads to an increase in heating and temperature, which lead to a further reduction in carrier concentration (with respect to the original temperature of the SOA). As mentioned before, this positive feedback loop leads to what is called thermal roll-over at some current level. An increase in current into an SOA leads to a lower output power of the device in which the SOA is used, instead of an increase.

A change in power dissipation in the SOA can be induced by a change in injection current, but also to a smaller extent by a change in input optical power. As a consequence of a change in power dissipation the temperature changes of the SOA. The heat is generated in the waveguide and the metal contact/semiconductor transition and the change when the dissipated power level changes starts immediately but it can take tens of milliseconds or more to get to a new thermal equilibrium. The exact timescales of temperature depends on the details of the thickness of the substrate, the mounting of the PIC, the Peltier cooler and its control, the heatsink of the Peltier cooling. Therefore a PIC can take a long time compared to the carrier dynamics time scale, to go to equilibrium again. It depends on the application how problematic such slow changes are.

optical pathlength

Effect of temperature on optical pathlength in a device and thus the phase of the transmitted light, is two-fold. The first is thermal expansion, which increases the path length, and the thermal dependent refractive index of the mode, which in turn is linked to the refractive index changes in the materials with temperature. This is usually the more important effect. For InP $\mathrm{d}n/\mathrm{d}T$ is approximately 2×10^{-4}. So for example the optical path length of an 800 µm SOA heating up 10 degrees will change by approximately 3.2 × 1.6 µm which is roughly 3.3 wavelengths. The numerical value of the change in modal index as function of the temperature for the SOA can be assumed to be similar to those of the shallow waveguide. Values for InP on the relevant thermal parameters can be found in [210], [211], [212].

Note that large temperature changes in the SOA can lead to significant strain inside the structure, which in turn influences the refractive index. Typically it is not practical to separate these effects for an SOA.

As mentioned before, the changes in phase of the transmitted light due to changes in temperature are slow, i.e. in the range of milliseconds. This means that the chirp due to temperature changes can often be neglected, except for low linewidth laser signals (approximately 1 kHz and below) passing through an SOA. For an SOA inside a laser the effects can be much larger and more significant. This is because the laser resonance frequencies will change due to the change in optical path length. Interferometric structures drift with approximately 0.12 nm/°C (at 1550 nm) in the InGaAsP/InP integration platform technology.

chirp

14.10 SOA in steady state - Polarization

The unstrained quantum well amplifiers work mainly for TE polarization. The gain in TM polarization is significantly lower such that it does not play a significant role in the amplifiers discussed here [199]. Some strain is typically used on the quantum wells to optimize the TE gain in the SOA [203] [213]. When looking at ASE from the amplifiers, e.g. from a device that is used to derive the unsaturated gain using the ASE, light from the TM polarization can play a role [199].

Polarization insensitive SOAs can be designed and are being considered for generic integration platforms. One of the possibilities is to use an SOA based on bulk gain material, i.e. a thick gain layer material of e.g. 120 nm with a suitable bandgap. Bulk material does not show a polarization dependence of the gain. Any difference between TE and TM polarization is caused by the waveguiding structure. In a standard ridge waveguide amplifier the TM losses are a bit higher, typically a few dB/cm^{-1}, which means that lasers with a bulk amplifier in the cavity will operate on TE. For shallow etched waveguides the effect of the waveguide can be compensated by bringing in a small difference between the gain for the two polarizations using a small amount of strain. Another solution is to use strained quantum wells [214] [215] and waveguide design [216].

polarization dependence

strained quantum wells

14.11 Time dependent signal amplification

14.11.1 Time scales

A time dependent signal sent into an SOA can lead to changes in the excess carrier concentration and can change the distribution of carriers over the energy levels in the valence and conduction band. These changes are time dependent and position dependent. One has to consider how fast the input signal changes are in relation to the time scales of the dynamics of the carriers. There are a number of different process in the carrier dynamics with different time scales. These are described below qualitatively.

carrier dynamics

The most important one is the lifetime of the excess carriers due to electron-hole recombination. This process has the longest time scale in an SOA, in the order of hundreds of picoseconds. Then there are the processes which relate to the redistribution of the carriers over the energy levels in the bands. Let us consider an ultra-short light pulse of say 50 fs. It interacts with electrons and holes in the SOA that have an energy

lifetime excess carriers

redistribution of carriers

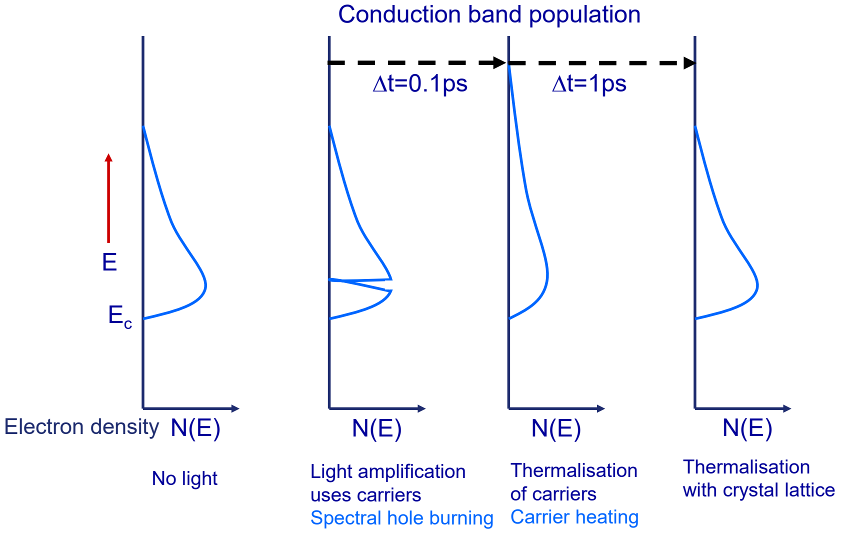


Figure 14.11: Schematic depiction of the electron concentration in the conduction band as a function of time for the amplification of an ultrashort optical pulse.

difference equal to the photon energy. So the pulse reduces the concentration of carriers at those levels. This makes that the distribution of carriers over the energy levels within the bands becomes non-thermal; it has a dip for the levels that had the carriers interacting with the light. This is indicated schematically in Fig. 14.11.

spectral hole burning
Fermi distribution

Such a process is called spectral hole burning. Now the carriers will redistribute themselves very quickly within the bands to get to a thermal Fermi distribution. A typical timescale for this is in the order of 100 fs. This new thermal distribution has however a different temperature than the surrounding ion lattice. Optical gain is usually available from carriers relatively close to the band edges. As a consequence the carriers used in the amplification process have an energy that is below the average energy of the carriers. The result is that after the pulse a thermal distribution is reached that tends to have a higher temperature. This is the process of carrier heating. The carriers are now out of thermal equilibrium with the lattice of the ions in the gain material. The interaction of the carriers with the ions in the lattice will make that the temperature will become equal again; this process takes place on the time scale of 1 picosecond.

carrier heating
thermal equilibrium

carrier lifetime

As mentioned earlier, one can use the steady state description of an amplifier for signals that are relatively slow compared to the carrier lifetime. As a rule of thumb that means the relevant frequency components in the signal due to modulation of the intensity have a frequency lower than approximately one third of the carrier recombination rate A introduced in Eq. 14.2. Since a typical number for the rate $A = 1/500$ ps, the upper frequency at which to use the steady state solutions is approximately 700 MHz. The exact criterion will depend on the accuracy required for the calculation.

14.11.2 Description of time dependent behaviour

carrier dynamics
excess carriers

In this chapter the SOA time dependent behaviour is described for pulses with a duration of 1 ps or longer, or signal frequency components up to approximately 1 THz. The carrier dynamics is described by the number of excess carriers in the conduction and valence bands with a recombination rate, in a way similar to Eqs. 14.1 and 14.2.

The time dependent behaviour of the spectral hole burning and carrier heating is not described. For many applications such a description is sufficient. The processes of carrier heating and spectral hole burning have an effect though. This is described by effective parameters such as the gain saturation parameter ε_1 in Eq. 14.2. For more detailed descriptions please see references [187], [188], [217].

We can describe the field of the mode $A(z,t)$ as a propagating wave moving in the positive z direction:

$$A(z,t) = E(z,t)\,\mathrm{e}^{i\varphi(z,t)}\mathrm{e}^{i(\omega_0 t - kz)} \tag{14.21}$$

where $E(z,t)$ is the electric field envelope that varies slowly in space and time compared to the $\mathrm{e}^{i(\omega_0 t - kz)}$ term; ω_0 is the optical frequency of the wave. Similarly the $\varphi(z,t)$ is a slowly varying phase in space and time of the envelope. Both $E(z,t)$ and $\varphi(z,t)$ are real valued functions. The normalization is such that the intensity $I(z,t) = |E(z,t)|^2$. The spectral information is in the time dependent phase and amplitudes. For a good description of an optical amplifier one needs to describe the light in a single transverse mode in two directions. Even when there is only input to the SOA into one end, the ASE will always travel in both directions.

Differential equations in time and place for the carrier density, photon densities and phase for light in a single transverse mode for both directions, can be used to describe time dependent behaviour. The PHIsim circuit simulator which is used to calculate the examples shown in this chapter, uses such equations. These differential equations are explained in detail on the PHIsim website Principles page [207]. The equations also include free carrier and two-photon absorption effects, effects resulting from carrier heating and spectral hole burning and a non-linear, i.e. intensity dependent refractive index (Kerr effect). These nonlinear effects can become relevant since pulsed operation can lead to high peak powers (typically > 100 mW) which will make these nonlinear effects more prominent. More advanced simulators are discussed in Chapter Simulation Methods.

14.11.3 Amplification of short pulses

We now will discuss and illustrate a number of effects that occur when amplifying optical pulses in the SOA. The examples are simulation results using PHIsimv3, the input parameters used are in Appendix 14A. Similar results are presented and discussed in [218]. Many of the effects that occur when amplifying short pulses are discussed using a specific example to illustrate.

The first example is the amplification of a single Gaussian shaped input pulse that is 5 ps long (FWHM) and a constant envelope phase. A single pulse means that the pulse is sent into an amplifier that is initially in steady state equilibrium with a fixed current level without any optical input. The amplifier in the example is 800 µm long, at 50 mA injection current which is a current density of 3.13 kA/cm^2. This is a relatively long amplifier operated at a moderate injection current density with an unsaturated gain of approximately 25 cm^{-1}. At such a current level the spontaneous emission intensity from the long amplifier without input is relatively low. The gain bandwidth at this injection current density (see Fig. 14.2) is much larger than the bandwidth of the 5 ps pulse.

To illustrate the important effects of the SOA on the pulse shape, i.e. the output power as a function of time, Fig. 14.12a depicts the output power as a function of time for a series of 5 ps pulses with pulse peak input power levels ranging from 10µW to 250mW.

Fig. 14.12b presents the same data with output power in dBm. In Fig. 14.12c the amplification of the energy in the output pulse is presented as a function of input pulse energy. The SOA is in equilibrium before the pulse is sent in. By comparing the curves in Fig. 14.12a for the four highest input pulse energies, one can see the effect of the saturation of the SOA being reached during the pulse. For the highest input pulse the maximum output power is reached earlier than for the lowest input power levels, in this example by approximately 2.5 ps. This is indicated by the vertical blue lines. It is as if the transit time of the pulse of 9.9 ps through the 800 µm SOA is reduced to 7.4 ps. This effect is due to the reshaping of the pulse by the saturation of the SOA. This saturation is being reached at an input power level (well) below that of the highest intensity point of the input pulse. The carrier density is decreased significantly before the input peak power level is reached and the amplification of the light arriving after saturation is reached is reduced. This distorts the pulse shape to a pulse that has a steeper rising edge, a longer falling edge and reaches peak power earlier.

intensity noise
timing noise

Since a change in intensity of the input pulse changes the transit time through the SOA when it starts to be saturated, intensity noise in a series of perfectly equally spaced pulses in time will lead to variations in the time that the maximum value of the pulses is reached. I.e. intensity noise will translate into timing noise, although the intensity noise and variations in energy per pulse may be reduced by the saturation of the SOA.

self-phase modulation
chirp

The change in shape of the pulse also has an effect on the spectral content of the pulse. But purely the change of the shape of the pulse by the SOA usually plays a minor role in the spectrum of the output pulse when it is in the picosecond range. The change in spectrum between input and output pulse is predominantly influenced by the self-phase modulation in the SOA. The shape of the pulse in time thus influences the change in its spectrum via the self-phase modulation. An example of the effect of self-phase modulation for a single pulse is shown in Fig. 14.13. The pulse has the same energy and shape as the most intense pulse in Fig. 14.12a (250 mW peak power). The phase of the input pulse is constant in time for the input pulse, that means the instantaneous frequency of the input pulse is constant. In Fig. 14.13b the shift in instantaneous frequency (the time derivative of the envelope phase), also called chirp, is depicted in GHz as a function of time.

When the intensity of the pulse is very low the output of the SOA is mainly ASE and the phase is very noisy and the calculated chirp becomes very large and varies a lot. It just means that the ASE is a very wide band signal. The chirp, as one can see, is down in frequency from the input frequency. The frequency goes down by over 250 GHz which is over 2 nm at 1550 nm central wavelength. The effect on the optical spectrum is presented in Figure 14.17. It shows the spectra of the same series of pulses as in Figure 14.15. One can see from this series how the spectrum changes with increasing pulse energy. The maximum shifts significantly with pulse energy. Very typical for this self-phase modulation effect is the occurrence of several maxima in the spectrum [218][219][220].

The ASE level out of the SOA is particularly important when amplifying pulses of only a few picoseconds that are time separated by more than a few times the carrier lifetime. The ASE power is constant if the amplifier is not saturated by the pulse and the average power of the ASE can swamp the amplified signal from the widely separated amplified short pulses. The ASE level that can be tolerated will depend on the details of the application. E.g., if a spectral filter can be used to filter out only the optical pulses one can tolerate more ASE from the SOA. Similarly if one utilizes time resolved detection, the ASE signal is much less of an issue. In general one needs to simulate designs to check such aspects.

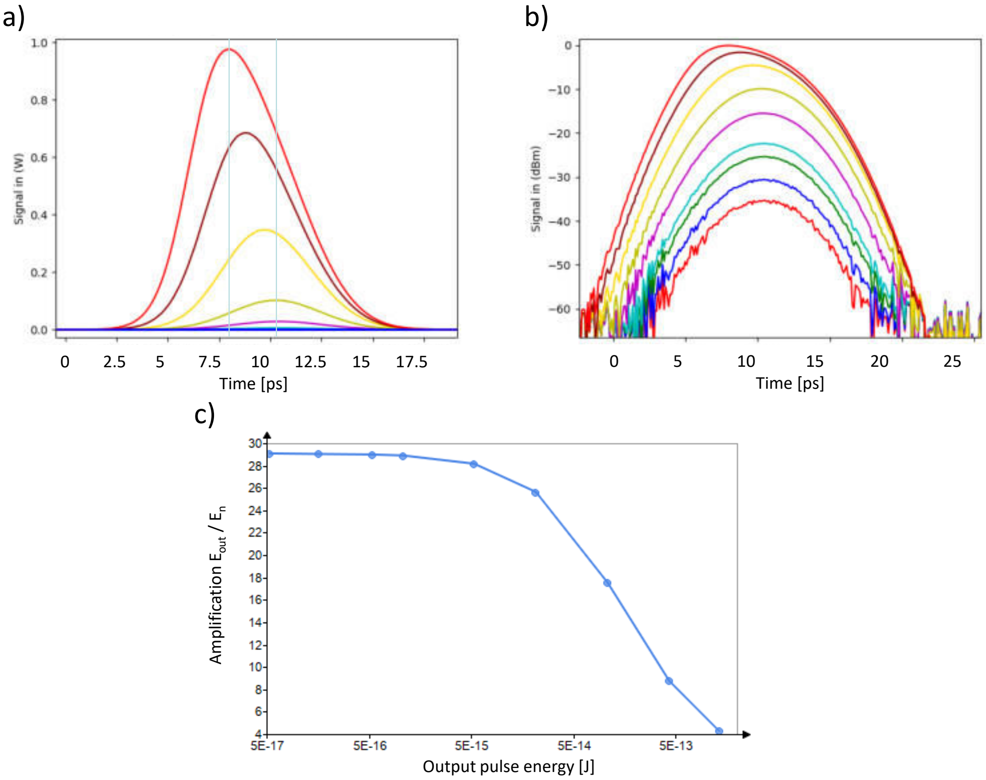


Figure 14.12: a) SOA output power as a function of time for a series of 5 ps (FHWM) input pulses with increasing peak power (10 µW red, 30 µW blue, 100 µW green, 200 µW cyan, 1 mW magenta, 4 mW light green, 20 mW yellow, 80 mW dark red, 250 mW red). The SOA is 800m long, injection current is 50 mA. b) The SOA output power data are plotted in dBm as function of time. Note the time axis is different from a). c) Output pulse energy amplification (linear scale) for single pulses as a function of input pulse energy (in J) for the 800 µm SOA operated at 50 mA.

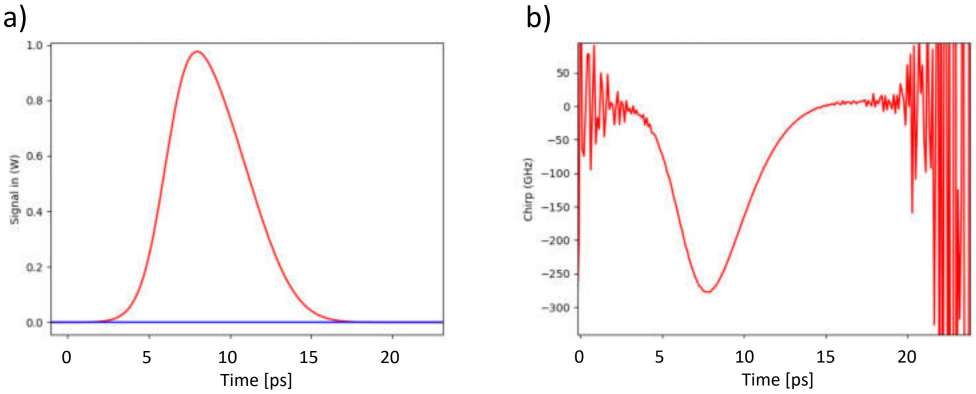


Figure 14.13: a) Output pulse after 800 µm amplifier, 250mW peak power, 5 ps input pulse. b) Instantaneous frequency (chirp) of the SOA output improve pictures in units of GHz as a function of time, illustrating the self-phase modulation. The large variations are at times when the main output is ASE and the phase is varying randomly.

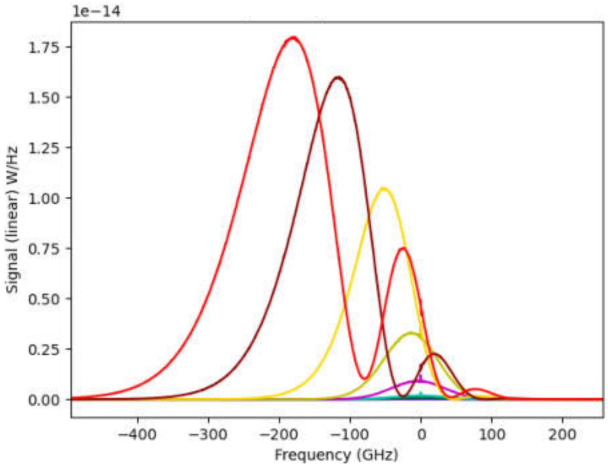


Figure 14.14: Spectra of amplified 5 ps pulses with varying pulse peak powers, (200 µW cyan, 1 mW magenta, 4 mW light green, 20 mW yellow, 80 mW dark red, 250 mW red) for the 800 µm SOA operated at 50 mA.

The examples presented here are for a relatively long SOA of 800 µm. This length was chosen to make the effects discussed more obvious in the figures. When a shorter amplifier at higher current density with 70 cm^{-1} gain is used ASE will be less due to the shorter SOA so this current density is feasible. It should also be pointed out at this point that the peak power level of the highest energy pulse is at the highest level one can expect. The peak power level of 1 Watt means that the intensity is in the order of 100 MW/cm^2 where non-linear effects start to dominate [221].

14.11.4 Amplification of pulse trains and pattern effects

If two or more light pulses go through the amplifier with less time between them than several times the carrier lifetime, the carrier depletion by the first pulse affects the gain in the second [222], [223]. If a burst of regularly spaced train of pulses goes through the SOA and the signal depletes the SOA to some extent, the first pulse will be amplified more than subsequent pulses. After a number of pulses an equilibrium sets in.

pattern effects

If an irregular train of closely spaced (i.e. less than the carrier lifetime) pulses is sent into an SOA one can observe in the output what is known in the telecommunication field as pattern effects. This is illustrated in Fig 14.18. A pulse pattern of 1550nm light pulses as presented in Fig. 14.18a is sent into an SOA (500-µm-long, operated at 5kA/cm^2). The individual Gaussian shaped pulses have a peak power of 10 mW, have a 17ps FWHM which means an energy of 180 fJ. The smallest distance between two pulses is 50 ps. If this was a regular pulse train the average power would be 3.55 mW. In Fig 14.18b the output signal is presented.

Initially the SOA is in equilibrium with no optical input. When the first pulse arrives it is amplified most. The first pulse reduces the carrier concentration and since the carrier lifetime is significantly longer than the time to the next pulse, the gain for the second pulse is lower. This then continues for a regular pulse train until an equilibrium is achieved. The higher the energy of the pulses, the faster equilibrium will be achieved. The equilibrium gain in average power will be close to the gain for a continuous signal, it will be somewhat lower due to the non-linear gain compression. In Fig. 14.15 one can see the effect of short interruptions in the pulse train (one, two or three pulses missing).

Problem 14.4: Estimate output pulse energy.

Problem: Consider an SOA with length $L_{\text{SOA}} = 800\,\mu\text{m}$, ridge width $w_{\text{ridge}} = 2\,\mu\text{m}$ and an injection current $I_{\text{SOA}} = 90\,\text{mA}$. The SOA is initially in a steady state with no optical input. Estimate the maximum pulse energy that can be extracted from the SOA when a pulse of 5 ps length is sent into the SOA that saturates the amplifier. The wavelength is 1530 nm. Use the information in Fig. 14.4, Fig. 14.6, and/or Table 14.1.

Solution: First the carrier density in the amplifier in equilibrium without optical input is calculated. The current density is $J = I_{\text{SOA}}/(L_{\text{SOA}} \cdot w_{\text{ridge}}) = 6\,k\text{A/cm}^2$. From Table 14.1 we can use the relation between the current density and the carrier density ratio N/N_t to calculate the carrier density N. Here $N_t = 6.5 \cdot 12^{23}\ m^{-3}$ is the transparency density at the bandgap energy (also in Table 14.1).

$$N/N_t = -0.102 + J_{\text{in}} \cdot 1.262 - J_{\text{in}}^2 \cdot 0.11629 + J_{\text{in}}^3 \cdot 4.091 \cdot 10^{-3}$$

Using this the carrier density becomes: $N = 2.709 \cdot 10^{24}\ \text{m}^{-3}$. If an incoming pulse saturates the amplifier completely, the carrier density immediately after the pulse has passed will be at the transparency density $N_{0_{1530nm}}$. At this density the gain = 0 cm^{-1} for light with a wavelength of $\lambda = 1530$ nm. This will be the case because the optical pulse is much shorter than the carrier lifetime and the period over which the carrier concentration is replenished to its equilibrium value will take approximately three times the carrier lifetime.
From Fig. 14.6 one can read-off a value for $N_{0_{1530nm}}$. Alternatively one can take Eq. 14.10 and solve it for $g_{mat}(N) = 0$ for N. Using Fig. 14.6 is an easier option and will give $N_{0_{1530nm}} = 8 \cdot 10^{23}\text{m}^{-3}$. The total volume V of the active gain material can be estimated at $V = L_{SOA} \cdot w_{ridge} \cdot h_{qw}$ where $h_{qw} = 26.5$ nm is the total thickness of the four quantum wells. That means the optical input pulse can take a number of carriers of at most $N_{out} = (N - N_{0_{1530nm}}) \cdot V = 7.59 \cdot 10^7$. Therefor the maximum increase in energy of the optical input pulse is estimated to be

$$N_{out} \cdot E_{ph} = N_{out} \cdot \frac{h \cdot c}{\lambda} = 9.85 \cdot 10^{-12}\text{J}$$

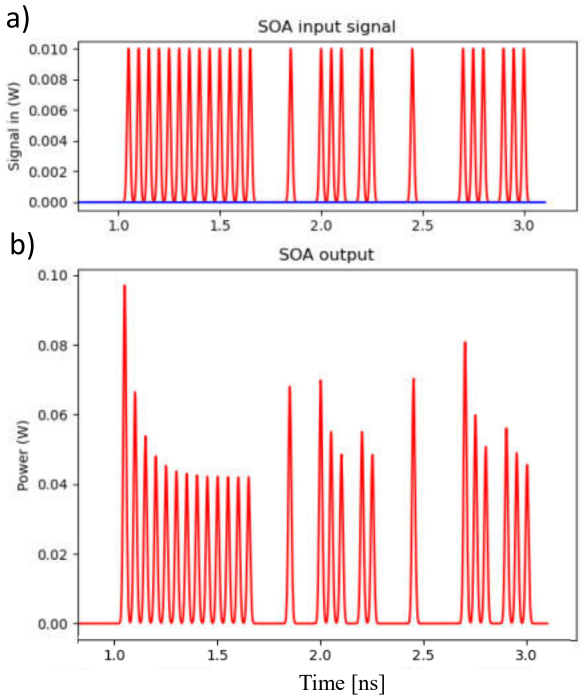


Figure 14.15: Pulse pattern effects calculated for a 500 μm long SOA operated at 50 mA. Top: Input pulse pattern, 17 ps FWHM Gaussian pulses, with 50 ps minimum spacing at 1550 nm. b) SOA output signal showing the variation in amplification of the different pulses.

The SOA will build up to a higher carrier concentration during those interruptions and the first pulse after it will be amplified more strongly again. So the pulse pattern leads to an uneven amplification for each pulse in the signal. This is what is called pattern effects. This does mean also that the spectral broadening, self-phase modulation and reshaping can vary from pulse to pulse. This also means there will be an increasing variation in timing of the pulses. How serious this is depends of course on many details.

Pattern effects can be minimized by operating the amplifier far from saturation. Then none of the pulses will change the carrier density noticeably and non-linear gain compression is also avoided. For the SOA and pulse duration discussed in the example above this input level is just over 1 mW of peak power in the pulses (18 fJ energy per pulse). The price to pay is that the SOA is operated in an energy inefficient way and that the signal level will be closer to the ASE noise from the amplifier. The signal will only take away a small amount of power relative to the saturated output power level.

There are several ways to reduce pattern effects [224], typically by influencing the carrier lifetime in some way, but this is outside the scope of this chapter.

14.11.5 Amplification of multiple signals

Amplifying multiple signals in the same amplifier leads to a number of different effects where the signals can influence each other. All signals use up carriers from the same reservoir of carriers (i.e. in quantum well and bulk amplifiers). When one of the signals is using relatively large amounts of carriers, it will affect the amplification of all other signals. The gain for all signals will be reduced due to this reduction in carriers and possible changes in the carrier distribution over the states (the effects of carrier heating and spectral hole burning). This effect is called cross-gain modulation.

cross-gain modulation

The change in carrier concentration and gain in the SOA also affects the refractive index so the phase of all signals is also influenced. This effect is called cross-phase modulation. Both cross-gain and cross-phase modulation effects can be used in circuits to e.g. transfer information from one channel to another [188], [224], [217], [187], e.g. to realize a wavelength converter. If the SOA is polarization dependent, also cross-polarization modulation can be observed.

cross-phase modulation

cross-polarization modulation

The amount by which one signal affects the other depends on the wavelength of the signals. The distance between the wavelengths is important as well as their values with respect to the gain spectrum of the SOA. When the two wavelengths are relatively close to each other (spaced less than 10 to 20 nm) and near the gain peak, the effects of mutual influence are more or less symmetric. When the wavelengths are further apart the response of carrier saturation by one wavelength on the gain and phase of the other signal will not be the same when the roles are reversed. This is due to the differences in optical gain and transparency carrier density for the different wavelengths.

When multiple signals with the same polarization are sent into an SOA the electric fields of the different signals interfere and the sum of the signals will be strongly time dependent and the SOA will respond to this. Even when the individual signals are purely CW or slowly varying in time, their beat frequency can be very high, i.e. much higher than the carrier lifetime (the electron-hole recombination time). The carrier heating and spectral hole burning processes have, as discussed in Sec. 14.11.1, much shorter time constants and the modulation in gain and refractive index due to these processes will follow the beating of the signals.

beat frequency

Since the SOA responds non-linearly due to the saturation and intensity dependent gain, additional optical frequencies will be produced. One such non-linear process is four wave mixing. This is coherent non-linear process between two optical fields. The

four wave mixing

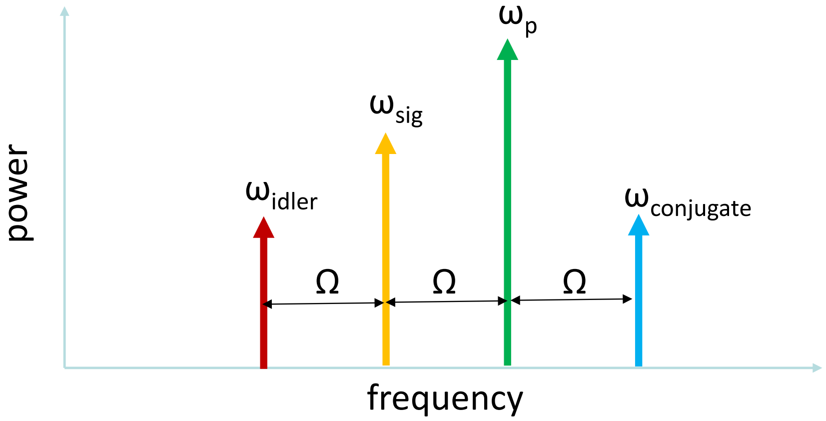


Figure 14.16: A schematic depiction of the optical frequencies in the four wave mixing process.

most intense of the fields is called the pump with optical frequency ω_p, the other field is called the signal (optical frequency ω_{sig}). The difference between the two frequencies is $|\omega_p - \omega_{sig}| = \Omega$. Two additional frequencies are generated in the four wave mixing process. One at $\omega_{sig} - \Omega$ which is called the idler and one at $\omega_p + \Omega$ which is called the conjugate. These signals can be used in devices for e.g. wavelength conversion. In laser systems with multiple wavelengths such signals can also be observed, see e.g. [225].

idler
conjugate
wavelength conversion

counter propagation

Multiple signals in an SOA can of course also be counter propagating. Then the effects discussed can also occur, but of course there are important differences. E.g. when the signals are short pulses, i.e. pulse that are physically less long than the SOA, the cross-gain and cross-phase effects can be different from the co-propagating situation. Another example is when two counter propagating CW signals of the same optical frequency are sent into the SOA, a standing wave will form a carrier density grating in the SOA. Light will be coupled from one direction into the other direction. This will lead to the effects of self and cross-gain saturation which will be different. When the frequencies of the two counter-propagating signals are different also four wave mixing will occur but this is far less efficient than for the co-propagating case. For a more complete discussion on non-linear effects in SOA and their use in applications the reader is referred to [217] and [187].

14.12 Using an amplifier in an integrated circuit and design aspects.

metal tracking

contact pads

Once you have chosen a specific foundry for the realization of your chip, the choices to be made related to the SOAs in your circuit design are limited to typically the length of an SOA and the position. In addition to that the metal tracking for the SOAs needs to be designed (together with all other tracking). Depending on the exact integration platform the SOA building block can include the required contact pads for the SOA, or these pads are options in the design software. In particular there are significant differences in requirements for space between the N-doped substrate chips and the semi-insulating substrate chips. The latter requires large contact openings to the buried n-doped layer.

For the choice of the length of the amplifier, one will have to set requirements and use simulations using the available information, starting with an estimate based on the available information on the platform.

When the length of the SOA has been decided, the positioning in the circuit layout needs to be considered. The n-contact in the semi-insulating platforms will take a considerable amount of space. The space requirement more than doubles compared to the n-doped substrate SOA.

When multiple SOAs are needed in a circuit, the relative position needs to be considered. The more closely together the SOAs are, the more likely that the operating temperature of the SOA will increase for the same injection current values. By how much will depend on details of the thermal design and location of the temperature sensor etc.. The main reason is that the heat flow density down through the chip will increase when the SOAs are more closely spaced and this will increase the temperature locally on the chip. For a discussion of such thermal aspects see e.g. [209], [226], [227].

resistance

In the design of the circuit one needs to pay attention to the metal tracking and contacts to the amplifier. The first thing that is important is the uniformity of the current injection in the SOA. In most building blocks uniform injection has been taken care of, but also SOA building blocks exist (in the Smart Photonics PDK) with only a minimum of metallization specified, e.g. with only a 30 µm wide and 300 nm thick gold layer. Such a gold track over the SOA has a resistance per unit length of at least 2.7 Ω/mm (using the resistivity of pure gold of $2.4 \cdot 10^{-8}$ Ωm). If we assume this 300 nm layer is evaporated gold, the real conductivity is expected to be close to the value of pure gold. The SOA is a forward biased diode so the current is highly sensitive to the voltage and such a resistance can cause an uneven current distribution over the length of the SOA. A simple model to look at this distribution is presented in Appendix 14B. When the plated gold track of 2.5 µm thickness is used even at 30 micron width, this solves the problem for most situations. There is also a report of a 450 nm thick sputtered gold layer as a first layer. In that case the resistance per unit length reported is approximately 1.5 times larger.

capacitance

For the metal tracking from a contact pad to the SOA one should keep the resistance in mind. In particular when multiple SOAs are connected to one pad. This situation can occur on the semi-insulating platform when connecting the different n-sides of the SOAs. A resistance in the common paths will lead to cross-talk from one SOA to the other. And again the fact that one is dealing with diodes makes that small voltage differences over the diode due to the small resistance in the common path, can lead to more significant current level changes. When modulated currents are used, the cross-talk can be stronger since the current levels can be higher than for DC circuits due to capacitances in the circuit. The higher peak current levels will lead to larger voltage drops over the diodes. The capacitances in the circuits tend to be dominated by the capacitance of the metal track and contacting pad. This is due to the relatively large surface area of the metal track compared to the SOA diode surface area and the small distance (in the order of 1 µm or even less) between the metal and n-doped layers. A crude estimate for a 20 µm wide track gives 0.6 pF per mm track length.

14.13 Mask design

Generic foundries for photonic integration that include active-passive integration technology, either monolithic or heterogeneous, will offer the optical amplifier as a basic component or building block in their Process Design Kit. Typically the user will only see the amplifier component in the design software with two optical connections to passive waveguides, and one or two electrical connections. All the details e.g. on the interface between the passive waveguide and the waveguide with the active quantum well layers, the metallization etcetera is hidden from the user. The amplifier building block is typically parametrized for its length and it can be used as described in the design manual of the PDK.

Three main aspects have to be taken into account when including an amplifier in a mask design. We will discuss here these aspects for a monolithic integration scheme on InP.

The first aspect is the choice for the length of the SOA. The optimal length can be derived from the specifications for the task of the SOA. The length of the building block is typically a bit larger (46 micrometer in the example presented in figure 14.18 below). This is due to the fact that the active-passive interface and a short section of passive waveguide is included in the component.

The second aspect is that the orientation of the SOA is described, it is always in the same direction with respect to the design space (the cell) on the wafer. Usually the orientation is such that the orientation is in the east-west direction. This has its origin in the crystal orientation of the substrate and requirements on the interface. Heterogeneous integration schemes may be more flexible on this point.

The third aspect is the choice for a conductive n-doped substrate or a semi-insulating substrate. When an n-doped substrate is used, the negative or ground contact is at the bottom of the substrate and thus common to all components on the chip. This can bring electrical issues caused by common grounding. The other issue is that metal tracks have a large capacitance due to their proximity to a (grounded) conductive plane. The SOA on n-doped substrate has only one top contact that can take the most surface area on the chip. In the example below this is a 130 micrometer wide thick plated gold contact that ensures good current spreading over the amplifier diode. Including the required distance between different metal contacts, this makes the total size of a 500-µm-long SOA building block 546 µm × 149 µm.

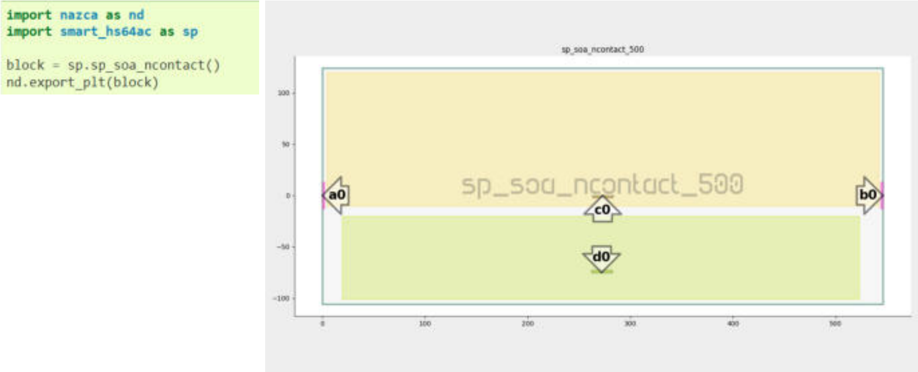
```
import nazca as nd
import smart_hs64ac as sp

block = sp.sp_soa()
nd.export_plt(block)
```

Figure 14.17: Python code snippet to display the image an SOA building block for a 500-µm-long SOA on n-doped substrate. The arrows indicate the connections for the passive waveguide and the electrical connection.

When a semi-insulating substrate is used there will be two contacts from the top. One for the p and one for the n -side of the diode. This enlarges the size of the building block. In the example shown in Figure 14.18 one can see a 500-µm-long SOA building block which has the dimensions on the chip of 546 µm by 230 µm. In addition to this one has to take care of tracking to the two contacts.

Some foundries may have other types of SOA available such as tapered amplifiers.

Other options like curved waveguides for the SOA or multiple contacts to different segments in the SOA are in principle possible, but typically not allowed by generic design



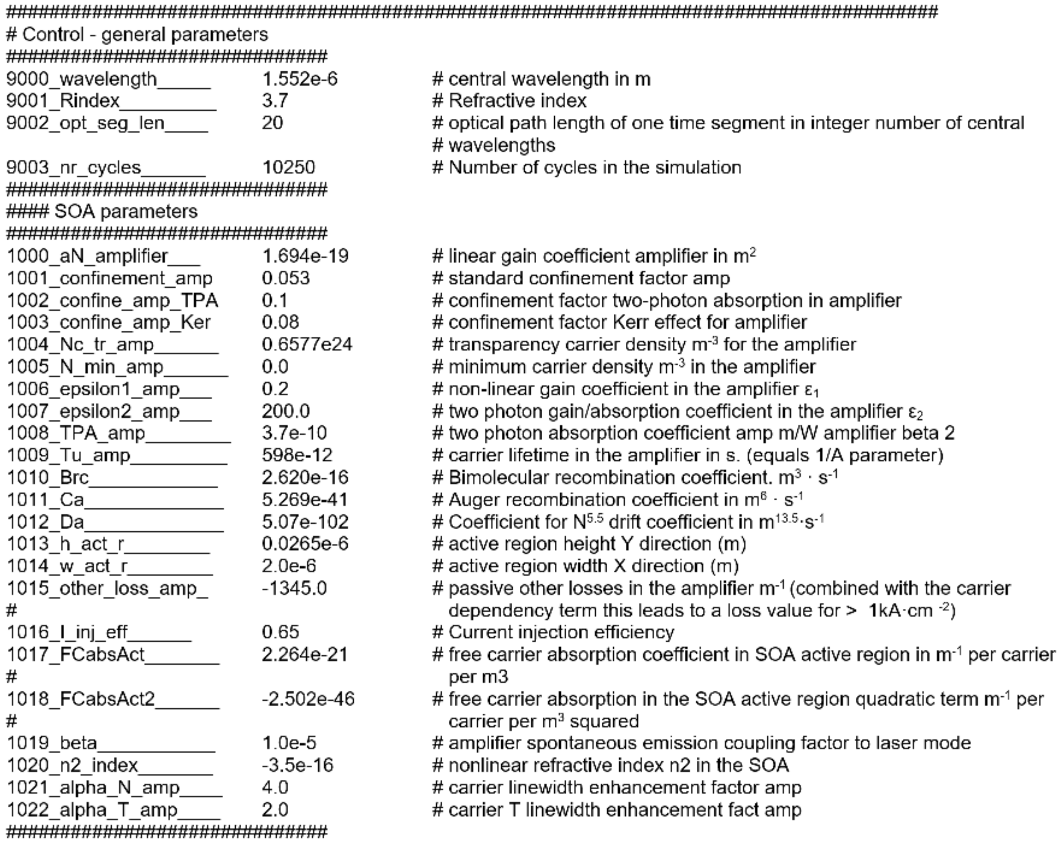
```
import nazca as nd
import smart_hs64ac as sp

block = sp.sp_soa_ncontact()
nd.export_plt(block)
```

Figure 14.18: Python code snippet to display an SOA building block for a 500-µm-long SOA on semi-insulating substrate. The arrows indicate the connections for the passive waveguide and the two electrical connections.

rules. If one wants to realize such devices one needs to contact the foundry and discuss ones wishes. Different deviations are sometimes possible when it involves only changes in mask design without changing the technology. The mask design features should not create any issues in the production for other circuits. In that case one maybe allowed changes in the mask design that go outside the design rules. This may also involve changing the building block itself which of course needs to be done in collaboration with the foundry. The price to pay will be that there is reduced or even absent performance guarantee of the non-generic component.

Appendix 14A

A number of figures in this chapter have been produced using the PHIsim simulator tool [207]. Below the relevant input parameters used in the simulations are given in the format of the parameter input file of the simulator. Each line starts with a string consisting of an identifier consisting of a four digit number and some text, followed by the value and a comment starting with #. A complete input file contains more parameters.



```
##########################################################################################
# Control - general parameters
##################################
9000_wavelength_____      1.552e-6        # central wavelength in m
9001_Rindex_________      3.7             # Refractive index
9002_opt_seg_len____      20              # optical path length of one time segment in integer number of central
                                          # wavelengths
9003_nr_cycles______      10250           # Number of cycles in the simulation
##################################
#### SOA parameters
##################################
1000_aN_amplifier___      1.694e-19       # linear gain coefficient amplifier in m^2
1001_confinement_amp      0.053           # standard confinement factor amp
1002_confine_amp_TPA      0.1             # confinement factor two-photon absorption in amplifier
1003_confine_amp_Ker      0.08            # confinement factor Kerr effect for amplifier
1004_Nc_tr_amp______      0.6577e24       # transparency carrier density m^-3 for the amplifier
1005_N_min_amp______      0.0             # minimum carrier density m^-3 in the amplifier
1006_epsilon1_amp___      0.2             # non-linear gain coefficient in the amplifier ε_1
1007_epsilon2_amp___      200.0           # two photon gain/absorption coefficient in the amplifier ε_2
1008_TPA_amp________      3.7e-10         # two photon absorption coefficient amp m/W amplifier beta 2
1009_Tu_amp_________      598e-12         # carrier lifetime in the amplifier in s. (equals 1/A parameter)
1010_Brc____________      2.620e-16       # Bimolecular recombination coefficient. m^3 · s^-1
1011_Ca_____________      5.269e-41       # Auger recombination coefficient in m^6 · s^-1
1012_Da_____________      5.07e-102       # Coefficient for N^5.5 drift coefficient in m^13.5·s^-1
1013_h_act_r________      0.0265e-6       # active region height Y direction (m)
1014_w_act_r________      2.0e-6          # active region width X direction (m)
1015_other_loss_amp_      -1345.0         # passive other losses in the amplifier m^-1 (combined with the carrier
#                                           dependency term this leads to a loss value for >  1kA·cm ^-2)
1016_I_inj_eff______      0.65            # Current injection efficiency
1017_FCabsAct_______      2.264e-21       # free carrier absorption coefficient in SOA active region in m^-1 per carrier
#                                           per m3
1018_FCabsAct2______      -2.502e-46      # free carrier absorption in the SOA active region quadratic term m^-1 per
#                                           carrier per m^3 squared
1019_beta___________      1.0e-5          # amplifier spontaneous emission coupling factor to laser mode
1020_n2_index_______      -3.5e-16        # nonlinear refractive index n2 in the SOA
1021_alpha_N_amp____      4.0             # carrier linewidth enhancement factor amp
1022_alpha_T_amp____      2.0             # carrier T linewidth enhancement fact amp
##################################
```

Appendix 14B

One can model the current distribution over the length of the SOA with a DC circuit model. The SOA is divided in a large number of segments (20 to 50). Each segment is represented by an ideal diode with series resistance. The series resistance is assumed to be mainly the contact resistance. The segments are contacted at the top by the gold track on top, the resistance of each section between two SOA segments is represented by a resistance. The total current is injected at the top at one end. Current injection in the middle just gives a symmetric profile so one can restrict oneself to describing only half the SOA. One can set a voltage and calculate the current levels in the different segments. The results are highly dependent on the exact value of the contact resistance. The example given below in Figure 14.20 shows what happens with the current density over half the total length of the SOA, for the case where the top gold track is 30 μm wide, 300 nm thick over the length of the SOA. The total current in the 500 μm half SOA is 68 mA. The contact resistance is assumed to be $1 \cdot 10^{-5}\Omega\text{cm}^2$. If this value becomes higher the current distribution becomes more even. For a ten times higher contact resistance, the variation of the current density reduces to just over 0.2 kA/cm^2.

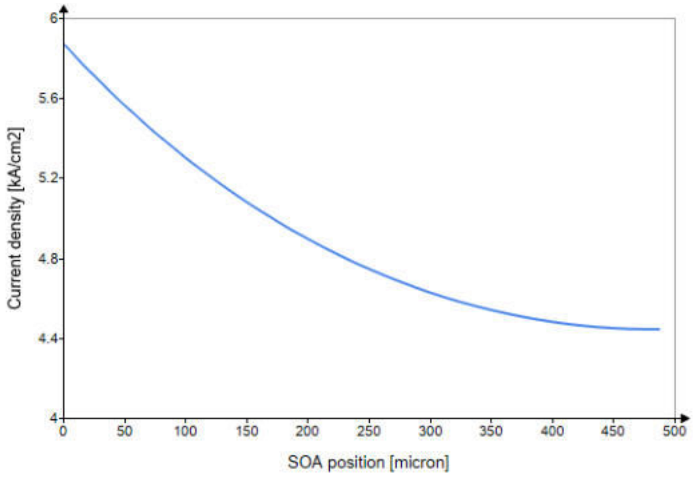


Figure 14.19: Calculated current density in an SOA as a function of distance from the current distribution point at x = 0. The total current is 49 mA. The contact resistance is $1.0 \cdot 10^{-5}$ Ωcm^2.