

Photonics II - Lectures

Lecturer: Rainer Kaltenbaek

Faculty of Mathematics and Physics
University of Ljubljana
IQOQI Vienna
Austrian Academy of Sciences

Ljubljana, Summer Semester 2022

1 Modulating light

In Photonics I, you heard about how light is generated, how it is focused, how it interferes, how its polarization can be modified, and how light is detected. In the following sections, we will discuss how we can modulate light by using electric or magnetic fields or by using sound waves or liquid crystals. In order to discuss these concepts, we will first quickly introduce a few concepts from nonlinear optics.

1.1 A brief introduction to Nonlinear optics

1.1.1 Susceptibility

Later we will discuss the modulation of optical signals by using electric or magnetic fields. Since (classical) light can be described as an electromagnetic wave, an electromagnetic modulation of light is a nonlinear effect. By nonlinear optics we refer to effects where electromagnetic waves modify the optical properties of a material. A well-known example is second-harmonic generation (SHG), where a beam of light with photons of frequency ω is incident on a nonlinear crystal, and part of that incident light is frequency “upconverted” to 2ω . For example, the incident light could have 780 nm and the outgoing wavelength 390 nm. We will discuss this effect in more detail later on.

To describe nonlinear effects more accurately, it is convenient to consider the polarization $P(t)$ of a material system in the presence of an electric field strength $E(t)$. For simplicity, we are only talking about scalar fields at the moment, not about vector fields.

If there is no nonlinearity, the polarization will be proportional to the electric field:

$$P(t) = \epsilon_0 \chi^{(1)} E(t). \quad (1.1)$$

ϵ_0 is the electric permittivity of vacuum, and $\chi^{(1)}$ is the linear **susceptibility**. If there are nonlinear effects, one can generalize equation 1.1 as follows:

$$\begin{aligned} P(t) &= \epsilon_0 [\chi^{(1)} E(t) + \chi^{(2)} E^2(t) + \chi^{(3)} E^3(t) + \dots] \\ &= P^{(1)}(t) + P^{(2)}(t) + P^{(3)}(t) + \dots \end{aligned} \quad (1.2)$$

$\chi^{(k)}$ is called the k -th order susceptibility. The most common effects we will discuss in nonlinear optics are based on the second-order nonlinear polarization. In general, when the field has a vectorial character, the $\chi^{(k)}$ will be tensors.

Equations 1.1 and 1.2 also assumed that the polarization at time t depends instantaneously on the field at time t . This can only be true if there is no loss and no dispersion in the material[1]. In general, there will be loss and dispersion, of course.

The polarization of the material can again be the source of an electromagnetic field (light). Specific examples for that are sum-frequency generation (SFG) and difference-frequency generation (DFG). In the special case where two photons of the incoming beam lead to the creation of one outgoing photon, one calls the effect not SFG but second-harmonic generation (SHG) or frequency doubling, which we mentioned above.

For example, let us consider two fields of frequencies ω_i ($I = 1, 2$) adding up in SFG:

$$E(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} + \text{c.c..} \quad (1.3)$$

Then the second-order contribution to the polarization in our nonlinear medium will be:

$$\begin{aligned} P^{(2)}(t) = & \epsilon_0 \chi^{(2)} E(t)^2 = \epsilon_0 \chi^{(2)} [E_1^2 e^{-2i\omega_1 t} + E_2^2 e^{-2i\omega_2 t} + 2E_1 E_2 e^{-i(\omega_1+\omega_2)t} + \\ & 2E_1 E_2^* e^{-i(\omega_1-\omega_2)t} + \text{c.c.}] + 2\epsilon_0 \chi^{(2)} [E_1 E_1^* + E_2 E_2^*]. \end{aligned} \quad (1.4)$$

If we compare this with the decomposition of the polarization into its frequency components (for simplicity, we denote these with P instead of $P^{(2)}$):

$$P^{(2)}(t) = \sum_n P(\omega_n) e^{-i\omega_n t}, \quad (1.5)$$

we can associate different effects to the various summands:

$$\begin{aligned} P(2\omega_i) &= \epsilon_0 \chi^{(2)} E_i^2 (\text{SHG}) \\ P(\omega_1 + \omega_2) &= \epsilon_0 \chi^{(2)} E_1 E_2 (\text{SFG}) \\ P(\omega_1 - \omega_2) &= \epsilon_0 \chi^{(2)} E_1 E_2^* (\text{DFG}) \\ P(0) &= 2\epsilon_0 \chi^{(2)} (E_1 E_1^* + E_2 E_2^*) (\text{OR}). \end{aligned} \quad (1.6)$$

We will discuss the effect of optical rectification (OR) in a later lecture. Essentially, the mixture of positive and negative frequency parts of the same fields, one gets an effective DC polarization. Usually, this has no noticeable effect, but in the case of ultra-fast pulses passing through a $\chi^{(2)}$ -nonlinear medium, the rapid building up of a DC field, followed by its rapid decay will lead to the emission of THz electromagnetic waves.

We should take note of two things: (1) the notation used for $E(t)$ and $P(t)$ often differ by a factor of 2 between different authors. (2) because $P(t)$ needs to be a real-valued function, $P(\omega)$ needs to fulfill the relation $P(-\omega) = P^*(\omega)$.

1.1.2 Order of magnitude estimates

To get an impression of the strength or weakness of nonlinear optical effects, let us quickly consider some rough numerical estimates.

The linear susceptibility $\chi^{(1)}$ is defined such that it will be of order 1 in condensed matter[1]. To estimate $P^{(2)}$, let us assume that it should be of order 1 if the field amplitude E is on the order of the typical atomic electric field strength $E_{\text{at}} = e/(4\pi\epsilon_0 a_0^2)$. Here, e is the electron charge, and $a = 4\pi\epsilon_0\hbar^2/(me^2)$ is the Bohr radius. This yields $E_{\text{at}} = 6.14 \times 10^{11}$ V/m, and we can then estimate [1]:

$$\chi^{(2)} \approx 1/E_{\text{at}} = 1.94 \times 10^{-12} \text{ m/V}. \quad (1.7)$$

Similarly, one can estimate that $\chi^{(3)}$ should be on the order of $1/E_{\text{at}}^2$ or:

$$\chi^{(3)} \approx 3.78 \times 10^{-24} \text{ m}^2/\text{V}^2. \quad (1.8)$$

1.1.3 The formal definition of the nonlinear susceptibility

The description above was strongly simplified because we only discussed the case of a scalar, real field with no losses and no dispersion. In general, the electric field will be a real function that we can write as a sum of complex frequency components:

$$\vec{E}(\vec{r}, t) = \sum_n \vec{E}_n(\vec{r}, t), \quad (1.9)$$

where the component \vec{E}_n varies with the angular frequency ω_n . Assuming that the field intensity varies on a time scale much slower than optical frequencies, we can write the field as a discrete sum of frequency components:

$$\vec{E}(\vec{r}, t) = \sum_n \vec{E}(\omega_n) e^{-i\omega_n t} = \sum_n \vec{A}(\omega_n) e^{i(\vec{k}_n \cdot \vec{r} - \omega_n t)}. \quad (1.10)$$

The last step assumes that \vec{A} is an amplitude that is varies only slowly in space[1].

We can decompose the polarization vector in a similar way:

$$\vec{P}(\vec{r}, t) = \sum_n \vec{P}(\omega_n) e^{-i\omega_n t}. \quad (1.11)$$

This latter notation is useful to see the frequency dependence between the incident and the outgoing fields. For example, consider two beams incident on a $\chi^{(2)}$ non-linear medium and that all other orders of susceptibility can be assumed to be zero. Then the components of the polarization will be[1]:

$$P_i(\omega_3) = \epsilon_0 \sum_{jk} \sum_{(nm)} \chi_{ijk}^{(2)}(\omega_n + \omega_m, \omega_n, \omega_m) E_j(\omega_n) E_k(\omega_m). \quad (1.12)$$

Here, the sum over n and m runs over all indices **but** under the restriction that $\omega_m + \omega_n$ be equal to the frequency ω_3 of the outgoing field for the case of SFG. To distinguish between the incoming and the outgoing fields, the susceptibility $\chi^{(2)}(\omega_3, \omega_2, \omega_1)$ is sometimes written as $\chi^{(2)}(\omega_3; \omega_2, \omega_1)$ or as $\chi^{(2)}(\omega_3 = \omega_2 + \omega_1)$.

1.1.4 The nonlinear susceptibility of a classical anharmonic oscillator

The **Lorentz model** of an atom provides a very good description of the **linear** optical response of some materials like e.g. atomic vapours or non-metallic solids. In that model, the electrons in an atom move in a harmonic oscillator potential. Here, we will discuss what we can learn about non-linear optical effects if we extend the Lorentz model by adding nonlinear terms to the potential of the electrons. An important observation is the influence of symmetry considerations on the nonlinear optical response of a medium. For example, the response depends on whether a medium is **centrosymmetric or noncentrosymmetric**. In both cases, we will consider the lowest-order nonlinear deviations from the harmonic potential.

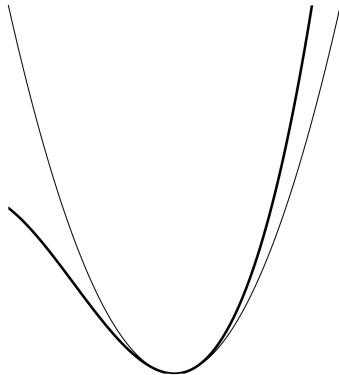


Figure 1: **A noncentrosymmetric potential.** The thin line shows a parabola, the thicker line shows a potential that also includes a third-order contribution. The latter is noncentrosymmetric.

Noncentrosymmetric media

We will see that a noncentrosymmetric potential will allow to describe nonlinear effects like SFG and DFG. Fig. 1 shows an example for a noncentrosymmetric potential. To the lowest order, the motion of an electron in a noncentrosymmetric potential is given by:

$$\frac{d^2x(t)}{dt^2} + 2\gamma \frac{dx(t)}{dt} + \omega_0^2 x(t) + ax^2(t) = -e \frac{E(t)}{m}. \quad (1.13)$$

e and m are the elementary charge and the mass of an electron, respectively. The parameters ω_0 and a describe the shape of the potential. ω_0 is the resonance frequency if we had a simple harmonic potential, and a describes the strength of the anharmonic term. γ describes how strongly the oscillations are damped.

For DFG/SFG, we overlap two fields in the crystal:

$$E(t) = E_1 e^{-i\omega_1 t} + E_2 e^{-i\omega_2 t} + \text{c.c.} \quad (1.14)$$

We can put that into equation 1.13 and solve it using a perturbation expansion if the fields are sufficiently weak. Then equation 1.13 becomes:

$$\frac{d^2x(t)}{dt^2} + 2\gamma \frac{dx(t)}{dt} + \omega_0^2 x(t) + ax^2(t) = -\lambda e \frac{E(t)}{m}, \quad (1.15)$$

where the small positive number λ is the strength of the perturbation. We then look for a solution that we can express as a power series in λ :

$$x(t) = \sum_{k=1}^{\infty} \lambda^k x^{(k)}(t). \quad (1.16)$$

We can insert this ansatz on the left-hand side of equation 1.15, and then the equation must be fulfilled separately for each order of λ^k . This yields a system of equations[1]:

$$\frac{d^2x^{(1)}(t)}{dt^2} + 2\gamma \frac{dx^{(1)}(t)}{dt} + \omega_0^2 x^{(1)}(t) = -e \frac{E(t)}{m} \quad (1.17a)$$

$$\frac{d^2x^{(2)}(t)}{dt^2} + 2\gamma \frac{dx^{(2)}(t)}{dt} + \omega_0^2 x^{(2)}(t) + a [x^{(1)}(t)]^2 = 0, \text{ etc.} \quad (1.17b)$$

The lowest-order equation 1.17a corresponds to a harmonic oscillator equation as in the Lorentz model. It describes the limiting case of the absence of non-linear effects. The steady-state solution (not including a decay of amplitudes due to attenuation) is:

$$x^{(1)}(t) = x^{(1)}(\omega_1) e^{-i\omega_1 t} + x^{(1)}(\omega_2) e^{-i\omega_2 t} + \text{c.c.}, \quad (1.18)$$

where the amplitudes can be written as[1]:

$$x^{(1)}(\omega_j) = -\frac{e}{m} \frac{E_j}{D(\omega_j)} \quad (1.19)$$

with $D(\omega_j) = \omega_0^2 - \omega_j^2 - 2i\omega_j\gamma$ being the “complex denominator function”[1].

Putting in that solution into equation 1.17b, one can get the lowest-order correction term $x^{(2)}$. The squaring of $x^{(1)}(t)$ in that equation will result in frequency contributions with $\pm 2\omega_1$, $\pm 2\omega_2$, $\pm(\omega_1 + \omega_2)$, $\pm(\omega_1 - \omega_2)$, and 0.

For example, one gets the following solution for SFG[1]:

$$x^{(2)}(\omega_1 + \omega_2) = -\frac{2a(e/m)^2 E_1 E_2}{D(\omega_1 + \omega_2) D(\omega_1) D(\omega_2)}. \quad (1.20a)$$

During today's exercises, we will express solutions of this kind in terms of the first and second-order susceptibilities by using the relation:

$$P^{(k)}(\omega) = -Nex^{(k)}(\omega). \quad (1.21)$$

The above considerations show that the lowest-order non-linear effects in a noncentrosymmetric medium are of second order.

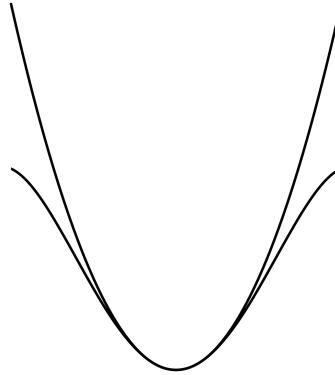


Figure 2: **A centrosymmetric potential.** The thin line shows a parabola, the thicker line shows a potential that also includes a fourth-order contribution. Both are centrosymmetric.

Centrosymmetric media

Let us now make similar considerations for centrosymmetric media. Fig. 2 shows an example for centrosymmetric potentials. To describe a nonlinear centrosymmetric medium, we will add a third-order term to the restoring force acting on our electron. The corresponding equation of motion will then be:

$$\frac{d^2x(t)}{dt^2} + 2\gamma \frac{dx(t)}{dt} + \omega_0^2 x(t) - bx^3(t) = -e \frac{E(t)}{m}. \quad (1.22)$$

b quantifies the strength of the nonlinearity.

In three dimensions, that becomes[1]:

$$\frac{d^2\vec{r}(t)}{dt^2} + 2\gamma \frac{d\vec{r}(t)}{dt} + \omega^2 \vec{r}(t) - b [\vec{r}(t) \cdot \vec{r}(t)] \vec{r}(t) = -\frac{e}{m} \vec{E}(t). \quad (1.23)$$

Now let us assume that we combine multiple electric fields inside our nonlinear medium. For simplicity, we assume that the fields have the same polarization but can have different frequencies:

$$\vec{E}(t) = \sum_n \vec{\omega}_n e^{-i\omega_n t}. \quad (1.24)$$

We can make a similar perturbation approach as we did before. For that purpose, one replaces $\vec{E}(t)$ in equation 1.22 with $\lambda \vec{E}(t)$. λ is the perturbation parameter.

Again, we assume the solution to be a power series in λ :

$$\vec{r}(t) = \sum_{k=1}^{\infty} \lambda^k \vec{r}^{(k)}, \quad (1.25)$$

and we get a system of differential equations like we did earlier. For the lowest orders:

$$\frac{d^2 \vec{r}^{(1)}(t)}{dt^2} + 2\gamma \frac{d\vec{r}^{(1)}(t)}{dt} + \omega_0^2 \vec{r}^{(1)}(t) = -\frac{e\vec{E}(t)}{m}, \quad (1.26a)$$

$$\frac{d^2 \vec{r}^{(2)}(t)}{dt^2} + 2\gamma \frac{d\vec{r}^{(2)}(t)}{dt} + \omega_0^2 \vec{r}^{(2)}(t) = 0, \text{ etc.} \quad (1.26b)$$

The first-order solution is similar to what we had in the noncentrosymmetric case, but the interesting difference is the form of equation 1.26b for the 2nd-orders solution. That equation represents a local oscillator with damping but without a driving force. For that reason, the steady-state solution will be $\vec{r}^{(2)}(t) = 0$. That means that the lowest-order nonlinear effects we will see in a centrosymmetric medium are of at least third order. That means **there are no second-order nonlinear optical effects in a centrosymmetric material**.

1.1.5 Properties of the nonlinear susceptibility

Symmetries in the nonlinear susceptibility help reducing the number of independent coefficients. That can be especially important for higher-order effects. For example, consider the 2nd order susceptibility $\chi^{(2)}$. The second-order nonlinear polarization is given by:

$$P_i(\omega_3) = \epsilon_0 \sum_{jk} \sum_{(nm)} \chi_{ijk}^{(2)}(\omega_n + \omega_m, \omega_n, \omega_m) E_j(\omega_n) E_k(\omega_m), \quad (1.27)$$

where (nm) again means that we require $\omega_3 = \omega_1 + \omega_2$. To characterize all interactions, we therefore need to determine the six tensors

$$\begin{aligned} \chi_{ijk}^{(2)}(\omega_1, \omega_3, -\omega_2), & \quad \chi_{ijk}^{(2)}(\omega_1, -\omega_2, \omega_3), & \chi_{ijk}^{(2)}(\omega_2, \omega_3, -\omega_1), \\ \chi_{ijk}^{(2)}(\omega_2, -\omega_1, \omega_3), & \quad \chi_{ijk}^{(2)}(\omega_3, \omega_1, \omega_2), & \chi_{ijk}^{(2)}(\omega_3, \omega_2, \omega_1), \end{aligned}$$

and six more tensors where all the frequencies have the opposite sign. Each of these 12 tensors has 27 components, which means we would have to know 324 complex numbers. For higher-order effects, that is even worse, but already for the 2nd order susceptibility we should take advantage of any symmetries we can find.

Reality of the fields

Because the polarization is a measurable quantity, it must be real, and therefore we must have:

$$P_i(-\omega_n - \omega_m) = P_i(\omega_n + \omega_m)^*. \quad (1.28)$$

A similar relation holds for the electric fields, and it follows that:

$$\chi_{ijk}^{(2)}(-\omega_n - \omega_m, -\omega_n, -\omega_m) = \chi_{ijk}^{(2)}(\omega_n + \omega_m, \omega_n, \omega_m)^*. \quad (1.29)$$

Intrinsic permutation symmetry

If we have products of multiple fields like $E_j(\omega_m)E_k(\omega_n)$, it is arbitrary, which of the fields we denote with $E_j(\omega_m)$ and which with $E_k(\omega_n)$. As an example, let us consider the contribution of the product:

$$\chi_{ijk}^{(2)}(\omega_n + \omega_m, \omega_n, \omega_m)E_j(\omega_n)E_k(\omega_m)$$

to the nonlinear polarization $P_i(\omega_n + \omega_m)$ in equation 1.27. Because j, k, n , and m are summation indices, we could also have written:

$$\chi_{ikj}^{(2)}(\omega_n + \omega_m, \omega_m, \omega_n)E_k(\omega_n)E_j(\omega_m).$$

These two terms will be equal if we require that the nonlinear susceptibility is symmetric under interchange of its last two frequency arguments and its last two Cartesian indices[1]

$$\chi_{ijk}^{(2)}(\omega_n + \omega_m, \omega_n, \omega_m) = \chi_{ikj}^{(2)}(\omega_n + \omega_m, \omega_m, \omega_n). \quad (1.30)$$

This *intrinsic permutation symmetry* means that it does not matter which field we denote as the first and which as the second in products like $E_j(\omega_n)E_k(\omega_m)$. Alternatively, we could have set one of the two components of $\chi^{(2)}$ equal to zero and doubled the value of the component with exchanged indices. The overall sum would still be the same[1].

Symmetries for lossless media

For lossless media, two additional symmetries occur. The first of these symmetries states that all of the components of $\chi_{ijk}^{(2)}(\omega_n + \omega_m, \omega_n, \omega_m)$ are real. That this is true for a lossless medium can, e.g., be shown by doing the appropriate limit in the quantum-

mechanical description[1]. Of course, no medium is perfectly lossless, but it will be a good approximation if the frequencies involved are far from resonance.

The second symmetry for lossless media is *full* permutation symmetry. This means that all the frequency arguments can be interchanged arbitrarily if one interchanges the indices of the susceptibility in the same order. One has to adapt the sign accordingly if one interchanges the frequency arguments to ensure that the first frequency argument remains the sum of the other two. For example:

$$\chi_{ijk}^{(2)}(\omega_3 = \omega_1 + \omega_2) = \chi_{jki}^{(2)}(-\omega_1 = \omega_2 - \omega_3). \quad (1.31)$$

Kleinman's symmetry

In many cases, the optical signals sent into the nonlinear medium have optical frequencies ω_i far from any relevant resonance frequency ω_0 of the material. If that is the case, the susceptibility will not depend on the frequency. That means there is no dispersion. Because we assume here that the optical frequencies are far from material resonances, the medium will be lossless, and full permutation symmetry will apply as in equation 1.31. In addition to that, if one assumes that the susceptibility does not depend on the frequency, we can interchange the indices of the susceptibility *without* interchanging the frequencies. This is known as *Kleinman's symmetry*:

$$\chi_{ijk}^{(2)} = \chi_{jki}^{(2)} = \chi_{kij}^{(2)} = \chi_{ikj}^{(2)} = \chi_{jik}^{(2)} = \chi_{kji}^{(2)}. \quad (1.32)$$

1.1.6 Contracted notation

Let us take Kleinman's symmetry into account in our notation and introduce the tensor:

$$d_{ijk} = \frac{1}{2}\chi_{ijk}^{(2)}. \quad (1.33)$$

The factor 1/2 is a historic convention[1].

Then we can write the nonlinear polarization as:

$$P_i(\omega_n + \omega_m) = \epsilon_0 \sum_{jk} \sum_{(nm)} 2d_{ijk} E_j(\omega_n) E_k(\omega_m). \quad (1.34)$$

If Kleinman's symmetry holds, one can assume that d_{ijk} is symmetric in its last two indices. This is always true in the case of SHG because we have $\omega_n = \omega_m$.

We can then contract our notation by writing d_{il} with only two indices instead of three, where the indices jk are combined into the index l using the following prescription:

$$\begin{array}{llllll} jk : & 11 & 22 & 33 & 23, 32 & 31, 13 & 12, 21 \\ l : & 1 & 2 & 3 & 4 & 5 & 6 \end{array} \quad (1.35)$$

Because of Kleinman's symmetry, not all components of d_{il} are independent. For example, one has:

$$\begin{aligned} d_{12} &\equiv d_{122} = d_{212} \equiv d_{26} \text{ and} \\ d_{14} &\equiv d_{123} = d_{213} \equiv d_{25}. \end{aligned} \quad (1.36)$$

By applying these arguments for all components, one can show that:

$$d_{il} = \begin{bmatrix} d_{11} & d_{12} & d_{13} & d_{14} & d_{15} & d_{16} \\ d_{16} & d_{22} & d_{23} & d_{24} & d_{14} & d_{12} \\ d_{15} & d_{24} & d_{33} & d_{23} & d_{13} & d_{14} \end{bmatrix}. \quad (1.37)$$

1.1.7 Effective dielectric constant

If the directions of propagation and the polarizations of the fields are fixed, we can simplify the expression for the nonlinear polarization for SFG and write:

$$P(\omega_3) = 4\epsilon_0 d_{\text{eff}} E(\omega_1) E(\omega_2). \quad (1.38)$$

For SHG, this simplifies further to:

$$P(2\omega) = 2\epsilon_0 d_{\text{eff}} E(\omega)^2, \quad (1.39)$$

where $E(\omega) = |\vec{E}(\omega)|$, and $P(\omega) = |\vec{P}(\omega)|$.

One can calculate d_{eff} for each crystal class. For example, in the case that the two lower-frequency waves have the same polarization and assuming that the medium is a negative uniaxial crystal of crystal class $3m$, one gets[1]:

$$d_{\text{eff}} = d_{31} \sin \theta - d_{22} \cos \theta \sin 3\phi. \quad (1.40)$$

θ is the angle between the direction of propagation and the crystal's z axis, and ϕ is the azimuthal angle between the propagation direction and the xz plane of the crystal.

1.1.8 Spatial symmetry of a nonlinear medium

Apart from the general symmetry considerations we discussed earlier, additional symmetries in the susceptibility tensor result from symmetry properties of the optical medium. All crystals can be categorized into 32 crystal classes, and the symmetry properties of each of these classes then result in corresponding symmetry properties for the susceptibility tensor. While we will not go into the detailed meaning of how these crystal classes are defined, suffice it to say that the crystal classes can be derived from group-theoretical considerations. In particular, the 32 crystal classes are related to the point group symmetry of crystals.

Triclinic	$\begin{bmatrix} xx & xy & xz \\ yx & yy & yz \\ zx & zy & zz \end{bmatrix}$
Monoclinic	$\begin{bmatrix} xx & 0 & xz \\ 0 & yy & 0 \\ zx & 0 & zz \end{bmatrix}$
Orthorhombic	$\begin{bmatrix} xx & 0 & 0 \\ 0 & yy & 0 \\ 0 & 0 & zz \end{bmatrix}$
Tetragonal	$\begin{bmatrix} xx & 0 & 0 \\ 0 & xx & 0 \\ 0 & 0 & zz \end{bmatrix}$
Trigonal	$\begin{bmatrix} xx & 0 & 0 \\ 0 & xx & 0 \\ 0 & 0 & xx \end{bmatrix}$
Hexagonal	$\begin{bmatrix} xx & 0 & 0 \\ 0 & xx & 0 \\ 0 & 0 & xx \end{bmatrix}$
cubic/isotropic	$\begin{bmatrix} xx & 0 & 0 \\ 0 & xx & 0 \\ 0 & 0 & xx \end{bmatrix}$

Table 1.1: The form of the linear susceptibility $\chi^{(1)}$ for all seven crystal classes as well as for isotropic media. Each non-vanishing component is denoted by its cartesian indices. The table structure and content reproduce a table in Ref. [1].

1.1.9 Influence of spatial symmetry on the linear optical properties

We will first discuss the influence of spatial symmetry on the first-order susceptibility. The symmetry properties of the crystal result in symmetries of the $\chi^{(1)}$ susceptibility tensor. Table 1.1 illustrates these symmetries for all seven crystal classes as well as for isotropic media.

1.1.10 Influence of spatial symmetry on 2nd-order nonlinearity

Similar to the consequences of inversion symmetry, which we discussed earlier, other symmetries of the various crystal classes will also result in symmetries of the nonlinear susceptibility. These properties are summarized in Table 1.2. One can also display these symmetry properties graphically. Table 1.3 shows a few examples of such graphical representations.

1.2 The electro-optic effect

In some materials, the refractive index can be changed by applying an electric field. This is called the linear electro-optic effect or the Pockels effect. It is a convenient way to quickly modulate the amplitude of polarized light or to generate frequency side-bands.

Crystal system	Crystal class	Nonvanishing tensor elements
Triclinic	$1 = C_1$	All elements are independent and nonzero
	$\bar{1} = S_2$	Each element vanishes
Monoclinic	$2 = C_2$	$xyz, xzy, xxy, xyx, yxx, yyy, yzz, yzx, yxz, zyz, zzy, xxy, zyx$ (twofold axis parallel to \hat{y})
	$m = C_{1h}$	$xxx, xyy, xzz, xzx, xxz, yyz, yzy, yxy, yyx, zxz, zyy, zzz, zxz$ (mirror plane perpendicular to \hat{y})
Orthorombic	$2/m = C_{2h}$	Each element vanishes
	$222 = D_2$	xyz, xzy, yzx, zxy, zyx
	$mm2 = C_{2v}$	$xzx, xxz, yyz, yzy, zxz, zyy, zzz$
	$mmm = D_{2h}$	Each element vanishes
Tetragonal	$4 = C_4$	$yyz = -yxz, xzy = -yzx, xzx = yzy, xxz = yyz, zxz = zyy, zzz, zxz = -zyy, zxy = -zyx$
	$\bar{4} = S_4$	$xyz = -yxz, xzy = yzx, xzx = -yzy, xxz = -yyz, zxz = -zyy, zxy = zyx$
	$422 = D_4$	$xyz = -yxz, xzy = -yzx, zxy = -zyx$
	$4mm = C_{4v}$	$xzx = yzy, xxz = yyz, zxz = zyy, zzz$
	$\bar{4}2m = D_{2d}$	$xyz = yxz, xzy = yzx, zxy = zyx$
	$4/m = C_{4h}$	Each element vanishes
	$4/mmm = D_{4h}$	Each element vanishes
Cubic	$432 = O$	$xyz = -xzy = yzx = -yxz = zxy = -zyx$
	$\bar{4}3m = T_d$	$xyz = xzy = yzx = yxz = zxy = zyx$
	$23 = T$	$xyz = yzx = zxy = xzy, xzy = yxz = zyx$
	$m3 = T_h, m3m = O_h$	Each element vanishes
Trigonal	$3 = C_3$	$zxz = xyy = -yyz = -xyy, xyz = -yxz, xzy = -yzx, xzx = yzy, xxz = yyz, zxz = zyy, zzz, zxz = -zyy, zxy = -zyx$
	$32 = D_3$	$zxz = -xyy = -yyx = -yxy, xyz = -yxz, xzy = -zyx, xyz = -yzx, zxy = -zyx$
	$3m = C_{3v}$	$xzx = yzy, xxz = yyz, zxz = zyy, zzz, yyy = -yxx = -xyy = -xyx$ (mirror plane perpendicular to \hat{x})
Hexagonal	$\bar{3} = S_6, \bar{3}m = D_{3d}$	Each element vanishes
	$6 = C_6$	$xyz = -yxz, xzy = -yzx, xzx = yzy, xxz = yyz, zxz = zyy, zzz, zxz = -zyy, zxy = -zyx$
	$\bar{6} = C_{3h}$	$xxx = -xyy = -yxy = -yyx, yyy = -yxx = -xyx = -xyy$
	$622 = D_6$	$xyz = -yxz, xzy = -yxz, zxy = -zyx$
	$6mm = C_{6v}$	$xzx = yzy, xxz = yyz, zxz = zyy, zzz$
	$\bar{6}m2 = D_{3h}$	$yyy = -yxx = -xyy = -xyx$
	$6/m = C_{6h}$	Each element vanishes
	$6/mmm = D_{6h}$	Each element vanishes

Table 1.2: Non-vanishing components for each of the 32 crystal classes. Table structure and content from Ref. [1].

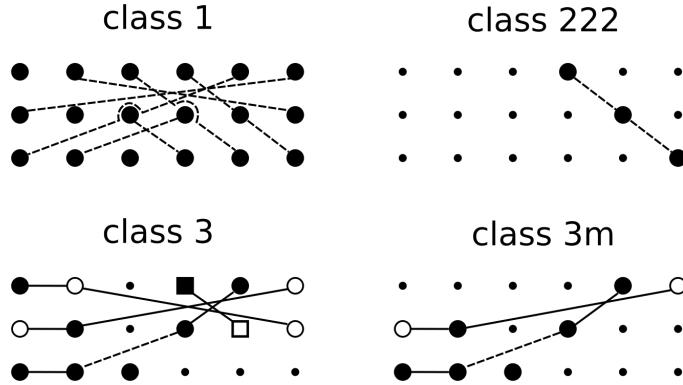


Table 1.3: Form of the d_{il} matrix for some of the crystal classes lacking inversion symmetry. The upper two classes correspond to biaxial crystals, the lower two classes are uniaxial crystals. Small dot: zero coefficient; large dot: nonzero coefficient; square: coefficient that is zero when Kleinman's symmetry is fulfilled; connected symbols: numerically equivalent coefficients, but if the symbol is empty, the sign is opposite. Dashed connections are only valid if Kleinmann's symmetry is valid. Figures based on Ref.[1].

We can describe the linear electro-optic effect using the 2nd-order susceptibility. If the applied field is DC, the polarization of the medium will be:

$$P_i(\omega) = 2\epsilon_0 \sum_{jk} \chi_{ijk}^{(2)}(\omega = \omega + 0) E_j(\omega) E_k(0). \quad (1.41)$$

As we discussed earlier, 2nd-order non-linear effects can only occur in noncentrosymmetric media. While it is possible to describe the linear electro-optic effect using this description, due to historic convention we will use a different formalism, which we will introduce in subsection 1.2.1.

1.2.1 The linear electro-optic effect

In an anisotropic medium, the electric field \vec{E} and the electric displacement field \vec{D} are related via the dielectric permeability tensor ϵ_{ij} :

$$D_i = \epsilon_0 \sum_j \epsilon_{ij} E_j. \quad (1.42)$$

For a lossless and not optically active medium, ϵ_{ij} is symmetric and real, and it will have 6 independent elements: ϵ_{xx} , ϵ_{yy} , ϵ_{zz} , $\epsilon_{xy} = \epsilon_{yx}$, $\epsilon_{xz} = \epsilon_{zx}$, and $\epsilon_{yz} = \epsilon_{zy}$. Sym-

metric, real matrices can be brought into diagonal form by using an orthogonal transformation to an appropriate coordinate system with coordinates (X, Y, Z) [1]. Equation 1.42 becomes:

$$\begin{bmatrix} D_X \\ D_Y \\ D_Z \end{bmatrix} = \epsilon_0 \begin{bmatrix} \epsilon_{XX} & 0 & 0 \\ 0 & \epsilon_{YY} & 0 \\ 0 & 0 & \epsilon_{ZZ} \end{bmatrix} \begin{bmatrix} E_X \\ E_Y \\ E_Z \end{bmatrix}. \quad (1.43)$$

This coordinate system is called the **principal-axis system**.

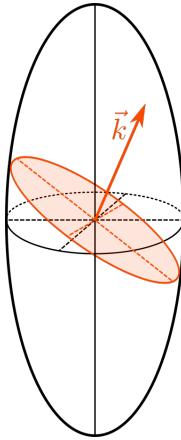


Figure 3: **The optical indicatrix.** Equation 1.44 defines an ellipsoid. If one cuts the ellipsoid with a plane that is perpendicular to \vec{k} and passes through the origin, one gets an ellipse (indicated in red). The semimajor and semiminor axes of the ellipse are drawn as red, dashed lines. They give the two allowed values of the refractive index for a given propagation direction and polarization.

Let us consider the energy density per unit volume inside the medium:

$$U = \frac{1}{2} \vec{D} \cdot \vec{E} = \frac{1}{2} \epsilon_0 \sum_{ij} \epsilon_{ij} E_i E_j = \frac{1}{2\epsilon_0} \left[\frac{D_X^2}{\epsilon_{XX}} + \frac{D_Y^2}{\epsilon_{YY}} + \frac{D_Z^2}{\epsilon_{ZZ}} \right]. \quad (1.44)$$

From this equation, we see that surfaces of constant energy density are ellipsoids. If we define the quantities:

$$X = \left(\frac{1}{2\epsilon_0 U} \right)^{1/2} D_X, Y = \left(\frac{1}{2\epsilon_0 U} \right)^{1/2} D_Y, \quad Z = \left(\frac{1}{2\epsilon_0 U} \right)^{1/2} D_Z, \quad (1.45)$$

equation 1.44 becomes:

$$\frac{X^2}{\epsilon_{XX}} + \frac{Y^2}{\epsilon_{YY}} + \frac{Z^2}{\epsilon_{ZZ}} = 1. \quad (1.46)$$

The ellipsoid defined by this equation is known as the *optical indicatrix* or *index ellipsoid*.

In a general coordinate system, we can write:

$$\begin{aligned} \left(\frac{1}{n^2}\right)_1 x^2 + \left(\frac{1}{n^2}\right)_2 y^2 + \left(\frac{1}{n^2}\right)_3 z^2 + \left(\frac{1}{n^2}\right)_4 yz + \\ \left(\frac{1}{n^2}\right)_5 xz + \left(\frac{1}{n^2}\right)_6 xy = 1, \end{aligned} \quad (1.47)$$

where the coefficients $(1/n^2)_i$ are optical constants describing the indicatrix in this new coordinate system. They can be written in terms of ϵ_{XX} , ϵ_{YY} , and ϵ_{ZZ} by using standard coordinate transformation rules. Fig. 3 illustrates the indicatrix.

If we apply an electric field to our material, the shape of the indicatrix can change due to the linear electro-optic effect. One can conveniently describe this change by using the *impermeability tensor* η_{ij} , which we define via the relation:

$$E_i = \frac{1}{\epsilon_0} \sum_j \eta_{ij} D_j. \quad (1.48)$$

This is the inverse of equation 1.42. That means, we have $\eta_{ij} = (\epsilon^{-1})_{ij}$.

Let us express our indicatrix in terms of the impermeability tensor. The energy density can be written as $U = (2\epsilon_0) \sum_{ij} \eta_{ij} D_i D_j$. If we transform to a new coordinate system x, y, z fulfilling $x = D_x/(2\epsilon_0 U)^{1/2}$ etc, we get:

$$1 = \eta_{11} x^2 + \eta_{22} y^2 + \eta_{33} z^2 + 2\eta_{12} xy + 2\eta_{23} yz + 2\eta_{13} xz. \quad (1.49)$$

If we compare this with our equation 1.47 for the indicatrix in a general coordinate system, we get:

$$\begin{aligned} \left(\frac{1}{n^2}\right)_1 = \eta_{11}, & \quad \left(\frac{1}{n^2}\right)_2 = \eta_{22}, \left(\frac{1}{n^2}\right)_2 = \eta_{33}, \\ \left(\frac{1}{n^2}\right)_4 = \eta_{23} = \eta_{32}, & \quad \left(\frac{1}{n^2}\right)_5 = \eta_{13} = \eta_{31}, \left(\frac{1}{n^2}\right)_6 = \eta_{12} = \eta_{21}. \end{aligned} \quad (1.50)$$

Let us express η_{ij} as a power series of the components E_k of the electric field:

$$\eta_{ij} = \eta_{ij}^{(0)} + \sum_k r_{ijk} E_k + \sum_{kl} s_{ijkl} E_k E_l + \dots \quad (1.51)$$

Here, we introduced the tensor r_{ijk} , which describes the linear electro-optic effect, and the tensor s_{ijkl} , which describes the quadratic electro-optic effect. We will only consider effects up to 2nd order.

Because ϵ_{ij} is real and symmetric, the same will be true for η_{ij} , and the tensor r_{ijk} must be symmetric in its first two indices. s_{ijkl} is symmetric in its first two and in its last two indices due to inherent permutation symmetry. Due to these symmetries, we

can again introduce a contracted notation, where we replace two indices ij with one index l according to the prescription we had in equation 1.35.

Then we can replace r_{ijk} with r_{lk} , and we can express deviations of the optical constants $(1/n^2)_l$ from their zero-order form as:

$$\Delta \left(\frac{1}{n^2} \right)_l = \sum_k r_{lk} E_k, \quad (1.52)$$

where the cartesian indices ij correspond to the contracted index l . This is true if we neglect contributions higher than linear order in equation 1.51. The quantities r_{lj} are called electro-optic coefficients and describe the rate with which the coefficients $(1/n^2)_l$ change as we apply an electric field to our material.

We noted earlier that symmetry considerations can help us in describing nonlinear optical effects. As an example, consider the form of the electro-optic coefficients for Lithium Niobate (LiNbO_3). This is a crystal of class $3m$ that is often used in optical experiments. Due to the symmetry properties of this crystal class, we have:

$$r_{lj} = \begin{bmatrix} 0 & -r_{22} & r_{13} \\ 0 & r_{22} & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{42} & 0 \\ r_{42} & 0 & 0 \\ r_{22} & 0 & 0 \end{bmatrix} \quad (\text{for class } 3m) \quad (1.53)$$

1.2.2 Electro-optic modulators

To illustrate the electro-optic effect, let us consider an example: let us use a Potassium dihydrogen phosphate (KH_2PO_4 or KDP) crystal, which is a uniaxial crystal of the crystal class $\bar{4}2m$. r_{lj} takes the form:

$$r_{lj} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ r_{41} & 0 & 0 \\ 0 & r_{41} & 0 \\ 0 & 0 & r_{63} \end{bmatrix} \quad (\text{for class } \bar{4}2m). \quad (1.54)$$

If we do not apply an electric field to the crystal, the indicatrix is given by the following equation in the principle-axis coordinate system:

$$\frac{X^2}{n_o^2} + \frac{Y^2}{n_o^2} + \frac{Z^2}{n_e^2} = 1. \quad (1.55)$$

n_o and n_e are the ordinary and extraordinary refractive indices of the crystal, respectively.

If we apply an electric field, the shape of the indicatrix will be given by the equation:

$$\frac{X^2}{n_o^2} + \frac{Y^2}{n_o^2} + \frac{Z^2}{n_e^2} + 2r_{41}E_XYZ + 2r_{41}E_YXZ + 2r_{63}E_ZXY = 1. \quad (1.56)$$

For KDP and $\lambda = 546.1$ nm, we have $r_{41} = 8.77$ pm/V, $r_{63} = 10.5$ pm/V, $n_o = 1.514$, and $n_e = 1.472$ [1].

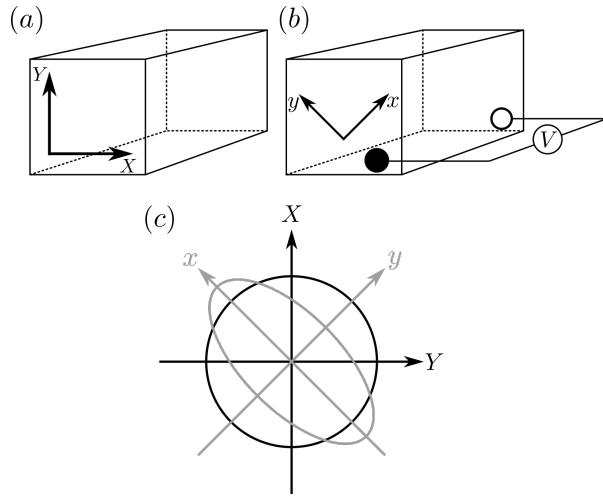


Figure 4: **Linear electro-optic effect in KDP.** (a) shows the orientation of the principal axes in the crystal. (b) When we apply an electric field, we can diagonalize the indicatrix in a rotated coordinate system, indicated by x, y . (c) This figure shows a cut through the indicatrix perpendicular to the Z (or z) axis. Without a field, the refractive index is independent of the direction as indicated by the black circle. With an applied field, the cut becomes an ellipse with refractive indices n_x and n_y along the semimajor and semiminor axis, respectively.

We can see that equation 1.57 contains off-diagonal terms. That means, applying an electric voltage does not just change n_o and n_e . The orientation of the indicatrix will depend on the orientation of the electric field applied. A special and commonly encountered case is that the field is applied along one of the principal axes of the crystal. For example, assume the E-field is applied along the Z axis. Then the equation for the indicatrix simplifies to:

$$\frac{X^2}{n_o^2} + \frac{Y^2}{n_o^2} + \frac{Z^2}{n_e^2} + 2r_{63}E_ZXY = 1. \quad (1.57)$$

In the exercises, we will find a new principal-axis system with coordinates x, y, z such

that equation 1.57 becomes:

$$\frac{x^2}{n_x^2} + \frac{y^2}{n_y^2} + \frac{z^2}{n_e^2} = 1. \quad (1.58)$$

The refractive indices along the three coordinate axes are given by $n_z = n_e$ and:

$$\begin{aligned} n_x &= n_o - \frac{1}{2} n_o^3 r_{63} E_z \\ n_y &= n_o + \frac{1}{2} n_o^3 r_{63} E_z. \end{aligned} \quad (1.59)$$

Fig. 4 illustrates the action of an applied field $E/V/L$ on the orientation of the principal axes, where V is the applied voltage, and L is the length of the crystal. The shape of the electrodes in Fig. 4(b) should not be taken literally. For the figure, it was assumed that $r_{63} E_z$ is negative. In that case, $n_y < n_x$.

Let us now consider an optical beam travelling through the crystal along its z axis. If the light has a polarization that is not along the x or y axis, the components of the E-field along these two axes will acquire a phase difference Γ because of the difference in the refractive indices:

$$\Gamma = (n_y - n_x) \frac{\omega L}{c}. \quad (1.60)$$

This phase difference is called retardation. It introduces a rotation in the polarization of the light sent through. If we insert the expressions from equation 1.59, we find:

$$\Gamma = \frac{n_0^3 r_{63} E_z \omega L}{c} = \frac{n_0^3 r_{63} V \omega}{c}. \quad (1.61)$$

That means, the retardation here depends only on the voltage applied.

Such devices are called electro-optic modulators (EOMs) or Pockels cells. If we apply the correct voltage to our crystal, we can achieve a π phase shift between the polarization components along x and y . This voltage is called the half-wave voltage:

$$V_{\lambda/2} = \frac{\pi c}{\omega n_o^3 r_{63}}. \quad (1.62)$$

The half-wave voltage for the configurations we discussed here are typically several kV or even 10 kV for visible or near-infrared light.

Intensity modulation using an EOM

In today's exercises, we will discuss one common application of EOMs. By placing an EOM between crossed polarizers, one can modulate the intensity of the light passing through by changing the voltage applied to the EOM.

Phase and frequency modulation with an EOM

Instead of using an EOM to change the polarization of light, we can set the polarization to be parallel to the principal axis of the crystal. In that case, the EOM will not modulate the intensity of the light, but it can modulate the phase or frequency.

The EOM will introduce a relative phase compared to if we would not apply any voltage. If our light is polarized along the x axis, that phase will be:

$$\Delta\phi = (n_x - n_o) \frac{\omega L}{c} = -\frac{\omega r_{63} n_o^3 V}{2c}. \quad (1.63)$$

Let us assume that the field we apply is oscillating at a angular frequency Ω :

$$E(t) = E_0 \sin(\Omega t) = \frac{V(t)}{L} = \frac{V_0}{L} \sin(\Omega t). \quad (1.64)$$

If we apply this varying electric field to our crystal, we will get the following:

$$\begin{aligned} \phi(t) &= \omega t + m \sin(\Omega t), \\ E(t) &= \operatorname{Re} \{ E_0 \exp(-i\omega t) \exp[-im \sin(\Omega t)] \}. \end{aligned} \quad (1.65)$$

Here, m is the modulation index, i.e., the amplitude of the modulation. It depends on the material parameters and the amplitude of the applied voltage:

$$m = \frac{\omega r_{63} n_o^3 V_0}{2c}. \quad (1.66)$$

Our EOM will also modulate the light's instantaneous frequency:

$$\omega(t) = \frac{d\phi(t)}{dt} = \omega + m\Omega \cos(\Omega t). \quad (1.67)$$

While we cannot clearly separate phase modulation (PM) from frequency modulation (FM), we can roughly catagorize the effect one will see in the following way[2]:

- for $m < 1$: PM
- for $m > 1$: FM

To illustrate that, let us decompose our electromagnetic wave into its frequency components[2]:

$$\begin{aligned} e^{-ims\sin(\Omega t)} &= J_0 + 2[J_2(m)\cos(2\Omega t) + J_4(m)\cos(4\Omega t) + \dots] - \\ &\quad 2i[J_1(m)\sin(\Omega t) + J_3(m)\sin(3\Omega t) + \dots]. \end{aligned}$$

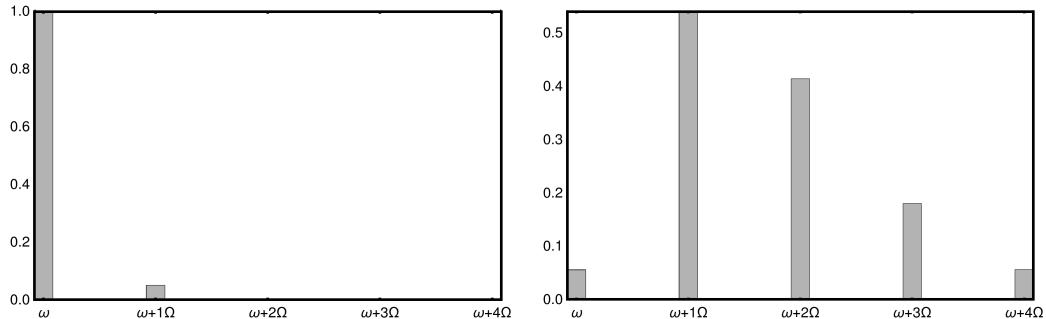


Figure 5: **Sideband amplitudes for $\omega + m\Omega$.** The figure is intended to distinguish the regimes where we mostly have phase modulation or significant frequency modulation. The left and right figures are for $m = 0.1$ and for $m = 2.3$, respectively. (left) For small m , most light will remain at the fundamental frequency. (right) for larger m , a significant amount of light will be in the sidebands. For simplicity, the figure does not show the negative sidebands.

J_n are Bessel functions of the 1st kind. Then we can write $E(t)$ as:

$$E(t) = E_0 e^{-i\omega t} [J_0(m) + J_1(m)(e^{-i\Omega t} - e^{i\Omega t}) + \dots] \quad (1.68)$$

$$J_2(m)(e^{-2i\Omega t} + e^{2i\Omega t}) + J_3(m)(e^{-3i\Omega t} - e^{3i\Omega t}) + \dots]. \quad (1.69)$$

This decomposition shows that the spectrum of the field will have a fundamental peak at ω , and it will have peaks at $\omega \pm k\Omega$ for $k \geq 1$. The corresponding signals are called *sidebands*. Fig.5 shows how the light becomes distributed into various frequency sidebands depending on modulation index m . For low m , most of the light remains at the fundamental (or “carrier”) frequency. For larger m , the carrier becomes suppressed, and for $m = 2.4$, the carrier vanishes and all the light intensity will be in the sidebands.

1.3 The magneto-optic effect

For some materials, it is possible to affect the light passing through a crystal by applying a magnetic field to the crystal. There are two such magneto-optic effects: (a) the **Faraday effect**, where the polarization of light is affected when it passes through material to which a magnetic field is applied. (b) the **magneto-optic Kerr effect**, which affects light reflected from the surface of a material to which a magnetic field is applied. Here, we will only briefly introduce the Faraday effect.

1.3.1 The Faraday effect

Consider light passing through a medium, and that a magnetic field is applied along the propagation direction (see Fig. 6). If we assume that the incident field is polarized linearly along \vec{E} , then the polarization \vec{E}' after the medium will be given by:

$$\vec{E}' = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \vec{E}. \quad (1.70)$$

The rotation angle α depends on the magnetic field strength and the length L of the medium:

$$\alpha = V|\vec{B}|L, \quad (1.71)$$

where V is the **Verdet constant**, which depends on the material chosen. An example for light of 589 nm is $209\text{ m}^{-1}\text{T}^{-1}$ for quartz[2].

The material results in different phase shifts for left-hand and right-hand circularly polarized light. Because these two polarizations form a basis for all polarizations, we get an effective rotation of polarization. A very interesting fact is that this effect is not directionally reversible. That means, if light is back-reflected after passing through such a material, the light does experience twice as much rotation. For that reason, the Faraday effect is often applied as a means to build *optical isolators* to prevent back reflection of light.

An example for an optical isolator works as follows: (1) incident light passes through a horizontal polarizer. (2) it is rotated by 45° in a Faraday rotator. (3) at the output, the light passes through a 45° polarizer. Light that is backreflected, will first have to pass through the 45° polarizer. Then it is rotated by 45° in the Faraday rotator and becomes vertically polarized. It then will be blocked by the horizontal polarizer.

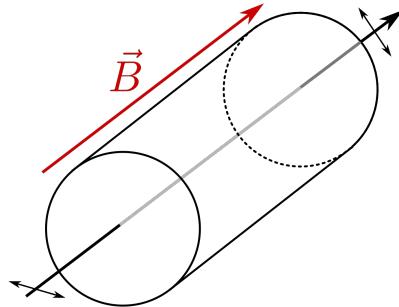


Figure 6: **Faraday effect.** A beam of light passes through a medium along the arrow indicated. In parallel, a magnetic field \vec{B} is applied to the medium. Due to the Faraday effect, the polarization of the beam will rotate as it passes through the medium.

1.4 Acousto-optics

In contrast to the electro-optic effect and the magneto-optic effect, this is not a non-linear optical effect, but our discussion here will form the basis for later discussions of non-linear optical effects. Assume we apply a sound-wave to a an optical medium - for example, via a piezo-electric transducer. Because sound waves inside a solid go along with changes to material density, the sound wave can influence the optical properties of the material.

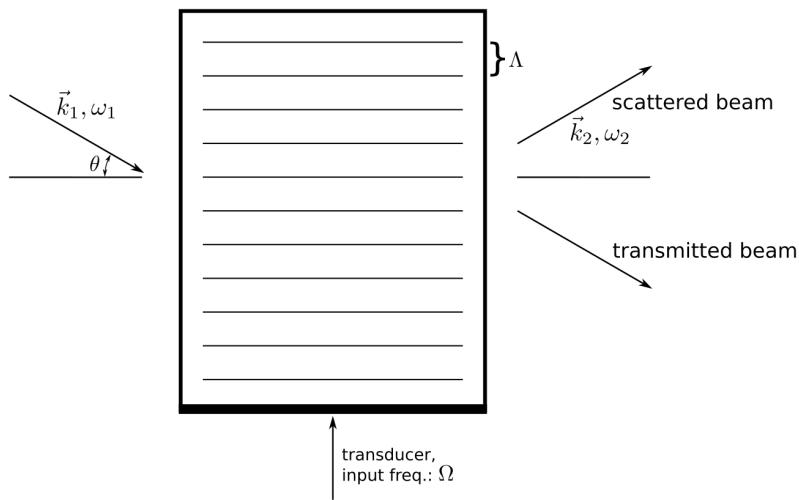


Figure 7: **Bragg acousto-optic modulation.** An rf signal with angular frequency Ω applied to a transducer generates an acoustic wave with wavelength Λ . An incoming beam with optical frequency ω_1 and wavevector \vec{k}_1 is diffracted. The scattered beam has ω_2 and \vec{k}_2 , part of the beam is transmitted.

One can classify acousto-optic devices into two main categories, depending on the typical length scales involved:

- Bragg scattering: The interaction length is long such that phase-matching conditions become important (we will discuss that later). In this case, one has a single diffracted beam.
- Raman-Nath scattering: The interaction length is short and phase-matching conditions are not important.

We will focus on the Bragg scattering regime, but first we will introduce some core concepts: let us assume that an acoustic sound wave passes through a medium, and that an input beam is scattered by the resulting modulation of the refractive index in the medium. We assume that the sound wave is a plane wave with a frequency Ω and a wavelength $\Lambda = 2\pi v/\Omega$. v is the velocity of sound in the medium, see Fig. 7.

Spontaneous Brioullin scattering

The acoustic wave will lead to a modulation $\Delta\chi$ of the susceptibility in the medium. That will lead to an additional polarization in a volume V of our medium:

$$\vec{P} = \Delta\chi\vec{E}_0 = \Delta\epsilon\vec{E}_0. \quad (1.72)$$

For simplicity, we represent the susceptibility as a scalar, and we use the relation $\epsilon = 1 + \chi$. \vec{P} leads to a dipole moment and to the emission of radiation: scattered light. The scattering of light from comparatively low-frequency acoustic waves is called **Brioullin scattering**. Later, we will also discuss stimulated Brioullin scattering in contrast to the “spontaneous” Brioullin scattering we discuss here.

Changes to the dielectric constant depend on the density ρ and the temperature T :

$$\Delta\epsilon = \left(\frac{\partial\epsilon}{\partial\rho}\right)_T \Delta\rho + \left(\frac{\partial\epsilon}{\partial T}\right)_\rho \Delta T. \quad (1.73)$$

Typically, the dielectric constant depends significantly more on variations of the pressure than on the temperature. Then we can neglect the second term and assume:

$$\Delta\epsilon \approx \left(\frac{\partial\epsilon}{\partial\rho}\right)_T \Delta\rho. \quad (1.74)$$

If we choose the pressure p and the entropy s as the independent thermodynamic variables, we can write changes to the density as:

$$\Delta\rho = \left(\frac{\partial\rho}{\partial p}\right)_s \Delta p + \left(\frac{\partial\rho}{\partial s}\right)_p \Delta s. \quad (1.75)$$

The first term describes (adiabatic) density fluctuations that represent acoustic waves[1].

The equation of motion for a pressure wave inside the medium is given by[1]:

$$\frac{\partial^2 \Delta p}{\partial t^2} - \Gamma \nabla^2 \frac{\partial \Delta p}{\partial t} - v^2 \nabla^2 \Delta p = 0, \quad (1.76)$$

where Γ is a damping parameter, and v is the speed of sound in the medium, given by:

$$v^2 = \frac{K_s}{\rho} = \frac{1}{C_s \rho} = \left(\frac{\gamma k_B T}{\mu}\right)^{1/2}. \quad (1.77)$$

C_s and K_s are the bulk modulus and the compressibility at constant entropy. μ is the molecular mass, k_B is Boltzmann's constant, and γ is the adiabatic index[1]:

$$\gamma = \frac{C_T}{C_s} = \frac{c_p}{c_V}. \quad (1.78)$$

C_T is the compressibility at constant temperature, and c_p and c_V are the specific heat capacities measured at constant pressure and at constant volume. The relation between the compressibility, the bulk modulus, the pressure and the material density is given by:

$$C = \frac{1}{K} - \frac{1}{V} \frac{\partial V}{\partial p} = \frac{1}{\rho} \frac{\partial \rho}{\partial p}. \quad (1.79)$$

Let us investigate the behaviour of a plane acoustic wave with wave vector \vec{q} :

$$\Delta p(\vec{r}, t) = \Delta p e^{i(\vec{q} \cdot \vec{r} - \Omega t)} + \text{c.c.} \quad (1.80)$$

if we apply equation 1.76, we get the dispersion relation $\Omega^2 = q^2(v^2 - i\Omega\Gamma)$.

Incoming light will be scattered by the acoustic wave. Let the incoming field be:

$$\vec{E}_0(\vec{r}, t) = \vec{E}_0 e^{i(\vec{k} \cdot \vec{r} - \omega t)} + \text{c.c..} \quad (1.81)$$

The field inside the medium has to fulfill the wave equation:

$$\nabla^2 \vec{E} - \frac{n^2}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \vec{P}}{\partial t^2}. \quad (1.82)$$

Let us assume that \vec{P} is given by equation 1.72, that $\Delta\epsilon$ is given by equation 1.74, and that the first term of equation 1.75 is dominant:

$$\Delta\rho \approx \left(\frac{\partial\rho}{\partial p} \right)_s \Delta p. \quad (1.83)$$

Then we get:

$$\vec{P}(\vec{r}, t) = \epsilon_0 \left(\frac{\partial\epsilon}{\partial\rho} \right) \left(\frac{\partial\rho}{\partial p} \right)_s \Delta p(\vec{r}, t) \vec{E}_0(\vec{r}, t) = \epsilon_0 \gamma_e C_s \Delta p(\vec{r}, t) \vec{E}_0(\vec{r}, t), \quad (1.84)$$

where ρ_0 is the undisturbed material density, and γ_e is the electrostrictive constant:

$$\gamma_e = \left(\rho \frac{\partial\epsilon}{\partial\rho} \right)_{\rho=\rho_0}. \quad (1.85)$$

If we put this into our wave equation for the electric field, we get:

$$\nabla^2 \vec{E} - \frac{n^2}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = - \frac{\gamma_e C_s}{c^2} \left[(\omega - \Omega)^2 E_0 \Delta p^* e^{i(\vec{k} - \vec{q}) \cdot \vec{r} - i(\omega - \Omega)t} + (\omega + \Omega)^2 E_0 \Delta p e^{i(\vec{k} + \vec{q}) \cdot \vec{r} - i(\omega + \Omega)t} \right]. \quad (1.86)$$

We can split the right-hand side into two separate parts corresponding to *Stokes Scat-*

tering and *Anti-Stokes Scattering*, respectively. We will describe these two effects below.

Stokes and Anti-Stokes scattering

For Stokes scattering, the wavevector \vec{k}' and the angular frequency ω' of the scattered field will fulfill:

$$\vec{k}' = \vec{k} - \vec{q}, \quad (1.87)$$

and

$$\omega' = \omega - \Omega. \quad (1.88)$$

The above equations are **not sufficient** to warrant efficient scattering of the incoming wave into the outgoing wave described by the equations above. The wave vector and the frequency of the scattered light will need to fulfill the dispersion relation:

$$\omega' = |\vec{k}'|c/n, \quad (1.89)$$

where n is the refractive index of our medium.

For Anti-Stokes scattering, the wavevector and the angular frequency fulfill:

$$\vec{k}' = \vec{k} + \vec{q}, \quad (1.90)$$

and

$$\omega' = \omega + \Omega. \quad (1.91)$$

Again, the dispersion relation in equation 1.89 has to be fulfilled for efficient scattering.

1.4.1 Bragg-scattering of light by acoustic waves

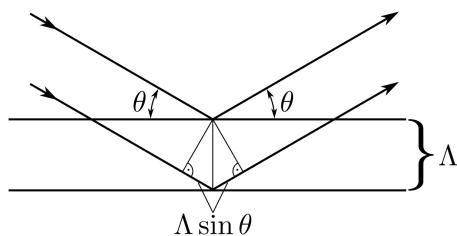


Figure 8: **Bragg condition.** The figure illustrates the condition under which constructive interference occurs between waves reflected at planes with a distance Λ .

Now let us get back to the situation described in Fig. 7. The incoming beam can be partially reflected at the wavefronts of the acoustic wave. For certain angles of incidence, the waves reflected from the various wavefronts can interfere constructively if the **Bragg**

condition holds:

$$\lambda = 2\Lambda \sin\theta. \quad (1.92)$$

We can derive this condition by considering Fig. 8. In particular, at an angle fulfilling the Bragg condition, the waves reflected from different crystal planes will have relative delays that are multiples of 2π .

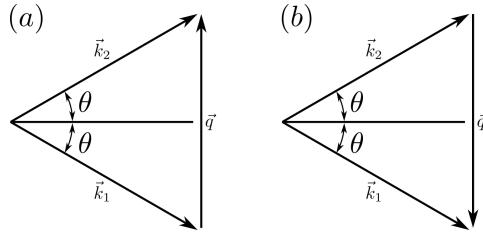


Figure 9: **Bragg condition & phase matching.** We can express the Bragg condition in terms of phase matching. Figures (a) and (b) illustrate that the sign of the wave vector \vec{q} of the acoustic wave must be adapted to the direction the sound wave is traveling.

The Bragg condition can also be understood as a **phase-matching condition**. That means, the phases of contributing waves have to match such that they can constructively interfere. If \vec{k}_1 and \vec{k}_2 represent the wave vectors of the incoming and the scattered wave as in Fig. 7, we can write the Bragg condition as:

$$\vec{k}_2 = \vec{k}_1 + \vec{q}. \quad (1.93)$$

Similar to equation 1.91, we will get:

$$\omega_2 = \omega_1 + \Omega. \quad (1.94)$$

Because $\Omega \ll \omega$, the two wave-vectors will fulfill: $|\vec{k}_2| \approx |\vec{k}_1|$. Equations 1.93 and 1.94 can be seen as equations for momentum and energy conservation, respectively.

If the sound waves travel in the opposite directions, we will instead get:

$$\begin{aligned} \vec{k}_2 &= \vec{k}_1 - \vec{q}, \text{ and} \\ \omega_2 &= \omega_1 - \Omega. \end{aligned} \quad (1.95)$$

To accurately describe the theory of Bragg scattering of light by sound, we need to investigate changes to the dielectric constant due to the acoustic waves as we did for Brioullin scattering. The description there only applies to isotropic media. For anisotropic media, the dielectric constant is a tensor related to the strain tensor S_{ij} and

the strain-optic tensor p_{ijkl} as follows:

$$[\Delta(\epsilon^{-1})_{ij}] = \sum_{kl} p_{ijkl} S_{kl}. \quad (1.96)$$

The strain tensor is given by:

$$S_{kl} = \frac{1}{2} \left(\frac{\partial d_k}{\partial x_l} + \frac{\partial d_l}{\partial x_k} \right) \quad (1.97)$$

From equation 1.96, it follows that:

$$(\Delta\epsilon)_{il} = - \sum_{jk} \epsilon_{ij} [\Delta(\epsilon^{-1})]_{jk} \epsilon_{kl}. \quad (1.98)$$

Let us assume a layout as in Fig. 7 with \vec{q} pointing in the z -direction, and x perpendicular to that in the plane and y perpendicular to the plane. Once this geometry is fixed, we can describe the incident and diffracted waves, respectively, as[1]:

$$\begin{aligned} E_1 &= A_1 e^{i(\vec{k}_1 \cdot \vec{r} - \omega_1 t)} + \text{c.c.}, \\ E_2 &= A_2 e^{i(\vec{k}_2 \cdot \vec{r} - \omega_2 t)} + \text{c.c..} \end{aligned} \quad (1.99)$$

$A_{1,2}$ are slowly varying amplitudes that only depend on x and z , but *not* on y .

If the interaction volume is sufficiently large, the diffracted beam will (approximately) fulfill Bragg phase matching:

$$\vec{k}_2 \approx \vec{k}_1 + \vec{q}. \quad (1.100)$$

The total electric field will be the superposition of the incoming and the diffracted field ($E = E_1 + E_2$), and it needs to fulfill the wave equation in our dielectric medium:

$$\nabla^2 E - \frac{n^2 + \Delta\epsilon}{c^2} \frac{\partial^2 E}{\partial t^2} = 0. \quad (1.101)$$

Solving this equation will give us some basic insight into the process of acoustic modulation. To facilitate finding a solution let us separate $\Delta\epsilon$ into a product of rapidly and slowly varying factors:

$$\Delta\epsilon = \Delta\tilde{\epsilon} e^{i\vec{q} \cdot \vec{r} - \Omega t} + \text{c.c.} \quad (1.102)$$

The exponential oscillates quickly in space and time while $\Delta\tilde{\epsilon}$ varies only slowly.

Now let us insert the field amplitudes from equation 1.99 into equation 1.101 and

collect only terms that vary periodically with ω_1 . Given that $\Omega = \omega_2 - \omega_1$, we get:

$$\begin{aligned} \frac{\partial^2 A_1}{\partial x^2} + \frac{\partial^2 A_1}{\partial z^2} + 2ik_{1x} \frac{\partial A_1}{\partial x} + 2ik_{1z} \frac{\partial A_1}{\partial z} - (k_{1x}^2 + k_{1z}^2)A_1 + \\ \frac{n^2\omega_1^2}{c^2}A_1 + \frac{\omega_2^2}{c^2}A_2\Delta\tilde{\epsilon}^*e^{i(\vec{k}_2-\vec{k}_1-\vec{q})\cdot\vec{r}} = 0. \end{aligned} \quad (1.103)$$

Since A_1 varies only slowly in x , we can neglect the 2nd-order derivatives. If we assume that A_1 does not depend on z , and that the optical dispersion relation holds for \vec{k}_1 and ω_1 , equation 1.103 becomes:

$$2ik_{1x} \frac{dA_1}{dx} = -\frac{-\omega_2^2}{c^2}A_2\Delta\tilde{\epsilon}^*e^{i(\vec{k}_2-\vec{k}_1-\vec{q})\cdot\vec{r}}. \quad (1.104)$$

If we assume that the extent of our medium is much larger in the z direction than in the x direction, we can neglect any phase mismatch in the z direction, and we will get:

$$(\vec{k}_2 - \vec{k}_1 - \vec{q}) \cdot \vec{r} = -\Delta kx. \quad (1.105)$$

Here we introduced $\Delta k = |\Delta\vec{k}|$ with $\Delta\vec{k} = \vec{k}_2 - \vec{k}_1 - \vec{q}$.

This will further simplify equation 1.103, and if we apply similar arguments to solve equation 1.101 for terms oscillating with ω_2 , we get the following coupled equations:

$$\begin{aligned} \frac{dA_1}{dx} &= \frac{i\omega_2^2\Delta\tilde{\epsilon}^*}{2k_{1x}c^2}A_2e^{-i\Delta kx} \\ \frac{dA_2}{dx} &= \frac{i\omega_1^2\Delta\tilde{\epsilon}}{2k_{2x}c^2}A_1e^{i\Delta kx}. \end{aligned} \quad (1.106)$$

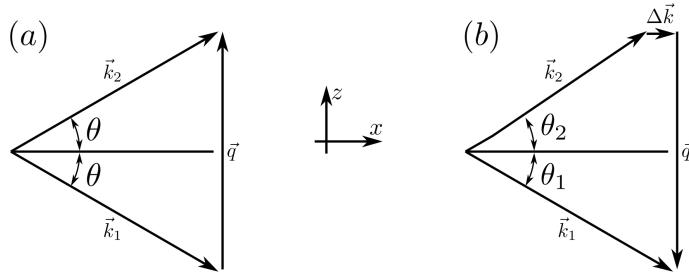


Figure 10: **Phase matching.** (a) shows perfect phase matching ($\Delta k = 0$). (b) shows imperfect phase matching, where there is a non-vanishing $\Delta\vec{k}$ necessary to add up the wave vectors to equal \vec{q} .

We can further simplify these equations if we assume $\omega_1 \approx \omega_2 \equiv \omega$ and $k_{1x} \approx k_{2x} \equiv k_x$:

$$\begin{aligned}\frac{dA_1}{dx} &= i\kappa A_2 e^{-i\Delta k x} \\ \frac{dA_2}{dx} &= i\kappa^* A_1 e^{i\Delta k x}.\end{aligned}\quad (1.107)$$

Here, we introduced the coupling constant

$$\kappa = \frac{\omega^2 \Delta \tilde{\epsilon}^*}{2k_x c^2}. \quad (1.108)$$

If the incoming beam is incident at the Bragg angle, the phase mismatch will vanish ($\Delta k = 0$), and the solutions of equation 1.107 become particularly simple. For our example here, the solutions become:

$$\begin{aligned}A_1(x) &= A_1(0) \cos(|\kappa|x), \\ A_2(x) &= \frac{i\kappa^*}{|\kappa|} A_1(0) \sin(|\kappa|x),\end{aligned}$$

where the solutions fulfill $|A_1(x)|^2 + |A_2(x)|^2 = |A_1(0)|^2$.

We can calculate the efficiency of the process using the ratio between the input intensity and the intensity of the scattered light:

$$\eta \equiv \frac{|A_2(L)|^2}{|A_1(0)|^2} = \sin^2(|\kappa|L). \quad (1.109)$$

To make numerical estimates given experimental parameters it is useful to express κ in terms of the intensity of the acoustic wave[1]:

$$I = Kv \frac{\langle \Delta \rho^2 \rangle}{\rho_0^2} = 2Kv |\Delta \tilde{\epsilon}|^2 / \gamma_e^2. \quad (1.110)$$

Then the coupling constant becomes:

$$|\kappa| = \frac{\omega \gamma_e}{2nccos\theta} \left(\frac{I}{2Kv} \right)^{1/2}. \quad (1.111)$$

In the simplified calculations above, we assumed that Δk vanishes. Fig. 10 illustrates

the general case where Δk can be non-zero. The vectors in Fig. 10(b) satisfy the relations:

$$k\cos\theta_1 - k\cos\theta_2 = \Delta k \quad (1.112)$$

$$k\sin\theta_1 + k\sin\theta_2 = q, \quad (1.113)$$

where we assumed that $k_1 \approx k_2 = k$.

If the incoming light is not incident at the Bragg angle θ_B , we will have $\theta_1 = \theta_B + \Delta\theta$. Typically, we will assume that $\Delta\theta \ll 1$. Equation 1.113 will be satisfied if we have $\theta_2 = \theta_B - \Delta\theta$. If we insert these expressions for $\theta_{1,2}$ into equation 1.112, we get:

$$\cos(\theta_B \pm \Delta\theta) = \cos\theta_B \mp (\sin\theta_B)\Delta\theta, \text{ and} \quad (1.114)$$

$$(2k\sin\theta_B)\Delta\theta = \Delta k, \quad (1.115)$$

resulting in:

$$\Delta k = -\Delta\theta q. \quad (1.116)$$

If we solve the differential equations 1.107 for $\Delta k \neq 0$ for the case in which the external field does not include ω_2 components, we get:

$$\begin{aligned} A_1(x) &= e^{-i\Delta kx/2} A_1(0) \left[\cos(xs) + i \frac{\Delta k}{2s} \sin(xs) \right] \\ A_2(x) &= ie^{i\Delta kx/2} A_1(0) \frac{\kappa^*}{s} \sin(xs). \end{aligned} \quad (1.117)$$

The parameter s is defined via the relation:

$$s^2 = |\kappa|^2 + (\Delta k/2)^2. \quad (1.118)$$

Using these relations, we can derive a more general relation for the diffraction efficiency:

$$\eta(\Delta k) \equiv \frac{|A_2(L)|^2}{|A_1(0)|^2} = \frac{|\kappa|^2}{|\kappa|^2 + (\Delta k/2)^2} \sin^2 \left\{ \left[|\kappa|^2 + (\Delta k/2)^2 \right]^{1/2} L \right\}. \quad (1.119)$$

1.5 Liquid crystal modulators

As an example, we will discuss a liquid crystal (LC) modulator using molecules in the nematic phase. There also exist other phases like the smectic and the cholesteric phase[2]. In the nematic phase, all the molecular slabs inside a liquid point in one direction. If they are exposed to a surface that imposes a preferred direction (e.g., grooves in the surface), then the molecular slabs will orient themselves along that direction. If such a liquid is placed between two such surfaces where the grooves on one surface are perpendicular to the grooves on the other surface, then the molecular slabs will rotate their orientation along the way from one surface to the other. This is illustrated in Fig. 11. If a voltage is

applied between the two surfaces, the molecular slabs orient themselves along the field lines. In the former case, the polarization of a light beam passing through the liquid will be rotated along with the molecular slabs. When the voltage is applied, this is not the case any longer, and the polarization of the light will be unchanged.

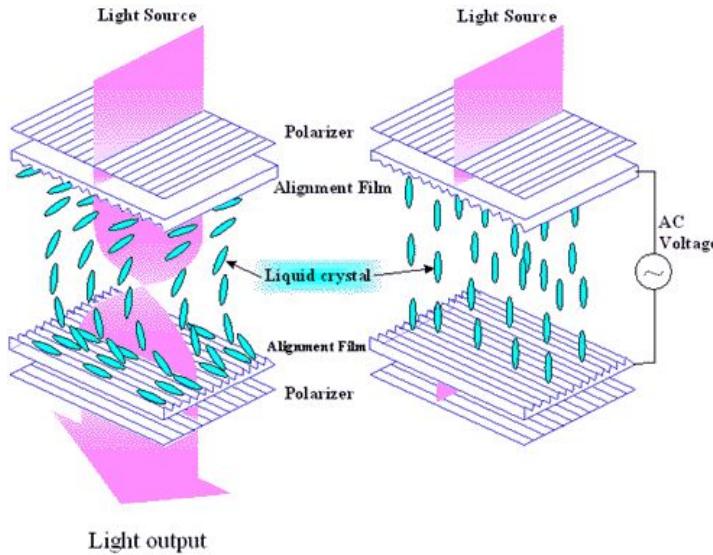


Figure 11: **Liquid crystal modulator.** (left) The molecular slabs orient themselves along the grooves on the surfaces. Between the two surfaces, the slabs reorient themselves. (right) if a voltage is applied, the molecular slabs orient themselves along the field lines. The figure is from Ref. [3].

Liquid crystal displays (LCDs) work based on the same principles, but they typically work in reflection.

2 Wave-equation description of nonlinear effects

In this chapter, we will discuss the theory describing some of the nonlinear effects we mentioned earlier. From the Lorentz model, we can conclude that, in order for there to be a significant amount of radiation exiting the nonlinear material, the dipoles contributing to the process have to emit radiation in phase with each other such that these contributions interfere constructively. This provides an intuitive understanding of the “phase matching” we will discuss.

2.1 The wave equation for nonlinear optical media

We will mostly be interested in nonmagnetic media that are free of charges and currents. The Maxwell equations in such a medium are:

$$\begin{aligned} \vec{\nabla} \cdot \vec{D} &= 0, & \vec{\nabla} \cdot \vec{B} &= 0, \\ \vec{\nabla} \times \vec{E} &= -\frac{\partial \vec{B}}{\partial t}, & \vec{\nabla} \times \vec{H} &= \frac{\partial \vec{D}}{\partial t}, \end{aligned} \quad (2.1)$$

where we used:

$$\vec{B} = \mu_0 \vec{H}, \quad \vec{D} = \epsilon_0 \vec{E} + \vec{P}. \quad (2.2)$$

The polarization vector \vec{P} can depend nonlinearly on the field \vec{E} .

These equations result in the following wave equation in nonlinear optical media:

$$\vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \vec{\nabla}^2 \vec{E} + \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \vec{E} = -\frac{1}{\epsilon_0 c^2} \frac{\partial^2 \vec{P}}{\partial t^2}. \quad (2.3)$$

Often the first term on the left-hand side vanishes (e.g. for plane waves), and we get:

$$\vec{\nabla}^2 \vec{E} - \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \vec{D}}{\partial t^2} = 0. \quad (2.4)$$

It is convenient to split \vec{D} and \vec{P} into linear and nonlinear parts as follows:

$$\begin{aligned}\vec{P} &= \vec{P}^{(1)} + \vec{P}^{\text{NL}} \\ \vec{D} &= \vec{D}^{(1)} + \vec{D}^{\text{NL}} = \vec{D}^{(1)} + \vec{P}^{\text{NL}},\end{aligned}$$

where $\vec{D}^{(1)} = \epsilon_0 \vec{E} + \vec{P}^{(1)}$.

Then the wave equation becomes:

$$\vec{\nabla}^2 \vec{E} - \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \vec{D}^{(1)}}{\partial t^2} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \vec{P}^{\text{NL}}}{\partial t^2}. \quad (2.5)$$

Let us split the field, the displacement vector and the polarization into distinct frequency components, and let us assume that the time dependencies are periodic with angular frequencies ω_n . For the field, we get:

$$\vec{E}(\vec{r}, t) = \sum_n \vec{E}_n(\vec{r}, t) = \sum_n (\vec{E}_n(\vec{r}) e^{-i\omega_n t} + \text{c.c.}). \quad (2.6)$$

The sum is over the positive frequency parts only. Similar equations hold for the displacement vector and the polarization.

If we introduce the following short-hand notation:

$$(\vec{\epsilon}^{(1)} \cdot \vec{E})_i = \sum_j \epsilon_{ij}^{(1)} E_j, \quad (2.7)$$

we can write:

$$\vec{D}_n^{(1)}(\vec{r}) = \epsilon_0 \vec{\epsilon}^{(1)}(\omega_n) \cdot \vec{E}. \quad (2.8)$$

All frequency components separately have to fulfill the wave equation:

$$\vec{\nabla}^2 \vec{E}_n(\vec{r}) + \frac{\omega_n^2}{c^2} \vec{\epsilon}^{(1)}(\omega_n) \cdot \vec{E}_n(\vec{r}) = -\frac{\omega_n^2}{\epsilon_0 c^2} \vec{P}_n^{\text{NL}}(\vec{r}). \quad (2.9)$$

2.2 The coupled wave equation for SFG

We will now describe non-linear optical processes in increasing detail. First, let us assume the case of two fields with frequencies $\omega_{1,2}$ incident on a $\chi^{(2)}$ nonlinear crystal such that they generate a field at the sum frequency $\omega_3 = \omega_1 + \omega_2$ (see Fig. 12).

If there is no nonlinear interaction, the nonlinear right-hand side of the wave equation will be zero, and the solution for ω_3 will be:

$$E_3(z, t) = A_3 e^{i(k_3 z - \omega_3 t)} + \text{c.c..} \quad (2.10)$$

If we know the polarizations of the fields involved, we can write E_3 as a scalar field. k_3

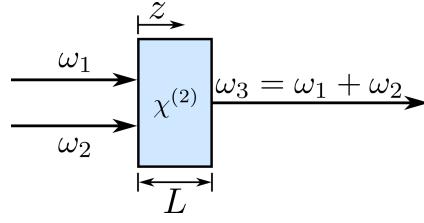


Figure 12: **Sum frequency generation (SFG)**. Two incident fields with frequencies ω_1 and ω_2 generate a field with frequency $\omega_3 = \omega_1 + \omega_2$.

and ω_3 are related by the dispersion relation $k_3 = \frac{n_3 \omega_3}{c}$, where $n_3^2 = \epsilon^{(1)}(\omega_3)$.

Now, let us assume that the amplitude A_3 in equation 2.10 depends (slowly) on the position z along the crystal, and that we can write the polarization at ω_3 as:

$$\begin{aligned} P_3(z, t) &= P_3(z) e^{-i\omega_3 t} + \text{c.c.} \\ P_3(z) &= 4\epsilon_0 d_{\text{eff}} E_1(z) E_2(z), \end{aligned} \quad (2.11)$$

where we assumed we can also separate quickly and slowly varying parts for the E fields:

$$\begin{aligned} E_i(z, t) &= E_i(z) e^{-i\omega_i t} + \text{c.c.}, \\ E_i(z) &= A_i(z) e^{ik_i z} + \text{c.c..} \end{aligned} \quad (2.12)$$

We assumed that our medium is lossless and has full permutation symmetry.

If we insert all that into the wave equation 2.9, we get:

$$\frac{d^2 A_3}{dz^2} + 2ik_3 \frac{dA_3}{dz} = -4 \frac{d_{\text{eff}} \omega_3^2}{c^2} A_1 A_2 e^{i(k_1 + k_2 - k_3)z}. \quad (2.13)$$

If the A_i are slowly varying functions of z , we can neglect the first term and get:

$$\frac{dA_3}{dz} = \frac{2id_{\text{eff}}\omega_3^2}{k_3 c^2} A_1 A_2 e^{i\Delta kz}, \quad (2.14)$$

with the wavevector mismatch $\Delta k = k_1 + k_2 - k_3$. This is called a coupled-amplitude equation. We can derive similar equations for the amplitudes $A_{1,2}$:

$$\begin{aligned} \frac{dA_1}{dz} &= \frac{2id_{\text{eff}}\omega_1^2}{k_1 c^2} A_3 A_2^* e^{-i\Delta kz}, \\ \frac{dA_2}{dz} &= \frac{2id_{\text{eff}}\omega_2^2}{k_2 c^2} A_3 A_1^* e^{-i\Delta kz}. \end{aligned} \quad (2.15)$$

2.2.1 General considerations about phase-matching

If we achieve $\Delta k = 0$, we have perfect *phase matching*. This ensures that the contributions from different emitters interfere constructively. In that case, and if we assume (for now) that A_1 and A_2 are constant, equation 2.14 will give us a linear increase in A_3 . That $A_{1,2}$ are constant would of course violate energy conservation, but for small A_3 we can often neglect the **depletion of power** in $A_{1,2}$.

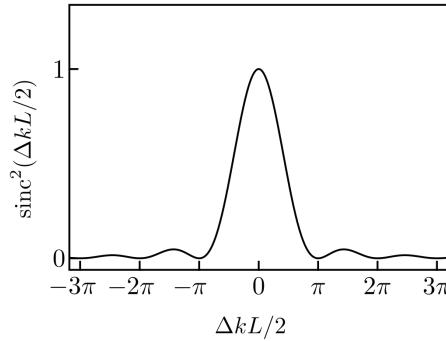


Figure 13: Phase mismatch factor. .

For $\Delta k \neq 0$ and no depletion, we can integrate equation 2.14 to get:

$$A_3(L) = \frac{2id_{\text{eff}}\omega_3 A_1 A_2}{c} \int_0^L dz e^{i\Delta kz} = \frac{2id_{\text{eff}}\omega_3 A_1 A_2}{c} \left(\frac{e^{i\Delta k L} - 1}{i\Delta k} \right), \quad (2.16)$$

where L is the length of the crystal.

The intensity of the SFG light will be proportional to:

$$\left| \frac{e^{i\Delta k L} - 1}{\Delta k} \right|^2 = L^2 \text{sinc}^2(\Delta k L / 2), \quad (2.17)$$

where $\text{sinc}(x) \equiv \sin(x)/x$. We plot this **phase mismatch factor** in Fig. 13.

The full expression for the intensity of the SFG field is:

$$I_3 = \frac{8d_{\text{eff}}^2 \omega_3^2 I_1 I_2}{n_1 n_2 n_3 \epsilon_0 c^2} L^2 \text{sinc}^2 \left(\frac{\Delta k L}{2} \right). \quad (2.18)$$

n_i are the refractive indices, and $I_i = 2n_i \epsilon_0 c |A_i|^2$ are the intensities for the three fields.

Fig. 13 shows that the phase mismatch factor is 1 for perfect phase matching, and then it vanishes quickly for increasing $|\Delta k L|$. Some oscillations occur because, for $L \gtrsim 1/\Delta k$, the output wave E_3 will get out of phase with the two input fields and power from E_3

field can flow back into the other frequency modes. The interaction builds up coherently over a length $L_{\text{coh}} = 2/\Delta k$, and we can also write the phase mismatch factor as:

$$\text{sinc}^2(L/L_{\text{coh}}). \quad (2.19)$$

2.3 Phase matching

The efficiency of nonlinear processes like SFG will drop drastically if the phasematching condition is not fulfilled. For SFG with $\omega_3 = \omega_1 + \omega_2$, the phase mismatch is:

$$\Delta k = k_1 + k_2 - k_3. \quad (2.20)$$

It is not straight forward to achieve perfect phase matching ($\Delta k = 0$). Let us take a closer look at what phase matching entails. We can also write equation 2.20 as:

$$\frac{n_1\omega_1}{c} + \frac{n_2\omega_2}{c} = \frac{n_3\omega_3}{c}. \quad (2.21)$$

If the polarization of all three fields were the same, the refractive indices would only depend on the frequencies of the light, such that:

$$\frac{n(\omega_1)\omega_1}{c} + \frac{n(\omega_2)\omega_2}{c} = \frac{n(\omega_3)\omega_3}{c}. \quad (2.22)$$

In materials with **normal dispersion**, $n(\omega)$ increases monotonically with ω . Now let us rewrite equation 2.20 as:

$$n_3 - n_2 = (n_1 - n_2) \frac{\omega_1}{\omega_3}. \quad (2.23)$$

Without loss of generality, we can assume $\omega_3 > \omega_2 \geq \omega_1$. For normal dispersion, we would then have $n_3 > n_2$ but also $n_2 > n_1$. This results in a contradiction for equation 2.23. This shows that phase matching is not possible in a material with normal dispersion if all fields have the same polarization.

We can overcome this limitation by using birefringent crystals. If the two input fields have the same polarization, orthogonal to the output field, we speak of *type-I phase matching*. If the two input fields have orthogonal polarizations, we speak of *type-II phase matching*. To achieve phase matching, we can use uniaxial or biaxial crystals. For simplicity, let us focus on uniaxial crystals. Polarization that is parallel to the plane formed by the crystal z axis and the wavevector \vec{k} is called *extraordinary* polarization. Light that is polarized perpendicular to that plane is said to have *ordinary* polarization. The z axis or *optic axis* is also often referred to as the c axis. The refractive index for ordinary polarization is n_o . For extraordinary polarization, the refractive index $n_e(\theta)$

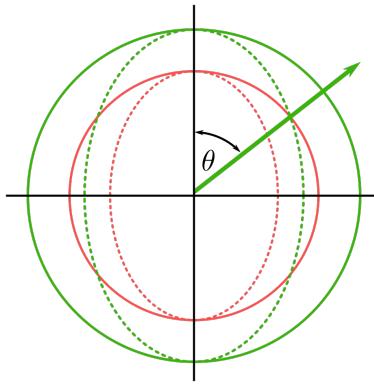


Figure 14: **Phase matching.** Second harmonic generation (SHG) in a negative uniaxial crystal. The ordinary refractive indices for $\omega_1 = \omega_2 = \omega$ and for $\omega_3 = 2\omega$ are solid circles in red and green, respectively. The extraordinary refractive indices for the two wavelengths are dashed ellipses. The phase-matching angle θ is given by the point where the extraordinary refractive index for 2ω matches the ordinary refractive index for ω (green narrow).

depends on the angle θ of \vec{k} relative to the z axis. It is defined by the relation:

$$\frac{1}{n_e^2(\theta)} = \frac{\sin^2\theta}{\bar{n}_e^2} + \frac{\cos^2\theta}{n_o^2}. \quad (2.24)$$

$\bar{n}_e = n_e(\pi/2)$ is the *principal value* of the extraordinary refractive index. By adjusting the angle θ , we can try to achieve phase matching as we illustrate in Fig. 14.

Uniaxial crystals can be positive ($n_e > n_o$) or negative ($n_e < n_o$). For SHG and a negative uniaxial crystal, the input field must have ordinary polarization, and we get phase matching if:

$$n_e(2\omega, \theta) = n_o(\omega). \quad (2.25)$$

Using this and equation 2.24, we get:

$$\frac{\sin^2\theta}{\bar{n}_e(2\omega)^2} + \frac{\cos^2\theta}{n_o(2\omega)^2} = \frac{1}{n_o(\omega)^2}. \quad (2.26)$$

A crystal will allow phase matching if a solution of this equation is possible.

A drawback of using birefringent phase matching is that, if the angle θ is not 0 or $\pi/2$, the Poynting vector \vec{S} and the wavevector \vec{k} will not point in the same direction. Then the input beams and the light generated will not travel in the same direction. This results in a *transverse walk-off* between these beams, and it limits the interaction volume for the non-linear effect because at some point the walk-off becomes larger than

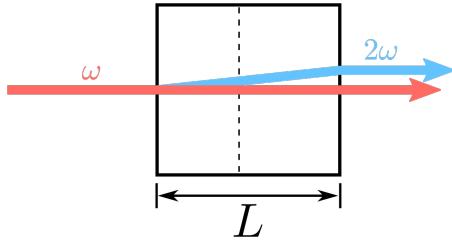


Figure 15: **Transverse walk-off in SHG.** If the incoming beam has ordinary polarization and the SHG beam has extraordinary polarization, the SHG beam will experience a transverse walk-off. After some distance indicated by the dashed line, there will essentially be no overlap between the two beams, and a longer crystal will not lead to increased SHG output.

the beams (see Fig. 15). If one uses a pulsed laser for non-linear processes, we may also encounter a *longitudinal walk-off* if the group velocities for the beams involved differ.

2.4 Quasi phasematching

Quasi phasematching (QPM) is a technique that allows to overcome several of the limitations we mentioned for “normal” phase matching. For example, it allows to achieve **type-0 phase matching**, where the polarizations of all fields are the same. In this case, one can make use of the d_{33} nonlinear coefficient, which is often significantly larger than other nonlinear coefficients.

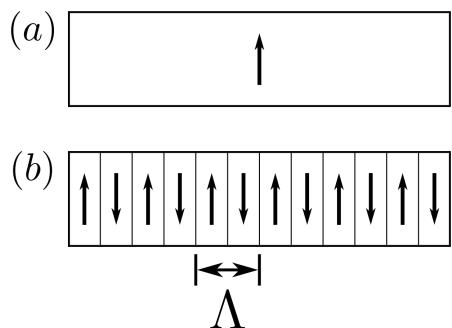


Figure 16: **Periodic poling.** (a) In a homogeneous nonlinear crystal, the crystal’s c axis points in the same direction everywhere in the crystal. (b) In a periodically poled crystal, the orientation of the c axis is periodically flipped. The poling period is Λ .

The idea is the following: earlier we introduced the L_{coh} as the length over which a coherent buildup of the frequency-converted field can occur. If one had the possibility to rotate the crystal by 180° after a distance comparable to L_{coh} , one could reestablish good phasematching before the interacting fields became out of phase. By repeatedly doing that in regular distances, one can achieve the phasematching even in very long crystals.

To achieve that, assume that the material used is ferroelectric. Then we can apply a periodically modulated field to the crystal during the fabrication process, and this will periodically invert the sign of d_{eff} (see Fig. 16). If $d(z)$ is the spatial dependence of the effective nonlinear coefficient along the crystal, we can write:

$$d(z) = d_{\text{eff}} \text{sign} \left[\cos \left(z \frac{2\pi}{\Lambda} \right) \right]. \quad (2.27)$$

Sometimes a more complicated spatial dependence is chosen, e.g., to achieve simultaneous phase matching for multiple processes. To describe QPM mathematically, let us replace d_{eff} by $d(z)$ in our coupled amplitude equations. To solve the equations, let us first decompose $d(z)$ into a Fourier series:

$$d(z) = d_{\text{eff}} \sum_{m=-\infty}^{\infty} G_m \exp(iq_m z). \quad (2.28)$$

q_m are the grating wave vectors, and the Fourier coefficients are:

$$G_m = \frac{2}{m\pi} \sin(m\pi/2). \quad (2.29)$$

If one assumes that the phasematching process is dominated by one particular component m in equation 2.28, our coupled equations become:

$$\begin{aligned} \frac{dA_1}{dz} &= \frac{2i\omega_1 d_Q}{n_1 c} A_3 A_2^* e^{-i(\Delta k_Q - 2q_m)z}, \\ \frac{dA_2}{dz} &= \frac{2i\omega_2 d_Q}{n_2 c} A_3 A_1^* e^{-i(\Delta k_Q - 2q_m)z}, \\ \frac{dA_3}{dz} &= \frac{2i\omega_3 d_Q}{n_3 c} A_1 A_2 e^{i\Delta k_Q z}. \end{aligned} \quad (2.30)$$

Here, we defined $d_Q = d_{\text{eff}} G_m$, and the wavevector mismatch for order m is:

$$\Delta k_Q = k_1 + k_2 - k_3 + q_m. \quad (2.31)$$

Compared to our earlier coupled amplitude equations, we have d_Q instead of d_{eff} and, most importantly, the wavevector mismatch is modified.

Because d_Q typically decreases with increasing m , usually one tries to establish QPM

using a first-order interaction ($m = -1$). In that case, we get:

$$\Delta k_Q = k_1 + k_2 - k_3 - \frac{2\pi}{\Lambda}, \quad d_Q = (2/\pi)d_{\text{eff}}. \quad (2.32)$$

For $\Delta k_Q = 0$, we can conclude that the optimal poling period Λ has to be:

$$\Lambda = 2L_{\text{coh}} = 2\pi/(k_1 + k_2 - k_3). \quad (2.33)$$

2.5 Frequency upconversion

Frequency upconversion is a special case of SFG. We assume the pump field with ω_2 is strong but the field with ω_1 is weak. Then SFG allows us to create a signal with $\omega_3 = \omega_1 + \omega_2$ where the strong pump allows us to convert the frequency of the input signal with ω_1 to ω_3 .

We can use equations 2.14 and 2.15 to describe this effect. Because A_2 is strong, we can assume it remains approximately constant during the process. Then we get:

$$\begin{aligned} \frac{dA_1}{dz} &= K_1 A_3 e^{-i\Delta kz}, \\ \frac{dA_3}{dz} &= K_3 A_1 e^{i\Delta kz}, \end{aligned} \quad (2.34)$$

with the coupling quantities:

$$K_1 = \frac{2i\omega_1^2 d_{\text{eff}}}{k_1 c^2} A_2^*, \quad K_3 = \frac{2i\omega_3^2 d_{\text{eff}}}{k_3 c^2} A_2. \quad (2.35)$$

For $\Delta k = 0$ and by eliminating A_3 , we get:

$$\frac{d^2 A_1}{dz^2} = -\kappa^2 A_1. \quad (2.36)$$

Here, we introduced the (positive) coupling coefficient κ^2 as:

$$\kappa^2 \equiv -K_1 K_3 = \frac{4\omega_1^2 \omega_3^2 d_{\text{eff}}^2 |A_2|^2}{k_1 k_3 c^4}. \quad (2.37)$$

The solutions of our coupled amplitude equations for this situation are:

$$\begin{aligned} A_1(z) &= B \cos(\kappa z) + C \sin(\kappa z), \\ A_3(z) &= -\frac{B\kappa}{K_1} \sin(\kappa z) + \frac{C\kappa}{K_1} \cos(\kappa z). \end{aligned} \quad (2.38)$$

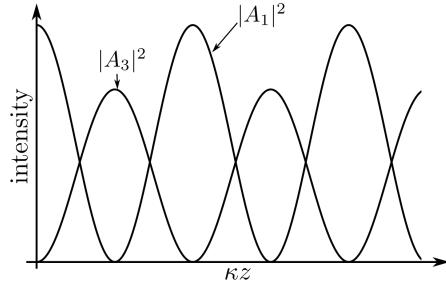


Figure 17: **Frequency conversion** allows converting the input field A_1 at ω_1 into an output field at ω_3 . The intensities of the two fields evolve along the crystal. For perfect phase matching, we get perfect conversion efficiency.

To accurately describe the frequency conversion of a weak signal, we can choose the integration constants B, C such that the solutions fulfill the boundary conditions. In particular, at the entry face of the crystal, the converted field should be zero - that means $A_3(0) = 0$. This implies that $C = 0$ and $B = A_1(0)$. Figure 17 shows how the intensities of the fields with ω_1 and ω_3 evolve along the crystal.

2.6 Second harmonic generation

For SHG, we have an input field with ω_1 , an output field with $\omega_2 = 2\omega_1$, and the total field is:

$$E(z, t) = e^{-i\omega_1 t + ik_1 z} A_1(z) + e^{-i\omega_2 t + ik_2 z} A_2(z).$$

The light propagates in the z direction, and the A_j vary slowly with z , and the individual frequency components of E have to fulfill the wave equations 2.9, which we write as:

$$\frac{\partial^2 E_j(z, t)}{\partial z^2} - \frac{\epsilon(1)(\omega_j)}{c^2} \frac{\partial^2 E_j(z, t)}{\partial t^2} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 P_j(z, t)}{\partial t^2}, \quad (2.39)$$

where $P_j(z, t) = P_j(z) e^{-i\omega_j t} + \text{c.c.}$ and

$$P^{\text{NL}}(z, t) = P_1(z, t) + P_2(z, t), \quad (2.40)$$

and we have:

$$\begin{aligned} P_1(z) &= 4\epsilon_0 d_{\text{eff}} E_2 E_1^* = 4\epsilon_0 d_{\text{eff}} A_2 A_1^* e^{i(k_2 - k_1)z}, \\ P_2(z) &= 2\epsilon_0 d_{\text{eff}} E_1^2 = 2\epsilon_0 d_{\text{eff}} A_1^2 e^{2ik_1 z}. \end{aligned} \quad (2.41)$$

Note the different degeneracy factors 2 and 4 in the above equation.

We then arrive at the following coupled amplitude equations for SHG:

$$\begin{aligned}\frac{dA_1}{dz} &= \frac{2i\omega_1^2 d_{\text{eff}}}{k_1 c^2} A_2 A_1^* e^{-i\Delta kz}, \\ \frac{dA_2}{dz} &= \frac{i\omega_2^2 d_{\text{eff}}}{k_2 c^2} A_1^2 e^{i\Delta kz},\end{aligned}\quad (2.42)$$

where $\Delta k = 2k_1 - k_2$.

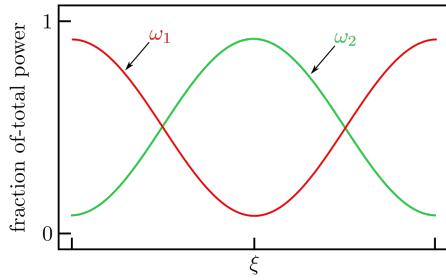


Figure 18: **Energy exchange.** For arbitrary but non-zero initial values of u_1 and u_2 , power from the fundamental frequency ω_1 is converted to $\omega_2 = 2\omega_1$ along the crystal over a typical length scale l . As ξ increases, the frequency conversion at some point saturates and then takes place from ω_2 to ω_1 instead.

To solve these equations, it is convenient to write the amplitudes as follows:

$$\begin{aligned}A_1 &= \left(\frac{I}{2n_1 \epsilon_0 c} \right)^{1/2} u_1 e^{i\phi_1}, \\ A_2 &= \left(\frac{I}{2n_2 \epsilon_0 c} \right)^{1/2} u_2 e^{i\phi_1}.\end{aligned}\quad (2.43)$$

We introduced the intensities $I = I_1 + I_2$ of the two fields:

$$I_j = 2n_j \epsilon_0 c |A_j|^2. \quad (2.44)$$

$u_{1,2}$ are slowly varying functions of z . It is convenient to introduce a normalized distance parameter $\xi = z/l$, where

$$l = \left(\frac{2n_1^2 n_2}{\epsilon_0 c I} \right)^{1/2} \frac{c}{2\omega_1 d_{\text{eff}}}. \quad (2.45)$$

This is the typical distance over which our two fields exchange energy. This exchange can be visualized as illustrated in Fig. 18.

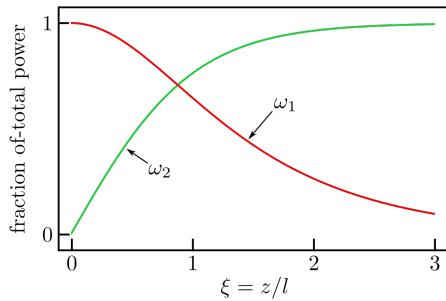


Figure 19: **Amplitudes of the fundamental field and the second harmonic field.**

If $u_1 = 1$ and $u_2 = 0$ in the beginning, then all the power will eventually be in u_2 for $\xi \rightarrow \infty$.

If one assumes the boundary conditions $u_1(0) = 1, u_2(0) = 0$, the solutions of these coupled differential equations are:

$$\begin{aligned} u_1(\xi) &= \operatorname{sech} \xi = (\cosh \xi)^{-1}, \\ u_2(\xi) &= \tanh \xi. \end{aligned} \quad (2.46)$$

This case is shown in Fig. 19.

2.7 Difference frequency generation and parametric amplification

Consider the case illustrated in Fig. 20. We will assume that the field with ω_3 is strong and that there is no depletion such that A_3 is constant. Using a similar approach as for SFG in section 2.2, one gets:

$$\begin{aligned} \frac{dA_1}{dz} &= \frac{2i\omega_1^2 d_{\text{eff}}}{k_1 c^2} A_3 A_2^* e^{i\Delta kz}, \\ \frac{dA_2}{dz} &= \frac{2i\omega_2^2 d_{\text{eff}}}{k_2 c^2} A_3 A_1^* e^{i\Delta kz}, \end{aligned} \quad (2.47)$$

where the wave vector mismatch is:

$$\Delta k = k_3 - k_1 - k_2. \quad (2.48)$$

For $\Delta k = 0$, one can combine the above equations to yield:

$$\frac{d^2 A_2}{dz^2} = \frac{4\omega_1^2\omega_2^2 d_{\text{eff}}^2}{k_1 k_2 c^4} A_3 A_3^* A_2 \equiv \kappa^2 A_2. \quad (2.49)$$

The real coupling constant κ is given by:

$$\kappa^2 = \frac{4d_{\text{eff}}^2\omega_1^2\omega_2^2}{k_1 k_2 c^4} |A_3|^2. \quad (2.50)$$

If one assumes the boundary condition $A_2(0) = 0$, one gets the solution:

$$\begin{aligned} A_1(z) &= A_1(0)\cosh(\kappa z) \\ A_2(z) &= i \left(\frac{n_1\omega_2}{n_2\omega_1} \right)^{1/2} \frac{A_3}{|A_3|} A_1^*(0) \sinh(\kappa z). \end{aligned} \quad (2.51)$$

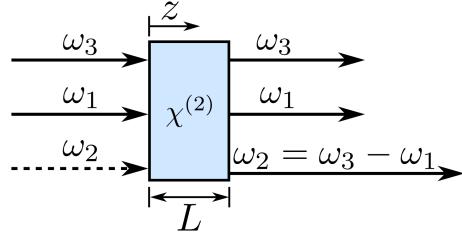


Figure 20: **Difference frequency generation (DFG).** Typically there is no input with ω_2 but only inputs at ω_1 and ω_3 .

We should note two things: (1) equations 2.51 show that A_1 retains the phase of the initial field but is amplified. The phase of the generated A_2 field depends on both input fields. (2) $A_1(z)$ and $A_2(z)$ increase monotonically. There is no oscillation between the fields like in the case of SFG or SHG. One can intuitively understand the second point by considering the DFG process in terms of transitions between energy levels as illustrated in Fig. 21. After an excitation by a photon of frequency ω_3 to a virtual energy level, the decay back to the ground state can happen in two ways: (a) a field at ω_1 can stimulate the according transition, leading to the generation of an ω_2 photon. Or (b), the presence of a field at ω_2 can stimulate the transition that leads to the emission of an ω_1 photon.

The interesting thing about this is that, a higher intensity at one frequency will lead to a higher intensity at the other frequency. Because the generation of the ω_1 field is amplified by the ω_2 field, which is generated in the parametric process of DFG, this is known as *optical parametric amplification*¹. In this context, one says that the signal

¹The difference between parametric and non-parametric processes is that the susceptibility in a para-

field at ω_1 is amplified by the nonlinear mixing process, and that the idler field at ω_2 is generated by this process. Here, I introduced the conventional terms “signal” and “idler” for the higher-frequency and the lower-frequency fields, respectively.

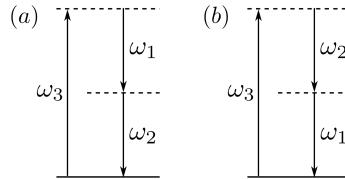


Figure 21: **Energy levels in DFG.** The decay from the excited state can occur in two ways shown in (a) and (b).

2.8 Optical parametric oscillators (OPOs)

If one places the nonlinear crystal used for parametric amplification inside a cavity resonant with ω_1 and/or ω_2 , one can enhance that amplification. Then only a pump field at ω_3 is necessary. The resulting device is then known as an optical parametric oscillator (OPO). Fig. 22 illustrates the schematic layout of an OPO and the corresponding energy transitions. At some point, the assumption that the field at ω_3 is approximately constant will not be valid any longer. But it is in principle possible to achieve very high efficiency in frequency conversion by “cavity-enhancing” nonlinear processes.

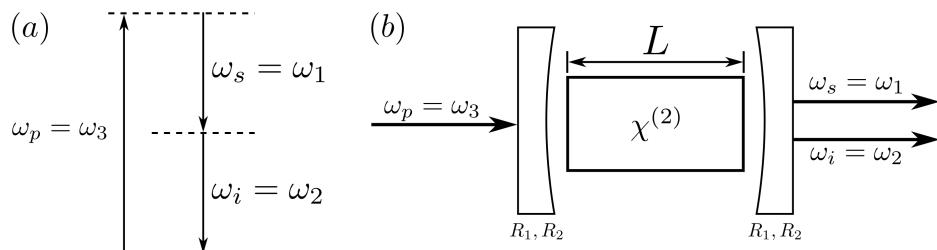


Figure 22: **Optical Parametric Oscillator.** (a) shows the energy levels as in DFG but with the notation of pump, signal and idler photons. (b) the nonlinear crystal incide a Fabry-Perot cavity with cavity mirrors that each have reflectivities $R_1 = |r_1|^2$ and $R_2 = |r_2|^2$ at the signal and idler wavelengths, respectively.

The coupled wave equations describing the evolution of the amplitudes along z in an OPO are the same as for DFG. The general solutions for these equations for $\Delta k =$

metric process is real and that the photon energy is conserved in the process[1].

$k_3 - k_1 - k_2$ are:

$$\begin{aligned} A_1(z) &= \left[A_1(0) \left(\cosh gz - \frac{i\Delta k}{2g} \sinh gz \right) + \frac{\kappa_1}{g} A_2^*(0) \sinh gz \right] e^{i\Delta kz/2}, \\ A_2(z) &= \left[A_2(0) \left(\cosh gz - \frac{i\Delta k}{2g} \sinh gz \right) + \frac{\kappa_2}{g} A_1^*(0) \sinh gz \right] e^{i\Delta kz/2}, \end{aligned} \quad (2.52)$$

with the gain g and the coupling constants κ_m given by:

$$g = [\kappa_1 \kappa_2^* - (\Delta k/2)^2]^{1/2} \quad \text{and } \kappa_m = \frac{2i\omega_m^2 d_{\text{eff}} A_3}{k_m c^2}. \quad (2.53)$$

For $\Delta k = 0$ and for $A_2(0) = 0$, the solutions become:

$$\begin{aligned} A_1(z) &= A_1(0) \cosh gz, \\ A_2(z) &= i \left(\frac{n_1 \omega_2}{n_2 \omega_1} \right) \frac{A_3}{|A_3|} A_1^*(0) \sinh gz. \end{aligned} \quad (2.54)$$

For large gz , both of these solutions asymptotically are proportional to $\exp(gz)$. That means we will have exponential growth. Of course, that is only true as long as the pump field is not significantly depleted.

An OPO will only have a steady output at ω_1 and ω_2 , if the gain inside the cavity is larger than the losses. This defines the *threshold for optical parametric oscillation*. To be more precise, at the threshold the field amplitudes $A_{1,2}$ should not change in one round trip in the cavity. This condition or the two fields can be written as:

$$\begin{aligned} A_1(0) &= \left[A_1(0) \cosh gL + \frac{\kappa_1}{g} A_2^*(0) \sinh gL \right] (1 - l_1), \\ A_2^*(0) &= \left[A_2^*(0) \cosh gL + \frac{\kappa_2^*}{g} A_1(0) \sinh gL \right] (1 - l_2), \end{aligned} \quad (2.55)$$

where $l_m = 1 - R_m e^{-\alpha_m L}$ is the fractional amplitude loss per round trip, α_m is the absorption coefficient for ω_m , and R_m are the reflectivities of the two cavity mirrors. Combining the equations above, one can conclude that:

$$\cosh gL = 1 + \frac{l_1 l_2}{2 - l_1 - l_2}. \quad (2.56)$$

3 Intensity-dependent refractive index

So far, we focused on second-order nonlinear optical effects, which we described using the $\chi^{(2)}$ nonlinearity. It is a very common effect, though, that light passing through a material, especially if it has high intensity, can change the optical properties of materials. This is also known as the *optical Kerr effect*. A well-known example is the self-focusing of a beam. Here, we will discuss how to describe such effects, and then we will discuss some prominent examples and applications.

3.1 Describing the intensity-dependent refractive index

In many cases, we can describe the dependence of the refractive index of a material on light passing through it via the relation

$$n = n_o + \bar{n}_2 \langle E^2 \rangle. \quad (3.1)$$

n_0 is the unmodified refractive index, and \bar{n}_2 is the second-order refractive index[1]. The bar on top of \bar{n}_2 is to avoid confusion with other quantities we will introduce[1]. The angular brackets denote the time average. If the electric field is a plane wave

$$E(t) = E(\omega)e^{-i\omega t} + \text{c.c.}, \quad (3.2)$$

we get:

$$n = n_o + 2\bar{n}_2|E(\omega)|^2. \quad (3.3)$$

In terms of the nonlinear polarization, we can describe the optical Kerr effect as:

$$P^{\text{NL}}(\omega) = 3\epsilon_0\chi^{(3)}(\omega = \omega + \omega - \omega)|E(\omega)|^2E(\omega). \quad (3.4)$$

We assumed that all fields are linearly polarized to suppress the tensor character of $\chi^{(3)}$.

If we neglect second-order effects, the total polarization up to third order will be:

$$P^{\text{TOT}}(\omega) = \epsilon_0\chi^{(1)}E(\omega) + 3\epsilon_0\chi^{(3)}|E(\omega)|^2E(\omega) \equiv \epsilon_0\chi_{\text{eff}}E(\omega) \quad (3.5)$$

with the effective susceptibility

$$\chi_{\text{eff}} = \chi^{(1)} + 3\chi^{(3)}|E(\omega)|^2. \quad (3.6)$$

If we write $n^2 = 1 + \chi_{\text{eff}}$, this results in:

$$\begin{aligned} n_0 &= (1 + \chi^{(1)})^{1/2}, \\ \bar{n}_2 &= \frac{3\chi^{(3)}}{4n_0}. \end{aligned} \quad (3.7)$$

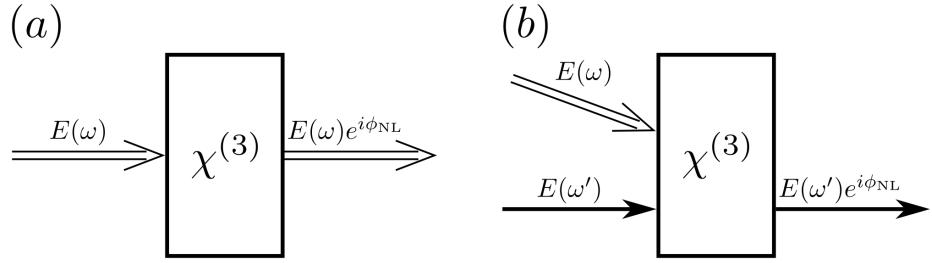


Figure 23: **Measuring the intensity-dependent refractive index.** (a) if the field $E(\omega)$ is sufficiently strong, it will modify the refractive index of a $\chi^{(3)}$ nonlinear medium and introduce a nonlinear phase shift ϕ_{NL} in the same field. (b) instead of observing changes to the strong field itself, one can measure the phase shift experienced by a non-collinear probe field $E(\omega')$ due to the change in refractive index caused by the strong field $E(\omega)$.

So far, we assumed that we observe changes to the refractive index due to the field $E(\omega)$ by studying its effect on the same field. Alternatively, we can use a weak probe field $E(\omega')$ to monitor changes of the refractive index caused by the stronger field $E(\omega)$ (see Fig. 23). The nonlinear polarization at the frequency ω' will then be:

$$P^{\text{NL}}(\omega') = 6\epsilon_0\chi^{(3)}(\omega' = \omega' + \omega - \omega)|E(\omega)|^2E(\omega'). \quad (3.8)$$

It is important to note that the degeneracy factor (6) in this case is larger than in equation 3.5. This is even true if $\omega = \omega'$ because the two modes propagate in different directions (see Fig. 23). The refractive index encountered by the probe will be:

$$n = n_0 + 2\bar{n}_2^{\text{(cross)}}|E(\omega)|^2, \quad (3.9)$$

where:

$$\bar{n}_2^{\text{(cross)}} = \frac{3\chi^{(3)}}{2n_0}. \quad (3.10)$$

This is twice as large as the \bar{n}_2 in Fig. 23(a). That means a strong wave will influence the refractive index of a weak wave twice as strongly as it affects its own refractive index. This phenomenon is known as “weak-wave retardation”[4].

One can also define the refractive index as a function of the light intensity instead of the electric field. For the case of a single beam as illustrated in Fig. 23, we can write:

$$n = n_0 + n_2 I = n_0 + 2n_2 n_0 \epsilon_0 c |E(\omega)|^2. \quad (3.11)$$

The relation between \bar{n}_2 and n_2 is:

$$n_2 = \frac{\bar{n}_2}{n_0 \epsilon_0 c} = \frac{3}{4n_0^2 \epsilon_0 c} \chi^{(3)}. \quad (3.12)$$

3.2 The third-order susceptibility tensor

So far, we described the intensity-dependent refractive index using the effective susceptibility. In a full treatment, $\chi_{ijkl}^{(3)}$ is a tensor with four indices with 81 separate elements. Many interesting third-order nonlinear effects occur even in isotropic media, though, like glass, air, liquids, or vapours. In such isotropic media, $\chi_{ijkl}^{(3)}$ has only three independent and non-zero elements.

Because all coordinate axes must be equivalent, we must have:

$$\begin{aligned} \chi_{1111} &= \chi_{2222} = \chi_{3333}, \\ \chi_{1122} &= \chi_{1133} = \chi_{2211} = \chi_{2233} = \chi_{3311} = \chi_{3322}, \\ \chi_{1212} &= \chi_{1313} = \chi_{2323} = \chi_{2121} = \chi_{3131} = \chi_{3232}, \\ \chi_{1221} &= \chi_{1331} = \chi_{2112} = \chi_{2332} = \chi_{3113} = \chi_{3223}. \end{aligned} \quad (3.13)$$

Also, a rotation around an arbitrary axes must lead to the same physical results. In particular, this need to be true for a rotation by 45 degrees, which implies:

$$\chi_{1111} = \chi_{1122} + \chi_{1212} + \chi_{1221}. \quad (3.14)$$

As a result, the nonlinear susceptibility can be written as:

$$\chi_{ijkl} = \chi_{1122} \delta_{ij} \delta_{kl} + \chi_{1212} \delta_{ik} \delta_{jl} + \chi_{1221} \delta_{il} \delta_{jk}. \quad (3.15)$$

For third-harmonic generation, this simplifies even further. Because of inherent permutation symmetries, we must have $\chi_{1122} = \chi_{1212} = \chi_{1221}$, and we get:

$$\chi_{ijkl}(3\omega = \omega + \omega + \omega) = \chi_{1122}(3\omega = \omega + \omega + \omega)(\delta_{ij} \delta_{kl} + \delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk}). \quad (3.16)$$

If we consider the the intrinsic permutation symmetry for the nonlinear refractive

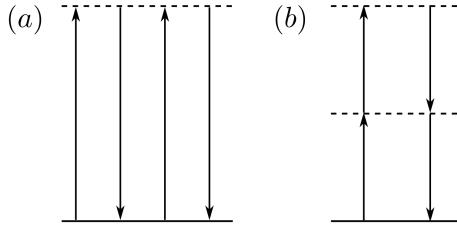


Figure 24: **Level diagrams for intensity-dependent refractive index.** Level representation of the two contributions to the nonlinear polarization for describing the intensity-dependent refractive index.

index, we must have $\chi_{1122} = \chi_{1212}$, such that:

$$\begin{aligned}\chi_{ijkl}(\omega = \omega + \omega - \omega) &= \chi_{1122}(\omega = \omega + \omega - \omega)(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl}) + \\ &\quad \chi_{1221}(\omega = \omega + \omega - \omega)\delta_{il}\delta_{jk}.\end{aligned}\quad (3.17)$$

Then the corresponding nonlinear polarization can be written as:

$$\begin{aligned}P_i^{\text{NL}}(\omega) &= 3\epsilon_0 \sum_{jkl} \chi_{ijkl}(\omega = \omega + \omega - \omega) E_j(\omega) E_k(\omega) E_l(-\omega) \\ &= 6\epsilon_0 \chi_{1122} E_i(\vec{E} \cdot \vec{E}^*) + 3\epsilon_0 \chi_{1221} E_i^*(\vec{E} \cdot \vec{E}).\end{aligned}\quad (3.18)$$

We can rewrite this as

$$\vec{P}^{\text{NL}} = \epsilon_0 A (\vec{E} \cdot \vec{E}^*) \vec{E} + \frac{1}{2} \epsilon_0 B (\vec{E} \cdot \vec{E}) \vec{E}^*, \quad (3.19)$$

where:

$$\begin{aligned}A &= 6\chi_{1122} = 3\chi_{1122} + 3\chi_{1212}, \\ B &= 6\chi_{1221}.\end{aligned}\quad (3.20)$$

Equation 3.19 has two distinct contributions for the A and B coefficients. One of these contributions behaves like \vec{E} and the other like \vec{E}^* . The “handedness” (right vs left) or *chirality* therefore remains the same for the A contribution but it switches for the B contribution. Fig. 24 illustrates one-photon and two-photon resonant contributions to the nonlinear polarization. The one-photon process only contributes to coefficient A . The two-photon process can, in principle, contribute to A and B , but under some circumstances it only contributes to B . We will learn more about this later.

3.2.1 Propagation through an isotropic nonlinear medium

Let us briefly discuss the propagation of light through such a nonlinear medium and how it affects the polarization of light. Given that the B contribution to the polarization in equation 3.19 has the opposite chirality than the incoming light, we can expect that there may be some phase shift between the different components of polarization in a beam transmitted through such a medium.

Let us start by writing our electric field in the basis of circular polarization:

$$\vec{E} = E_+ \hat{\sigma}_+ + E_- \hat{\sigma}_-, \quad (3.21)$$

where the unit vectors for right-hand ($\hat{\sigma}_+$) and left-hand polarization ($\hat{\sigma}_-$) are given by:

$$\hat{\sigma}_\pm = \frac{\hat{x} \pm i\hat{y}}{\sqrt{2}}. \quad (3.22)$$

\hat{x} and \hat{y} denote the unit vectors along the x and y axis, respectively.

Using the definition in equation 3.21, we can write equation 3.19 as:

$$\vec{P}^{\text{NL}} = \epsilon_0 A(|E_+|^2 + |E_-|^2) \vec{E} + \epsilon_0 B(E_+ E_-) \vec{E}^*. \quad (3.23)$$

If we decompose the polarization into a right-hand and a left-hand part as

$$\vec{P}^{\text{NL}} = P_+ \hat{\sigma}_+ + P_- \hat{\sigma}_-, \quad (3.24)$$

then we can write:

$$P_\pm \equiv \epsilon_0 \chi_\pm^{\text{NL}} E_\pm. \quad (3.25)$$

Here, we introduced the effective nonlinear susceptibilities

$$\chi_\pm^{\text{NL}} = A|E_\pm|^2 + (A+B)|E_\mp|^2. \quad (3.26)$$

Let us use these definitions in the wave equation in our nonlinear medium:

$$\nabla^2 \vec{E}(z, t) = \frac{\epsilon^{(1)}}{c^2} \frac{\partial^2 \vec{E}(z, t)}{\partial t^2} + \frac{1}{\epsilon_0 c^2} \frac{\partial^2 \vec{P}^{\text{NL}}}{\partial t^2}. \quad (3.27)$$

We get the following two differential equations:

$$\nabla^2 E_\pm(z, t) = \frac{\epsilon_\pm^{\text{eff}}}{c^2} \frac{\partial^2 E_\pm(z, t)}{\partial t^2}, \quad (3.28)$$

where

$$\epsilon_\pm^{\text{eff}} = \epsilon^{(1)} + \chi_\pm^{\text{NL}}. \quad (3.29)$$

Equation 3.28 has plane-wave solutions travelling with phase velocities c/n^\pm with

$n_{\pm}^2 = \epsilon_{\pm}^{\text{eff}}$. Plane waves with those polarizations experience different refractive indices:

$$\begin{aligned} n_{\pm}^2 &= n_0^2 + \chi_{\pm}^{\text{NL}} = n_0^2 \left(1 + \frac{1}{n_0^2} [A|E_{\pm}|^2 + (A+B)|E_{\mp}|^2] \right) \\ &\approx \left\{ n_0 + \frac{1}{2n_0} [A|E_{\pm}|^2 + (A+B)|E_{\mp}|^2] \right\}^2. \end{aligned} \quad (3.30)$$

There will be a relative phase shift due to a difference Δn between refractive indices:

$$\Delta n = n_+ - n_- = \frac{B}{2n_0} (|E_-|^2 - |E_+|^2). \quad (3.31)$$

From this equation we can conclude: (a) linear polarization and circular polarization are unaffected (why?) while light with elliptical polarization will experience a rotation of polarization. (b) this effect only depends on the B coefficient. That means only two-photon processes (Fig. 24) contribute to this effect. The rotation angle will be:

$$\theta = \frac{1}{2} \Delta n \frac{\omega}{c} z. \quad (3.32)$$

3.3 Self-focusing of light and other self-action effects

The change in the refractive index of a nonlinear material due to high-intensity light can lead to self-focusing. Because the refractive index will be higher in regions of higher intensity, the intense light generates a lens-like refractive-index distribution. At the same time, the diffraction of light will lead to a growth in the beam diameter. When these two effects balance exactly, the beam is said to trap itself. This occurs at the critical power P_{cr} [1]:

$$P_{\text{cr}} = \frac{\pi(0.61)^2 \lambda_0^2}{8n_0 n_2}. \quad (3.33)$$

λ_0 is the wavelength in vacuum. Self-focusing occurs for $P \geq P_{\text{cr}}$. For $P \gg P_{\text{cr}}$, the beam will split into filaments due to imperfections in the laser front. Each filament will carry a power of approximately P_{cr} . We will now discuss these effects in more detail.

Let us first consider the self-focusing of light in the ray-optics regime following the description in Ref. [1]. That means, we neglect diffraction effects in this context. This approximation can be made if the beam diameter or the beam intensity is sufficiently large. The argument is illustrated in Fig. 25. Let us assume that the linear refractive index is n_0 , and that the nonlinear refractive index along the central ray is $n_0 + n_2 I_0$, where I_0 is the intensity of the beam. According to Fermat's principle, we require:

$$(n_0 + n_2 I) z_{\text{sf}} = n_0 z_{\text{sf}} / \cos \theta_{\text{sf}}. \quad (3.34)$$

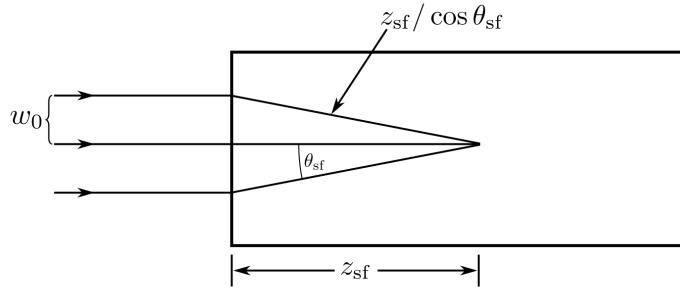


Figure 25: **Self-focusing.** Assuming ray optics propagation, we can predict the focal distance for self-focusing using Fermat's principle.

If we assume that $\cos(\theta_{sf}) \approx 1 - \theta_{sf}^2$, we can solve equation 3.34 to yield:

$$\theta_{sf} = \sqrt{2n_2 I / n_0}. \quad (3.35)$$

This “self-focusing angle” is typically a small quantity is the typical angle by which a beam is deflected due to self-focusing. The self-focusing distance for $P \gg P_{cr}$ is:

$$z_{sf} = w_0 \sqrt{\frac{n_0}{2n_2 I}} = \frac{2n_0 w_0^2}{\lambda_0} \frac{1}{\sqrt{P/P_{cr}}}. \quad (3.36)$$

We assumed $P \gg P_{cr}$ to ensure that self-focusing is much stronger than diffraction. If one takes diffraction into account, diffraction will counter-act the self-focusing such that the overall focusing angle becomes approximately $\theta = (\theta_{sf}^2 - \theta_{diff}^2)^{1/2}$, where

$$\theta_{diff} = \frac{0.61 \lambda_0}{n_0 d} \quad (3.37)$$

is the diffraction angle for a beam with diameter d . Then the focusing distance becomes[1]:

$$z_{sf} = \frac{2n_0 w_0^2}{\lambda_0} \frac{1}{\sqrt{P/P_{cr} - 1}}. \quad (3.38)$$

For arbitrary powers and arbitrary beam-waist position, one can show that [1]:

$$z_{sf} = \frac{kw^2/2}{(P/P_{cr} - 1)^{1/2} + 2z_{min}/kw_0^2}, \quad (3.39)$$

where $k = n_0 \omega/c$, w is the waist of the beam when it enters the crystal, and z_{min} is the distance from the crystal face to the beam waist if there was no medium.

3.3.1 Self-trapping of light

Roughly speaking, self-trapping of light occurs if we have

$$\theta_{\text{diff}} = \theta_{\text{sf}}, \quad (3.40)$$

i.e., the diffraction of the beam is balanced by the self-focusing of the light.

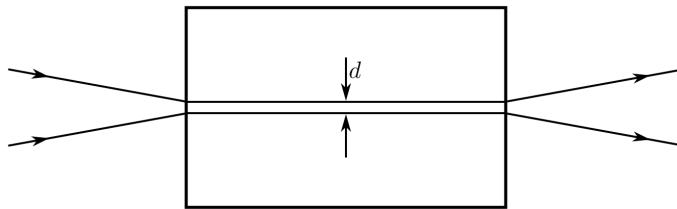


Figure 26: **Self-trapping.** If the incoming beam has the critical power P_{cr} , the light will trap itself transversally and travel along the crystal with some constant diameter d .

If we use equations 3.35 and 3.37, the intensity required for self trapping becomes:

$$I = \frac{(0.61)^2 \lambda_0^2}{2n_2 n_0 d^2}. \quad (3.41)$$

d is the diameter of the self-trapped beam (Fig. 26). If we express the intensity in terms of power, we find that self-trapping occurs if:

$$P = P_{\text{cr}} = \frac{\pi(0.61)^2 \lambda_0^2}{8n_0 n_2} \approx \frac{\lambda_0^2}{8n_0 n_2}. \quad (3.42)$$

Here, d cancelled out. I.e., self-trapping only depends on the power but not on the intensity. The beam diameter will depend on how we focus that power into the material.

A possible alternative way to understand this effect was illustrated by Chiao et al.[5, 1]. If we assume that the laser beam has a flat-top intensity distribution, it will lead to a modified refractive index $n_0 + \delta n$ inside the medium. This will create a waveguide inside which light will be trapped if it fulfills the condition required for total internal reflection (see Fig. 27). Self-trapping occurs if the angle θ is less than the critical angle θ_{tot} for total internal reflection, where:

$$\cos(\theta_{\text{tot}}) = \frac{n_0}{n_0 + \delta n}. \quad (3.43)$$

Because δn will typically be small, we can Taylor expand the cosine up to second

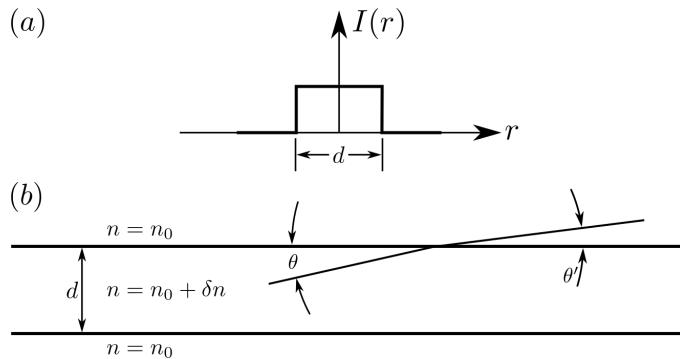


Figure 27: **Self-trapping of a beam with flat-topped intensity.** (a) the flat-top radial intensity distribution. (b) a ray of light originating inside the higher-index region is diffracted into the low-index region at an angle $\theta' < \theta$. Total reflection can trap the eay inside the waveguide.

order and get:

$$\theta_{\text{tot}} = \left(\frac{2\delta n}{n_0} \right)^{1/2}. \quad (3.44)$$

If we have a beam of light with diameter \$d\$, the rays inside that beam will have an angular distribution with angles $\leq \theta_{\text{diff}}$. Setting this equal to θ_{tot} , we can conclude that self-trapping will occur if the deviation of the refractive index fulfills:

$$\delta n = \frac{1}{2} n_0 \left(\frac{0.61\lambda_0}{dn_0} \right)^2. \quad (3.45)$$

On the other hand, if δn is given, then this becomes a condition on the beam diameter:

$$d = \frac{0.61\lambda_0}{\sqrt{2n_0\delta n}}. \quad (3.46)$$

If we replace δn with $n_2 I$, and if we use equation 3.46, we get the corresponding beam power:

$$P_{\text{cr}} = \frac{\pi}{4} d^2 I = \frac{\pi(0.61)^2 \lambda_0^2}{8n_0 n_2}. \quad (3.47)$$

Mathematically describing self-action effects

To describe the effects of self-focusing and self-trapping more rigorously, let us look at the paraxial wave equation. This is appropriate because the angles resulting from self-focusing or self-trapping will be very small, and therefore the transverse components of

the wavevectors will be much smaller than the longitudinal component along the beam axis z . The paraxial wave equation for the slowly varying amplitude of our beam will be:

$$2ik_0 \frac{\partial A}{\partial z} + \nabla_T^2 A = -\frac{\omega^2}{\epsilon_0 c^2} p_{NL}. \quad (3.48)$$

Here, we used p_{NL} to denote the slowly varying amplitude of the nonlinear polarization in order to distinguish it from the expression for the full nonlinear polarization P^{NL} .

For a purely third-order nonlinear response, this slowly varying amplitude of the nonlinear polarization will be[1]:

$$p_{NL} = 3\epsilon_0 \chi^{(3)} |A|^2 A. \quad (3.49)$$

This equation can be used to find steady-state solutions for self trapping - that means, this applies for continuous-wave laser beams, not for pulsed laser beams.

Let us consider that wave equation for a case where we only have one transverse dimension. That means, the light is propagating, e.g., in a planar waveguide. Then equation 3.48 simplifies to:

$$2ik_0 \frac{\partial A}{\partial z} + \frac{\partial^2 A}{\partial x^2} = -3\chi^{(3)} \frac{\omega^2}{c^2} |A|^2 A. \quad (3.50)$$

Here, A is a function only of x and z .

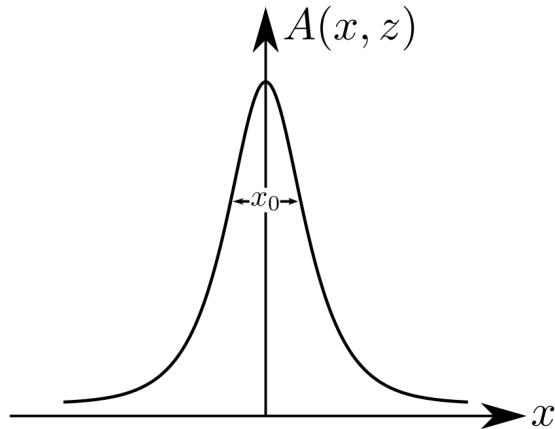


Figure 28: **Spatial soliton.** If there is only one transverse dimension x , then one solution of the paraxial wave equation is a spatial soliton as plotted here for an arbitrary amplitude, and for an arbitrary position along z . The width of the soliton is determined by x_0 .

A specific solution of equation 3.50 is:

$$A(x, z) = A_0 \operatorname{sech}(x/x_0) e^{i\gamma z}, \quad (3.51)$$

where $\operatorname{sech}(x) = (\cosh x)^{-1}$. The width of this peaked function is determined by x_0 , which is given by the relation:

$$x_0 = \frac{1}{k_0} \left(\frac{n_0}{2\bar{n}_2 |A_0|^2} \right)^{1/2}, \quad (3.52)$$

where $\bar{n}_2 = 3\chi^{(3)}/4n_0$ is the same as in equation 3.7.

The solution given in equation 3.51 is sometimes called a “transverse soliton”. Fig. 28 shows the shape of this wavepacket. It can travel long distances without broadening. While the width will stay nearly constant, equation 3.51 shows that the wavepacket will acquire a phase during its propagation through the nonlinear medium. The rate of phase acquisition is given by:

$$\gamma = \frac{k_0 \bar{n}_2 |A_0|^2}{n_0}. \quad (3.53)$$

In higher dimensions, finding such solutions is not straight forward. They are often unstable. For a cylindrically symmetric beam, a solution was found by Chiao et al.[5].

3.3.2 Laser beam breakup

When a beam travels through a medium, impurities in that medium can cause distortions to the beam’s wavefront. If the intensity of the beam is much larger than P_{cr} , then the distorted parts of the wavefront can still have a larger power than P_{cr} and then that part of the beam will self-trap. For that reason, impurities in a nonlinear crystal can lead to the breaking up of a high-power optical beam into clearly separate parts that will then be self-trapped. Effectively, this looks as if the total beam would break apart into separate filaments. We illustrate this phenomenon in Fig. 29.

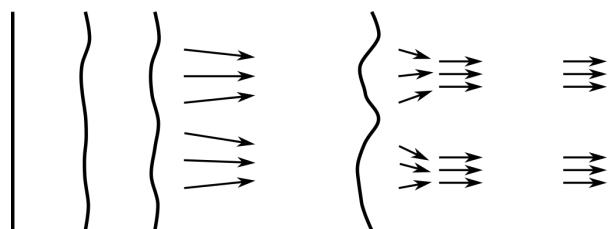


Figure 29: **Beam break-up.** Impurities in the crystal can distort the wavefront. If parts of that distorted wavefront still have sufficient power, these parts can self trap into separate filaments.

3.4 Optical phase conjugation

Aberration is an effect where different parts of an electromagnetic wavefront experience different delays as they pass through a medium or an optical system. The effects of aberration can be overcome by using optical phase conjugation. The principle behind phase conjugation is illustrated in Fig. 30. Part (a) of that figure shows a distorted wavefront reflected by a normal mirror. Parts of the wavefront that are preceding other parts will still remain foremost after the reflection. In this sense, a normal mirror mirrors the wavefront. The goal of a phase-conjugated mirror (PCM) is to reflect the wavefront such that parts that were lagging before the reflection will still be lagging behind the other parts of the wavefront after the reflection. This is illustrated in part (b) of Fig. 30.

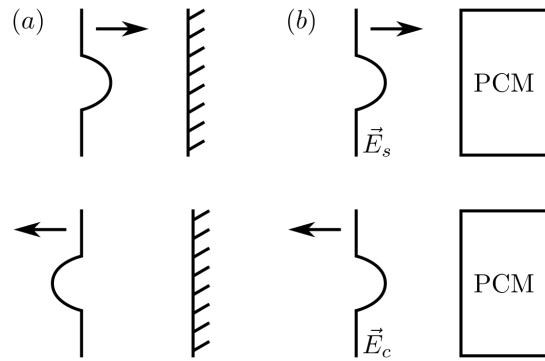


Figure 30: **Phase-conjugated mirrors (PCMs)**. (a) shows the reflection of a distorted wavefront from a normal mirror. (b) shows the reflection of the same wavefront from a PCM. In the latter case, parts of the wavefront that were lagging behind still do so after the reflection from the PCM. In (a), this is not the case.

Let us assume that the incoming field can be described by the vector:

$$\vec{E}_s(\vec{r}, t) = \vec{E}_s(\vec{r})e^{-i\omega t} + \text{c.c.} \quad (3.54)$$

After the reflection from the PCM, the field should then be:

$$\vec{E}_c(\vec{r}, t) = r\vec{E}_s^*(\vec{r})e^{-i\omega t} + \text{c.c.}, \quad (3.55)$$

where r is the reflection coefficient of the PCM.

To better understand the meaning of taking the conjugate of \vec{E}_s in equation 3.55, let us split the field into its amplitude, which is only slowly varying in space, the field's

polarization, and a part that is quickly oscillating in space:

$$\vec{E}_s(\vec{r}, t) = \hat{\epsilon}_s A_s(\vec{r}) e^{i\vec{k}_s \cdot \vec{r}}. \quad (3.56)$$

Here, $\hat{\epsilon}_s$ denotes the polarization unit vector, and \vec{k}_s is the wave vector of the incoming light.

Taking the complex conjugate gives:

$$\vec{E}_s^*(\vec{r}, t) = \hat{\epsilon}_s^* A_s^*(\vec{r}) e^{-i\vec{k}_s \cdot \vec{r}}. \quad (3.57)$$

From that we can conclude that the action of an ideal PCM implies the following:

- the polarization unit vector is replaced by its conjugate. That means, circularly polarized light will remain its chirality upon reflection: right-hand circular polarized light will remain right-hand polarized etc. Upon reflection from a metallic mirror, the chirality is reversed.
- $A_s(\vec{r})$ is replaced by its conjugate. This implies that the wavefront is reversed as illustrated in Fig. 30.
- \vec{k}_s is replaced with $-\vec{k}_s$. That means the wave is reflected back. In terms of ray optics, this means that each incident ray is reflected back exactly the way it came. In other words, the reflected angle is the inverse of the angle of incidence, not the same as would be in a normal mirror.

From equation 3.55, we can conclude that, apart from the reflection coefficient, the action of a PCM is essentially a time reversal because:

$$\vec{E}_c(\vec{r}, t) = \vec{E}_s(\vec{r}, -t). \quad (3.58)$$

This may already make clear, that it is no simple feat to realize a perfect PCM. Typically, real elements that are referred to as PCMs do not fulfill some of the characteristics of a perfect PCM. For example, they might not act on polarization as a perfect PCM would.

3.4.1 Aberration correction using phase conjugation

Assume that we have a plane wave travelling through a possibly inhomogeneous optical medium. This will introduce varying delays to different parts of the wavefront. A simple argument why the reflection at a PCM can correct this aberration is the following: the PCM will time-reverse the incoming wavefront, and then it will travel backwards through the same optical medium, experiencing the same inhomogeneities but in the opposite direction. The wave exiting the medium will then look exactly like the original incoming wave. We illustrate this in Fig. 31.

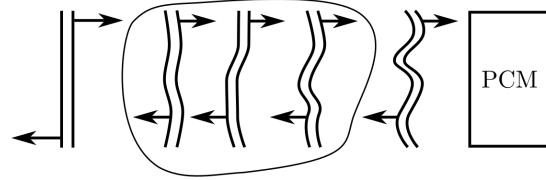


Figure 31: **Aberration correction by a PCM.** An incoming plane wavefront is distorted by its propagation through an inhomogeneous medium. After a PCM reflects the wavefront, it travels backwards through the same medium, resulting in the correction of the aberration it experienced.

In more formal terms, let us assume that the medium the incoming wave passes through has the non-homogeneous refractive index $n(\vec{r}) = \epsilon(\vec{r})^{1/2}$, and that our field satisfies the scalar waveequation:

$$\nabla^2 E - \frac{\epsilon(\vec{r})}{c^2} \frac{\partial^2 E}{\partial t^2} = 0. \quad (3.59)$$

Let us write our incoming field as:

$$E(\vec{r}, t) = A(\vec{r}) e^{i(kz - \omega t)} + \text{c.c..} \quad (3.60)$$

We assume that the field amplitude $A(\vec{r})$ varies slowly in space. If our field is travelling in the z direction, and if the transverse components of the wavevector are sufficiently small, it is convenient to use the paraxial form of the wave equation. To do so, let us will split the Laplacian into its longitudinal and transverse parts and neglect the second derivative in z . Then the wave equation becomes:

$$\nabla_T^2 A + \left[\frac{\omega^2 \epsilon(\vec{r})}{c^2} - k^2 \right] A + 2ik \frac{\partial A}{\partial z} = 0. \quad (3.61)$$

If this equation holds, so must its complex conjugate. That means, we must have:

$$\nabla_T^2 A^* + \left[\frac{\omega^2 \epsilon(\vec{r})}{c^2} - k^2 \right] A^* - 2ik \frac{\partial A^*}{\partial z} = 0, \quad (3.62)$$

where we assumed that $\epsilon(\vec{r})$ is real (no loss).

A solution to this equation is:

$$E_c(\vec{r}, t) = A^*(\vec{r}) e^{i(-kz - \omega t)} + \text{c.c..} \quad (3.63)$$

Given that the amplitude of this field agrees with the incoming wave at any one time,

this boundary condition will ensure that the above wave is the complex conjugate of the incoming wave *everywhere*. The boundary condition could for example be that the two solutions agree on the surface of the PCM.

3.4.2 Phase conjugation by degenerate four-wave mixing

While it is difficult to realize a perfect phase-conjugated mirror (PCM), people found ways to realize devices that implement some of its features if not all already quite some time ago. An example for that is phase conjugation via degenerate four-wave mixing. We will first describe the rough idea of how the process works before looking at it in a bit more detail. Fig. 32 illustrates the basic layout of realizing a PCM using degenerate four-wave mixing (DFWM) in a $\chi^{(3)}$ nonlinear medium. This approach has already been suggested by Hellwarth et al and by Yariv et al in 1977[6, 7]. Bloom and Bjorklund performed an experimental demonstration in the same year[8].

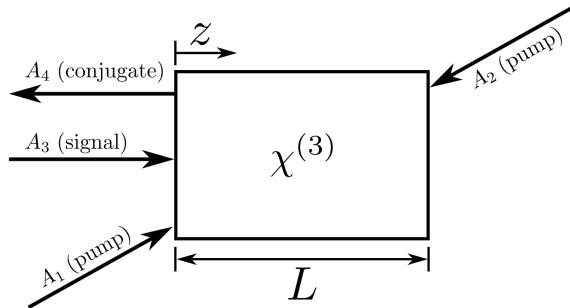


Figure 32: **Realizing a PCM via four-wave mixing.** A crystal is pumped by two fields $A_{1,2}$, which propagate at exactly opposite directions. The input field (signal) is the field A_3 . If the geometry is chosen properly, the output field A_4 will be the conjugate of the input field. The figure is based on Ref. [1].

We will write the four fields as we are used to:

$$E_i(\vec{r}, t) = E_i(\vec{r})e^{-i\omega t} + \text{c.c.} = A_i(\vec{r})e^{i(\vec{k}_i \cdot \vec{r} - \omega t)} + \text{c.c.}, \quad (3.64)$$

where $i = 1, 2, 3, 4$, and the A_i are varying only slowly. We assume scalar quantities because the geometry will be fixed for the purpose of DFWM. We will focus on the following term in the nonlinear polarization:

$$P^{\text{NL}} = 6\epsilon_0\chi^{(3)}E_1E_2E_3^* = 6\epsilon_0\chi^{(3)}A_1A_2A_3^*e^{i(\vec{k}_1 + \vec{k}_2 - \vec{k}_3) \cdot \vec{r}}. \quad (3.65)$$

This already assumes that the four frequencies match. such that $\omega \equiv \omega_i$, and that we can write $\chi^{(3)} = \chi^{(3)}(\omega = \omega + \omega - \omega)$. Then the time-dependent part in the exponent will vanish.

If E_1 and E_2 are counter-propagating, we will have:

$$\vec{k}_1 + \vec{k}_2 = 0, \quad (3.66)$$

and then the term above becomes:

$$P^{\text{NL}} = 6\epsilon_0\chi^{(3)}A_1A_2A_3^*e^{-i\vec{k}_3\cdot\vec{r}}. \quad (3.67)$$

This shows that the nonlinear polarization contains a term that can act as a phase-matched source for the output field E_4 with a wave vector that is the conjugate of the wavevector of the input field:

$$\vec{k}_4 = -\vec{k}_3. \quad (3.68)$$

The amplitude of the output field will be proportional to $A_1A_2A_3^*$. If A_1 and A_2 are constant in space (plane waves), or if $A_1 = A_2^*$, then the product A_1A_2 will be proportional to the real quantity $|A_1|^2$.

Another way of looking at this process is that the interference of A_3 with one of the pump fields (e.g. A_1) will create an interference pattern inside the nonlinear crystal. This then will create a 3D grating of varying refractive index inside the material. The second pump field will then diffract from that grating to create the output wave A_4 .

Let us now look at the process in more detail. The total electric field will be the sum of the four fields E_i , and this field will generate the third-order nonlinear polarization:

$$P = 3\epsilon_0\chi^{(3)}E^2E^*. \quad (3.69)$$

This will have 4^3 different terms, but the ones that are most interesting to us are the ones with the spatial dependence

$$e^{i\vec{k}_i\cdot\vec{r}} \quad (3.70)$$

for $i = 1, 2, 3, 4$ because they act as phase-matched source terms for the output field. One example of the corresponding polarization amplitudes for these phase-matched contributions is [1]:

$$P_4 = 3\epsilon_0\chi^{(3)}[E_4^2E_4^* + 2E_4E_1E_1^* + 2E_4E_2E_2^* + 2E_4E_3E_3^* + 2E_1E_2E_3^*]. \quad (3.71)$$

To simplify our treatment, we will assume that the pump fields $E_{1,2}$ are much stronger than the signal and output fields $E_{3,4}$. Using this assumption, we can neglect all terms in equation 3.71 that depend on more than one of these weak fields, and we get:

$$P_4 = 3\epsilon_0\chi^{(3)}[2E_4E_1E_1^* + 2E_4E_2E_2^* + 2E_1E_2E_3^*].$$

Because the corresponding equations for all four fields are much simpler, we can first concentrate on how the propagation through the medium affects $E_{1,2}$. Once we have solutions for that, we can use them to analyze the evolution of the weaker fields $E_{3,4}$.

Each of the fields has to fulfill the wave equation:

$$\nabla^2 E_i - \frac{\epsilon}{c^2} \frac{\partial^2 E_i}{\partial t^2} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 P_i}{\partial t^2}. \quad (3.72)$$

If we introduce a new spatial coordinate z' along the direction of \vec{k}_1 , and if we assume that the pump fields $E_{1,2}$ are plane waves, the amplitude A_1 must fulfill the equation:

$$\begin{aligned} & \left[\left(-k_1^2 + 2ik_1 \frac{d}{dz'} + \frac{\epsilon\omega^2}{c^2} \right) A_1 \right] e^{i(k_1 z' - \omega t)} \\ &= -\frac{\omega^2}{c^2} 3\chi^{(3)} [|A_1|^2 + 2|A_2|^2] A_1 e^{i(k_1 z' - \omega t)} \end{aligned} \quad (3.73)$$

Doing the equivalent thing for A_2 , one arrives at two first-order ordinary differential equations for A_1 and A_2 . The solutions are:

$$A_1(z') = A_1(0)e^{i\kappa_1 z'}, \quad A_2(z') = A_2(0)e^{-i\kappa_2 z'}. \quad (3.74)$$

This should not be terribly surprising since we assumed that the two pump fields are plane waves. The quantities $\kappa_{1,2}$ are coupling constants. If the two pump fields have equal intensities, we will have $\kappa \equiv \kappa_1 = \kappa_2$, and the product of the two pump amplitudes will then also be constant:

$$A_1(z')A_2(z') = A_1(0)A_2(0). \quad (3.75)$$

If we enter our solutions for $A_{1,2}$ into our wave equation 3.72, we arrive at the following coupled differential equations for $A_{3,4}$:

$$\begin{aligned} \frac{dA_3}{dz} &= i\kappa_3 A_3 + i\kappa A_4^*, \\ \frac{dA_4}{dz} &= -i\kappa_3 A_4 - i\kappa A_3^*, \end{aligned} \quad (3.76)$$

where we introduced the coupling coefficients:

$$\begin{aligned} \kappa_3 &= \frac{3\omega}{nc} \chi^{(3)} (|A_1|^2 + |A_2|^2), \\ \kappa &= \frac{3\omega}{nc} \chi^{(3)} A_1 A_2. \end{aligned} \quad (3.77)$$

We can simplify these equations by introducing new functions $A'_{3,4}$ defined via the rela-

tions:

$$\begin{aligned} A_3 &= A'_3 e^{i\kappa_3 z}, \\ A_4 &= A'_4 e^{-i\kappa_3 z}. \end{aligned} \quad (3.78)$$

Then equations 3.76 become:

$$\begin{aligned} \frac{dA'_3}{dz} &= i\kappa A'^*_4, \\ \frac{dA'_4}{dz} &= -i\kappa A'^*_3. \end{aligned} \quad (3.79)$$

From the last equation, we can see why DFWM leads to phase conjugation. The field amplitude A'_4 is only driven by the input field amplitude.

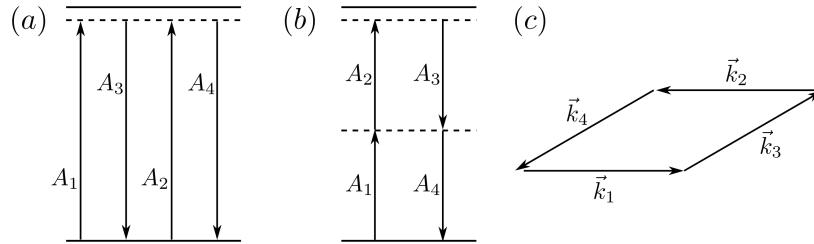


Figure 33: **Level diagrams and phase matching for DFWM.** (a) and (b) are the level diagrams for two possible interactions leading to the DFWM process described in the text. In (a) $A_{1,2}$ pump from the ground state close to a one-photon transition, and A_3 and A_4 correspond to the one-photon decay from that near resonance back to the ground state. (b) describes the case where we are near resonance with a two-photon transition. Two photons from A_1 and A_2 pump close to that transition, and then we have a two-photon decay generating the fields A_3 and A_4 . (c) shows the phase matching of the four wave vectors.

One can combine the two equations 3.79 to find solutions for the field amplitudes $A'_{3,4}$. If we assume that we know the values of $A'^*_3(0)$ and $A'_4(L)$, then the solutions will be:

$$\begin{aligned} A'^*_3(z) &= -\frac{i|\kappa|}{\kappa} \frac{\sin|\kappa|z}{\cos|\kappa|L} A'_4(L) + \frac{\cos[|\kappa|(z-L)]}{\cos|\kappa|L} A'^*_3(0) \\ A'_4(z) &= \frac{\cos|\kappa|z}{\cos|\kappa|L} A'_4(L) - \frac{i\kappa}{|\kappa|} \frac{\sin[|\kappa|(z-L)]}{\cos|\kappa|L} A'^*_3(0) \end{aligned} \quad (3.80)$$

If we want to achieve optical phase conjugation, we usually will not insert a field A_4 at the end face of the crystal. That means, we have $A'(L) = 0$, and what we are mostly

interested in are the values of the fields as they leave the crystal:

$$\begin{aligned} A_3'^*(L) &= \frac{A_3'^*(0)}{\cos|\kappa|L}, \\ A_4'(0) &= \frac{i\kappa}{|\kappa|} (\tan|\kappa|L) A_3'^*(0). \end{aligned} \quad (3.81)$$

That means the output field A_4' will be the conjugate of the input field A_3' . Another interesting thing to note is that the intensities of the output fields can take on any value from zero to infinity. This is due to the fact that we are continuously pumping the crystal with the fields $E_{1,2}$. Of course, our results here are only valid as long as our initial assumption holds that the fields $E_{3,4}$ are much weaker than the fields $E_{1,2}$.

3.4.3 Phase conjugation - the conclusion

In the preceding sub-section we showed that it is possible to realize a PCM fulfilling the condition $\vec{k} \rightarrow -\vec{k}$ and that the output wave will have a conjugated wavefront. So far, we have not discussed the matter of polarization. There have been theoretical works as well as experiments showing that one can realize a PCM that also fulfills the condition of compensating any non-uniform changes to the polarization of the incoming wavefront. To achieve that, one performs four-wave mixing in an atomic system utilizing two-photon resonances. An important prerequisite in this context is that the two counter-propagating pump fields $E_{1,2}$ have to be circularly polarized and counter-rotating[1].

4 Nonlinear effects in optical fibers

So far, we have studied nonlinear optical effects for light passing through bulk media. If light is travelling along a waveguide, all its power can pass through a very small cross section. Consider, for example, a case where we have 1 W of light in an optical fiber with a fiber core diameter of 10 μm . In this case, the intensity of the light will be 1 MW/cm². In itself, this is nothing special. We could just take a beam and focus it tightly to a waist of that size. The important thing is that this power will travel along the waveguide, and the field will maintain that high intensity all the way. This makes optical fibers and, more in general, optical waveguides an ideal playground for nonlinear optics.

4.1 A figure of merit for nonlinear effects in fibers

We can make this comparison between nonlinear effects in bulk media and in fibers more precise and derive a figure of merit for comparing these two cases. In both cases, we see that the nonlinear effects will become stronger if we have higher intensity I and if we have a longer interaction length L . We can therefore say that we want to maximize the product $I \cdot L$ in order to maximize nonlinear effects.

In the case of bulk optics, simply increasing the length of the nonlinear medium is not sufficient because diffraction will lead to the beam waist to change along the direction of beam propagation (z axis). In particular, the waist $w(z)$ of a Gaussian beam behaves as follows:

$$w(z) = w_0 \sqrt{1 + \left(\frac{\lambda z}{\pi n w_0^2}\right)^2}. \quad (4.1)$$

Then the optical intensity at position z will be:

$$I_B(z) = \frac{P}{\pi w(z)^2}, \quad (4.2)$$

where B denotes the case of bulk optics. The case for fibers, we will denote with an F . Then we can define the product $[I \cdot L]$ for our figure of merit by integrating over z :

$$[I \cdot L]_B = \int_{-\infty}^{\infty} dz I_B(z) = \pi \frac{n P}{\lambda}. \quad (4.3)$$

From this, we can see that the only way to increase the magnitude of our nonlinear

interaction will be to increase the power. Of course, this is not taking into account the possibility of optimizing the material parameters. But these we can also optimize for waveguides.

Now let us compare that with the case where we have our light in an optical fiber. In that case, the waist w_0 will stay constant over the interaction length, but we should also take into account the attenuation along the fiber. The latter is determined by the attenuation coefficient α . The intensity of light inside the fiber will be:

$$I_F(t) = \frac{P \exp(-\alpha z)}{\pi w_0^2}, \quad (4.4)$$

and our product of intensity and length becomes:

$$[I \cdot L]_F = \int_{-\infty}^{\infty} dz I_F(z) = \frac{P}{\pi w_0^2} L_{\text{eff}}. \quad (4.5)$$

Here, we defined the effective interaction length

$$L_{\text{eff}} = \frac{[1 - \exp(-\alpha L)]}{\alpha}. \quad (4.6)$$

If the interaction length L becomes very large, we can write $L_{\text{eff}} \approx 1/\alpha$, and the ratio between those two $[I \cdot L]$ products will become:

$$\frac{[I \cdot L]_F}{[I \cdot L]_B} = \frac{\lambda}{(\pi w_0)^2 n \alpha}. \quad (4.7)$$

As an example, consider the case where we have the following experimental parameters: $P = 1 \text{ W}$, $w_0 = 10 \mu\text{m}$, $n = 1.5$, $\lambda = 1 \mu\text{m}$, and $\alpha = 2.3 \times 10^{-4} \text{ m}^{-1} = 1 \text{ dB/km}$. For these parameters, the ratio in equation 4.7 becomes approximately 5×10^7 . That means, we could expect the same nonlinear effects in a fiber with 1 W of power as in a bulk crystal with 50 MW. This is assuming that we could even have a nonlinear bulk crystal as long as a fiber, and we also assumed that there is no loss in the bulk crystal.

4.2 The intensity-dependent refractive index in an optical fiber

Given the figure of merit we just discussed, some nonlinear optical effects we encountered earlier but dismissed as small effects may become significantly more relevant. For that reason, let us consider how the optical Kerr effect can influence our optical signals if we send them through a fiber. Often, the optical fibers we will be interested in consist of silica glass (SiO_2). The nonlinearity for that material is typically around $\bar{n}_2 = 3.18 \times$

$10^{-20} \text{ m}^2/\text{W}$. There are several interesting nonlinear effects we can discuss in the context of silica fibers: (a) optical solitons, (b) optical pulse compression, and (c) modulational instabilities. We will focus on the first two.

4.2.1 Self-phase modulation

Consider a short optical pulse we send through a fiber. The instantaneous phase of the pulse will change rapidly as it travels along the fiber. Due to the optical Kerr effect, the refractive index of the medium will change according to:

$$n(\omega_0, |E|^2) = \frac{c\beta}{\omega_0} = n_0(\omega_0) + \bar{n}_2|E|^2., \quad (4.8)$$

where we used β to denote the (instantaneous) wavenumber of the light inside the fiber. The phase of our optical pulse then is defined as $\Phi = \omega_0 t - \beta z$, and the instantaneous frequency of our pulse will be:

$$\omega(t) = \frac{\partial \Phi}{\partial t} = \omega - \frac{\omega_0 \bar{n}_2}{c} z \frac{\partial |E|^2}{\partial t}. \quad (4.9)$$

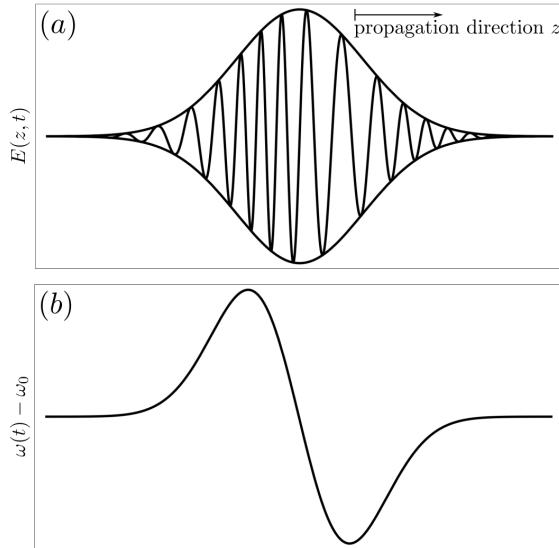


Figure 34: **Self-phase modulation of an optical pulse.** (a) shows the oscillation of the field amplitude and the slowly varying envelop. (b) shows the change in the instantaneous frequency.

Equation 4.9 shows that, and a non-zero nonlinear coefficient \bar{n}_2 will lead to a variation of the frequency as the pulse travels along. This is called self-phase modulation (SPM),

and we illustrate the effect in Fig. 34. Part (b) of that figure shows that the instantaneous frequency of the light at first decreases for the leading part of the pulse, and then the frequency increases for the trailing part of the pulse. In telecommunication one often uses zero-dispersion fibers. Such fibers are denoted as dispersion-shifted fibers (DSFs). These will be designed such that there is zero dispersion around a given design wavelength λ_0 . Typical design wavelengths are 1300 nm and 1550 nm. The wavelength range with $\lambda < \lambda_0$ is called the *normal dispersion region*. In that regime, the dispersion in the fiber will be normal (group velocity increases with wavelength). The region with $\lambda > \lambda_0$ is called the *anomalous dispersion region*, where the group velocity decreases with increasing wavelength.

Because the frequency decreases for the leading edge of the pulse, the group velocity will also decrease. At the same time, for the trailing edge of the pulse the frequency will be larger than ω_0 , and that corresponds to a higher group velocity. That means the leading edge will be delayed while the trailing edge will catch up. Effectively, this means that the pulse will be compressed due to this SPM effect. If this compression of the pulse exactly cancels the pulse broadening due to dispersion, the pulse width will remain constant as the pulse travels along the fiber. That means, we will have an *optical soliton*. To be more precise, the envelope of the pulse will keep its shape. Such solitons are therefore sometimes called *envelope solitons*.

In the normal dispersion region, there is also another solution to the wave equation: a *dark soliton*. This describes a solitary wave that results from cutting a portion from the continuous wave, which we will discuss in subsection 4.3.1.

4.2.2 Temporal pulse spreading and the nonlinear Schrödinger equation

Now let us describe the electromagnetic field in a nonlinear optical fiber a bit more precisely. We will see, that the equation governing that propagation can be reduced to an equation often referred to as the nonlinear Schrödinger equation. For simplicity, let us only take into account the propagation direction z and time t for now and not the transverse dimensions x, y . We want to use this equation to describe the propagation of an optical pulse through a nonlinear medium. Let us write the electric field as:

$$E(z, t) = A(z, t)e^{i(k_0 z - \omega_0 t)} + \text{c.c.}, \quad (4.10)$$

where $A(z, t)$ here describes the envelope of the pulse. The wave equation describing the propagation of the pulse through our nonlinear medium is:

$$\frac{\partial^2 E}{\partial z^2} - \frac{1}{\epsilon_0 c^2} \frac{\partial^2 D}{\partial t^2} = 0, \quad (4.11)$$

where D is the total displacement including linear and nonlinear contributions. Let us express $E(z, t)$ and $D(z, t)$ in terms of their Fourier transforms:

$$E(z, t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} E(z, \omega), \quad D(z, t) = \int_{-\infty}^{\infty} \frac{d\omega}{2\pi} e^{-i\omega t} D(z, \omega).$$

Then the Fourier components of E and D fulfill the relation:

$$D(z, \omega) = \epsilon_0 \epsilon(\omega) E(z, \omega)$$

with $\epsilon(\omega)$ being the effective dielectric constant.

Using these expressions, we can write the wave equation as:

$$\frac{\partial^2 E(z, \omega)}{\partial z^2} + \epsilon(\omega) \frac{\omega^2}{c^2} E(z, \omega) = 0. \quad (4.12)$$

If we express the Fourier transform of the field in terms of the Fourier transform of the slowly varying field amplitude $A(z, t)$, we get:

$$\begin{aligned} E(z, \omega) &= A(z, \omega - \omega_0) e^{ik_0 z} + A(z, \omega + \omega_0) e^{-ik_0 z} \\ &\approx A(z, \omega - \omega_0) e^{ik_0 z}. \end{aligned} \quad (4.13)$$

The approximation holds true if the amplitude $A(z, t)$ is slowly varying, because in that case there cannot be high-frequency components at $\omega + \omega_0$. If we insert that into the wave equation 4.12, and if we drop the second derivatives in z , we get:

$$2ik_0 \frac{\partial A}{\partial z} + (k^2 - k_0^2) A = 0, \quad (4.14)$$

where the frequency dependent wavenumber fulfills:

$$k(\omega) = \sqrt{\epsilon(\omega)} \omega / c. \quad (4.15)$$

If k is close to k_0 , we can write $k^2 - k_0^2 \approx 2k_0(k - k_0)$, which will reduce 4.14 to the *reduced wave equation*:

$$\frac{\partial A(z, \omega - \omega_0)}{\partial z} - i(k - k_0) A(z, \omega - \omega_0) = 0. \quad (4.16)$$

It is important to keep in mind that k is not simply a variable in the equation above. The wavenumber will depend on the refractive index and therefore on the intensity of the light field. We can take this into account by expanding k in a power series:

$$k = k_0 + \Delta k_{\text{NL}} + k_1(\omega - \omega_0) + \frac{1}{2} k_2 (\omega - \omega_0)^2, \quad (4.17)$$

where we introduced the following quantitites:

$$\begin{aligned}\Delta k_{\text{NL}} &= \Delta n_{\text{NL}} \omega_0 / c = n_2 I \omega_0 / c, \\ k_1 &= \left. \frac{dk}{d\omega} \right|_{\omega=\omega_0} = \frac{1}{c} \left[n_{\text{lin}} + \omega \frac{dn_{\text{lin}}}{d\omega} \right]_{\omega=\omega_0} \equiv \frac{1}{v_g}(\omega_0), \\ k_2 &= \left. \frac{d^2k}{d\omega^2} \right|_{\omega=\omega_0} = \left. \frac{d}{d\omega} \left[\frac{1}{v_g(\omega)} \right] \right|_{\omega=\omega_0} = \left[-\frac{1}{v_g^2} \frac{dv_g}{d\omega} \right]_{\omega=\omega_0}.\end{aligned}\quad (4.18)$$

k_1 is the inverse of the *group velocity*, and k_2 is the *group velocity dispersion*.

We then insert that into the reduced wave equation, and then we transform the equation to the time domain by multiplying both sides with $\exp[-i(\omega - \omega_0)t]$ and then integrating everything over all values of $\omega - \omega_0$. This results in the following equation:

$$\frac{\partial A}{\partial z} + k_1 \frac{\partial A}{\partial t} + \frac{1}{2} i k_2 \frac{\partial^2 A}{\partial t^2} - i \Delta k_{\text{NL}} A = 0. \quad (4.19)$$

The third term on the left-hand side describes the spreading of the pulse due to group-velocity dispersion, while the last term describes the spreading of the pulse due to self-phase modulation. As we stated in the beginning of this section, this equation is sometimes referred to as the nonlinear Schrödinger equation.

As a next step, let us do a coordinate transformation introducing the “retarded time” τ :

$$\tau = t - \frac{z}{v_g} = t - k_1 z. \quad (4.20)$$

If we introduce a new function $A_s(z, \tau) = A(z, t)$ to describe our pulse, then equation 4.19 becomes:

$$\frac{\partial A_s}{\partial z} + \frac{1}{2} i k_2 \frac{\partial^2 A_s}{\partial \tau^2} - i \Delta k_{\text{NL}} A_s = 0. \quad (4.21)$$

By rewriting the nonlinear contribution to the propagation constant as:

$$\Delta k_{\text{NL}} = n_2 \frac{\omega_0}{c} I = 2n_0 \epsilon_0 n_2 \omega_0 |A_s|^2 \equiv \gamma |A_s|^2, \quad (4.22)$$

we can rewrite equation 4.21 as:

$$\frac{\partial A_s}{\partial z} + \frac{1}{2} i k_2 \frac{\partial^2 A_s}{\partial \tau^2} = i \gamma |A_s|^2 A_s. \quad (4.23)$$

4.3 Temporal solitons in a fiber

So far, we have treated this as a 1-dimensional problem. Even in a fiber, however, the field will have radial components of course. Let us quickly investigate how this affects our analysis. If we are considering a *single-mode fiber* (SMF), the fiber will only support

HE_{11}^x modes. This is simply a Gaussian mode. In this case, the H and E fields satisfy:

$$\begin{aligned} H_y &= n_c \epsilon_0 c E_x, \\ E_y &= H_x = 0, \end{aligned} \quad (4.24)$$

where we assumed $k = k_0 n_c$ (light in the core of the fiber, and we use n_{cl} for the fiber cladding).

The (linear) propagation of light in step-index fibers was covered in the Photonics I course, but to be consistent with the notation here, we need to define a few quantities. In particular, the relative refractive index difference between the refractive index of the core (n_c) and the cladding (n_{cl}) is given by:

$$\Delta = \frac{n_c^2 - n_{cl}^2}{2n_c^2} \approx \frac{n_c - n_{cl}}{n_c}, \quad (4.25)$$

which is often expressed in percent. The following relation holds between the numerical aperture NA of the fiber and the maximum angle θ_{\max} a ray can have with the propagation direction and still be guided along the fiber:

$$\text{NA} = \theta_{\max} \approx n_c \sqrt{2\Delta}. \quad (4.26)$$

Within the fiber, the maximum angle for a ray is $\phi_{\max} \approx \theta_{\max}/n_c \approx \sqrt{2\Delta}$.

For a ray that has an angle ϕ with the z axis inside the fiber, the *propagation constants* along the propagation direction (z) and perpendicular to that (x) are defined as:

$$\begin{aligned} \beta &= k_0 n_c \cos \phi \\ \kappa &= k_0 n_c \sin \phi \end{aligned} \quad (4.27)$$

If a mode of order m is to be guided along the fiber, it has to fulfill the following phase-matching condition:

$$v \equiv k_0 n_c a \sqrt{2\Delta} = \frac{\cos^{-1} \xi + m\pi/2}{\xi}, \quad (4.28)$$

where v is called the normalized frequency, and we use the definition:

$$\xi = \frac{\sin \phi}{\sqrt{2\Delta}}. \quad (4.29)$$

Then the instantaneous peak power of our mode is:

$$P = \int_0^{2\pi} \int_0^\infty dr d\theta r (E_x H_y^* - E_y H_x^*) = n_c \epsilon_0 c \int_0^{2\pi} \int_0^\infty dr d\theta r |E_x|^2. \quad (4.30)$$

If we assume that the peak power is achieved for $\tau = 0$, then we can replace E_x with $R(r, \theta)A_s(0, 0)$ in the above equation, and we get:

$$P = n_c \epsilon_0 c |A_s^0|^2 \int_0^{2\pi} \int_0^\infty dr d\theta r |R|^2.$$

This is related to the cross-sectional area occupied by the light field in the fiber. In particular, this *effective area* is defined as[9]:

$$A_{\text{eff}} = \frac{\left[\int_0^{2\pi} \int_0^\infty dr d\theta r |R|^2 \right]^2}{\int_0^{2\pi} \int_0^\infty dr d\theta r |R|^4}. \quad (4.31)$$

If we use this definition, and if we define:

$$\eta = \frac{\int_0^{2\pi} \int_0^\infty dr d\theta r |R|^4}{\int_0^{2\pi} \int_0^\infty dr d\theta r |R|^2}, \quad (4.32)$$

then we can write P as:

$$P = \eta n_c \epsilon_0 c |A_p|^2 A_{\text{eff}} \approx \frac{1}{2} n_c \epsilon_0 c |A_s^0|^2 A_{\text{eff}}, \quad (4.33)$$

where we introduced the peak amplitude $A_s^0 = A_s(0, 0)$. The approximation on the right-hand side assumes that $\eta \approx 0.5$. This is the case for the “normal” operation of optical fibers . In that case, the normalized frequency v is between 1.5 and 2.5. That is the range of single-mode operation (or at least not very multi-mode).

Let us now reconsider the nonlinear Schrödinger equation 4.23 we derived. In the context of self trapping, we derived equation 3.50, which has similar features to the nonlinear Schrödinger equation and had spatial solitons as a solution. It should therefore not surprise us that equation 4.23 has a similar solution of the form:

$$A_s(z, \tau) = A_s^0 \operatorname{sech}(\tau/\tau_0) e^{i\zeta z}, \quad (4.34)$$

where the pulse amplitude A_s^0 and the pulse width τ_0 must fulfill:

$$|A_s^0|^2 = \frac{-k_2}{\gamma \tau_0^2} = \frac{-k_2}{2n_0 \epsilon_0 n_2 \omega_0 \tau_0^2}, \quad (4.35)$$

$$\zeta = -k_2/2\tau_0^2 = \frac{1}{2} \gamma |A_s^0|^2. \quad (4.36)$$

ζ is the phase shift experienced by the soliton pulse during propagation - similar to the phase shift we had in the case of the spatial soliton. Equation 4.35 is sometimes

called the soliton condition[9] because it defines how high the amplitude of our pulse has to be in order for a soliton to form. This is similar to the case of self trapping, where an optical beam has to have the critical power such that diffractive broadening is compensated by self-focusing. Similarly, here the dispersive broadening of the pulse has to be compensated by the self-phase modulation.

The solution discussed above is known as the *fundamental soliton*, which can form if the soliton condition is fulfilled. If we send more power into the fiber, higher-order solitons can form. In general, one uses the parameter[9]:

$$\nu = |A_s^0| \tau_0 \sqrt{-\frac{k\bar{n}_2}{k_2}}, \quad (4.37)$$

to analyze the generation of higher-order solitons. In particular, $\nu = 1$ is equivalent to the (fundamental) soliton condition. A fundamental soliton can form as long as $0.5 \leq \nu \leq 1.5$ [9]. For $\nu \neq 1$, the efficiency for exciting the soliton becomes low, however, and a significant amount of the input light will pump non-soliton modes.

If the pump power is higher ($\nu \geq 1.5$), higher-order solitons can be excited. Such higher-order solitons periodically change their pulse shape with a period [9]:

$$z_0 = \frac{\pi \tau_0^2}{2|k_2|}. \quad (4.38)$$

After that period, the original pulse shape is restored with the initial width. For $z < z_0$, the pulse will be compressed because, in this case, the self-phase modulation will dominate over the dispersion. This “soliton compression” can be used for optical pulse compression[9].

4.3.1 Dark solitons

In several equations in the previous subsection showed square roots, with arguments that had a minus sign before the dispersion k_2 in the fiber. This shows that the soliton solutions we discussed there can only occur in media where $k_2 < 0$. That means, such *bright solitons* cannot occur in media with normal dispersion ($k_2 > 0$). *Dark solitons*, on the other hand, can also occur in normal dispersion regions. They do not represent a travelling pulse of light but rather a *dip* in the amplitude of a continuous wave.

For $k_2 > 0$, the solution of the nonlinear Schrödinger equation 4.23 is of the form:

$$A(z, \tau) = A_s^0 \tanh \left(\frac{\tau}{\tau_0} \right) e^{i\zeta z}, \quad (4.39)$$

where the definition of ζ is the same as in equations 4.36. The equation relating the

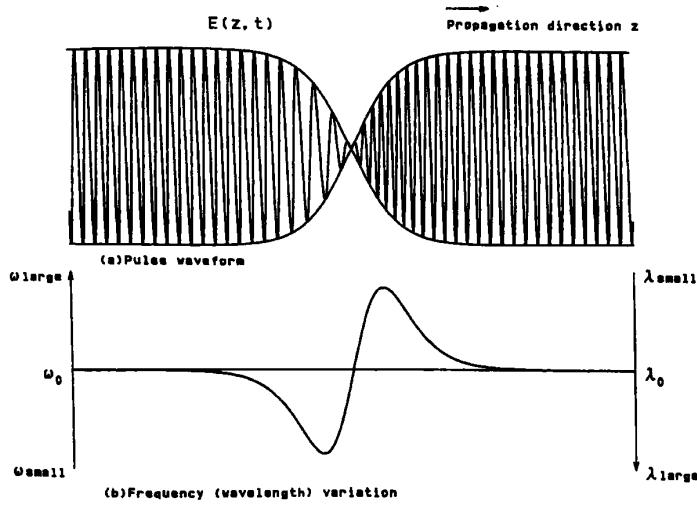


Figure 35: **Dark-soliton pulse waveform.** (top) the oscillation of the field amplitude and the envelop clearly show the shape of the dark soliton as well as the change in the oscillation period close to the amplitude dip. At the minimum of the dip, the phase flips. (b) shows the change in the instantaneous frequency on the same scale along the propagation axis z . The figure is taken from Ref. [9].

pulse width and the pulse amplitude is the same as in 4.35 apart from a minus sign:

$$|A_s^0|^2 = \frac{k_2}{\gamma \tau_0^2} = \frac{k_2}{2n_0 \epsilon_0 n_2 \omega_0 \tau_0^2}. \quad (4.40)$$

Fig. 35 is a copy of Fig. 5.12 in Ref. [9] that illustrates a dark soliton propagating along the z direction as a dip in the electric field amplitude of a plane wave. I should note that I could not successfully reproduce this figure, and so I copied it from Ref. [9].

4.4 Optical pulse compression

Pulsed lasers have become a common and very useful tool in optical physics. For example, short laser pulses are needed for fast optical communication, ultrafast optical imaging, and in nonlinear optics, where the high peak powers are beneficial to harness nonlinear optical effects. Due to the uncertainty principle, the minimum duration Δt of a pulse is related to the bandwidth of the pulse $\Delta\nu$ via the relation $\Delta\nu\Delta t \geq 1$. If this product fulfills $\Delta\nu\Delta t = 1$, one calls the pulse *Fourier limited*. When short pulses travel through dispersive media, the dispersion will typically increase the pulse width such that $\Delta t > 1/\Delta\nu$. In this section, we will first discuss a way to compress such pulses again in order to compensate the effects of dispersion by using a pair of gratings. After

that, we will discuss how one can use nonlinear optical effects in a fiber in combination with two such gratings in order to actually compress the pulse width below the original pulse width.

4.4.1 Pulse compression using gratings

Consider a collimated beam diffracted by two identical diffraction gratings A and B that are oriented parallel to each other as illustrated in Fig. 36. If a ray of light that contains multiple wavelengths passes through this setup, the different wavelengths will exit along parallel beams after the second grating. Because the paths of the different wavelengths through this setup are different, the various wavelengths will experience a delay relative to each other. In particular, if the center frequency is ω_0 , light with a frequency ω will experience a relative time delay τ given by:

$$\tau = \tau_G - \frac{\omega - \omega_0}{\mu} + O(\omega - \omega_0)^2. \quad (4.41)$$

We defined τ_G and μ by the relations[9]:

$$\tau_G = \left(\frac{b}{c} \right) (1 + \cos\theta), \quad (4.42)$$

$$\mu = \frac{\omega^3 d^2 \left\{ 1 - [(2\pi c/\omega d) - \sin\gamma]^2 \right\}}{4\pi^2 cb}. \quad (4.43)$$

τ_G is the time light of frequency ω_0 takes through the grating setup, c is the speed of light, b is the distance between the gratings, and d is the grating pitch (the period). γ is the angle of incidence of the light on grating A .

For the first order of diffraction, the angle of incidence and the diffraction angle are related via the equation:

$$\sin(\gamma - \theta) = \frac{\lambda}{d} - \sin\gamma. \quad (4.44)$$

Equation 4.43 can be derived by considering the propagation of rays through the grating setup. This equation shows that the propagation through the setup will lead to a delay of lower frequencies, while higher frequencies experience less delay than light at the center frequency ω_0 . This is equivalent to a passage through a material of anomalous dispersion. That means, while the passage through a material with normal dispersion leads to the broadening of short pulses, the passage through a two-grating setup can actually compress pulses. Pulses where the higher-frequency parts are delayed are called “chirped”. The opposite case, which is realized by anomalous dispersion, is sometimes denoted as “anti-chirp”.

As we noted earlier, if a pulse travels through a material of normal dispersion $k_2 > 0$, the pulse will be broadened, such that it will no longer be Fourier limited (if it was so before). If one knows the amount of dispersion the pulse experiences, one can compress

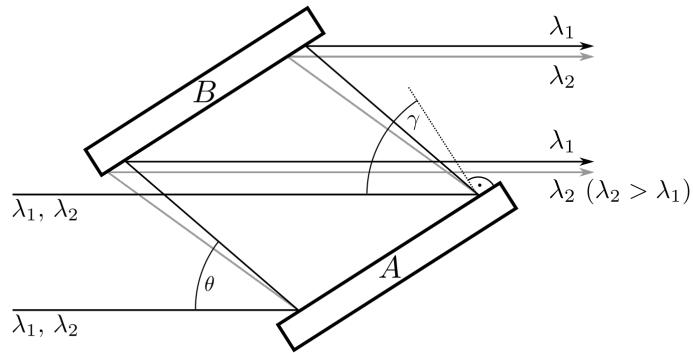


Figure 36: **Grating pair for pulse compression.** This figure based on Ref. [9] illustrates two rays of light diffracted first by grating A , then by grating B . The incident light is assumed to carry two wavelengths $\lambda_{1,2}$ with $\lambda_2 > \lambda_1$. The angle of incidence is γ , the angle of diffraction for λ is θ .

the pulse again afterwards, such that it becomes Fourier limited again. Alternatively, one can also “pre-compress” the pulse by sending it through a two-grating setup. That means, the pulse will be anti-chirped before entering the material, and the dispersion in the material will counter-act the anti-chirp.

If the bandwidth of the pulse is high, the higher-order terms in equation 4.41 may become non-negligible. One can compensate the higher-order terms by inserting a pair of prisms between the two gratings.

4.4.2 Pulse compression harnessing nonlinear optics

Now, assume that we combine the pulse compression technique using two gratings with the (nonlinear) self-phase modulation (SPM) in an optical fiber. Let us begin with a Fourier-limited pulse with bandwidth $\Delta\omega$ and pulse width Δt . If we send such a pulse through a fiber, SPM can lead to a broadening of the bandwidth to $\Delta\omega' > \Delta\omega$, and it may also lead to a broadening of the pulse width. However, because the bandwidth of the pulse is now larger, we can then use a pair of gratings to compress the pulse width ideally down to:

$$\Delta t' = \frac{1}{\Delta\omega'} < \frac{1}{\Delta\omega} = \Delta t. \quad (4.45)$$

5 Stimulated Raman and Brillouin scattering

In this chapter, we will discuss stimulated as opposed to spontaneous scattering processes. In spontaneous scattering processes, the properties of the material system scattering the electromagnetic field are unmodified. Spontaneous scattering processes are:

- *Raman scattering*: scattering from vibrational modes in the medium (optical phonons).
- *Brioullin scattering*: scattering of light from sound waves - that means, from propagating pressure variations in the optical medium (acoustic phonons).
- *Rayleigh scattering*: scattering from constant density variations in the medium.
- *Rayleigh-wing scattering*: scattering from variations in the orientation of anisotropic molecules. This process is spectrally broad because the reorientation process of molecules is very rapid.

5.1 Fluctuations as the origin of light scattering

Scattering can only occur in a medium if it is inhomogeneous. If a medium were homogeneous, and if there were light scattered from any infinitesimal volume dV_1 , then for any non-zero angle θ there would be destructive interference with light scattered by another infinitesimal volume dV_2 as shown in Fig. 37. For $\theta = 0$, coherent forward scattering can occur even in a homogeneous medium. This is the origin of the refractive index[1].

This argument that no scattering can occur in a homogeneous medium requires that the medium is *completely* homogeneous. If there is any inhomogeneity, scattering **can** occur. Temporal or spatial inhomogeneities will lead to variations $\Delta\epsilon_{ik}$ from the mean dielectric constant:

$$\epsilon_{ik} = \bar{\epsilon}\delta_{ik} + \Delta\epsilon_{ik}. \quad (5.1)$$

Let us split the dielectric tensor into a scalar contribution and a tensor contribution:

$$\epsilon_{ik} = \Delta\epsilon\delta_{ik} + \Delta\epsilon_{ik}^{(t)}, \quad (5.2)$$

where we assumed the contribution of “tensor light scattering” to be traceless such that $\sum_i \epsilon_{ii}^{(t)} = 0$. Any contributions that are not traceless, will be part of the scalar

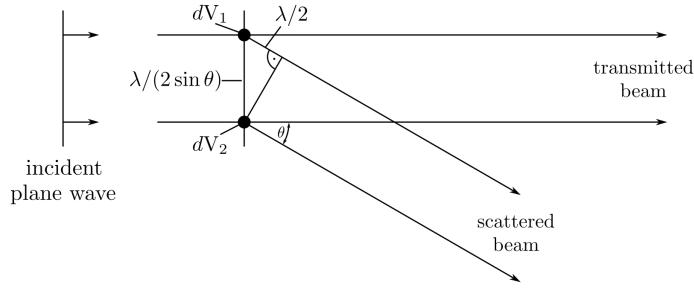


Figure 37: **No scattering in a homogeneous medium.** If light were scattered by a volume dV_1 , there would be a volume dV_2 such that light scattered from there would interfere destructively with the light scattered from dV_1 . This is true for any angle $\theta \neq 0$. This figure is adapted from Ref. [1].

contribution. Scattering that results from these scalar contributions is called “scalar light scattering”. Examples are Brioullin scattering and Rayleigh scattering.

The tensorial part can be split further into a symmetric and an antisymmetric part:

$$\epsilon_{ik}^{(t)} = \epsilon_{ik}^{(s)} + \epsilon_{ik}^{(a)}. \quad (5.3)$$

The symmetric part ($\epsilon_{ik}^{(s)} = \epsilon_{ki}^{(s)}$) is responsible for Rayleigh-wing scattering. The anti-symmetric part ($\epsilon_{ik}^{(a)} = -\epsilon_{ki}^{(a)}$) gives rise to Raman scattering.

5.2 Scattering coefficient and scattering cross section

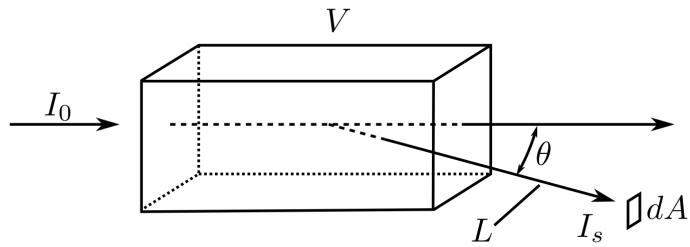


Figure 38: **Definitions for scattering coefficient.** Incident light of intensity I_0 is scattered from a volume V . The intensity of the scattered light over an area dA is I_s at a scattering angle θ and at a distance L from the scattering volume. This figure is based on a figure in Ref. [1].

We describe the efficiency of the scattering process using the scattering coefficient R .

It is defined in terms of the quantities introduced in Fig. 38:

$$I_s = \frac{I_0 RV}{L^2}. \quad (5.4)$$

This relation is the inverse-square relation between the intensity of the scattered light and the distance L from the volume, and it introduces a linear dependence of the scattered light on the incoming light intensity and the scattering volume.

The (total) scattering cross section σ relates the power P of scattered light to the incident power:

$$P = \sigma I_0. \quad (5.5)$$

If we detect the scattered light using a detector with a sensitive area dA , we have $P = I_s dA$. From the scattering volume, the area of the detector will subtend a solid angle $d\Omega = dA/L^2$, and the scattered power per unit solid angle will be:

$$\frac{dP}{d\Omega} = I_s L^2 = I_0 RV = I_0 \frac{d\sigma}{d\Omega}. \quad (5.6)$$

Integrating the differential cross section $d\sigma/d\Omega$ yields the total cross section:

$$\sigma = \int_{4\pi} d\Omega \frac{d\sigma}{d\Omega}. \quad (5.7)$$

One can relate the quantities describing scattering on the macroscopic level to many scattering processes on the microscopic level. If the light scattered from our medium results from light being scattered from \mathcal{N} molecules in the volume V , if each of the molecules scatters according to equation 5.6, and if the scattering processes are independent from each other, then the total power scattered will be \mathcal{N} times the power scattered by an individual molecule. We can conclude that:

$$R = \frac{\mathcal{N}}{V} \frac{d\sigma}{d\Omega}. \quad (5.8)$$

5.3 Stimulated scattering processes

Scattering only occurs if there are fluctuations in the optical processes of an optical medium. The scattering process is called spontaneous if it is caused by thermal or quantum-mechanical zero-point effects[1]. Stimulated scattering effects take place if the fluctuations leading to scattering result themselves from the presence of a light field. That means, they are nonlinear processes. Stimulated scattering can be *much more efficient* than spontaneous scattering. For example the spontaneous scattering of light in water is so small that only a millionth part will be scattered. For stimulated processes, up to 100% can be scattered. See exercise 1 in today's problems.

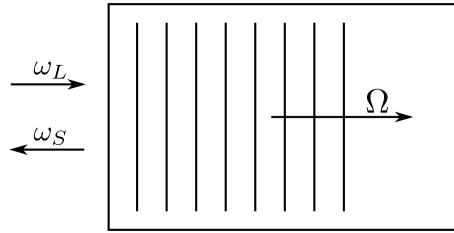


Figure 39: **Stimulated Brillouin scattering (SBS).** An incoming wave with frequency ω_L induces an acoustic wave of frequency Ω . Light is scattered from the acoustic wave. The figure is adapted from Ref. [1].

One of the most important examples is stimulated Brioullin scattering (SBS). We illustrate the effect in Fig. 39. An incoming wave with frequency ω_L induces an acoustic wave of frequency Ω travelling in the same direction as the incident light. Light scattered back from that wave will have the Stokes frequency $\omega_S = \omega_L - \Omega$. This interaction can lead to stimulated light scattering because the incident and the scattered light wave will interfere, resulting in a frequency component at the difference frequency $\omega_L - \omega_S = \Omega$. This frequency part can increase the amplitude of the sound wave, which will lead to more scattering etc. The beating between the incident frequency and the acoustic wave reinforces the Stokes wave, and the beating between the incident light and the Stokes wave reinforces the acoustic wave. This positive feedback can lead to an exponential growth in the amplitude of the Stokes wave.

The incident light and the Stokes wave can drive the acoustic wave via electrostriction or absorption. We focus on electrostriction, which results from materials becoming more dense in the presence of intense optical fields.

5.4 Electrostriction

Apart from stimulated Brillouin scattering, electrostriction is also of interest as a mechanism leading to a third-order nonlinear response.

Let us first consider the situation from the macroscopic perspective illustrated in Fig.40 (a). A small rod of a dielectric material is positioned close to the edges of two capacitor plates and experiences the evanescent field illustrated in Fig.40(b). The potential energy of the dielectric material of our rod per unit volume in the presence of an electric field E compared to the case when there is no field present will be higher by the following amount [1]:

$$u = \frac{1}{2} \epsilon \epsilon_0 E^2, \quad (5.9)$$

ϵ is the dielectric constant of the material. The total energy of our material $\int u dV$ is

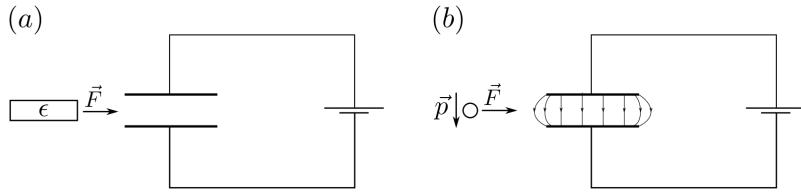


Figure 40: **The origin of electrostriction.** (a) macroscopic description using a dielectric rod close to a capacitor. (b) microscopic description using a polarizable molecule close to the field of a capacitor. Figure adapted from Ref. [1].

maximized if the rod is between the capacitor plates. A force \vec{F} pointing towards the maximum field strength will act on the rod in order to *maximize the stored energy*.

Now let us consider the same situation but on a microscopic level as illustrated in Fig. 40 (b). An individual molecule is located in the vicinity of the field \vec{E} between two capacitor plates. The presence of the field leads to the development of a dipole moment in the molecule: $\vec{p} = \epsilon_0 \alpha \vec{E}$. α is the polarizability of the molecule. The energy stored in the polarization of the molecule will be:

$$U = - \int_0^{\vec{E}} d\vec{E}' \cdot \vec{p} = -\epsilon_0 \alpha \int_0^{\vec{E}} d\vec{E}' \cdot \vec{E}' = -\frac{1}{2} \epsilon_0 \alpha \vec{E} \cdot \vec{E} = -\frac{1}{2} \epsilon_0 \alpha E^2. \quad (5.10)$$

The force experienced by the molecule will then be:

$$\vec{F} = -\nabla U = \frac{1}{2} \epsilon_0 \alpha \nabla (E^2). \quad (5.11)$$

Each molecule experiences a force drawing it towards increasing field strength.

Let us now assume we place the capacitor into a dielectric liquid as in Fig. 41. Due to the force in equation 5.11, the molecules in the liquid will move towards the intensity maximum between the capacitor plates. As a result, the mass density in that region will increase, and the resulting increase in pressure will act to move particles away from that volume. A balance will be achieved if we have the following increase in material density:

$$\Delta\rho = \frac{1}{2} \epsilon_0 \rho C \gamma_e \langle \vec{E} \cdot \vec{E} \rangle. \quad (5.12)$$

$\gamma_e = \rho(\partial\epsilon/\partial\rho)$ is the electrostrictive constant, $C = \rho^{-1}(\partial\rho/\partial p)$ is the compressibility. Here, γ_e and C should be calculated for constant entropy.

In nonlinear optical materials, electrostriction can result in a change in the suscepti-

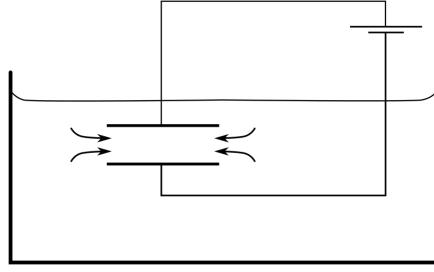


Figure 41: **Electrostriction in a dielectric liquid.** If we apply a field between capacitor plates within a dielectric liquid, the dielectric molecules will flow towards the intensity maximum between the capacitors. The figure is based on a figure in Ref. [1].

bility in the presence of an optical field:

$$\Delta\chi = \Delta\epsilon = \frac{\partial\epsilon}{\partial\rho}\Delta\rho = \frac{1}{2}\epsilon_0 C\gamma_e^2 \langle \vec{E} \cdot \vec{E} \rangle. \quad (5.13)$$

Let us assume that the applied field is monochromatic:

$$\vec{E}(t) = \vec{E}e^{-i\omega t} + \text{c.c.} \quad (5.14)$$

Because $\langle \vec{E} \cdot \vec{E} \rangle = 2\vec{E} \cdot \vec{E}^*$, we get:

$$\Delta\chi = \epsilon_0 C_T \gamma_e^2 \vec{E} \cdot \vec{E}^* = \epsilon_0 C_T \gamma_e^2 |\vec{E}|^2, \quad (5.15)$$

where the isothermal compressibility C_T is defined as[1]:

$$C_T = -\frac{1}{V} \left(\frac{\partial V}{\partial p} \right)_T. \quad (5.16)$$

Then the complex amplitude of the nonlinear polarization will be:

$$\vec{P} = \Delta\chi \vec{E} = \epsilon_0 C_T \gamma_e^2 |\vec{E}|^2 \vec{E}. \quad (5.17)$$

If we compare this with our usual notation for a third-order susceptibility, we get:

$$\chi^{(3)}(\omega = \omega + \omega - \omega) = \frac{1}{3} \epsilon_0 C_T \gamma_e^2. \quad (5.18)$$

If we compare that with the coefficients A and B we introduced in section 3.2, we get $A = C_T \gamma_e^2$ and $B = 0$. These represent the contributions to the third-order process due

to one-photon and two-photon transitions.

5.5 Stimulated Brillouin scattering induced by electrostriction

If the input light is intense enough, the scattered light can also become quite intense. The beating between the incident and the scattered light can lead to density and pressure variations in the material due to electrostriction. In turn, the incident light can scatter off these variations, and the resulting light will be at the Stokes frequency. In a positive feedback, this light can add constructively to the Stokes radiation that produced the density variations in the first place. Fig. 42 illustrates this stimulated Brillouin scattering (SBS) process.

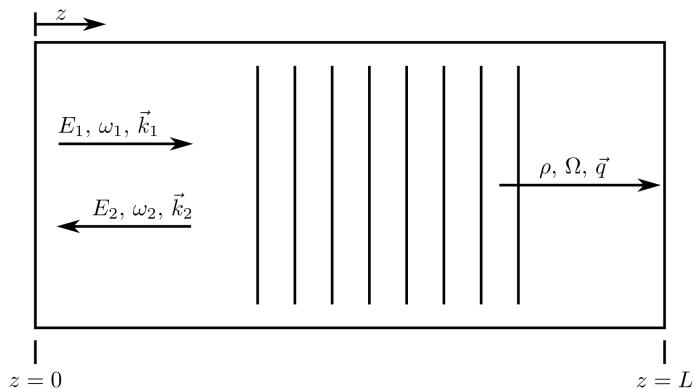


Figure 42: **Stimulated Brillouin scattering** in a crystal of length L . The incident light with frequency ω_1 and wave vector \vec{k}_1 induces an acoustic wave with frequency Ω and wave vector \vec{q} . The scattered field E_2 has frequency ω_2 and wave vector \vec{k}_2 . The figure is adapted from Ref. [1].

The incident light with frequency ω_1 and wave vector \vec{k}_1 scatters off a retreating sound wave with wavevector \vec{q} and frequency Ω . The frequency of the scattered light will be:

$$\omega_2 = \omega_1 - \Omega_B, \quad (5.19)$$

where Ω_B is called the Brillouin frequency.

The Brillouin frequency is related to the acoustic wave vector and the velocity of sound v via the phonon dispersion relation:

$$\Omega_B = |\vec{q}_B|v. \quad (5.20)$$

Beating between the incident light and the Stokes field drives the sound wave such that:

$$\vec{q}_B = \vec{k}_1 - \vec{k}_2. \quad (5.21)$$

Using the standard dispersion relations $|\vec{k}_i| = n\omega_i/c$, we get:

$$\Omega_B = \frac{nv}{c}(\omega_1 + \omega_2). \quad (5.22)$$

If we combine this equation with equation 5.19, we get the Brillouin frequency:

$$\Omega_B = \frac{\frac{2v}{c/n}\omega_1}{1 + \frac{v}{c/n}} \approx \frac{2nv}{c}\omega_1. \quad (5.23)$$

The approximation assumes $v \ll c$. At the same, very good approximation gives[1]:

$$\vec{q}_B \approx 2\vec{k}_1. \quad (5.24)$$

If we additionally pump the process from the opposite direction with a counter-propagating beam at frequency ω_2 , the acoustic wave generated will have the frequency:

$$\Omega = \omega_1 - \omega_2. \quad (5.25)$$

We will see that this process is only efficient if Ω lies within the “Brillouin linewidth” Γ_B around the Brillouin frequency.

Let us look more closely at the theoretical treatment of SBS. We will assume the two fields $E_{1,2}$ to be scalar quantities (well defined crystal orientation and field polarizations). The total field will be $E(z, t) = E_1(z, t) + E_2(z, t)$ with:

$$E_j(z, t) = A_j(z, t)e^{i(k_j z - \omega_j t)} + \text{c.c..} \quad (5.26)$$

The material density is described as:

$$\rho(z, t) = \rho_0 + [\rho(z, t)e^{i(qz - \Omega t)} + \text{c.c.}], \quad (5.27)$$

where $\Omega = \omega_1 - \omega_2$, $q = 2k_1$, and ρ_0 is the undisturbed material density.

Like for AOMs, we assume the material density to satisfy the acoustic wave equation:

$$\frac{\partial^2 \rho}{\partial t^2} - \Gamma' \nabla^2 \frac{\partial \rho}{\partial t} - v^2 \nabla^2 \rho = \nabla \cdot \vec{f}, \quad (5.28)$$

where Γ' is the damping parameter

$$\Gamma' = \frac{1}{\rho} \left[\frac{4}{3} \eta_s + \eta_b + \frac{\kappa}{C_p} (\gamma - 1) \right]. \quad (5.29)$$

η_s is the shear viscosity coefficient, η_b is the bulk viscosity coefficient, κ is the thermal conductivity, and γ is the adiabatic index given by:

$$\gamma = \frac{C_T}{C_S} = \frac{c_p}{c_V}. \quad (5.30)$$

$C_{T,S}$ express the compressibility at constant temperature or constant entropy, the $c_{p,V}$ express the heat capacity at constant pressure or constant volume.

The source term on the right-hand side of equation 5.28 is the divergence of the force per unit volume \vec{f} :

$$\vec{f} = \nabla p_{\text{st}}, \quad (5.31)$$

where $p_{\text{st}} = -\frac{1}{2}\epsilon_0\gamma_e \langle E^2 \rangle$. If we assume fields of the form given in equation 5.26, we get:

$$\nabla \cdot \vec{f} = \epsilon_0\gamma_e q^2 [A_1 A_2^* e^{i(qz-\Omega t)} + \text{c.c.}] . \quad (5.32)$$

If we use all these definitions, then the acoustic wave equation 5.28 becomes:

$$-2i\Omega \frac{\partial \rho}{\partial t} + (\Omega_B^2 - \Omega^2 - i\Omega\Gamma_B)\rho - 2iqv^2 \frac{\partial \rho}{\partial z} = \epsilon_0\gamma_e q^2 A_1 A_2^*. \quad (5.33)$$

Here, we introduced the Brillouin linewidth Γ_B we mentioned earlier:

$$\Gamma_B = q^2\Gamma'. \quad (5.34)$$

The inverse of the Brillouin linewidth $\tau_p = \Gamma_B^{-1}$ is the phonon lifetime.

We will see in the exercises that one can usually neglect the last term on the left-hand side of equation 5.28 because hypersonic phonons are usually absorbed very rapidly. That means, we can neglect $\partial\rho/\partial z$ in SBS. Sometimes, this approximation will fail.

If we also assume steady-state conditions such that $\partial\rho/\partial t$ vanishes, we get:

$$\rho(z, t) = \epsilon_0\gamma_e q^2 \frac{A_1 A_2^*}{\Omega^2_B - \Omega^2 - i\Omega\Gamma_B}. \quad (5.35)$$

The two contributions ($j = 1, 2$) to the electrical field have to fulfill the wave equations:

$$\frac{\partial^2 E_j}{\partial z^2} - \frac{n^2}{c^2} \frac{\partial^2 E_j}{\partial t^2} = \frac{1}{\epsilon_0 c^2} \frac{\partial^2 P_j}{\partial t^2}. \quad (5.36)$$

The source terms in this equation result from the total polarization, which is given by:

$$P = \epsilon_0 \Delta \chi E = \epsilon_0 \Delta \epsilon E = \epsilon_0 \rho_0^{-1} \gamma_e \rho E. \quad (5.37)$$

Only some parts of P can act as phase-matched source terms. These are given by:

$$\begin{aligned} P_j &= p_j e^{i(k_j z - \omega_j t)} + \text{c.c.}, \text{ where} \\ p_1 &= \epsilon_0 \gamma_e \rho_0^{-1} \rho A_2, \text{ and} \\ p_2 &= \epsilon_0 \gamma_e \rho_0^{-1} \rho^* A_1. \end{aligned} \quad (5.38)$$

Putting that into the wave equations 5.36 and assuming slowly varying amplitudes, we get the coupled wave equations:

$$\begin{aligned} \frac{\partial A_1}{\partial z} + \frac{n}{c} \frac{\partial A_1}{\partial t} &= \frac{i\omega \gamma_e}{2nc\rho_0} \rho A_2 \\ -\frac{\partial A_2}{\partial z} + \frac{n}{c} \frac{\partial A_2}{\partial t} &= \frac{i\omega \gamma_e}{2nc\rho_0} \rho^* A_1. \end{aligned} \quad (5.39)$$

For the steady-state case this simplifies to:

$$\begin{aligned} \frac{dA_1}{dz} &= \frac{i\epsilon_0 \omega q^2 \gamma_e^2}{2nc\rho_0} \frac{|A_2|^2 A_1}{\Omega_B^2 - \Omega^2 - i\Omega\Gamma_B} \\ \frac{dA_2}{dz} &= \frac{-i\epsilon_0 \omega q^2 \gamma_e^2}{2nc\rho_0} \frac{|A_1|^2 A_2}{\Omega_B^2 - \Omega^2 + i\Omega\Gamma_B}. \end{aligned} \quad (5.40)$$

These equations describe a “pure gain process”[1] with exponential decay or growth. Only the gain depends on the amplitude of the other field. Such a process is automatically phase matched, and we can rewrite the coupled amplitude equations in terms of the intensities $I_j = 2n\epsilon_0 c |A_j|^2$:

$$\frac{dI_1}{dz} = \frac{dI_2}{dz} = -g I_1 I_2. \quad (5.41)$$

g is the *SBS gain factor*, which can be written to good approximation[1] as:

$$g = g_0 \frac{(\Gamma_B/2)^2}{(\Omega_B - \Omega)^2 + (\Gamma_B/2)^2}, \quad (5.42)$$

where g_0 is the line-center gain factor:

$$g_0 = \frac{\gamma_e^2 \omega^2}{nvc^3 \rho_0 \Gamma_B}. \quad (5.43)$$

If the gain is not too strong (low intensities, short crystals) and the incident pump is not depleted, one can assume I_1 to remain constant, yielding exponential growth in I_2 :

$$I_2(z) = I_2(L) e^{gI_1(L-z)}. \quad (5.44)$$

That means, if we inject a Stokes field from the backside of the medium ($z = L$), the Stokes field will experience an exponential growth as it propagates through the medium. This describes an *SBS amplifier*, given that there is no pump depletion. A more general description including pump depletion can be found in Ref. [1]. In any case, an SBS amplifier describes a medium that is pumped from one side by a beam with intensity I_1 , and a Stokes beam with intensity I_2 enters through the other end of the medium. This Stokes beam is exponentially amplified via SBS as it travels through the nonlinear medium. This allows very strong amplification even for weak intensities I_2 .

An interesting thing to note is that the line-center gain factor g_0 in equation 5.43 is actually independent of the frequency ω because Γ_B is proportional to ω^2 .

5.5.1 SBS Generator

We just saw that SBS amplifiers can amplify even weak Stoke signals entering from the opposite side of the gain medium. A special case is if we do **not** apply a Stokes field at frequency ω_2 . We can still get a significant output at the Stokes frequency $\omega_2 = \omega_1 - \omega_B$ because the incident field can result in spontaneous Brillouin scattering close to the end of the medium ($z = L$). On the way back through the medium, this will be amplified.

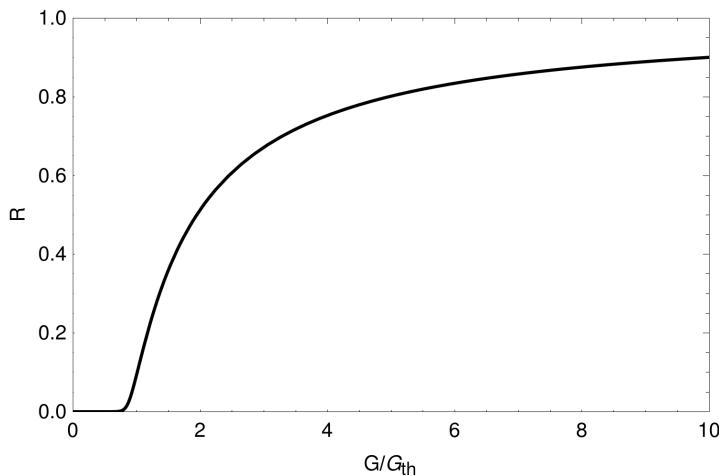


Figure 43: **SBS reflectivity vs gain.** The figure shows the SBS reflectivity as a function of the SBS gain in units of the threshold gain. We assumed $G_{\text{th}} = 25$.

Let us assume that the intensity at ω_2 is proportional to the input intensity at ω_1 :

$$I_2(L) = f I_1(L). \quad (5.45)$$

How large will f be? Let us assume that the “SBS reflectivity” $R = I_2(0)/I_1(0)$ is much smaller than unity. Then we can assume that $I_1(z)$ is constant (no depletion), and the

Stokes output intensity will be $I_2(0) = I_2(L)e^G$, where $G = gI_1(0)L$. That means:

$$R \equiv \frac{I_2(0)}{I_1(0)} = fe^G. \quad (5.46)$$

Experiments show that significant SBS only sets in if the gain is beyond a certain threshold G_{th} that typically is between 25 and 30. Using equation 5.46, we conclude that f should be between 10^{-12} and 10^{-11} .

One can show that the following relation holds[1]:

$$\frac{G}{G_{\text{th}}} = \frac{G_{\text{th}}^{-1} \ln R + 1}{1 - R}. \quad (5.47)$$

Fig. 43 illustrates how the SBS reflectivity depends on the gain.

5.6 Stimulated Raman scattering

We mentioned earlier that Raman scattering is due to the interaction of light with vibrational modes in the medium. While spontaneous Raman scattering typically is a very weak process, we will see that a significant amount of the incident light can be scattered into Stokes or anti-Stokes modes in stimulated Raman scattering.

5.6.1 Spontaneous Raman scattering

If scattering occurs in a medium, the scattered light will have components at the Stokes frequency ω_S , and at the anti-Stokes frequency ω_a . As we will see, the anti-Stokes component will usually be much weaker than the Stokes component. This situation is depicted in Fig. 44, but an energy-level diagram will give us a better understanding (see Fig. 45). The Stokes and anti-Stokes Raman processes are both two-photon processes resulting from a transition either from a ground state g to a higher energy state n (Stokes) or the other way round (anti-Stokes). In both cases, the transition occurs via an intermediate virtual level associated with an excited state n' . The anti-Stokes component is typically weaker than the Stokes component because the energy-level occupation number in thermal equilibrium is lower for state n than for state g by the Boltzmann factor $\exp(-\hbar\omega_{ng}/k_B T)$.

An interesting thing to note about the Raman effect is that it can occur between energy levels where a one-photon transition is normally forbidden. For example, the Raman transitions we illustrate in Fig. 45 can only occur if the matrix elements $\langle g|\vec{r}|n'\rangle$ and $\langle n'|\vec{r}|n\rangle$ are nonzero[1]. If our medium is inversion symmetric, this implies that the states g and n must have the same parity. The matrix element $\langle g|\vec{r}|n\rangle$ therefore must vanish because \vec{r} has odd parity. Since the single-photon dipole operator is proportional to \vec{r} , this means that there cannot be any single-photon transitions $g \rightarrow n$.

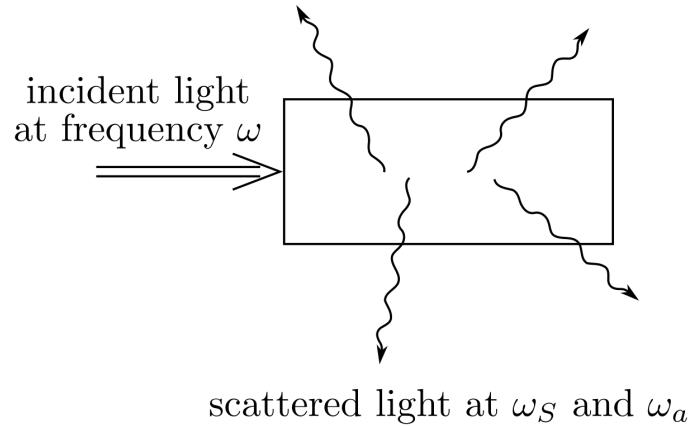


Figure 44: **Spontaneous Raman scattering.** If light of frequency ω is Raman scattered, the scattered light will have components at the Stokes frequency ω_S and at the anti-Stokes frequency ω_a .

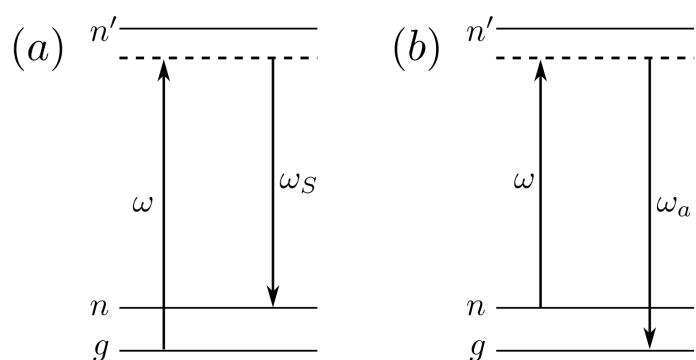


Figure 45: **Energy level diagrams for Raman scattering.** (a) the Raman Stokes scattering process, (b) the Raman anti-Stokes scattering process.

5.6.2 Spontaneous vs stimulated Raman scattering

The process of spontaneous Raman scattering is typically a very weak effect. The probability for an incident photon to generate a Raman Stokes scattered photon as it travels through 1 cm of a solid medium is usually on the order of 10^{-6} [1]. If the incident light is very intense, however, stimulated Raman scattering may occur. The efficiency of this process can be 10% or more at the Stokes frequency[1]. Another difference between spontaneous and stimulated Raman scattering is the spatial distribution of the scattered photons. For spontaneous Raman scattering, it follows a dipole radiation pattern covering a large solid angle. Stimulated Raman scattering leads to the emission of narrow cones in the forward and backward directions.

We can use a very simple argument to describe the different efficacies of spontaneous and stimulated Raman scattering. Using quantum mechanics one can show that the probability per unit time for a photon to be emitted into the Stokes mode S is given by[1]:

$$P_S = Dm_L(m_S + 1). \quad (5.48)$$

D is a material constant, m_L is the number of photons in the pump beam, and m_S is the number of photons in the Stokes mode. As we would expect, the probability is proportional to the number of pump photons. The second factor consists of two contributions. As we will see, the first summand m_S relates to stimulated Raman scattering, and the second summand relates to spontaneous Raman scattering.

Given that P_S is the probability per unit time for a photon to be emitted into the Stokes mode, we can write:

$$\frac{dm_S}{dt} = P_S = Dm_L(m_S + 1). \quad (5.49)$$

If we assume that our Stokes mode corresponds to a wave travelling in the positive z direction with the phase velocity c/n , we can rewrite equation 5.49 in terms of z :

$$\frac{dm_S}{dz} = \frac{n}{c} \frac{dm_S}{dt} = \frac{n}{c} Dm_L(m_S + 1). \quad (5.50)$$

Let us look at this equation for two limiting cases. First, let us assume that $m_S \ll 1$. In that case, the differential equation becomes:

$$\frac{dm_S}{dz} = \frac{n}{c} Dm_L \text{ (for } m_S \ll 1\text{).} \quad (5.51)$$

That means there will be a linear growth of Stokes photons along the z direction:

$$m_S(z) = m_S(0) + \frac{n}{c} Dm_L z \text{ (for } m_S \ll 1\text{).} \quad (5.52)$$

This corresponds to the case of spontaneous Raman scattering.

For the second limiting case, we assume $m_S \gg 1$. Then the equation 5.50 becomes:

$$\frac{dm_S}{dz} = \frac{n}{c} D m_L m_S \quad (\text{for } m_S \gg 1), \quad (5.53)$$

which corresponds to an exponential growth in the number of Stokes photons:

$$m_S(z) = m_S(0) e^{Gz} \quad (\text{for } m_S \ll 1), \quad (5.54)$$

where the Raman gain coefficient is given by:

$$G = \frac{n D m_L}{c}. \quad (5.55)$$

In both cases, $m_S(0)$ is the number of Stokes photons at the input of the Raman medium. If there is no pump light, we can get Stokes photons due to quantum fluctuations.

So far, we assumed that there would be no pump depletion as we create Stokes photons. In general, one pump photon is lost for every Stokes photon that is created. That means, we should have $dm_L/dz = -dm_S/dz$. Usually, however, Stokes photons will not only be emitted into one mode but into many modes. If we take that into account, we get:

$$\frac{dm_L}{dz} = -M b \frac{dm_S}{dz} = -\frac{D m_L M b}{c/n}. \quad (5.56)$$

Here, we introduced M to denote the number of modes we scatter Stokes photons into. and b is a geometrical factor taking into account the non-uniformity of the scattered radiation. In particular, b is the ratio between the angularly averaged Stokes emission rate and the rate in the direction of the particular Stokes modes S , we are interested in. If $|f(\theta, \phi)|^2$ is the angular distribution of the Stokes radiation, we have:

$$b = \frac{\int d\Omega |f(\theta, \phi)|^2}{4\pi |f(\theta_S, \phi_S)|^2}. \quad (5.57)$$

θ_S and ϕ_S define the direction of the Stokes mode S .

5.6.3 Describing stimulated Raman scattering using the nonlinear polarization

We will give a simplified, classical description of stimulated Raman scattering. Still, this derivation will hopefully provide an intuition how this process works. Consider a molecule consisting of two parts with a mean distance q_0 . The distance of the two parts can vary in time as $q_0 + q(t)$ (see Fig. 46). We describe the evolution q as a harmonic oscillator:

$$\frac{d^2q}{dt^2} + 2\gamma \frac{dq}{dt} + \omega_v^2 q = \frac{F(t)}{m}. \quad (5.58)$$

γ is a damping constant, and ω_v is the vibration frequency. We assumed that the parts of the molecule have the nuclear mass m .

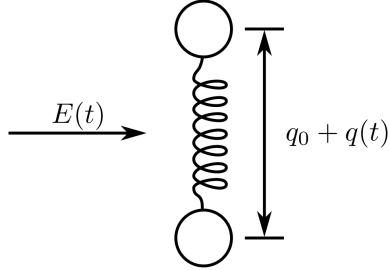


Figure 46: **Vibrations of a molecule.** We use a simple oscillator model describing the vibrations of two parts of a molecule as illustrated here. q_0 is the mean distance between the two parts, $q(t)$ describes the displacement with respect to that mean. $E(t)$ denotes the electric field.

Let us assume that the polarizability of our molecule depends then is given by:

$$\alpha(t) = \alpha_0 + \left(\frac{\partial \alpha}{\partial q} \right)_0 q(t). \quad (5.59)$$

α_0 is the undisturbed equilibrium polarization of the molecule. Periodic vibrations of the molecule will result in corresponding variations of the polarization. If our material consists of N molecules, the refractive index will change accordingly:

$$n(t) = \sqrt{\epsilon(t)} = [1 + N\alpha(t)]^{1/2}. \quad (5.60)$$

This modulation of the refractive index in time will result in Raman scattering. Let us assume that the optical field vector is $\vec{E}(z, t)$. In the presence of this field, every molecule will become polarized, and the induced dipole moment of a molecule located at position z will be[1]:

$$\vec{p}(z, t) = \epsilon \alpha \vec{E}(z, t). \quad (5.61)$$

To establish this oscillating dipole, the field has to exert some work W :

$$W = \frac{1}{2} \langle \vec{p}(z, t) \cdot \vec{E}(z, t) \rangle = \frac{1}{2} \epsilon_0 \alpha \langle E^2(z, t) \rangle, \quad (5.62)$$

where the angular brackets denote the time average. The applied optical field will exert a force F on our oscillator, which we can write as:

$$F = \frac{dW}{dq} = \frac{\epsilon_0}{2} \left(\frac{d\alpha}{dq} \right)_0 \langle E^2(z, t) \rangle. \quad (5.63)$$

Because the force is proportional to the square of the field, the molecule will experience a time-varying force at the beat frequency between different components of the field.

If the incident light has the frequency ω_L , and the molecules vibrate with frequency ω_v , this will create sidebands at the frequencies $\omega_S = \omega_L \pm \omega_v$. If the Stokes light beats with the laser frequency, the total intensity will be modulated as:

$$I(t) = I_0 + I_1 \cos(\omega_L - \omega_S)t = I_0 + I_1 \cos(\omega_v)t. \quad (5.64)$$

That means, the modulation will resonantly drive the molecular vibration, leading to a higher amplitude in the Stokes sideband. This positive feedback can lead to an exponential growth of the scattered field.

To describe the effect in more detail, let us assume that we can write our total field in terms of slowly varying amplitudes $A_{L,S}$:

$$E(z, t) = A_L e^{i(k_L z - \omega_L t)} + A_S e^{i(k_S z - \omega_S t)} + \text{c.c.} \quad (5.65)$$

Then the time-varying part of the applied force in equation 5.63 becomes:

$$F(z, t) = \epsilon_0 \left(\frac{\partial \alpha}{\partial q} \right)_0 [A_L A_S^* e^{i(Kz - \Omega t)} + \text{c.c.}], \quad (5.66)$$

where we used the following definitions:

$$K = k_L - k_S \text{ and } \Omega = \omega_L - \omega_S. \quad (5.67)$$

Let us put that into equation 5.58 and try to find a solution for the displacement $q(t)$ using the “ansatz”:

$$q(t) = q(\Omega) e^{i(Kz - \Omega t)} + \text{c.c..} \quad (5.68)$$

Here, $q(\Omega)$ is the amplitude of the oscillation at frequency Ω . One can then show that:

$$q(\Omega) = \frac{(\epsilon_0/m)(\partial \alpha / \partial q)_0 A_L A_S^*}{\omega_v^2 - \Omega^2 - 2i\Omega\gamma}. \quad (5.69)$$

If we use our earlier assumption that our medium consists of N molecules, we can derive the nonlinear polarization. In particular, we will be interested in the part of the resulting expression that oscillates at the Stokes frequency:

$$P_S^{\text{NL}}(z, t) = P(\omega_S) e^{-i\omega_S t} + \text{c.c..} \quad (5.70)$$

During the exercises, we will show that the Stokes polarization is given by:

$$P(\omega_S) = 6\epsilon_0 \chi_R(\omega_S) |A_L|^2 A_S e^{ik_S z}, \quad (5.71)$$

where we defined the *Raman susceptibility* $\chi_R(\omega_S) \equiv \chi^{(3)}(\omega_S = \omega_S + \omega_L - \omega_L)$ as:

$$\chi_R(\omega_S) = \frac{\epsilon_0(N/6m)(\partial\alpha/\partial q)_0^2}{\omega_v^2 - (\omega_L - \omega_S)^2 + 2i(\omega_L - \omega_S)\gamma}. \quad (5.72)$$

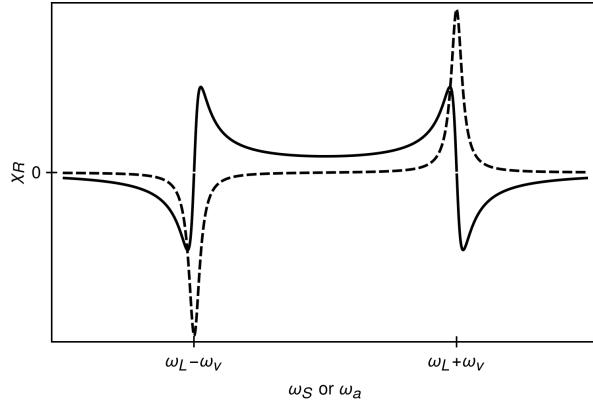


Figure 47: **Raman susceptibility.** The solid curve shows the real part of the Raman susceptibility, and the dashed curve shows the imaginary part. The Stokes scattering leads to a dip in the imaginary part at $\omega_L - \omega_v$, and anti-Stokes scattering leads to a peak of the imaginary part at $\omega_L + \omega_v$. The real part becomes zero at $\omega_L \pm \omega_v$ for these two cases.

Similarly, one can get the anti-Stokes susceptibility:

$$\chi_R(\omega_a) = \frac{\epsilon_0(N/6m)(\partial\alpha/\partial q)_0^2}{\omega_v^2 - (\omega_L - \omega_a)^2 + 2i(\omega_L - \omega_a)\gamma}. \quad (5.73)$$

Since this expression is of exactly the same form as the expression in equation 5.72, we can plot the behaviour of the overall susceptibility in Fig. 47 by using ω_S or ω_a in these expressions as the plot argument. However, as we mentioned earlier, usually the amplitude for Stokes scattering will be significantly higher than for anti-Stokes scattering.

To describe the spatial evolution of the Stokes (or anti-Stokes) wave, let us put the nonlinear polarization into the driven wave equation 2.5. For a lossless, dispersionless and isotropic medium, we can assume:

$$\vec{D}^{(1)} = \epsilon_0 \epsilon^{(1)} \vec{E}. \quad (5.74)$$

The resulting simplified driven wave equation will be:

$$-\vec{\nabla}^2 \vec{E} - \frac{\epsilon^{(1)}}{c^2} \frac{\partial^2 \vec{E}}{\partial t^2} = -\frac{1}{\epsilon_0 c^2} \frac{\partial^2 \vec{P}^{\text{NL}}}{\partial t^2}. \quad (5.75)$$

We can further simplify that assuming slowly varying amplitudes. Then we get the following equation for the amplitude of the Stokes field:

$$\frac{dA_S}{dz} = -\alpha_S A_S, \quad (5.76)$$

where we introduced the Stokes wave “absorption” coefficient:

$$\alpha_S = -3i \frac{\omega_S}{n_S c} \chi_R(\omega_S) |A_L|^2. \quad (5.77)$$

Because the imaginary part of $\chi_R(\omega_S)$ is negative, the real part of the absorption coefficient α_S will be negative. That implies that we actually get an exponential growth of the Stokes amplitude.

For the amplitude of the anti-Stokes wave, we get:

$$\frac{dA_a}{dz} = -\alpha_a A_a, \quad (5.78)$$

where the anti-Stokes wave absorption coefficient is:

$$\alpha_a = -3i \frac{\omega_a}{n_a c} \chi_R(\omega_a) |A_L|^2. \quad (5.79)$$

Because the imaginary part of $\chi_R(\omega_a)$ is positive, the real part of the absorption coefficient α_a will be positive. This implies that we get an exponential attenuation of the anti-Stokes amplitude.

It turns out, however, that this is not the complete picture. There is an additional contribution to the anti-Stokes generation beyond what is apparent from the Raman susceptibility given in equation 5.73. In particular, if one looks at the terms in P^{NL} , one finds that there is a contribution to the anti-Stokes polarization

$$P_a^{\text{NL}} = P(\omega_a) e^{-i\omega_a t} + \text{c.c.} \quad (5.80)$$

that depends on the Stokes amplitude. It is given by:

$$P(\omega_a) = N \epsilon_0 \left(\frac{\partial \alpha}{\partial q} \right)_0 q(\Omega) A_L = \frac{(N \epsilon_0^2 / m) (\frac{\partial \alpha}{\partial q})_0 A_L^2 A_S^*}{\omega_v^2 - \Omega^2 - 2i\Omega\gamma} e^{i(2k_L - k_S)z}. \quad (5.81)$$

This contribution can be described in terms of a four-wave mixing susceptibility $\chi_F(\omega_a) \equiv \chi^{(3)}(\omega_a = \omega_L + \omega_L - \omega_S)$, which we can define via the following relation:

$$P(\omega_a) = 3\epsilon_0 \chi_F(\omega_a) A_L^2 A_S^* e^{i(2k_L - k_S)z}. \quad (5.82)$$

That means, we have:

$$\chi_F(\omega_a) = 2\chi_R(\omega_a) = \frac{\epsilon_0(N/3m)(\partial\alpha/\partial q)_0^2}{\omega_v^2 - (\omega_L - \omega_a)^2 + 2i(\omega_L - \omega_a)\gamma}. \quad (5.83)$$

The total polarization at the anti-Stokes frequency then is the sum of this four-wave mixing contribution and the contribution we got earlier:

$$P(\omega_a) = 6\epsilon_0\chi_R(\omega_a)|A_L|^2 A_a e^{ik_a z} + 3\epsilon_0\chi_F(\omega_a)A_L^2 A_S^* e^{i(2k_L - k_S)z}. \quad (5.84)$$

For the Stokes amplitude, we also get a contribution from four-wave mixing, and the total polarization at the Stokes frequency then is:

$$P(\omega_S) = 6\epsilon_0\chi_R(\omega_S)|A_L|^2 A_S e^{ik_S z} + 3\epsilon_0\chi_F(\omega_S)A_L^2 A_a^* e^{i(2k_L - k_a)z}. \quad (5.85)$$

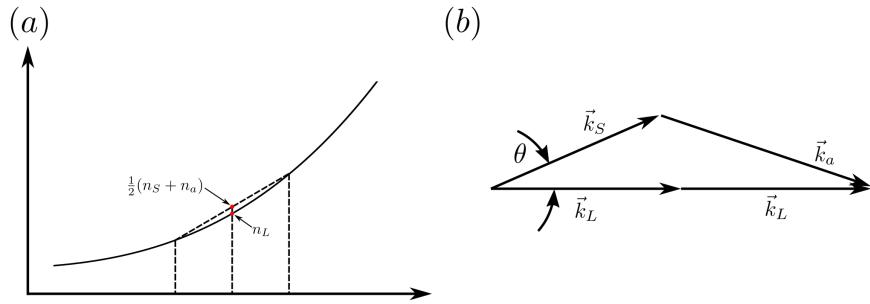


Figure 48: **Phase-matching considerations for Stokes and anti-Stokes coupling.**
 (a) Refractive index in a medium with normal dispersion. This shows that phase-matching can be fulfilled in principle. (b) In order to fulfill phase-matching the Stokes and anti-Stokes fields need to have a small angle θ relative to the pump beam.

If one puts this polarization into the driven wave equation 5.75 and takes the slowly-varying amplitude approximation, one arrives at the following coupled differential equations for the Stokes and anti-Stokes field amplitudes:

$$\begin{aligned} \frac{dA_S}{dz} &= -\alpha_S A_S + \kappa_S A_a^* e^{i\Delta kz}, \\ \frac{dA_a}{dz} &= -\alpha_a A_a + \kappa_a A_S^* e^{i\Delta kz}. \end{aligned} \quad (5.86)$$

Here, we introduced the following nonlinear absorption and coupling coefficients:

$$\begin{aligned}\alpha_j &= \frac{-3i\omega_j}{n_j c} \chi_R(\omega_j) |A_L|^2, \quad j = S, a \\ \kappa_j &= \frac{3i\omega_j}{2n_j c} \chi_F(\omega_j) A_L^2, \quad j = S, a.\end{aligned}\quad (5.87)$$

The wavevector mismatch is given by:

$$\Delta k = \Delta \vec{k} \cdot \hat{\vec{z}} = (2\vec{k}_L - \vec{k}_S - \vec{k}_a) \cdot \hat{\vec{z}}. \quad (5.88)$$

The coupled amplitude equations 5.86 shows that the Stokes and anti-Stokes amplitudes are driven by a Raman gain or loss term (the first terms in the two differential equations) and by a phase-matched four-wave mixing term (the second terms in the two differential equations). Fig.48 (a) shows that the refractive index n_L at the laser wavelength is always lower than the average of the refractive indices at the Stokes and anti-Stokes frequencies if we have normal dispersion. For that reason, one can always achieve perfect phase matching if the Stokes and anti-Stokes fields have some nonzero angle θ with the pump field. This is illustrated in Fig.48 (b). On the other hand, if Δk is large and phase-matching is not fulfilled, only the first terms on the right-hand sides of the coupled amplitude equations 5.86 are relevant. Then the Stokes amplitude experiences amplification while the anti-Stokes amplitude experiences loss.

5.6.4 Stokes–anti-Stokes coupling

Now, let us assume that Δk is sufficiently small such that we have to take into account all terms in the coupled amplitude equations 5.86, but let us rewrite these equations in a more general form that one often encounters in nonlinear optics - for example, in any forward four-wave mixing process in the constant-pump approximation:

$$\begin{aligned}\frac{dA_1}{dz} &= -\alpha_1 A_1 + \kappa_1 A_2^* e^{i\Delta kz}, \\ \frac{dA_2^*}{dz} &= -\alpha_2^* A_2^* + \kappa_2^* A_1 e^{-i\Delta kz}.\end{aligned}\quad (5.89)$$

Next, we will simplify these equations by first writing them in the following form:

$$\begin{aligned}e^{-i\Delta kz/2} \left(\frac{dA_1}{dz} + \alpha_1 A_1 \right) &= \kappa_1 A_2^* e^{i\Delta kz/2}, \\ e^{i\Delta kz/2} \left(\frac{dA_2^*}{dz} + \alpha_2^* A_2^* \right) &= \kappa_2^* A_1 e^{-i\Delta kz/2}.\end{aligned}$$

Then we can rewrite the coupled equation as:

$$\begin{aligned} \left(\frac{d}{dz} + \alpha_1 + \frac{i\Delta k}{2} \right) A_1 e^{-i\Delta kz/2} &= \kappa_1 A_2^* e^{i\Delta kz/2}, \\ \left(\frac{d}{dz} + \alpha_2^* - \frac{i\Delta k}{2} \right) A_2^* e^{i\Delta kz/2} &= \kappa_2^* A_1 e^{-i\Delta kz/2}. \end{aligned}$$

We can then introduce new variables $F_j = A_j \exp(-i\Delta kz/2)$, where $j = 1, 2$. Then we can write the coupled equations as:

$$\begin{aligned} \left(\frac{d}{dz} + \alpha_1 + \frac{i\Delta k}{2} \right) F_1 &= \kappa_1 F_2^*, \\ \left(\frac{d}{dz} + \alpha_2^* - \frac{i\Delta k}{2} \right) F_2^* &= \kappa_2^* F_1. \end{aligned}$$

One can then eliminate F_2^* and end up with a single equation instead:

$$\left(\frac{d}{dz} + \alpha_2^* - \frac{i\Delta k}{2} \right) \left(\frac{d}{dz} + \alpha_1 + \frac{i\Delta k}{2} \right) F_1 = \kappa_1 \kappa_2^* F_1. \quad (5.90)$$

An ansatz for the solution of that equation is:

$$F_1(z) = F_1(0) e^{gz}. \quad (5.91)$$

With this ansatz and equation 5.90, we get the following equation for the growth rate g :

$$\left(g + \alpha_2^* - \frac{i\Delta k}{2} \right) \left(g + \alpha_1 + \frac{i\Delta k}{2} \right) = \kappa_1 \kappa_2^*. \quad (5.92)$$

This equation has the following two solutions:

$$g_{\pm} = -\frac{1}{2}(\alpha_1 + \alpha_2^*) \pm \frac{1}{2} [(\alpha_1 - \alpha_2^* + i\Delta k)^2 + 4\kappa_1 \kappa_2^*]^{1/2}. \quad (5.93)$$

Except for special values of $\alpha_{1,2}$, $\kappa_{1,2}$, and Δk , the gains g_{\pm} have two distinct values. In that case, we get two independent solutions of our differential equation, and the general solution for F can be written as:

$$F_1 = F_1^+(0) e^{g_+ z} + F_1^-(0) e^{g_- z}. \quad (5.94)$$

Using this expression, one can write down the general solution for A_1 :

$$A_1(z) = (A_1^+ e^{g_+ z} + A_1^- e^{g_- z}) e^{i\Delta kz/2}, \quad (5.95)$$

where A_1^\pm are integration constants we can determine from the boundary conditions. From this, we can conclude that the general solution for A_2^* is given by:

$$A_2^*(z) = (A_2^{+*} e^{g_+ z} + A_2^{-*} e^{g_- z}) e^{-i\Delta k z/2}. \quad (5.96)$$

We will show in the exercises that these amplitudes are related in the following way:

$$\frac{A_2^{\pm*}}{A_1^\pm} = \frac{g_\pm + \alpha_1 + i\Delta k/2}{\kappa_1} = \left[\frac{\kappa_2^*(g_\pm + \alpha_1 + i\Delta k/2)}{\kappa_1(g_\pm + \alpha_2^* - i\Delta k/2)} \right]^{1/2}. \quad (5.97)$$

The last equation shows the symmetry between the field with ω_1 and ω_2 .

5.6.5 Using these result for Raman scattering

Let us now apply what we just derived to stimulated Raman scattering. To this end, we identify $\omega_1 = \omega_S$ and $\omega_2 = \omega_a$. Earlier, we introduced the absorption and coupling coefficients in equation 5.87, and we then introduced expressions for the Stokes and anti-Stokes susceptibilities. In addition, we saw that:

$$\chi_F(\omega_S) = 2\chi_R(\omega_S), \text{ and } \chi_F(\omega_S) = \chi_F(\omega_a)^*. \quad (5.98)$$

Using these relations, one can derive the following:

$$\begin{aligned} \alpha_a &= -\alpha_S^* \left(\frac{n_S \omega_a}{n_a \omega_S} \right), \\ \kappa_S &= -\alpha_S e^{2i\phi_L}, \\ \kappa_a &= \alpha_S^* \left(\frac{n_S \omega_a}{n_a \omega_S} \right) e^{2i\phi_L}. \end{aligned} \quad (5.99)$$

Here, we introduced the phase of the pump field ϕ_L as $A_L = |A_L|e^{i\phi_L}$. During the exercises, we will use these expressions to calculate the Stokes amplitude absorption coefficient α_S and the two gain eigenvalues g_\pm . Using the usually very good approximation that $n_S \omega_L \approx n_a \omega_S$, the expressions for the gain eigenvalues simplifies to:

$$g_\pm = \pm [i\alpha_S \Delta k - (\Delta k/2)^2]. \quad (5.100)$$

If we look at the gain as plotted in Fig. 49, it is interesting to note that the gain vanishes for perfect phase matching ($\Delta k = 0$). The reason for that is that in this case the coupling between the Stokes and anti-Stokes fields is so good, that the generation of the anti-Stokes field prevents the Stokes field from experiencing growth. If the phase

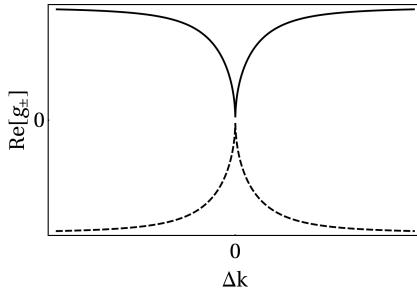


Figure 49: **Coupled gain as a function of phase mismatch.** The figure shows the real parts of g_+ (solid) and g_- (dashed) as a function of Δk .

mismatch becomes very large ($\Delta k \gg |\alpha_S|$), the coupled gain becomes:

$$g_{\pm} = \pm i \frac{\Delta k}{2} \left(1 - \frac{4i\alpha_S}{\Delta k} \right) \approx \pm (\alpha_S + \frac{1}{2}i\Delta k). \quad (5.101)$$

5.7 Applications of Stimulated Broullin and Raman scattering

In the preceding sections of this chapter, we discussed the principled of stimulated Brioullin and Raman scattering, and we investigated how to treat these effects mathematically. Here, we will provide an overview of the applications of these effects. Since the discovery of stimulated Raman and soon after of stimulated Brioullin scattering, the effects have led to a multitude of applications. We will focus on a few examples to provide a taste of the usefulness of these effects.

5.7.1 Applications of stimulated Brioullin scattering

A nice overview of the applications of SBS was given in a review of SBS in celebration of the 50th anniversary of its discovery[10]. We may come across SBS again in alter parts of this lecture where SBS has useful applications. Some of these applications we already heard about. For example, SBS provides another way to implement phase-conjugate mirrors. The example we discussed in more detail with respect to the realization of phase-conjugate mirrors was degenerate four-wave mixing, but SBS provides another possible realization. During the exercises, we also came across another application of SBS: it can be used for pulse shortening. SBS also allows for low-noise signal amplification. The low noise is due to the fact that SBS does not originate from spontaneous emission but it results from thermal vibrations, which in many cases represent the lower limit on noise[10]. Another interesting application is SBS-induced transparency. Usually, one talks about induced transparency in atomic systems, when sufficiently powerful light

drives half the atoms into the excited state such that the absorption at the respective frequency is reduced. SBS can achieve a similar effect in microresonators, but with acoustic excitations (phonons) instead of atomic excitations. SBS also finds applications in the generation of frequency combs. Here, we will only discuss a few selected examples.

Phase conjugation via SBS

In contrast to ordinary mirrors, where the phase of the reflected beam keeps evolving as if there was no mirror, an SBS mirror will shift the phase of the reflected beam by π . That means, the sign of the reflected phase factor reverses. For that reason, the beam seems to propagate exactly backwards as is intended in the case of a phase-conjugate mirror.

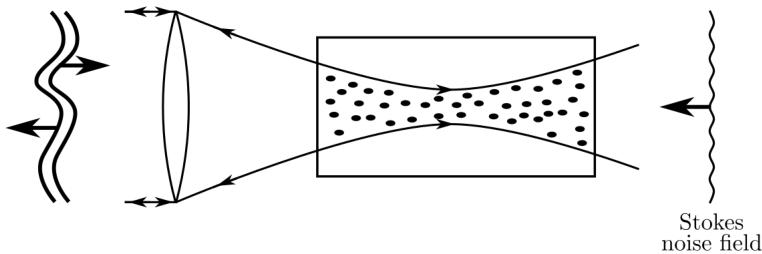


Figure 50: **SBS phase conjugation.** A highly aberrated input beam is focused into our non-linear optical medium. The aberration leads to a speckle pattern and therefore a non-uniform intensity distribution inside the medium. The SBS process amplifies fluctuations in the field at the end of the crystal. The gain of this amplification will be non-uniform as the speckle pattern. The SBS reflected light will therefore have the same wave front as the input light. The figure is based on a figure in Ref. [1].

To be more precise, in SBS the Stokes radiation is emitted in a highly collimated beam that is travelling in the backward direction compared to the incoming light[1]. The reason why SBS leads to the Stokes field being the phase-conjugate of the input field is a bit more subtle than in the case of the phase conjugation due to degenerate four-wave mixing. In particular, the Stokes field amplitude in SBS is driven by a term proportional to $|A_1|^2 A_2$ [1]. From this, it would not seem that the Stokes field should carry any phase information on A_1 and, in particular, not about the conjugate A_1^* .

Why SBS will lead to a phase-conjugate output in the SBS reflected field is illustrated in Fig. 50. In particular, let us assume that the incoming light field is highly aberrated. The non-uniform phase front in the input beam will result in a speckle pattern inside our medium. The SBS reflected light results from the amplification of fluctuations at

the end of the medium. These fluctuations are indicated to the right in Fig. 50. These fluctuations will contain all spatial Fourier components. Those parts of these fluctuations that best match the intensity fluctuations inside the medium will receive the maximum amplification. For that reason, the SBS reflected light will have the same wave front as the incoming light. That means, it will be the phase conjugate of the input field.

Let us make this argument more formal. We know that the intensity I_S of the Stokes field satisfies the following equation:

$$\frac{dI_S}{dz} = -gI_L I_S, \quad (5.102)$$

where I_L is the intensity of the input light. To take into account a non-uniform distribution of the intensity, let us consider the total power at some fixed position z :

$$P_L = \int dA I_L, \text{ and } P_S = \int dA I_S. \quad (5.103)$$

We assume that we integrate over a sufficiently large area to include essentially all the beam power (remember that, e.g., Gaussian beam profiles in principle go to infinity in the transverse direction).

In terms of these powers, we can then rewrite equation 5.102 as:

$$\frac{dP_S}{dz} = -g \frac{P_L P_S}{A} C, \quad (5.104)$$

where $A = \int dA$, and:

$$C = \frac{\langle I_L I_S \rangle}{\langle I_L \rangle \langle I_S \rangle}. \quad (5.105)$$

C represents the spatial cross-correlation function of the Stokes field and the input field. The angular brackets represent the spatial average such that $\langle x \rangle = \int dA x / A$. In this case x stands for $I_{L,S}$ or the product of these intensities.

The important thing to note is the following: if I_L and I_S are completely uncorrelated or if both are constant, then C will be equal to unity. On the other hand, if I_L and I_S are correlated, C can become larger than unity. In contrast to that, let us look at a particular limiting case where the input field is so badly aberrated that the transverse variations in the complex field amplitude obey Gaussian statistics[1]. In this case, one can write down the explicit expressions for the moments of the probability density function for the laser intensity. For the second moment, we get:

$$\langle I^2 \rangle = 2\langle I \rangle^2. \quad (5.106)$$

That means if we have such high aberration, and if the Stokes field is the phase conjugate of the input field, we will have $C = 2$ (**think:** why?). That means that the exponential

gain $G \equiv gP_LCL/A$ experienced by the phase-conjugate portion of the noise field will be twice as high as the gain experienced by other, uncorrelated parts of the noise field. Because the threshold for SBS is on the order of 30 (e.g., we often assumed 25), the phase-conjugate part of the noise field will experience an amplification that is around $\exp(15)$ times higher than the amplification of the uncorrelated parts (**think:** why?).

An interesting observation in this context is that SBS will lead to a better phase-conjugate signal if there is **a lot** of aberration in the incoming light. The factor 2 we got above was for the maximum amount of aberration. If there is no aberration at all, we have a plane wavefront, and we get $C = 1$. In that case, the phase-conjugate parts of the noise field will not experience more amplification than the other parts.

Frequency-comb generation via SBS

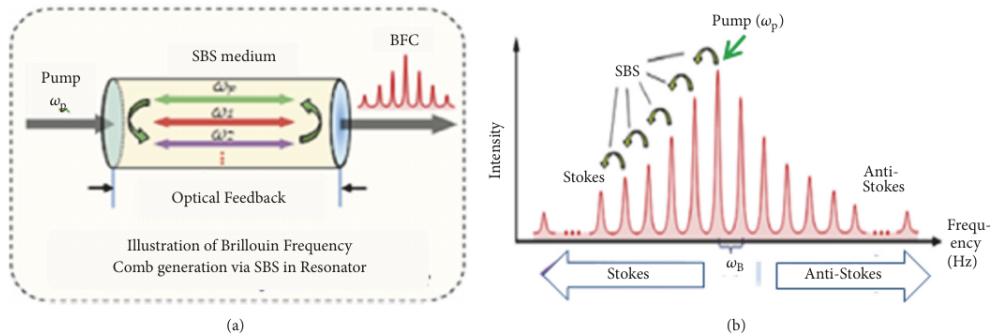


Figure 51: **Frequency comb generation based on SBS.** A laser pumps an SBS medium inside an optical cavity. The figure is from Ref. [10].

5.7.2 Applications of stimulated Raman scattering

Stimulated Raman scattering (SRS) has resulted in a host of applications like biological imaging (e.g., SRS microscopy), Raman fiber amplifiers, Raman lasers, the generation of supercontinuous light, and Raman spectroscopy. Important differences between SBS and SRS are that the bandwidth of SRS Stokes (or anti-Stokes) light has a very high bandwidth, while we saw that the bandwidth in the case of SBS is very narrow. This high bandwidth of SRS is beneficial in several of the applications mentioned before. Raman fiber amplifiers can amplify signals with a much higher bandwidth than many other amplifiers like, e.g., Erbium-doped fiber amplifiers (EDFAs). On the other hand, a drawbacks of Raman fiber amplifiers is that one often needs to overcome a higher power threshold than in the case of EDFAs. In modern-day telecommunication, where high bandwidth is a driving concern, the benefits of the broad bandwidth of a Raman

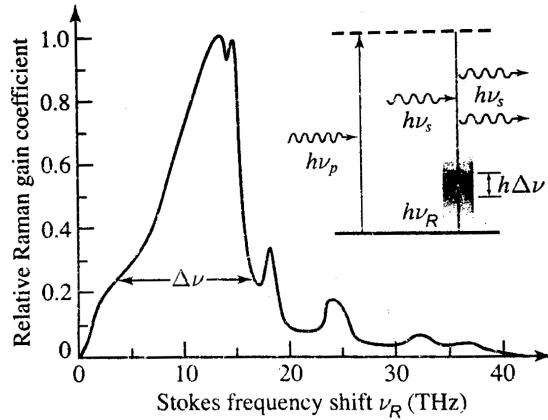


Figure 52: SRS spectrum in a single-mode silica fiber. Typical SRS spectra are very broad because there are many different vibrational levels in the molecules a medium consists of. For that reason, the Stokes and anti-Stokes frequencies $\omega_p \pm \Omega$ can vary over a large bandwidth $\Delta\Omega = 2\pi\Delta\nu$ that is often several THz. The figure is from Ref. [11].

fiber amplifier often overcomes the disadvantage of the need for higher pump powers. The high bandwidth is of course also beneficial in the generation of supercontinuous (high-bandwidth) light. Because the very basis of Raman scattering is the coupling to vibrational levels of the constituents of matter, Raman scattering is an essential component of any spectroscopic analysis of materials. The advantage of SRS is that the scattering of light from particular vibration levels in molecules will be significantly enhanced compared to spontaneous scattering. One can therefore expect a significantly higher signal-to-noise ratio when taking advantage of SRS. As an example, let us take a quick look at SRS microscopy.

Because Raman scattering results from molecular vibrations, optically imaging samples using a microscope in principle allows to resolve differences in the constitution of the imaged material. In particular, one can decide to only look at light that corresponds to the vibrations occurring in very specific molecules one is interested in. The disadvantage is that if one wants to image biological samples, the optical power one can use without destroying the sample is very limited. Combining sufficiently low pump power with the low probability of spontaneous Raman scattering means that it can be very difficult to resolve the optical signal of interest from background noise. Stimulated Raman scattering offers the possibility to achieve significantly stronger optical signals, and in recent years SRS microscopy has attracted significant interest. As an example, we will consider one of the early experiments introducing this technique in 2008 [12]. In that work, the authors achieved 3D imaging of biological tissue using SRS microscopy. Illustrations from their paper are shown in Fig. 53.

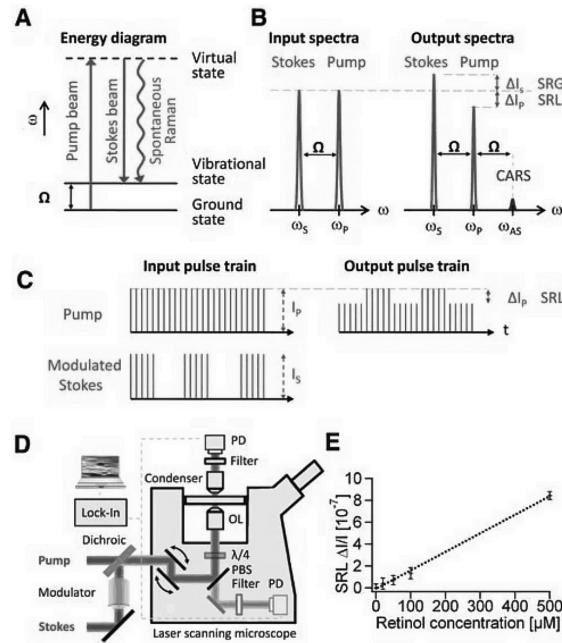


Figure 53: **SRS spectrum in a single-mode silica fiber.** Picosecond pulses at the pump frequency ω_p and the Stokes frequency ω_S illuminate the sample. The Stokes beam is intensity modulated such that a Lock-in amplifier can be used to improve the signal-to-noise ratio. Stimulated Raman scattering will lead to a “Stimulated Raman gain”(SRG) at the Stokes frequency and a corresponding “stimulated Raman loss”(SRL) at the laser frequency. The figure is from Ref. [12].

The central idea of the method is that one does not only illuminate the sample with the pump light at frequency ω_p , but one also shines light at the Stokes frequency ω_S onto the sample. This will lead to stimulated Raman scattering for any molecular transition with a frequency $\Omega = \omega_p - \omega_S$. Because the resulting signal is still much weaker than the pump light and the incident Stokes light, the authors modulated the intensity of the Stokes light with a high frequency of 1.7 MHz. The reason for the choice of this frequency was that the laser noise at such high frequencies becomes much lower compared to low frequencies. One can then use a lock-in amplifier for detecting the signal. That means, one essentially only amplifies those components with a frequency close to the modulation frequency.

6 Electromagnetically induced transparency and the two-level approximation

Electromagnetically induced transparency (EIT) is an interesting technique with applications, e.g., in atomic and optomechanical systems. EIT is also a promising candidate for quantum memories and for the storage of light. The effect allows to render a medium transparent to some resonant laser radiation while the material will still exhibit the non-linear optical properties associated with the resonant response of the material. EIT is also the first example where we encounter a saturable absorber. This is significant because the classical treatment with field and amplitudes can break down in this case. That means, we will have to use a quantum description in order to capture all the characteristics of EIT. We will discuss that in the context of the two-level approximation of the interaction between an optical field and an optical medium.

6.1 Electromagnetically induced transparency

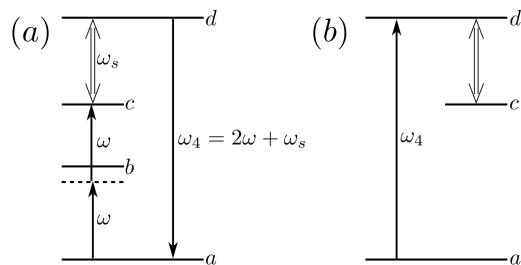


Figure 54: **Typical EIT energy-level diagram.** Fields with frequencies ω and ω_s are applied to (e.g.) atoms with the indicated level structure to generate a field with frequency $\omega_4 = 2\omega + \omega_s$. EIT allows to reduce or to eliminate absorption at the $a \rightarrow d$ transition frequency. The figure was adapted from Ref. [1].

Fig. 54 illustrates a typical situation in EIT. Let us for now assume that we have a medium (e.g.. a gas of atoms) with the level structure indicated. We pump the

medium with two frequencies ω and ω_s . The goal is to generate a field at the sum frequency $\omega_4 = \omega + \omega_s$ via the process shown in Fig. 54 (a). Under normal circumstances, absorption due to the transition $d \rightarrow a$ would significantly limit the emission of light at the sum frequency ω_4 . Using EIT, we can reduce or eliminate such transitions and therefore significantly reduce the absorption at ω_4 . In this way, we can use our medium for SFG despite this absorption we would usually expect. We will first discuss how we can reduce or eliminate the absorption due to some particular resonance, and then we will come back to our example for SFG.

To describe this process, let us look at the wavefunction describing our atomic state in the interaction picture, which we can use to describe the time evolution of states in the presence of interaction. For comparison, in the Schrödinger picture, the evolution of a physical system is described by a time-dependent state $|\psi(t)\rangle_s$ that fulfills the Schrödinger equation:

$$\frac{d}{dt}|\psi(t)\rangle_s = -\frac{i}{\hbar}\hat{H}_s|\psi(t)\rangle_s, \quad (6.1)$$

where the index s indicates the Schrödinger picture. In the Heisenberg picture, the evolution of the system is described by time-varying operators instead. For some operator \hat{A}_h (h indicates the Heisenberg picture), we have:

$$\frac{d}{dt}\hat{A}_h = -\frac{i}{\hbar}[\hat{A}_h, \hat{H}_h]. \quad (6.2)$$

The state is constant: $|\psi\rangle_h = |\psi(0)\rangle_h = |\psi(0)\rangle_s$. The operators in the Heisenberg picture are related to the ones in the Schrödinger picture via the relation:

$$\hat{A}_h = U^\dagger(t, t_0)\hat{A}_s U(t, t_0), \quad (6.3)$$

where the unitary U results from integrating the Schrödinger equation:

$$U(t, t_0) = \int dt' e^{-\frac{i}{\hbar}\hat{H}_s t'}. \quad (6.4)$$

The interaction picture is useful if the Hamiltonian consists of a non-interacting part \hat{H}_0 and an interaction part \hat{H}_i . In this representation, the state evolves as:

$$\frac{d}{dt}|\psi(t)\rangle_i = -\frac{i}{\hbar}\hat{H}_i|\psi(t)\rangle_i, \quad (6.5)$$

and the interaction Hamiltonian evolves according to the equation:

$$\hat{H}_i(t) = e^{\frac{i}{\hbar}\hat{H}_0 t}\hat{H}_i(0)e^{-\frac{i}{\hbar}\hat{H}_0 t}. \quad (6.6)$$

Operators evolve according to:

$$\frac{d}{dt}\hat{A}_i = -\frac{i}{\hbar}[\hat{A}_i, \hat{H}_i]. \quad (6.7)$$

The advantage of the interaction picture is that one can describe the time evolution of the states and the operators using only the interaction Hamiltonian instead of having to always take into account the full Hamiltonian.

Now, let us get back to EIT. We will assume the interaction picture unless we say otherwise and drop the index i from now on. We can write our atomic state for the level diagram in Fig. 54 as[1]:

$$\psi(\vec{r}, t) = C_a(t)u_a(\vec{r}) + C_d(t)u_d(\vec{r})e^{-i\omega_4 t} + C_c(t)u_c(\vec{r})e^{-i(\omega_4 - \omega_s)t}. \quad (6.8)$$

That means our atomic state is a superposition of energy eigenstates $u_{a,d,c}$. We write our Hamiltonian as $\hat{H} = \hat{H}_0 + \hat{V}$, where \hat{V} represents the interaction energy. The eigenstates fulfill $\hat{H}_0 u_j = \hbar\omega_j u_j$. The operator \hat{V} can be written as¹:

$$\hat{V} = -\hat{\mu}(E_4 e^{-i\omega_4 t} + E_s^* e^{i\omega_s t}). \quad (6.9)$$

We require that the state of our atomic system fulfills the Schrödinger equation:

$$\frac{d}{dt}|\psi(t)\rangle = -\frac{i}{\hbar}\hat{H}|\psi(t)\rangle. \quad (6.10)$$

In one of the exercises we will show that if we put our state into this equation, and if we assume that the eigenstates $u_{a,d,c}$ are orthonormal functions, we can obtain the following three equations[1]:

$$\begin{aligned} i\hbar\dot{C}_a &= \hbar\omega_a C_a + V_{ad}C_d e^{-i\omega_4 t}, \\ i\hbar[\dot{C}_d e^{-i\omega_4 t} - i\omega_4 C_d e^{-i\omega_4 t}] &= \hbar\omega_d C_d e^{-i\omega_4 t} + V_{da}C_a + V_{dc}C_c e^{-i(\omega_4 - \omega_s)t}, \\ i\hbar[\dot{C}_c e^{-i(\omega_4 - \omega_s)t} - i(\omega_4 - \omega_s)C_c e^{-i(\omega_4 - \omega_s)t}] &= \hbar\omega_c C_c e^{-i(\omega_4 - \omega_s)t} + V_{cd}C_d e^{-i\omega_4 t}. \end{aligned} \quad (6.11)$$

For simplification, let us introduce the explicit forms of the matrix elements of \hat{V} :

$$\begin{aligned} V_{ad}^* &= V_{da} = -\mu_{da}E_4 e^{-i\omega_4 t} \\ V_{dc}^* &= V_{cd} = -\mu_{cd}E_s^* e^{i\omega_s t}. \end{aligned}$$

¹I will not go into details here, but this expression assumes the rotating-wave approximation and the electric-dipole approximation[1].

Let us measure our energies with respect to the ground-state of our atoms, such that[1]:

$$\hbar\omega_a \rightarrow 0, \hbar\omega_d \rightarrow \omega_{da}, \quad \hbar\omega_c \rightarrow \omega_{ca}. \quad (6.12)$$

If we introduce the Rabi frequencies $\Omega = \mu_{da}E_4/\hbar$ and $\Omega_s^* = \mu_{cd}E_s^*/\hbar$, the equations 6.11 become:

$$\begin{aligned} i\hbar\dot{C}_a &= iC_d\Omega^*, \\ \dot{C}_d - i\delta C_d &= iC_a\Omega + iC_c\Omega_s, \\ \dot{C}_c - i(\delta - \Delta)C_c &= iC_d\Omega_s^*. \end{aligned} \quad (6.13)$$

Here, we used the definitions $\delta \equiv \omega - \omega_{da}$ and $\Delta \equiv \omega_s - \omega_{dc}$ [1].

Our goal is to solve these equations to lowest order in Ω . To this end, we introduce a (real) strength parameter λ and replace Ω with $\lambda\Omega$. We also expand the amplitudes of our atomic state into power series in λ :

$$C_j = \sum_{l=0}^{\infty} \lambda^l C_j^{(l)}. \quad (6.14)$$

If we apply these changes to the equations 6.13, and if we take into account that the resulting equation have to be fulfilled for any value of λ , we get new sets of equations for each order of λ . During the exercises, we will derive the resulting equations for λ^0 and for λ^1 . In particular, the 0th order will result in the following equations:

$$\begin{aligned} \dot{C}_a^{(0)} &= 0, \\ \dot{C}_d^{(0)} - i\delta C_d^{(0)} &= iC_a^{(0)}\Omega + iC_c^{(0)}\Omega_s, \\ \dot{C}_c^{(0)} - i(\delta - \Delta)C_c^{(0)} &= iC_d^{(0)}\Omega_s^*. \end{aligned} \quad (6.15)$$

We will assume that the solution of these equations will correspond to fulfilling the initial conditions (boundary conditions). This means:

$$C_a^{(0)} = 1 \text{ and } C_d^{(0)} = C_c^{(0)} = 0 \quad (6.16)$$

for all times. In other words, we start with our atoms in the ground state.

For the order λ^1 , we will show that the following equations need to hold:

$$\begin{aligned} \dot{C}_a^{(1)} &= 0, \\ \dot{C}_d^{(1)} - i\delta C_d^{(1)} &= i\Omega + iC_c^{(1)}\Omega_s, \\ \dot{C}_c^{(1)} - i(\delta - \Delta)C_c^{(1)} &= iC_d^{(1)}\Omega_s^*. \end{aligned} \quad (6.17)$$

We see that $C_a^{(1)}$ is constant, and we set $C_a^{(1)} = 0$ [1]. If we drop the superscript (1) for now, the other equations become:

$$\begin{aligned}\dot{C}_d - i\delta C_d &= i\Omega + iC_c \Omega_s, \\ \dot{C}_c - i(\delta - \Delta)C_c &= iC_d \Omega_s^*.\end{aligned}\quad (6.18)$$

For the steady-state solution, we can set the time-derivatives equal to zero and get:

$$\begin{aligned}0 &= \Omega + \delta C_d + \Omega_s C_c, \\ 0 &= \Omega_s^2 C_d + (\delta - \Delta)C_c.\end{aligned}\quad (6.19)$$

One can algebraically solve these equations to get an expression for C_d :

$$C_d = \frac{\Omega(\delta - \Delta)}{|\Omega_s|^2 - \delta(\delta - \Delta)}. \quad (6.20)$$

We are mostly interested in the induced dipole moment because its various matrix elements will determine the linear and nonlinear susceptibilities.

The expectation value of the dipole moment to first order in λ is (**think:** why?):

$$\begin{aligned}\tilde{p} &= \langle \psi | \hat{\mu} | \psi \rangle = \langle \psi^{(0)} | \hat{\mu} | \psi^{(1)} \rangle + \langle \psi^{(1)} | \hat{\mu} | \psi^{(0)} \rangle = \\ &= \langle a | \hat{\mu} | d \rangle C_d e^{-i\omega_4 t} + \text{c.c.} = \mu_{ad} C_d e^{-i\omega_4 t} + \text{c.c.}\end{aligned}\quad (6.21)$$

Here, the tilde symbolizes that the expression included quickly oscillating factors like $e^{i\omega_4 t}$. The slowly varying amplitude is p , and the overall polarization is:

$$P = Np = \epsilon_0 \chi^{(1)} E, \quad (6.22)$$

where N is the number density of our atomic gas. We can conclude that:

$$\chi^{(1)} = \frac{N |\mu_{da}|^2}{\epsilon_0 \hbar} \frac{(\delta - \Delta)}{|\Omega_s|^2 - (\delta - \Delta)\delta}. \quad (6.23)$$

To include damping, we can replace δ with $\delta + i\gamma_d$ and Δ with $\Delta - i(\gamma_c - \gamma_d)$ [1]. The linear susceptibility then becomes:

$$\chi^{(1)} = \frac{N}{\epsilon_0 \hbar} \frac{|\mu_{da}|^2 (\delta - \Delta + i\gamma_c)}{|\Omega_s|^2 - (\delta + i\gamma_d)(\delta - \Delta + i\gamma_c)}. \quad (6.24)$$

In the case, where both fields are exactly on resonance ($\Delta = \delta = 0$), the linear susceptibility becomes:

$$\chi^{(1)} = \frac{N}{\epsilon_0 \hbar} \frac{|\mu_{da}|^2 i\gamma_c}{|\Omega_s|^2 + \gamma_c \gamma_d}. \quad (6.25)$$

This shows that the linear susceptibility is purely imaginary. That means there is only absorption, but the important thing is that we can make this absorptive susceptibility arbitrarily small by saturating the field intensity $|\Omega_s|^2$.

We promised earlier, that we will first discuss how EIT can be used to reduce or eliminate absorption at some resonant frequency, and that we will then show that we can realize EIT but still see nonlinear effects like SFG. To this end, let us now use similar considerations as above to calculate an expression for the third-order susceptibility $\chi^{(3)}$ in SFG. We will again treat the evolution of the atomic state in the interaction picture. The wavefunction in the context of SFG can then be written as:

$$\begin{aligned}\psi(\vec{r}, t) = & C_a(t)u_a(\vec{r}) + C_b(t)u_b(\vec{r})e^{-i\omega t} \\ & + C_c(t)u_c(\vec{r})e^{-i2\omega t} + C_d(t)u_d(\vec{r})e^{-i(2\omega+\omega_c)t}.\end{aligned}\quad (6.26)$$

This wavefunction will also need to satisfy the Schrödinger equation 6.10, and if we assume that the energy eigenfunctions from an orthonormal set, this will again lead to a set of equations we will derive in one of the exercises. We can again simplify these equations by inserting the following expressions for the matrix elements of our interaction Hamiltonian:

$$\begin{aligned}V_{ba} = V_{ab}^* &= -\mu_{ba}Ee^{-i\omega t} = -\hbar\Omega_{ba}e^{-i\omega t} \\ V_{cb} = V_{bc}^* &= -\mu_{cb}Ee^{-i\omega t} = -\hbar\Omega_{cb}e^{-i\omega t} \\ V_{dc} = V_{cd}^* &= -\mu_{dc}E_s e^{-i\omega_s t} = -\hbar\Omega_{dc}e^{-i\omega_s t},\end{aligned}\quad (6.27)$$

and by introducing the detuning factors:

$$\delta_1 = \omega - \omega_{ba}, \delta_2 = 2\omega - \omega_{ca} \quad \Delta = \omega_s - \omega_{dc}. \quad (6.28)$$

The simplified equations then will be:

$$\dot{C}_a = iC_b\Omega_{ba}^*, \quad (6.29)$$

$$\dot{C}_b - iC_b\delta_1 = iC_a\Omega_{ba} + iC_c\Omega_{cb}^*, \quad (6.30)$$

$$\dot{C}_c - iC_c\delta_2 = iC_a\Omega_{cb} + iC_d\Omega_{dc}^*, \quad (6.31)$$

$$\dot{C}_d - iC_d(\delta_2 + \delta) = iC_c\Omega_{dc}. \quad (6.32)$$

Our goal is to solve this system of equations perturbatively in Ω_{ba} and Ω_{cb} in a similar way we did earlier in the context of deriving $\chi^{(1)}$ for EIT. The first simplification we can make is by noticing that we can simply ignore equation 6.29 because we can assume that the occupation of level b will always be much smaller than the occupation of the ground level a . That means $|C_b| \ll |C_a|$. Then the right-hand side of this equation is essentially zero, and we can assume $C_a \approx 1$ as before. In equation 6.30, we can drop the last term because we can assume that $|C_c| \ll |C_b|$. If we set $C_a = 1$ and $\dot{C}_b = 0$ for the

steady-state solution, we get the following relation:

$$C_b = -\Omega_{ba}/\delta_1. \quad (6.33)$$

Since we are interested in the steady-state solution, we can then set the time derivatives zero in equations 6.31 and 6.32. This yields:

$$\begin{aligned} -C_c &= \frac{C_b \Omega_{cb}}{\delta_2} + \frac{C_d \Omega_{dc}^*}{\delta_2}, \\ -C_d &= \frac{C_c \Omega_{dc}}{\delta_2 + \Delta}. \end{aligned} \quad (6.34)$$

One of the exercises will be to use these equations in combination with equation 6.33 to get an expression for C_d .

We can then use that to calculate the induced dipole moment:

$$\tilde{p} = \langle \psi | \hat{\mu} | \psi \rangle = \mu_{ad} C_d + \text{c.c.} \quad (6.35)$$

The slowly varying part of this then becomes:

$$\begin{aligned} p &= \frac{-\mu_{ad} \Omega_{dc} \Omega_{cb} \Omega_{ba}}{\delta_1 [\delta_2(\delta_2 + \Delta) - |\Omega_{dc}|^2]} = \frac{-\mu_{ad} \mu_{dc} \mu_{cb} \mu_{ba} E^2 E_s}{\hbar^3 \delta_1 [\delta_2(\delta_2 + \Delta) - |\Omega_{dc}|^2]} \\ &\equiv \frac{3\epsilon_0 \chi^{(3)} E^2 E_c}{N}. \end{aligned} \quad (6.36)$$

From this, we can conclude that:

$$\chi^{(3)} = \frac{-N \mu_{ad} \mu_{dc} \mu_{cb} \mu_{ba}}{3\epsilon_0 \hbar \delta_1 [\delta_2(\delta_2 + \Delta) - |\Omega_{dc}|^2]}. \quad (6.37)$$

In the limit $|\Omega_{dc}| \rightarrow 0$ this expression converges against the usual expression one gets for the resonant contribution to the third-order susceptibility. We can include the effects of damping by replacing δ_2 with $\delta_2 + i\gamma_c$, and by replacing $\delta_2 + \Delta$ with $\delta_2 + \Delta + i\gamma_d$.

Figure 55 illustrates the behaviour of $\chi^{(1)}$ and $\chi^{(3)}$ as functions of frequency. Panels (a) and (b) of that figure show that we can get a nearly perfect suppression of absorption near the center frequency - for strong and weak saturating fields Ω_s . In panels (c) and (d), we see that in both cases we still have a significant non-zero nonlinearity for SFG. This means, we can suppress the absorption at a specific frequency and yet still harness the nonlinearity at that frequency.

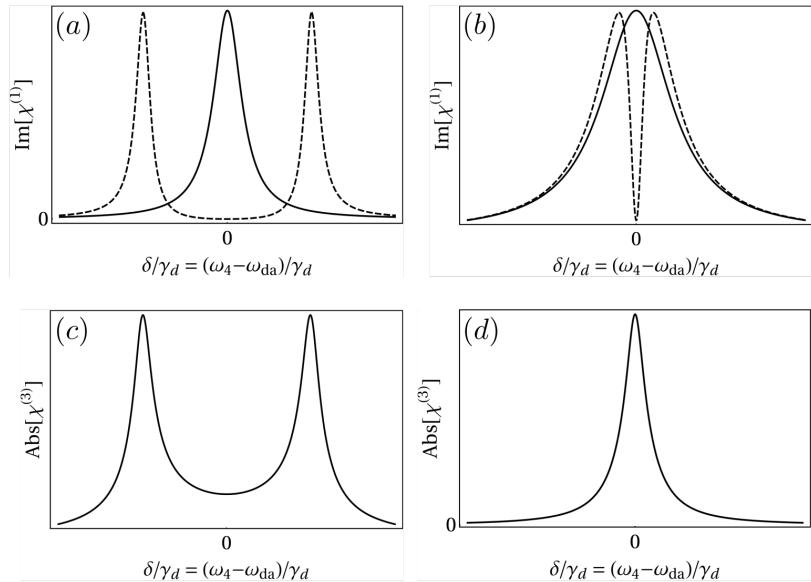


Figure 55: Examples of predicted susceptibilities. Panels (a) and (b) show the behaviour of the linear susceptibility $\chi^{(1)}$ for two cases of Ω_s that represent a strong and a weak saturating field, respectively (dashed lines). For comparison, the solid line indicates the absorption in the absence of a saturating field (Ω_s). These figures indicated that one can nearly completely suppress absorption at the center frequency. The lower panels (c) and (d) show the third-order susceptibility $\chi^{(3)}$, which leads to SFG. In both cases, the non-linear susceptibility can have significant, non-vanishing values at the center frequency. The figure is based on Ref. [1].

6.2 The limits of the perturbative treatment

In the beginning of this chapter, I indicated that the classical approach we used so far to describe nonlinear effects will break down under some circumstances, and that EIT is a good example for that. This is also why we treated EIT using a quantum description. Before that, we have assumed that we can write the nonlinear polarization in a power-series expansion like this:

$$P(t) = \epsilon_0 \chi^{(1)} E(t) + \epsilon_0 \chi^{(2)} E(t)^2 + \epsilon_0 \chi^{(3)} E(t)^3 + \dots \quad (6.38)$$

When this expansion does not converge, the perturbative treatment we have been using fails. This is the case for a saturable absorber, where the absorption coefficient α depends on the field intensity via the relation:

$$\alpha = \frac{\alpha_0}{1 + I/I_s}, \quad (6.39)$$

where α_0 is the absorption coefficient for weak fields, and I_s is the saturation intensity. If we expand the expression for α as a power series, we get:

$$\alpha = \alpha_0 [1 - (I/I_s) + (I/I_s)^2 - (I/I_s)^3 + \dots]. \quad (6.40)$$

This expansion will only converge as long as $I < I_s$. That means, we can only describe this effect in terms of equation 6.38 if $I < I_s$.

We will typically encounter this problem if a transition in our medium is resonantly excited. But even in these cases, one does not need to treat each full atom (or molecule) quantum mechanically, but it is often sufficient if one focuses on the two levels that are involved in such a resonant excitation. Of course, it is also possible to describe more complicated situations quantum mechanically, but for simplicity we will concentrate on this two-level case. It will allow us to introduce many concepts that are also useful in a more general quantum treatment of (nonlinear) optical effects.

6.3 Density-matrix equations for a two-level atom

Let us first discuss the density-matrix equations of motion for a two-level atom in the absence of any damping effects. We will assume a level structure as illustrated in Fig. 56, and that the Hamiltonian of this two-level system is given by:

$$\hat{H} = \hat{H}_0 + \hat{V}(t). \quad (6.41)$$

Here, \hat{H}_0 is the atomic Hamiltonian in the absence of an electromagnetic field, and $\hat{V}(t)$ corresponds to the energy of the interaction of the atom with the electromagnetic field.

We assume the two energy levels to be a and b , and the eigenvalues of \hat{H}_0 are:

$$E_a = \hbar\omega_a, \text{ and} \quad E_b = \hbar\omega_b. \quad (6.42)$$

Then the Hamiltonian can be represented by a diagonal matrix with elements $H_{0,nm} = E_n \delta_{nm}$. If we can describe the interaction using the electric dipole approximation, we can write the interaction Hamiltonian as:

$$\hat{V}(t) = -\hat{\mu}E(t). \quad (6.43)$$

Let us also assume that the two energy levels have a definite parity such that the diagonal matrix elements of $\hat{\mu}$ vanish, which will also cause the diagonal elements of the interaction Hamiltonian to vanish:

$$\mu_{aa} = \mu_{bb} = V_{aa} = V_{bb} = 0. \quad (6.44)$$

The only non-vanishing matrix elements of the interaction Hamiltonian then are:

$$V_{ba} = V_{ab}^* = -\mu_{ba}E(t). \quad (6.45)$$

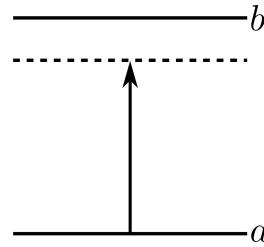


Figure 56: **Two-level system.** Our two levels a and b , and a near-resonant virtual level that is $\hbar\omega$ above the lower energy level a . The figure is adapted from Ref. [1].

We can describe the evolution of our two-level system by using a density matrix $\hat{\rho}$ with 2×2 elements:

$$\hat{\rho} = \begin{bmatrix} \rho_{aa} & \rho_{ab} \\ \rho_{ba} & \rho_{bb} \end{bmatrix}. \quad (6.46)$$

Because the time evolution of the density matrix is given by:

$$\frac{d}{dt}\hat{\rho} = -\frac{i}{\hbar} [\hat{H}, \hat{\rho}], \quad (6.47)$$

we get the following equations of motion for the elements of the density matrix (**think:**

why only three?):

$$\begin{aligned}\dot{\rho}_{ba} &= -i\omega_{ba}\rho_{ba} + \frac{i}{\hbar}V_{ba}(\rho_{bb} - \rho_{aa}), \\ \dot{\rho}_{bb} &= \frac{i}{\hbar}(V_{ba}\rho_{ab} - \rho_{ba}V_{ab}), \\ \dot{\rho}_{aa} &= \frac{i}{\hbar}(V_{ab}\rho_{ba} - \rho_{ab}V_{ba}),\end{aligned}\tag{6.48}$$

where we introduced the transition frequency $\omega_{nm} = (E_n - E_m)/\hbar$. By combining these equations, we see that $\dot{\rho}_{aa} + \dot{\rho}_{bb} = 0$, and we know from the properties of density matrices that $\rho_{aa} + \rho_{bb} = 1$.

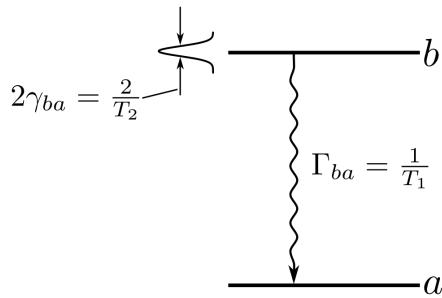


Figure 57: **Relaxation processes in a closed two-level system.** The decay rate from level b to level a is Γ_{ba} . The dipole moment can dephase on a time-scale T_2 , corresponding to a characteristic width γ_{ba} . The figure is adapted from Ref. [1].

The equations 6.48 describe the evolution of the density matrix well in the absence of relaxation processes or if they can be neglected. An example is if the exciting light is pulsed with a pulse-width much shorter than the relevant relaxation time scales. In the following, we will discuss various relaxation processes in two distinct cases: the case of a “closed two-level atom” and the case of an “open two-level atom.”

6.3.1 Closed two-level atom

We speak of a closed two-level atom if the only possible transitions are between the levels a and b . The situation is illustrated in Fig. 57. We denote the decay rate with which the atom relaxes from the excited state b into the ground state a as $\Gamma_{ba} = T_1^{-1}$. T_1 is the lifetime of the upper level. In addition to that, the dipole moment will dephase in a characteristic time T_2 , which is related to a characteristic width $\gamma_{ba} = T_2^{-1}$.

These relaxation processes can be described by adding decay terms to our equations

6.48:

$$\begin{aligned}\dot{\rho}_{ba} &= -(i\omega_{ba} + T_2^{-1}) \rho_{ba} + \frac{i}{\hbar} V_{ba} (\rho_{bb} - \rho_{aa}), \\ \dot{\rho}_{bb} &= -\frac{\rho_{bb}}{T_1} - \frac{i}{\hbar} (V_{ba} \rho_{ab} - \rho_{ba} V_{ab}), \\ \dot{\rho}_{aa} &= \frac{\rho_{bb}}{T_1} + \frac{i}{\hbar} (V_{ab} \rho_{ba} - \rho_{ab} V_{ba}),\end{aligned}\tag{6.49}$$

In the exercises, we will solve these equations when no field is present ($V_{ba} = 0$) and show that a population inversion [$\rho_{bb}(t) - \rho_{aa}(t)$] will relax from its initial value [$\rho_{bb}(0) - \rho_{aa}(0)$] to its equilibrium value [$\rho_{bb} - \rho_{aa}$]^(eq) in a time of the order of T_1 . The time T_1 is therefore called the population relaxation time[1]. Similarly, we will show that the induced dipole moment evolves as follows:

$$\langle \mu(t) \rangle = [\mu_{ab} \rho_{ba}(0) e^{-\omega_{ba} t} + \text{c.c.}] e^{-t/T_2}.\tag{6.50}$$

This shows that the dipole moment oscillates with the frequency ω_{ba} , and that the amplitude of the induced dipole moment decays in the characteristic time T_2 . This time is known as the dipole dephasing time.

The times T_1 and T_2 are not completely independent. That is to say that the damping rates for the off-diagonal elements of the density matrix (T_2) is not completely independent from the damping rate for the diagonal elements of the density matrix (T_1). In particular, they are related via the following equation:

$$\frac{1}{T_2} = \frac{1}{2T_1} + \gamma_c.\tag{6.51}$$

If our medium is an atomic vapour, γ_c can be described very accurately via the relation:

$$\gamma_c = C_s N + C_f N_f,\tag{6.52}$$

where N is the number density of atoms with the resonance frequency ω_{ba} , and N_f is the number density of “foreign” atoms of a different atomic species that have a different resonance frequency. C_s describes self-broadening of our linewidth, and C_f describes foreign-gas broadening - that means, broadening due to collisions with gas molecules of a different kind.

6.3.2 Open two-level atom

Of course, a closed two-level system can only provide a rough approximation to the response of a real atom because any atom will typically contain significantly more than two levels, for example some magnetic sublevels or hyperfine levels. The transitions to

some of these levels may also not be accompanied by the exchange of an optical photon with the environment. For example, energy levels might be coupled via phonons or other means. In order to treat such situations, let us consider a more general model, which we will denote as an “open two-level atom”. This model is often used in laser theory. We call it “open” because the model will not only allow transitions between the two levels we consider but also exchange of energy with reservoir levels. We will assume that our two levels a and b acquire population due to some pump rates $\lambda_{a,b}$. The situation is schematically illustrated in Fig. 58.

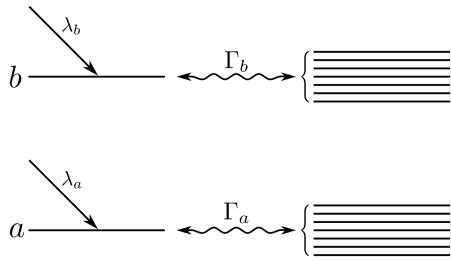


Figure 58: **Relaxation in an open two-level atom.** The population in levels a and b is pumped with rates λ_a and λ_b , respectively. The two levels can relax to reservoir levels with decay rates $\Gamma_{a,b}$. The figure is adapted from Ref. [1].

In order to describe the pumping of the two levels and the relaxation into reservoir levels, we need to adapt our differential equations 6.48. These become:

$$\begin{aligned}\dot{\rho}_{ba} &= - \left(i\omega_{ba} + \frac{1}{T_2} \right) \rho_{ba} + \frac{i}{\hbar} V_{ba} (\rho_{bb} - \rho_{aa}), \\ \dot{\rho}_{bb} &= \lambda_b - \Gamma_b (\rho_{bb} - \rho_{aa})^{\text{eq}} - \frac{i}{\hbar} (V_{ba} \rho_{ab} - \rho_{ba} V_{ab}), \\ \dot{\rho}_{aa} &= \lambda_a - \Gamma_a (\rho_{aa} - \rho_{bb})^{\text{eq}} + \frac{i}{\hbar} (V_{ab} \rho_{ba} - \rho_{ab} V_{ba}).\end{aligned}\quad (6.53)$$

In contrast to the closed two-level system, the total population in levels a and b is not conserved in this case. Therefore, one typically has to take into account all three of these equations. The relaxation rates are related to the collisional dephasing rate γ_c and the population rates $\Gamma_{a,b}$ via the equation:

$$\frac{1}{T_2} = \frac{1}{2} (\Gamma_b + \Gamma_a) + \gamma_c. \quad (6.54)$$

6.3.3 A two-level atom non-radiatively coupled to a third level

The closed two-level approach sometimes is adapted to include non-radiative coupling to a third level. For example, this is useful in describing saturable absorbers. A schematic of the corresponding level diagram is shown in Fig. 59. Population in the excited level decays at a total rate $\Gamma_{ba} + \Gamma_{bc}$. That means the population can decay to level c with rate Γ_{bc} or to level a with rate Γ_{ba} . The intermediate level c acts as a “trap level”[1]. In addition, any dipole moment associated with a transition between a and b is damped at a rate γ_{ba} . One can take these modifications to the two-level system into account by modifying the equations 6.48 as follows:

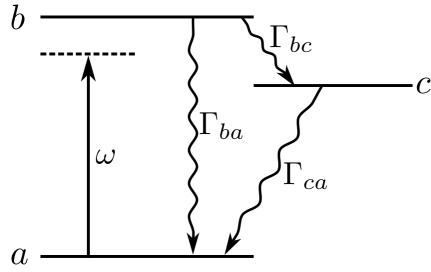


Figure 59: **Relaxation for a two-level atom non-radiatively coupled to a third level.** States in the excited level b can either decay directly to a or first to c and from there to a . The figure is adapted from Ref. [1].

$$\begin{aligned}\dot{\rho}_{ba} &= -(i\omega_{ba} + \gamma_{ba})\rho_{ba} + \frac{i}{\hbar}V_{ba}(\rho_{bb} - \rho_{aa}), \\ \dot{\rho}_{bb} &= -(\Gamma_{ba} + \Gamma_{bc})\rho_{bb} - \frac{i}{\hbar}(V_{ba}\rho_{ab} - \rho_{ba}V_{ab}), \\ \dot{\rho}_{cc} &= \Gamma_{bc}\rho_{bb} - \Gamma_{ca}\rho_{cc}, \\ \dot{\rho}_{aa} &= \Gamma_{ba}\rho_{bb} + \Gamma_{ca}\rho_{cc} + \frac{i}{\hbar}(V_{ba}\rho_{ab} - \rho_{ba}V_{ab}).\end{aligned}\quad (6.55)$$

By adding the three equations, one can see that the total population number is conserved:

$$\dot{\rho}_{aa} + \dot{\rho}_{bb} + \dot{\rho}_{cc} = 0. \quad (6.56)$$

6.4 Steady-state response of a two-level atom to a monochromatic field

Let us consider the case where two-level atoms interact with a monochromatic, steady-state field. Our goal will be to derive the corresponding susceptibility. For simplicity, we will treat our atoms as closed two-level systems. The equations of motion for our density matrix will be given by:

$$\begin{aligned}\dot{\rho}_{ba} &= -\left(i\omega_{ba} + \frac{1}{T_2}\right)\rho_{ba} + \frac{i}{\hbar}V_{ba}(\rho_{bb} - \rho_{aa}), \\ \dot{\rho}_{bb} - \dot{\rho}_{aa} &= -\frac{(\rho_{bb} - \rho_{aa}) - (\rho_{bb} - \rho_{aa})^{(\text{eq})}}{T_1} - \frac{2i}{\hbar}(V_{ba}\rho_{ab} - \rho_{ba}V_{ab}).\end{aligned}\quad (6.57)$$

In the electric dipole approximation, where the wavelength of the electromagnetic radiation is much larger than the size of the electron wavefunctions, the interaction Hamiltonian for the application of a monochromatic plane wave with frequency ω is:

$$\hat{V} = -\hat{\mu}E(t) = -\hat{\mu}(Ee^{-i\omega t} + E^*e^{i\omega t}), \quad (6.58)$$

and the corresponding matrix elements are given by:

$$V_{ba} = -\mu_{ba}(Ee^{-i\omega t} + E^*e^{i\omega t}). \quad (6.59)$$

While one cannot solve equations 6.57 exactly under these conditions, one can do so if one assumes the rotating wave approximation (RWA). In particular, in the absence of a driving field, ρ_{ba} will evolve in time as $\exp(-i\omega_{ba}t)$. If ω is close to ω_{ba} , then the part of V_{ba} evolving as $e^{-i\omega t}$ will act as a much more effective driving term than the part of V_{ba} that evolves as $e^{i\omega t}$. Loosely speaking, the difference frequency in the latter case is approximately $2\omega_{ba}$, and this term will oscillate so quickly that it will effectively average to zero. For the other term, the difference frequency will be very low, and we get a more significant contribution.

For that reason, in the RWA, we assume that we can write:

$$V_{ba} = -\mu_{ba}Ee^{-i\omega t}. \quad (6.60)$$

Using this approximation, our differential equations for the density matrix 6.57 become:

$$\begin{aligned}\dot{\rho}_{ba} &= -\left(i\omega_{ba} + \frac{1}{T_2}\right)\rho_{ba} - \frac{i}{\hbar}\mu_{ba}Ee^{-i\omega t}(\rho_{bb} - \rho_{aa}), \\ \dot{\rho}_{bb} - \dot{\rho}_{aa} &= -\frac{(\rho_{bb} - \rho_{aa}) - (\rho_{bb} - \rho_{aa})^{(\text{eq})}}{T_1} + \frac{2i}{\hbar}(\mu_{ba}Ee^{-i\omega t}\rho_{ab} - \mu_{ab}E^*e^{i\omega t}\rho_{ba}).\end{aligned}\quad (6.61)$$

It is interesting to note here that, in the RWA, ρ_{ba} is only driven close to its resonance frequency ω_{ba} , while $\rho_{bb} - \rho_{aa}$ is only driven close to zero frequency, which is its natural frequency.

Let us now find the steady-state solution of equations 6.61. That means a solution that is valid long after initial variations due to switching on the driving field have equilibrated. Let us assume a solution of the following form:

$$\rho_{ba}(t) = \sigma_{ba}(t)e^{-i\omega t} \quad (6.62)$$

with a slowly varying quantity σ_{ba} . Then our differential equations become:

$$\begin{aligned} \dot{\sigma}_{ba} &= \left[i(\omega - \omega_{ba}) - \frac{1}{T_2} \right] \sigma_{ba} - \frac{i}{\hbar} \mu_{ba} E ((\rho_{bb} - \rho_{aa})^{(\text{eq})}), \\ \dot{\rho}_{bb} - \dot{\rho}_{aa} &= - \frac{(\rho_{bb} - \rho_{aa}) - (\rho_{bb} - \rho_{aa})^{(\text{eq})}}{T_1} \end{aligned} \quad (6.63)$$

$$+ \frac{2i}{\hbar} (\mu_{ba} E \sigma_{ab} - \mu_{ab} E^* \sigma_{ba}). \quad (6.64)$$

In the steady state, the left-hand sides of these equations vanish (no more change of the populations in time), and we get:

$$\begin{aligned} \mu_{bb} - \mu_{aa} &= \frac{(\rho_{bb} - \rho_{aa})^{(\text{eq})}[1 + (\omega - \omega_{ba})^2 T_2^2]}{1 + (\omega - \omega_{ba})^2 T_2^2 + (4/\hbar^2) |\mu_{ba}|^2 |E|^2 T_1 T_2}, \\ \rho_{ba} &= \sigma_{ba} e^{-i\omega t} = \frac{\mu_{ba} E e^{-i\omega t} (\rho_{bb} - \rho_{aa})}{\hbar (\omega - \omega_{ba} + i/T_2)}. \end{aligned} \quad (6.65)$$

We can use this result to calculate the polarization, i.e., the dipole moment per unit volume:

$$\tilde{P}(t) = N \langle \hat{\mu} \rangle = N \text{Tr}(\hat{\rho} \hat{\mu}) = N (\mu_{ab} \rho_{ba} + \mu_{ba} \rho_{ab}). \quad (6.66)$$

N is the number density of atoms. The tilde signifies that the quantity oscillates quickly in time. We can define the complex amplitude of the polarization P via the relation:

$$\tilde{P}(t) = P e^{-i\omega t} + \text{c.c.} \quad (6.67)$$

By relating P with the slowly varying amplitude E of the optical field as $P = \epsilon_0 \chi E$, we can define the susceptibility:

$$\chi = \frac{N |\mu_{ba}|^2 (\rho_{bb} - \rho_{aa})}{\epsilon_0 \hbar (\omega - \omega_{ba} + i/T_2)}. \quad (6.68)$$

If we replace the quantity $\rho_{bb} - \rho_{aa}$ using equation 6.65, we finally get:

$$\chi = \frac{N(\rho_{bb} - \rho_{aa})^{(\text{eq})} |\mu_{ba}|^2 (\omega - \omega_{ba} - i/T_2) T_2^2 / \epsilon_0 \hbar}{1 + (\omega - \omega_{ba})^2 T_2^2 + (4/\hbar^2) |\mu_{ba}|^2 |E|^2 T_1 T_2}. \quad (6.69)$$

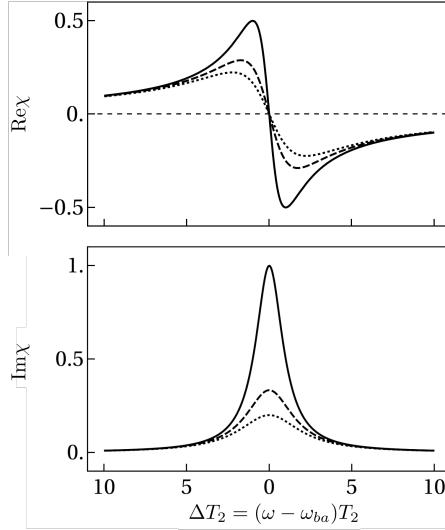


Figure 60: **Real and Imaginary part of the susceptibility.** We plot $\text{Re}(\chi)$ and $\text{Im}(\chi)$ for different values of the saturation parameter $\Omega^2 T_1 T_2$ and as a function of the optical frequency ω . The figure is adapted from Ref. [1].

This is the total susceptibility containing the linear *and* the nonlinear response. Our goal will be to simplify this expression and later to separate the linear and nonlinear contributions. Simplifying the expression for the susceptibility results in:

$$\chi = -\frac{\alpha_0(0)}{\omega_{ba}/c} \frac{\Delta T_2 - i}{1 + \Delta^2 T_2^2 + \Omega^2 T_1 T_2}, \quad (6.70)$$

where $\Delta = \omega - \omega_{ba}$ is the detuning factor, and

$$\Omega = 2|\mu_{ba}||E|/\hbar \quad (6.71)$$

is the “on-resonance Rabi frequency”. $\alpha_0(\Delta)$ denotes the absorption coefficient experienced by a *weak* optical field that is detuned from the atomic resonance by Δ . In general, the absorption coefficient in this context is given by:

$$\alpha = \frac{2\omega}{c} \text{Im}\chi. \quad (6.72)$$

Next, let us split the susceptibility in its real and imaginary parts as $\chi = \chi' + i\chi''$ with:

$$\chi' = -\frac{\alpha_0(0)}{\omega_{ba}/c} \frac{1}{\sqrt{1 + \Omega^2 T_1 T_2}} \frac{\Delta T_2 / \sqrt{1 + \Omega^2 T_1 T_2}}{1 + \Delta^2 T_2^2 / (1 + \Omega^2 T_1 T_2)}, \quad (6.73)$$

$$\chi'' = \frac{\alpha_0(0)}{\omega_{ba}/c} \left(\frac{1}{1 + \Omega^2 T_1 T_2} \right) \frac{1}{1 + \Delta^2 T_2^2 / (1 + \Omega^2 T_1 T_2)}. \quad (6.74)$$

The real part χ' has a “standard” dispersive lineshape, and χ'' has a Lorentzian lineshape. Each of these lines is broadened with respect to the weak-field case by a factor $(1 + \Omega^2 T_1 T_2)^{1/2}$. In particular, we get the following expression for the width of the absorption line:

$$\Delta\omega_{\text{FWHM}} = \frac{2}{T_2} (1 + \Omega^2 T_1 T_2)^{1/2}. \quad (6.75)$$

That spectral lines broaden if one measures them using intense optical fields is known as *power broadening*. One can also deduce from equation 6.73 that the peak value of χ'' and therefore of α decreases with respect to its weak-field value by the factor $(1 + \Omega^2 T_1 T_2)^{1/2}$. This tendency of the absorption to decrease with increased intensity is known as *saturation*. The dependence of the real and the imaginary parts of the susceptibility on the detuning and on the optical power is illustrated in Fig. 60.

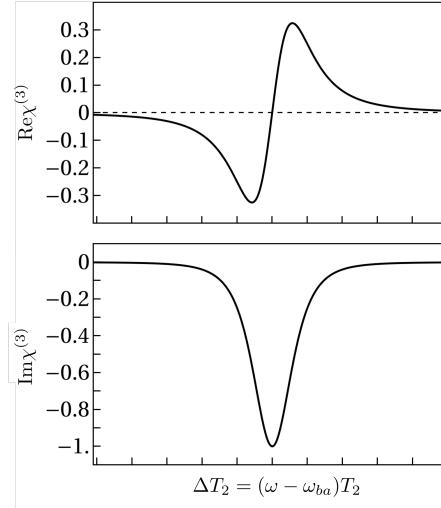


Figure 61: **Real and Imaginary part of the third-order susceptibility.** We plot $Re(\chi^{(3)})$ and $Im(\chi^{(3)})$ as a function of the optical frequency ω . The figure is adapted from Ref. [1].

Let us finally separate the linear and the nonlinear contributions to the susceptibility.

We can do so by performing a power-series expansion of the susceptibility in the quantity $|E|^2/|E_s^0|^2$, where we define the line-center *saturation field strength* E_s^0 via the relation:

$$\Omega^2 T_1 T_2 = \frac{|E|^2}{|E_s^0|^2}. \quad (6.76)$$

Then the power series expansion if we only take into account the first and second terms becomes:

$$\chi \approx \frac{-\alpha_0(0)}{\omega_{ba}/c} \left(\frac{\Delta T_2 - i}{1 + \Delta^2 T_2^2} \right) \left(1 + \frac{1}{1 + \Delta^2 T_2^2} \frac{|E|^2}{|E_s^0|^2} \right). \quad (6.77)$$

Now let us compare this expression with the usual power series expansion of the susceptibility $\chi = \chi^{(1)} + 3\chi^{(3)}|E|^2$, where $\chi^{(3)} \equiv \chi^{(3)}(\omega = \omega + \omega - \omega)$. Then we get:

$$\begin{aligned} \chi^{(1)} &= \frac{-\alpha_0(0)}{\omega_{ba}/c} \frac{\Delta T_2 - i}{1 + \Delta^2 T_2^2}, \\ \chi^{(3)} &= \frac{\alpha_0(0)}{3\omega_{ba}/c} \left[\frac{\Delta T_2 - i}{(1 + \Delta^2 T_2^2)^2} \right] \frac{1}{|E_s^0|^2}. \end{aligned} \quad (6.78)$$

It is interesting to note that we can relate the first-order and the third-order susceptibilities via the relation:

$$\chi^{(3)} = \frac{-\chi^{(1)}}{3(1 + \Delta^2 T_2^2)|E_s^0|^2} = \frac{-\chi^{(1)}}{3|E_s^\Delta|^2}, \quad (6.79)$$

where we introduced the saturation field strength E_s^Δ for an arbitrary detuning Δ :

$$|E_s^\Delta|^2 = |E_s^0|^2(1 + \Delta^2 T_2^2). \quad (6.80)$$

7 Optical coupling

In the context of nonlinear optics, we discussed how optical fields, electrical, or acoustic waves can influence the propagation of light. Using such effects, it is possible to realize optical logic gates, where, e.g., the presence of light of a certain state or intensity in one mode can influence light in another mode. All-optical logic gates promise the realization of ultra-fast optical computation circuits because they take advantage of the high velocity of light compared to comparatively slow electronic circuits. In this chapter, we will discuss some basic elements for realizing such optical logic circuits and some examples of the current state of the art.

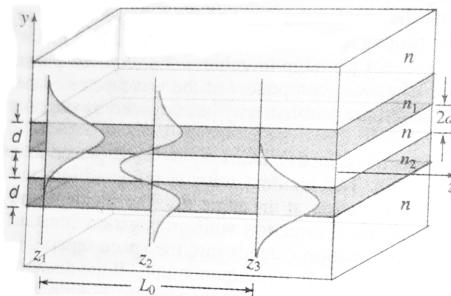


Figure 62: **Coupling between two planar waveguides.** Light propagates along the z axis. At position z_1 , we assume that the light is mostly in waveguide 1. At position z_2 , the light is distributed equally among the two waveguides, and at position z_3 , nearly all light is now in waveguide 2. The distance over which this exchange happens is $L_0 = z_3 - z_1$. The figure is from Ref. [11].

7.1 An introduction to optical couplers

A central element in any optical circuits is optical coupling between waveguides. Consider we have two optical waveguides, and that an optical field travels along one of these waveguides. If we bring the two waveguides close to each other for a certain distance, the evanescent field from one optical waveguide can couple light into the second waveguide. This effect can be used to realize optical couplers or optical switches.

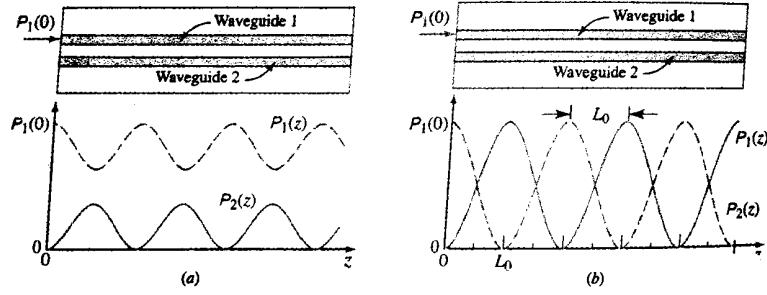


Figure 63: **Periodic exchange of power between our two waveguides.** (a) shows the case when there is a phase mismatch between the two waveguides. (b) assumes that phasematching is fulfilled. In this case, there can be a complete exchange of power. The figure is from Ref. [11].

Consider two planar waveguides as illustrated in Fig.62. Each of them is a slab of thickness d , and they are separated by a distance $2a$. An exact treatment of this situation would require to solve Maxwell's equations for the different regions and to use the boundary conditions between the waveguides. In many cases, a simpler approach using **coupled mode theory** is sufficient. In this approach, one first considers the mode in each waveguide in absence of the second waveguide. One then assumes that both waveguides are present but that the presence of the other waveguide only leads to small changes in the mode. In particular, the coupling is assumed to only modify the z -dependent amplitudes of the two modes:

$$a_i(z)u_i(y) \exp(-i\beta_i z), \quad (7.1)$$

where $i = 1, 2$, β_i are the propagation constants in the two waveguides, and u_i are the amplitudes along y . The z -dependent amplitudes $a_i(z)$ are assumed to vary slowly with z . The coupling between the two modes can be regarded as a scattering effect, and this will lead to two coupled first-order differential equations:

$$\begin{aligned} \frac{da_1}{dz} &= -iC_{21} \exp(i\Delta\beta z) a_2(z), \\ \frac{da_2}{dz} &= -iC_{12} \exp(-i\Delta\beta z) a_1(z), \end{aligned} \quad (7.2)$$

where $\Delta\beta = \beta_1 - \beta_2$ is the phase mismatch per unit length. The coupling coefficients

are given by:

$$\begin{aligned} C_{21} &= \frac{1}{2}(n_2^2 - n^2) \frac{k_o^2}{\beta_1} \int_a^{a+d} u_1(y) u_2(y) dy, \\ C_{12} &= \frac{1}{2}(n_1^2 - n^2) \frac{k_o^2}{\beta_2} \int_a^{a+d} u_2(y) u_1(y) dy. \end{aligned} \quad (7.3)$$

The rate of variation of a_1 along z is proportional to the rate of change of a_2 . The coefficient of proportionality is the product of the coupling coefficients and the phase mismatch factor $\exp(i\Delta\beta z)$.

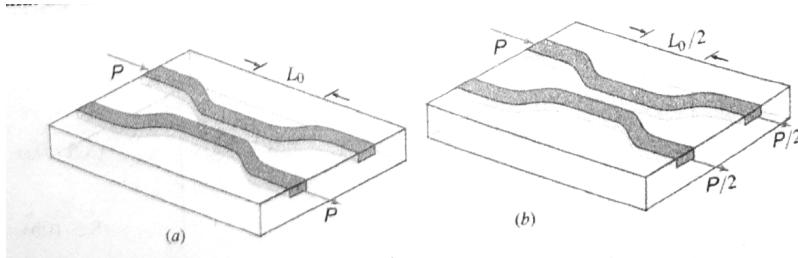


Figure 64: **Optical couplers.** (a) shows a complete transfer of power from one waveguide to the other. (b) shows the case where the power is split evenly. This is called a 3dB coupler or a beam splitter. The figure is from Ref. [11].

The solutions for equation 7.2 are:

$$\begin{pmatrix} a_1(z) \\ a_2(z) \end{pmatrix} = \begin{pmatrix} A(z) & B(z) \\ C(z) & D(z) \end{pmatrix} \cdot \begin{pmatrix} a_1(0) \\ a_2(0) \end{pmatrix} \equiv T \cdot \begin{pmatrix} a_1(z) \\ a_2(z) \end{pmatrix}, \quad (7.4)$$

where we introduced the 2×2 transmission matrix T , and:

$$\begin{aligned} A(z) &= D^*(z) = \exp\left(\frac{i\Delta\beta z}{2}\right) \left(\cos\gamma z - i\frac{\Delta\beta}{2\gamma} \sin\gamma z\right), \\ B(z) &= \frac{C_{21}}{i\gamma} \exp\left(i\frac{\Delta\beta z}{2}\right) \sin\gamma z, \\ C(z) &= \frac{C_{12}}{i\gamma} \exp\left(-i\frac{\Delta\beta z}{2}\right) \sin\gamma z, \end{aligned} \quad (7.5)$$

We will derive that and an expression for γ in the exercises.

The optical powers in the two waveguides are $P_i(z) \propto |a_i(z)|^2$. If the two waveguides are identical ($n_1 = n_2$, $\beta_1 = \beta_2$, and $\Delta\beta = 0$), they are said to be phase matched. In

that case, $\gamma = \mathcal{C} = C_{12} = C_{21}$, the powers in the waveguides behave as follows:

$$\begin{aligned} P_1(z) &= P_1(0)\cos^2(\mathcal{C}z) \\ P_2(z) &= P_2(0)\sin^2(\mathcal{C}z). \end{aligned} \quad (7.6)$$

In this case, there can be a complete exchange of power between the two waveguides. This complete exchange of power from one waveguide to the other happens over the coupling distance (or transfer distance) L_0 . If one brings two waveguides close to each other only for half that distance, half the power is coupled from one waveguide to the other. That means, we can use this approach to realize a 50/50 beam splitter. This is illustrated in Fig. 64.

7.2 Coupled mode theory

In the last section, we saw a simplified theoretical description of the coupling process using coupled mode theory. In particular, we just said that one “can show” that the slowly varying amplitudes a_j fulfill the coupled differential equations 7.2 with the coupling coefficients given in equations 7.3. We will now discuss in more detail how to derive equations like that. The example we will consider here is actually a bit more complicated than the one in the previous section. The waveguide geometry we will consider is illustrated in Fig. 65. In Ref. [9], they label this the directional coupling of waveguides because the light travels in the same direction in both waveguides.

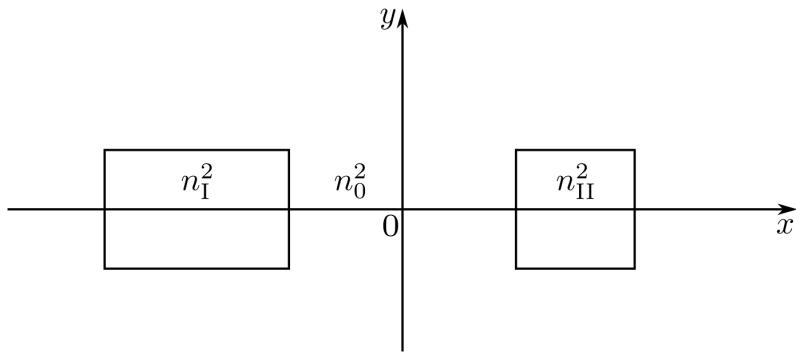


Figure 65: **Directionally coupled waveguides.** Two separate waveguides I and II with respective refractive indices n_I and n_{II} . The dielectric constant in the waveguides is of course proportional to the square of the refractive indices. The waveguides are embedded in a material with refractive index n_0 . The figure is adapted from Ref. [9].

In order to describe the situation depicted in Fig. 65, we denote the eigenmodes in each

of the two waveguides as \mathbf{E}_p and \mathbf{H}_p with $p = 1, 2$. They satisfy Maxwell's equations as follows:

$$\begin{aligned}\nabla \times \tilde{\mathbf{E}}_p &= -i\omega\mu_0 \tilde{\mathbf{H}}_p \\ \nabla \times \tilde{\mathbf{H}}_p &= i\omega\epsilon_0 N_p^2 \tilde{\mathbf{E}}_p,\end{aligned}\quad (7.7)$$

where $N_p(x, y)$ denotes the refractive index distribution of each waveguide. The *tilde* in this notation denotes that we are dealing with the total fields in all three dimensions. This will become useful later when we split the spatial dependence in the propagation direction z from the spatial dependence in the plane perpendicular to that. We assume that the fields of the coupled waveguides can be expressed as the sum of the eigenmodes in each waveguide:

$$\begin{aligned}\tilde{\mathbf{E}} &= A(z)\tilde{\mathbf{E}}_1 + B(z)\tilde{\mathbf{E}}_2 \\ \tilde{\mathbf{H}} &= A(z)\tilde{\mathbf{H}}_1 + B(z)\tilde{\mathbf{H}}_2.\end{aligned}\quad (7.8)$$

If we insert that into the Maxwell equations 7.7, we get:

$$\begin{aligned}\nabla \times \tilde{\mathbf{E}} &= -i\omega\mu_0 \tilde{\mathbf{H}} \\ \nabla \times \tilde{\mathbf{H}} &= i\omega\epsilon_0 N^2 \tilde{\mathbf{E}},\end{aligned}\quad (7.9)$$

where $N(x, y)$ denotes the distribution of the refractive index in the complete coupled waveguide structure.

If we then use the Maxwell equations 7.7 and the vector formula:

$$\nabla \times (A\tilde{\mathbf{E}}) = A\nabla \times \tilde{\mathbf{E}} + \nabla A \times \tilde{\mathbf{E}} = A\nabla \times \tilde{\mathbf{E}} + \frac{dA}{dz} \mathbf{u}_z \times \tilde{\mathbf{E}}, \quad (7.10)$$

where A denotes a function of z and \mathbf{u}_z is the unit vector along the z axis, then we get:

$$(\mathbf{u}_z \times \tilde{\mathbf{E}}_1) \frac{dA}{dz} + (\mathbf{u}_z \times \tilde{\mathbf{E}}_2) \frac{dB}{dz} = 0, \quad (7.11)$$

$$\begin{aligned}(\mathbf{u}_z \times \tilde{\mathbf{H}}_1) \frac{dA}{dz} - i\omega\epsilon_0 (N^2 - N_1^2) A \tilde{\mathbf{E}}_1 \\ + (\mathbf{u}_z \times \tilde{\mathbf{H}}_2) \frac{dB}{dz} - i\omega\epsilon_0 (N^2 - N_2^2) B \tilde{\mathbf{E}}_2 = 0.\end{aligned}\quad (7.12)$$

If one then inserts equations 7.11 and 7.12 into the following integral equations:

$$\begin{aligned} \int_{-\infty}^{\infty} dx dy [\tilde{\mathbf{E}}_1^* \cdot (7.12) - \tilde{\mathbf{H}}_1^* \cdot (7.11)] &= 0 \\ \int_{-\infty}^{\infty} dx dy [\tilde{\mathbf{E}}_2^* \cdot (7.12) - \tilde{\mathbf{H}}_2^* \cdot (7.11)] &= 0, \end{aligned} \quad (7.13)$$

then one gets the following equations:

$$\begin{aligned} \frac{dA}{dz} + \frac{dB}{dz} \frac{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\tilde{\mathbf{E}}_1^* \times \tilde{\mathbf{H}}_2 + \tilde{\mathbf{E}}_2 \times \tilde{\mathbf{H}}_1^*)}{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\tilde{\mathbf{E}}_1^* \times \tilde{\mathbf{H}}_1 + \tilde{\mathbf{E}}_1 \times \tilde{\mathbf{H}}_1^*)} \\ + iA \frac{\omega \epsilon_0 \int_{-\infty}^{\infty} dx dy (N^2 - N_1^2) \tilde{\mathbf{E}}_1^* \cdot \tilde{\mathbf{E}}_1}{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\tilde{\mathbf{E}}_1^* \times \tilde{\mathbf{H}}_1 + \tilde{\mathbf{E}}_1 \times \tilde{\mathbf{H}}_1^*)} \\ + iB \frac{\omega \epsilon_0 \int_{-\infty}^{\infty} dx dy (N^2 - N_2^2) \tilde{\mathbf{E}}_1^* \cdot \tilde{\mathbf{E}}_2}{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\tilde{\mathbf{E}}_1^* \times \tilde{\mathbf{H}}_1 + \tilde{\mathbf{E}}_1 \times \tilde{\mathbf{H}}_1^*)} = 0, \end{aligned} \quad (7.14)$$

$$\begin{aligned} \frac{dB}{dz} + \frac{dA}{dz} \frac{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\tilde{\mathbf{E}}_2^* \times \tilde{\mathbf{H}}_1 + \tilde{\mathbf{E}}_1 \times \tilde{\mathbf{H}}_2)}{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\tilde{\mathbf{E}}_2^* \times \tilde{\mathbf{H}}_2 + \tilde{\mathbf{E}}_2 \times \tilde{\mathbf{H}}_2^*)} \\ + iA \frac{\omega \epsilon_0 \int_{-\infty}^{\infty} dx dy (N^2 - N_1^2) \tilde{\mathbf{E}}_2^* \cdot \tilde{\mathbf{E}}_1}{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\tilde{\mathbf{E}}_2^* \times \tilde{\mathbf{H}}_2 + \tilde{\mathbf{E}}_2 \times \tilde{\mathbf{H}}_2^*)} \\ + iB \frac{\omega \epsilon_0 \int_{-\infty}^{\infty} dx dy (N^2 - N_2^2) \tilde{\mathbf{E}}_2^* \cdot \tilde{\mathbf{E}}_2}{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\tilde{\mathbf{E}}_2^* \times \tilde{\mathbf{H}}_2 + \tilde{\mathbf{E}}_2 \times \tilde{\mathbf{H}}_2^*)} = 0. \end{aligned} \quad (7.15)$$

We can significantly simplify these equations if we split our fields into a transverse part and an axial part along the propagation direction as follows:

$$\begin{aligned} \tilde{\mathbf{E}}_p &= \mathbf{E}_p e^{-i\beta_p z}, \\ \tilde{\mathbf{H}}_p &= \mathbf{H}_p e^{-i\beta_p z}. \end{aligned} \quad (7.16)$$

Then the equations 7.14 and 7.15 become:

$$\begin{aligned} \frac{dA}{dz} + c_{12} \frac{dB}{dz} e^{-i(\beta_2 - \beta_1)z} + i\chi_1 A + i\kappa_{12} B e^{-i(\beta_2 - \beta_1)z} &= 0, \\ \frac{dB}{dz} + c_{21} \frac{dA}{dz} e^{i(\beta_2 - \beta_1)z} + i\chi_2 B + i\kappa_{21} A e^{i(\beta_2 - \beta_1)z} &= 0 \end{aligned} \quad (7.17)$$

Here, we defined the mode coupling coefficients κ_{pq} and the butt coupling coefficients c_{pq} . Why c_{pq} is called like that, we will explain below. The constants χ_p describe a spatial frequency of modulation of the amplitude along z . These coefficients are given

by the following expressions:

$$\begin{aligned}\kappa_{pq} &= \frac{\omega\epsilon_0 \int_{-\infty}^{\infty} dx dy (N^2 - N_q^2) \mathbf{E}_p^* \cdot \mathbf{E}_q}{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\mathbf{E}_p^* \times \mathbf{H}_p + \mathbf{E}_p \times \mathbf{H}_p^*)} \\ c_{pq} &= \frac{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\mathbf{E}_p^* \times \mathbf{H}_q + \mathbf{E}_q \times \mathbf{H}_p^*)}{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\mathbf{E}_p^* \times \mathbf{H}_p + \mathbf{E}_p \times \mathbf{H}_p^*)} \\ \chi_p &= \frac{\omega\epsilon_0 \int_{-\infty}^{\infty} dx dy (N^2 - N_p^2) \mathbf{E}_p^* \cdot \mathbf{E}_p}{\int_{-\infty}^{\infty} dx dy \mathbf{u}_z \cdot (\mathbf{E}_p^* \times \mathbf{H}_p + \mathbf{E}_p \times \mathbf{H}_p^*)}. \end{aligned} \quad (7.18)$$

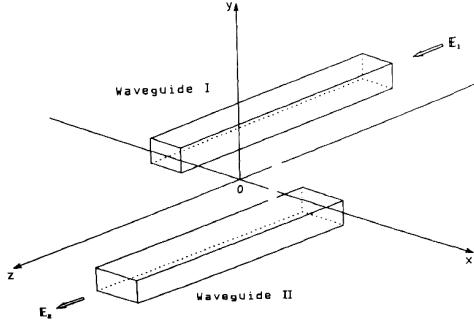


Figure 66: **Butt coupled waveguides.** Waveguides I and II only exist in the regions $z < 0$ and $z \geq 0$, respectively. At the plane $z = 0$, some light can be coupled from waveguide I to waveguide II. This is known as butt coupling. The figure is from Ref. [9].

In order to understand the meaning of the butt coupling coefficients, consider the following situation: assume that waveguide I only exists for $z < 0$, and the waveguide II only exists for $z \geq 0$. We send light into waveguide I. At $z = 0$ some of the light will be coupled into waveguide II and then exit from the end of that waveguide. The efficiency for the field in waveguide I to excite a field in waveguide II at position $z = 0$ is called the butt coupling efficiency. The situation is illustrated in Fig. 66.

It is instructive to make an order of magnitude comparison between the coupling coefficients κ_{pq} and c_{pq} . Let us assume that some portion of the field in waveguide I will couple into waveguide II at the plane $z = 0$, and let us denote that in the form $|\mathbf{E}_2| = \eta |\mathbf{E}_1|$ with $\eta \ll 1$ ¹. To understand this intuitively: it is unlikely for light to come out of waveguide I, then bend around and again enter into waveguide II in the same direction of propagation. The only option for coupling is via some evanescent fields that

¹Of course, this does not mean that the two fields are proportional to each other everywhere. It is only a rough description that we couple some small part of the field from one waveguide into the other.

reach beyond waveguide I and then excite the eigenmode in waveguide II. Now let us look at the numerator of the expression for κ_{12} in equation 7.18. The integral term of the numerator should be about $(n_1^2 - n_0^2)\eta$ because at the exit of waveguide I, the refractive index is still n_1 , and the refractive index of waveguide II at that position is n_0 because we are outside of that waveguide. For comparison, the integral for the numerator of χ_1 is carried out in the waveguide II, where the difference $(N^2 - N_1^2)$ is non-zero. Then the magnitude of that integral term is about $(n_1^2 - n_0^2)\eta^2$. That illustrates that χ_p is about η times smaller than κ_{pq} . This is a huge difference if $\eta \ll 1$. The butt-coupling coefficients c_{pq} are typically also much smaller than κ_{pq} . For these reasons, one often assumes that $c_{pq} = \chi_p = 0$. In a strict treatment, however, one has to take these effects into account.

The optical power carried by the waveguide eigenmodes is:

$$P_p = \frac{1}{2} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} dx dy (\mathbf{E}_p \times \mathbf{H}_p^*). \quad (7.19)$$

Then the denominator in the equations 7.18 are all equal to $4P$. In the following, we will assume that this is normalized to units. That means, $4P = 1$. From this, it follows that $c_{21} = c_{12}^*$ and $\chi_p = \chi_p^*$.

The difference between the propagation constants in the two waveguides I and II can be expressed by:

$$\delta = \frac{\beta_2 - \beta_1}{2}. \quad (7.20)$$

During the exercises, we will show that the coupling coefficients for loss-less waveguides have to fulfill:

$$\kappa_{21} = \kappa_{12}^* + 2\delta c_{12}^*. \quad (7.21)$$

Often, one simply assumes $\kappa_{21} = \kappa_{12}^*$, but that is only true if (1) the two waveguides have the same propagation constants ($\delta = 0$) or if (2) the waveguides are far enough separated such that $c_{12} \approx 0$. Discussing concrete usage scenarios and realizations of different couplers becomes a bit difficult if one treats the coupling in a very general way. Things become significantly easier to handle, if one assumes $c_{pq} = \chi_p = 0$. Then the coupled mode equations describing the coupler become equivalent to the equations we had in the previous section:

$$\begin{aligned} \frac{dA}{dz} &= -i\kappa_{12} B \exp[-i(\beta_2 - \beta_1)z], \\ \frac{dB}{dz} &= -i\kappa_{21} A \exp[i(\beta_2 - \beta_1)z]. \end{aligned} \quad (7.22)$$

If one assumes reciprocity, then $\kappa \equiv \kappa_{12} = \kappa_{21}$. In most directional couplers, κ is real. For contradirectional couplers, the reciprocity relation instead is $\kappa_{21} = -\kappa_{12}^*$.

During one of the exercises, we will enter trial solutions into the coupled equations

above. These trial solutions have the form:

$$\begin{aligned} A(z) &= [a_1 e^{iqz} + a_2 e^{-iqz}] e^{-i\delta z}, \\ B(z) &= [b_1 e^{iqz} + b_2 e^{-iqz}] e^{i\delta z}. \end{aligned} \quad (7.23)$$

In that exercise, we will show that the parameter q is given by:

$$q = \sqrt{\kappa^2 + \delta^2}. \quad (7.24)$$

The parameters a_p, b_b ($p = 1, 2$) will be determined from the initial conditions for $A(0)$ and $B(0)$. The parameter q determines the length over which there will be a periodic exchange of power between the coupled waveguides. The maximum power-coupling efficiency is defined by:

$$F = \left(\frac{\kappa}{q}\right)^2 = \frac{1}{1 + (\delta/\kappa)^2}. \quad (7.25)$$

As we go along the waveguide, the power will periodically change between the waveguides, and the coupling efficiency from waveguide I to II will reach its maximum whenever we have:

$$z = \frac{\pi}{2q}(2m + 1) \quad (7.26)$$

for some integer m . The distance we get for $m = 0$ is called the *coupling length* L_c :

$$L_c = \frac{\pi}{2q} = \frac{\pi}{2\sqrt{\kappa^2 + \delta^2}}. \quad (7.27)$$

As we already saw last time, if there is perfect phase matching ($\beta_1 = \beta_2$ and $\delta = 0$), we can get 100% power coupling. In this case, we have $L_c = \pi/2\kappa$.

7.3 Contradirectional coupling in corrugated waveguides

In the previous section we often implicitly assumed that the light in both coupled waveguides is travelling in the same direction although most of the treatment also applies for the case where the light in the two waveguides is counterpropagating. In the case where the light propagates in the same direction, we can get perfect power transfer from one waveguide to the other. If the light in the two waveguides is counterpropagating, then there typically is **no** coupling between these waveguides **unless** one introduces some disturbance - either in at least one of the waveguides or in the medium between the waveguides such that light is scattered in the opposite direction.

To demonstrate this effect, let us assume that the medium between the waveguides is periodically disturbed with a period Λ such that the coupling between the waveguides

takes form:

$$\kappa_{12}(z) = \kappa_G \exp\left[-i \frac{2\pi}{\Lambda} z\right]. \quad (7.28)$$

Taking into account the reciprocity relation for counterpropagating waves:

$$\kappa_{21} = -\kappa_{12}^* = -\kappa_G \exp\left(i \frac{2\pi}{\Lambda} z\right), \quad (7.29)$$

our coupled differential equations 7.22 become:

$$\begin{aligned} \frac{dA}{dz} &= -i\kappa_G B \exp\left[i \left(\beta_1 - \beta_2 - \frac{2\pi}{\Lambda}\right) z\right], \\ \frac{dB}{dz} &= i\kappa_G A \exp\left[-i \left(\beta_1 - \beta_2 - \frac{2\pi}{\Lambda}\right) z\right]. \end{aligned} \quad (7.30)$$

We will derive the reciprocity relation 7.29 from power conservation in the exercises.

For co-propagating light, we got the most efficient coupling if we had phase matching ($\beta_1 = \beta_2$). For counter-propagating light, we will also get a phase matching condition to get efficient coupling. It will be useful to first define the following parameter:

$$\phi = \frac{(\beta_1 - \beta_2 - 2\pi/\Lambda)}{2}. \quad (7.31)$$

We will assume that the coupling between the waveguides is periodically modulated for $z \in [0, L]$ (see Fig. 67). The boundary conditions are $A(0) = A_0$ (some initial power) and $B(L) = 0$ (no reflected light from the undisturbed region in $z > L$).

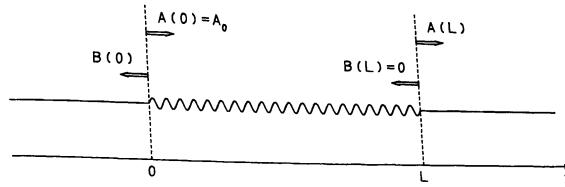


Figure 67: **Bragg optical waveguide.** Contradirectional coupling in a corrugated Bragg waveguide. The figure is from Ref. [9].

For light to be coupled between the waveguides, we will have to fulfill phase-matching conditions. In particular, depending on the value of ϕ , we will get the following solutions for A and B :

(A) $|\phi| > \kappa_G$:

$$\begin{aligned} A(z) &= A_0 \frac{\rho \cos[\rho(z-L)] - i\phi \sin[\rho(z-L)]}{\rho \cos(\rho L) + i\phi \sin(\rho L)} \exp(i\phi z), \\ B(z) &= A_0 \frac{i\kappa_G \sin[\rho(z-L)]}{\rho \cos(\rho L) + i\phi \sin(\rho L)} \exp(-i\phi z), \\ \rho &= \sqrt{\phi^2 - \kappa_G^2}. \end{aligned} \quad (7.32)$$

(B) $|\phi| = \kappa_G$:

$$\begin{aligned} A(z) &= A_0 \frac{1 - i\phi(z-L)}{1 + i\phi L} \exp(i\phi z), \\ B(z) &= A_0 \frac{i\kappa_G(z-L)}{1 + i\phi L} \exp(-i\phi z). \end{aligned} \quad (7.33)$$

(C) $|\phi| < \kappa_G$:

$$\begin{aligned} A(z) &= A_0 \frac{\alpha \cosh[\alpha(z-L)] - i\phi \sinh[\alpha(z-L)]}{\alpha \cosh(\alpha L) + i\phi \sinh(\alpha L)} \exp(i\phi z), \\ B(z) &= A_0 \frac{i\kappa_G \sinh[\alpha(z-L)]}{\alpha \cosh(\alpha L) + i\phi \sinh(\alpha L)} \exp(-i\phi z), \\ \alpha &= \sqrt{\kappa_G^2 - \phi^2}. \end{aligned} \quad (7.34)$$

Let us discuss a very important example: a **Bragg waveguide** as shown in Fig. 67. To realize such a waveguide, the refractive index is modulated periodically either in the core or in the cladding. One option is also to use geometrical corrugations as is shown in the example of Fig. 67. For example, Bragg waveguides are used to implement bandwidth filters or in various lasers.

In a Bragg waveguide, the waveguides I and II we talked about so far are the same physical waveguide, but the waves A and B are counterpropagating. For that reason, we have:

$$\beta_1 = -\beta_2 = kn_{\text{eff}}. \quad (7.35)$$

The phase-matching parameter ϕ becomes:

$$\phi = kn_{\text{eff}} - \frac{\pi}{\Lambda}. \quad (7.36)$$

Using this in combination with the solutions 7.32-7.34, we can write the normalized transmitted (forward) and the normalized reflected (backward) optical powers as: (A)

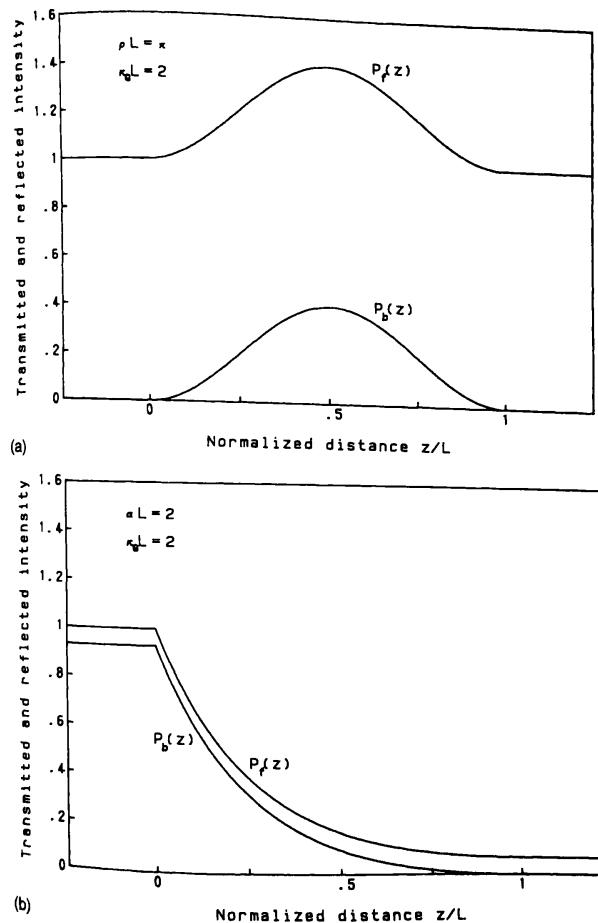


Figure 68: **Normalized forward and backwards optical power.** (a) forward and backwards power for $\rho = \pi/L$ and $\kappa_G = 2/L$ ($|\phi| > \kappa_G$). (b) forward and backwards power for $\alpha = \kappa_G = 2/L$ and $\kappa_G = 2/L$ ($|\phi| = 0$). The figure is from Ref. [9].

$|\phi| > \kappa_G$:

$$\begin{aligned} P_f(z) &= \frac{|A(z)|^2}{|A_0|^2} = \frac{\rho^2 + \kappa_G^2 \sin^2[\rho(z - L)]}{\rho^2 + \kappa_G^2 \sin^2(\rho L)}, \\ P_b(z) &= \frac{|B(z)|^2}{|A_0|^2} = \frac{\kappa_G^2 \sin^2[\rho(z - L)]}{\rho^2 + \kappa_G^2 \sin^2(\rho L)}, \\ \rho &= \sqrt{\phi^2 - \kappa_G^2}. \end{aligned} \quad (7.37)$$

(B) $|\phi| = \kappa_G$:

$$\begin{aligned} P_f(z) &= \frac{|A(z)|^2}{|A_0|^2} = \frac{1 + \kappa_G^2(z - L)^2}{1 + \kappa_G^2 L^2}, \\ P_b(z) &= \frac{|B(z)|^2}{|A_0|^2} = \frac{\kappa_G^2(z - L)^2}{1 + \kappa_G^2 L^2}. \end{aligned} \quad (7.38)$$

(C) $|\phi| < \kappa_G$:

$$\begin{aligned} P_f(z) &= \frac{|A(z)|^2}{|A_0|^2} = \frac{\alpha^2 + \kappa_G^2 \sinh^2[\alpha(z - L)]}{\alpha^2 + \kappa_G^2 \sinh^2(\alpha L)}, \\ P_b(z) &= \frac{|B(z)|^2}{|A_0|^2} = \frac{\kappa_G^2 \sinh^2[\alpha(z - L)]}{\alpha^2 + \kappa_G^2 \sinh^2(\alpha L)} \\ \alpha &= \sqrt{\kappa_G^2 - \phi^2}. \end{aligned} \quad (7.39)$$

Let us look at two different examples to illustrate the behaviour of these curves as shown in Fig. 68. Light is reflected back, but this back-reflected light is again reflected in the forward direction. Effectively, a Bragg waveguide will transmit light in this case, when $|\phi| > \kappa_G$. This defines a range of frequencies called a *pass band*:

$$\begin{aligned} \frac{\omega}{c} n_{\text{eff}} &< \frac{\pi}{\Lambda} - \kappa_G, \\ \frac{\omega}{c} n_{\text{eff}} &> \frac{\pi}{\Lambda} + \kappa_G. \end{aligned} \quad (7.40)$$

In the opposite case, when $|\phi| < \kappa_G$, the light decays exponentially inside the waveguide as shown in Fig. 68(b). Most of the light is reflected backwards. The phase-matching condition $|\phi| < \kappa_G$ then defines another frequency range for which this reflection occurs, and this is known as the *stop band*:

$$\frac{\pi}{\Lambda} - \kappa_G < \frac{\omega}{c} n_{\text{eff}} < \frac{\pi}{\Lambda} + \kappa_G. \quad (7.41)$$

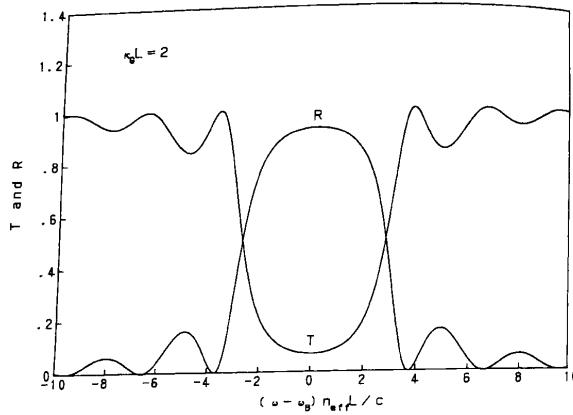


Figure 69: **Bragg bandpass filter.** Transmittance (T) and reflectivity (R) of a Bragg waveguide. Within the stop band, most light is reflected, whereas little is transmitted. In the surrounding pass band, the opposite holds true. The figure is from Ref. [9].

The optical wavelength in the middle of this band and satisfying $(\omega/c)n_{\text{eff}} = \pi/\Lambda$ ($\phi = 0$) is called the *Bragg wavelength*:

$$\lambda_B = 2n_{\text{eff}}\Lambda. \quad (7.42)$$

To get an impression of the transmission/reflection characteristics of such a Bragg filter, let us look at the following quantities as a function of the detuning $\phi L = (\omega - \omega_B)n_{\text{eff}}L/c$ from the Bragg angular frequency $\omega_B = 2\pi c/\lambda_B$:

$$T = \frac{|A(L)|^2}{|A_0|^2},$$

$$R = \frac{|B(0)|^2}{|A_0|^2}. \quad (7.43)$$

That frequency dependence is shown in Fig. 69. In one of today's exercises, we will look at different cases of the reflectivity R of a Bragg waveguide. In this exercise, we will also calculate the reflectance at the Bragg wavelength to be:

$$R = \tanh^2(\kappa_G L). \quad (7.44)$$

For example, this yields $R = 0.93$ for $\kappa_G L = 2$ [9].

7.4 Optical waveguide devices with directional couplers

Let us discuss some examples of useful waveguide devices that are common in integrated photonic circuits. By coupling waveguides to each other and by allowing a dynamic change of that coupling, it is possible to implement photonic switches, waveguide resonators and more.

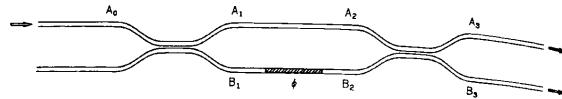


Figure 70: **Mach-Zehnder interferometer**. If two waveguides are brought close to each other two times, each photon can be in a superposition of taking the upper or the lower path. This can lead to interference in the output ports. The phase difference accumulated can be tuned if a variable phase shift ϕ is applied to one of the arms. The figure is from Ref. [9].

7.4.1 Mach-Zehnder interferometers

One of the big advantages of integrated photonic circuits is that they are small, and that it is therefore relatively simple to keep them thermally stable. For example, consider an integrated Mach-Zehnder interferometer. Usually great care has to be taken to ensure that the phase difference between the two interferometer arms does not vary too much due to thermal variations or due to air currents. In an integrated circuit, these effects pose a significantly smaller problem. The device we will consider here is depicted in Fig.

We will assume that both waveguides have the same optical properties such that $\delta = 0$ and $q = \kappa$ (see equation 7.24). If we take the boundary conditions as $A(0) = A_0$ and $B(0) = 0$ (all the light in the lower input), then the solutions $A(z)$ and $B(z)$ from equations 7.23 become:

$$\begin{aligned} A_1 &= A_0 \cos(\kappa l), \\ B_1 &= -iA_0 \sin(\kappa l). \end{aligned} \tag{7.45}$$

The index 1 for these solutions indicates that we are concerned with the solutions at the first of the two couplers. l denotes an *effective* straight coupling length because there will not only be coupling directly in the straight part but also in the curved parts as they get closer. l will therefore be longer than the actual straight section. If the first coupling strength is chosen such that $\kappa l = \pi/4$, we say that it is a 3 dB coupler or a 50 : 50 beamsplitter. Then the coupling ratios are: $A_1 = A_0/\sqrt{2}$ and $B_1 = -iA_0/\sqrt{2}$. After passing through the two straight arms of the interferometer, our field amplitudes

A_2 and B_2 will be:

$$\begin{aligned} A_2 &= A_1 \exp(-i\beta L) = \frac{A_0}{\sqrt{2}} \exp(-i\beta L), \\ B_2 &= B_1 \exp(-i\beta L + i\phi) = -i \frac{A_0}{\sqrt{2}} \exp(-i\beta L + i\phi). \end{aligned} \quad (7.46)$$

The excess phase shift ϕ might be tunable, for example by applying a voltage, if the material is responsive to that.

If the second coupler is a 50 : 50 beamsplitter, the output amplitudes will be:

$$\begin{aligned} A_3 &= -iA_0 \sin\left(\frac{\phi}{2}\right) \exp(-i\beta L + \frac{i\phi}{2}), \\ B_3 &= -iA_0 \cos\left(\frac{\phi}{2}\right) \exp(-i\beta L + \frac{i\phi}{2}). \end{aligned} \quad (7.47)$$

Then the power at the two outputs will be:

$$\begin{aligned} |A_3|^2 &= |A_0|^2 \sin^2\left(\frac{\phi}{2}\right), \\ |B_3|^2 &= |A_0|^2 \cos^2\left(\frac{\phi}{2}\right). \end{aligned} \quad (7.48)$$

By modulating the phase ϕ , we can switch the power to one of the outputs or have the light exiting in a superposition.

7.4.2 Ring resonators

Ring resonators, like the one illustrated in Fig. 71, are common elements in photonic circuits. With the tools we discussed today, we can readily describe the behaviour of light in such a device. We will assume that both waveguides have the same propagation constant β such that $\delta = 0$. κ is the mode-coupling coefficient and l is the coupling length. L is the length of the ring. Additionally, the new quantity γ is the intensity-insertion-loss coefficient. The steady-state input-output relations for a ring resonator are[9]:

$$\begin{aligned} A &= (1 - \gamma)^{1/2} [A_0 \cos(\kappa l) - iB_0 \sin(\kappa l)], \\ B &= (1 - \gamma)^{1/2} [-iA_0 \sin(\kappa l) + B_0 \cos(\kappa l)]. \end{aligned} \quad (7.49)$$

If we assume that the optical power decays exponentially in the resonator with some attenuation parameter ρ , then we can relate the amplitude B of the field in the ring

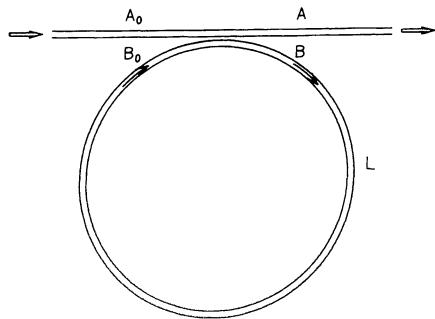


Figure 71: **Optical waveguide ring resonator.** We assume that a linear waveguide is coupled to a second waveguide that forms a ring. The ring-shaped waveguide will act as a resonator we can couple light into and out of. The figure is from Ref. [9].

resonator after the coupler with the amplitude B_0 of the field before the coupler:

$$B_0 = \exp\left(-\frac{\rho}{2}L - i\beta L\right). \quad (7.50)$$

If we put this into the equations 7.49, we can solve for A :

$$\frac{A}{A_0} = (1 - \gamma)^{1/2} \left[\frac{\cos(\kappa l) - (1 - \gamma)^{1/2} \exp(-\frac{\rho}{2}L - i\beta L)}{1 - (1 - \gamma)^{1/2} \cos(\kappa l) \exp(-\frac{\rho}{2}L - i\beta L)} \right]. \quad (7.51)$$

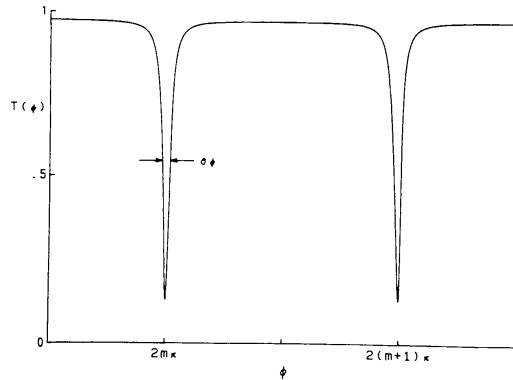


Figure 72: **Transmittance of a waveguide ring resonator.** The transmittance shows clear resonance dips with a width $\Delta\phi$ at integer multiples of 2κ . The figure is from Ref. [9].

We can simplify that if we introduce the following new parameters:

$$\begin{aligned} x &= (1 - \gamma)^{1/2} \exp\left(-\frac{\rho}{2}L\right), \\ y &= \cos(\kappa l), \\ \phi &= \beta L. \end{aligned} \quad (7.52)$$

Then the intensity transmittance of the ring resonator becomes:

$$T(\phi) = \left| \frac{A}{A_0} \right|^2 = (1 - \gamma) \left[1 - \frac{(1 - x^2)(1 - y^2)}{(1 - xy)^2 + 4xysin^2(\phi/2)} \right]. \quad (7.53)$$

Fig. 72 shows how the transmittance depends on ϕ . Resonance peaks occur whenever we have $\phi = \beta L = 2m\pi$. The FWHM $\Delta\phi$ of these dips is given by:

$$\Delta\phi = \frac{2(1 - xy)}{\sqrt{xy}}. \quad (7.54)$$

The finesse of a resonator is always the ratio between the distance between resonances and the width of the resonance. Usually, this is specified in terms of frequency. Then the distance between peaks is known as the *free spectral range* (FSR), and if the width of the resonance is $\Delta\nu$, the finesse would be:

$$\mathcal{F} = \frac{\text{FSR}}{\Delta\nu}. \quad (7.55)$$

The FSR corresponds to the frequency difference between light of two frequencies that differ in their wavenumber by an amount Δk such that a full round trip through the cavity gives a phase difference of 2π . To get the finesse, one then has to compare the frequency difference corresponding to Δk to the frequency bandwidth of the resonance. In the case of a ring resonator, we have the advantage that we already know the FWHM of the resonator in units of phase $\Delta\phi$ given by equation 7.54. At the same time, a full round trip in the cavity corresponds to a phase shift of 2π . Therefore the finesse of our ring resonator has to be:

$$\mathcal{F} = \frac{2\pi}{\Delta\phi} = \frac{\pi\sqrt{xy}}{1 - xy}. \quad (7.56)$$

During the one of the exercises, we will show that the minimum and maximum transmission of a ring resonator will be:

$$T_{\max} = (1 - \gamma) \frac{(x + y)^2}{(1 + xy)^2}, \quad T_{\min} = (1 - \gamma) \frac{(x - y)^2}{(1 - xy)^2}. \quad (7.57)$$

The minimum transmission can reach zero if $x = y$ or, equivalently:

$$\cos(\kappa l) = (1 - \gamma)^{1/2} \exp\left(-\frac{\rho}{2}L\right). \quad (7.58)$$

We will also show for which ϕ one achieves the minimum transmission T_{\min} , and we will look at a few examples for the length of a ring resonator, the optical losses, and the finesse of such resonators.

Let us consider the difference in frequency between neighbouring reference peaks in the transmission of a ring resonator, and let us denote the wavenumber for one peak as k and the wavenumber for the neighbouring peak as $k + \Delta k$. These two peaks correspond to phases $\phi = 2m\pi$ and $\phi + \Delta\phi = 2(m+1)\pi$, where m is some integer that is typically very large (**think**: why?). Because m is large, $|\Delta k|$ will be small compared to k . The difference between the propagation constants corresponding to the neighbouring resonances will be:

$$\beta(k + \Delta k) - \beta(k) = \frac{2\pi}{L}. \quad (7.59)$$

Because $|\Delta k| \ll k$, we can rewrite that as:

$$\frac{d\beta}{dk} \Delta k = \frac{2\pi}{L}. \quad (7.60)$$

If we substitute $\beta = kn$, where n is the effective refractive index, we can write:

$$\frac{d\beta}{dk} = n + k \frac{dn}{dk} = n - \lambda \frac{dn}{d\lambda} \equiv N. \quad (7.61)$$

Here, we introduced the *group refractive index* N . Using these relations, we will be able to show in one of the exercises how to express the distance between resonance peaks as well as their FWHM in terms of frequency and wavelength.

8 Photonic crystals

When we discussed quasi-phase matching in section 2.4, we already saw a good example how periodic modifications of a material can be used to tailor specific optical properties of that material. Other examples we already encountered are optical fibers and waveguides that can be used to confine light in specific dimensions in order to guide it. In the last chapter, we discussed how such waveguides can be coupled to each other, or how one can realize fiber Bragg gratings or optical switches.

In this chapter, we will go a bit further and discuss how one can modify the structure of materials in order to realize photonic crystal structures. Photonic crystals have become a very versatile tool. The interesting thing is, that one can micro- or nano-structure materials in order to guide and confine light very closely on photonic chips without having to resort to bulk optics or to rely on varying material properties.

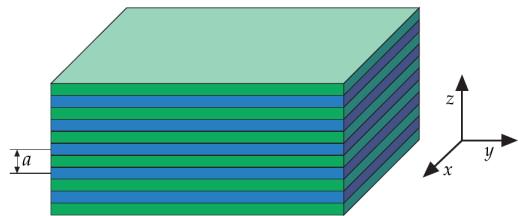


Figure 73: **A multilayer film.** Multilayer films, e.g., dielectric coatings are wellknown examples of 1D photonic crystal structures. The figure is from Ref. [13].

8.1 Photonic crystals

The central idea behind photonic crystals is to modulate the optical properties of a material such that band-gaps form similar to what we know from crystalline solids. A well-known example of 1D photonic crystals are multilayer films - for example, in dielectric mirrors (see e.g. Fig. 73). By stacking multiple layers of different dielectric materials such that the refractive index changes periodically, one can realize very good mirrors with reflectivities of 99.999% or even higher. The waves that are reflected from the various layers can interfere constructively if certain conditions are fulfilled. In the case of our earlier example of fiber Bragg gratings, this resulted in high reflectivity close to the Bragg wavelength (see section 7.3). Usually, a simple stack of identical

layers will lead to a sharply-peaked high reflectivity at a specific design wavelength. However, there are also broadband dielectric coatings. In all of these coatings, one has to keep impurities and unevenness as small as possible to achieve high reflectivities and to minimize wavefront distortions.

While multilayer films are good examples for how the periodic modulation of material properties can be used to shape the optical response of a material, it is still based on actual changes of material properties. One can achieve similar results, by modulating the *structure*. Instead of the optical properties depending on a periodic assembly of atoms or molecules that scatter light which then interferes, one can modify the structure of a material such that the scattering centres are local modifications of the material structure like, e.g., an array of holes in an otherwise unaltered optical material. Of course, the simplest example is again in 1D. For example, if one takes a waveguide and alters its structure such that it has a periodic array of small holes. This again results in a change of the optical response of the material because of the difference in how light is scattered from smaller holes compared to the scattering from larger holes.

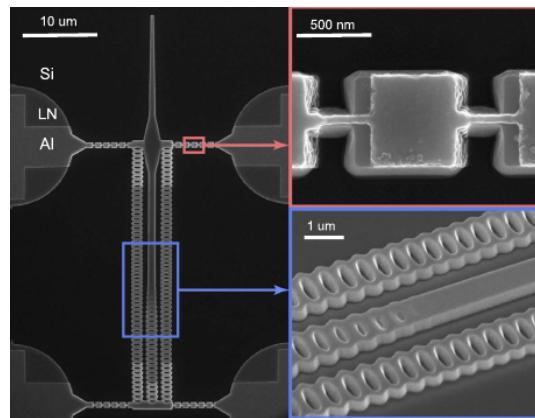


Figure 74: A photonic-crystal cavity as an optomechanical resonator. 1D photonic crystal structures formed in Lithium Niobate. The top right inset shows how the structure is isolated from vibrations by structuring the support struts. The figure is from Ref. [14].

Fig. 74 shows an example from a work by Jean et al. [14], where the authors combined multiple approaches that show the power of photonic crystal structures. In this particular example, the waveguide is free-standing over some distance, such that the interaction between light and the vibrational modes of this structure can be more easily addressed. At the same time, the regions where the waveguide contacts the remaining bulk material, is formed in such a way as to suppress the coupling of thermal vibrational modes. The material used is Lithium Niobate, a common piezoelectric material. That means, the mechanical motion of the structure also couples to electric fields. This renders

such structures interesting candidates for *transducers* between different forms of energy. For example, such devices could facilitate the coupling between microwave radiation and optical fields. This is of high interest in the context of quantum communication technologies because many of the currently most advanced candidates for a quantum computer operate in the microwave regime. At the same time, the most advanced form of long-distance quantum communication is via optical fibers. Achieving a platform to translate between these two regimes of interest promises to allow connecting distant quantum computers.

8.1.1 Electromagnetism in mixed dielectric media

Since photonic crystals are formed by varying the material properties of an optical medium, we have to describe the propagation of electromagnetic waves in a mixed dielectric medium using Maxwell's equations:

$$\begin{aligned} \nabla \cdot \mathbf{B} &= 0, & \nabla \times \mathbf{E} + \frac{\partial \mathbf{B}}{\partial t} &= 0, \\ \nabla \cdot \mathbf{D} &= \rho, & \nabla \times \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} &= \mathbf{J}. \end{aligned} \quad (8.1)$$

Here, \mathbf{E} and \mathbf{H} are the macroscopic electric and magnetic fields, \mathbf{D} and \mathbf{B} are the displacement and magnetic induction fields, and ρ and \mathbf{J} are the free charge and current densities[13]. Let us assume that our medium does not contain any charges or sources of light, such that $\rho = 0$ and $\mathbf{J} = 0$. The dielectric properties will vary as a function of the spatial coordinates, and there may be many different regions of homogeneous dielectric material properties. These variations do not necessarily have to be periodic.

For simplicity, we will neglect nonlinear effects, and we will assume that the material is isotropic and that there are no variations of the material on length scales small compared to the wavelength of light. Given these assumptions, we will have $\mathbf{D}(\mathbf{r}) = \epsilon_0 \epsilon(\mathbf{r}) \mathbf{E}(\mathbf{r})$. The relative permittivity $\epsilon(\mathbf{r}, \omega)$ can depend on the spatial coordinates and the optical frequency. For now, we will neglect material dispersion such that the relative permittivity is constant over the frequency range we are interested in. We will also assume that our material is transparent. Then $\epsilon(\mathbf{r})$ will be real and positive.

For the magnetic field, we can write $\mathbf{B}(\mathbf{r}) = \mu_0 \mu(\mathbf{r}) \mathbf{H}(\mathbf{r})$. In most cases, the relative magnetic permeability $\mu(\mathbf{r})$ is close to unity, and we can make the approximation $\mathbf{B}(\mathbf{r}) = \mu_0 \mathbf{H}(\mathbf{r})$. Then ϵ is the square of the refractive index n (in general: $n = \sqrt{\epsilon \mu}$), and Maxwell's equations 8.1 simplify to[13]:

$$\begin{aligned} \nabla \cdot \mathbf{H}(\mathbf{r}, t) &= 0, & \nabla \times \mathbf{E}(\mathbf{r}, t) + \mu_0 \frac{\partial \mathbf{H}(\mathbf{r}, t)}{\partial t} &= 0, \\ \nabla \cdot [\epsilon(\mathbf{r}) \mathbf{E}(\mathbf{r}, t)] &= 0, & \nabla \times \mathbf{H}(\mathbf{r}, t) - \epsilon_0 \epsilon(\mathbf{r}) \frac{\partial \mathbf{E}(\mathbf{r}, t)}{\partial t} &= 0. \end{aligned} \quad (8.2)$$

Even with this simplified approach one can describe many interesting phenomena one encounters in photonic crystals.

Because the Maxwell equations are linear, we can separate the time dependence from the spatial dependence and achieve a separation between time and space variables by expanding the fields into a set of harmonic modes. For now, we will consider only a sinusoidal variation in time:

$$\begin{aligned}\mathbf{H}(\mathbf{r}, t) &= \mathbf{H}(\mathbf{r})e^{-i\omega t}, \\ \mathbf{E}(\mathbf{r}, t) &= \mathbf{E}(\mathbf{r})e^{-i\omega t}.\end{aligned}\quad (8.3)$$

If we enter that into our simplified Maxwell's equations 8.2, we get the following two equations from the divergence equations:

$$\nabla \cdot \mathbf{H}(\mathbf{r}) = 0, \quad \nabla \cdot [\epsilon(\mathbf{r})\mathbf{E}(\mathbf{r})] = 0. \quad (8.4)$$

This means that there are no point sources or sinks of displacement fields or magnetic fields in our optical medium[13] and that we are dealing with transverse electromagnetic waves. For example, if we have $\mathbf{H}(\mathbf{r}) = \exp(i\mathbf{k} \cdot \mathbf{r})\mathbf{a}$, then equation 8.4 results in $\mathbf{a} \cdot \mathbf{k} = 0$.

The remaining Maxwell equations relate $\mathbf{E}(\mathbf{r})$ and $\mathbf{H}(\mathbf{r})$:

$$\begin{aligned}\nabla \times \mathbf{E}(\mathbf{r}) - i\omega\mu_0\mathbf{H}(\mathbf{r}) &= 0, \\ \nabla \times \mathbf{H}(\mathbf{r}) + i\omega\epsilon_0\epsilon(\mathbf{r})\mathbf{E}(\mathbf{r}) &= 0.\end{aligned}\quad (8.5)$$

One can combine the equations 8.5 into:

$$\nabla \times \left[\frac{1}{\epsilon(\mathbf{r})} \nabla \times \mathbf{H}(\mathbf{r}) \right] = \left(\frac{\omega}{c} \right)^2 \mathbf{H}(\mathbf{r}). \quad (8.6)$$

If we combine this **master equation** with equation 8.4, and if we know $\epsilon(\mathbf{r})$ describing the dielectric properties of our material, we can calculate $\mathbf{H}(\mathbf{r})$. Once we have $\mathbf{H}(\mathbf{r})$, we can calculate $\mathbf{E}(\mathbf{r})$ by using the Maxwell equations 8.5 to get:

$$\mathbf{E}(\mathbf{r}) = \frac{i}{\omega\epsilon_0\epsilon(\mathbf{r})} \nabla \times \mathbf{H}(\mathbf{r}). \quad (8.7)$$

The advantage of this approach is that $\mathbf{E}(\mathbf{r})$ will automatically fulfill the condition that the field vector needs to be perpendicular to the propagation direction because the divergence of a curl is always zero[13]. In principle, one can also start by determining $\mathbf{E}(\mathbf{r})$, but starting with $\mathbf{H}(\mathbf{r})$ can be more convenient[13].

The master equation 8.6 is an eigenvalue equation for $\mathbf{H}(\mathbf{r})$. We can write:

$$\hat{\Theta}\mathbf{H}(\mathbf{r}) = \left(\frac{\omega}{c} \right)^2 \mathbf{H}(\mathbf{r}), \quad (8.8)$$

where $\hat{\Theta}$ is a differential operator defined via the relation:

$$\hat{\Theta}\mathbf{H}(\mathbf{r}) = \nabla \times \left[\frac{1}{\epsilon(\mathbf{r})} \nabla \times \mathbf{H}(\mathbf{r}) \right]. \quad (8.9)$$

To get a more intuitive understanding of equation 8.9, consider the following: the eigenvectors $\mathbf{H}(\mathbf{r})$ will describe the spatial patterns of the harmonic modes, and ω^2/c^2 is proportional to the square of the frequencies of these modes. Because $\hat{\Theta}$ is a linear operator, any linear combination of solutions of equation 8.9 will also solve that equation. All solutions that differ only by a constant factor are effectively the same solution.

The mathematical description here is similar to that of quantum physics. Because the operator $\hat{\Theta}$ is Hermitian, there are some interesting similarities between the eigenmodes we are interested in here and the solutions to the Schröinger equation. In particular, the solutions of equation 8.9 will form an orthonormal set with real eigenvalues. Let us look at the inner product of two solutions \mathbf{F} and \mathbf{G} of equation 8.9:

$$(\mathbf{F}, \mathbf{G}) = \int d^3\mathbf{r} \mathbf{F}^*(\mathbf{r}) \cdot \mathbf{G}(\mathbf{r}), \quad (8.10)$$

similar to the scalar product of a quantum state $\langle \psi | \psi \rangle$. The scalar product of a solution with itself is always positive. Like in quantum physics, we can *normalize* solutions to equation 8.9 such that $(\mathbf{F}, \mathbf{F}) = 1$. This is, of course, only true if one is only interested in the spatial structure of a solution and not the overall energy contained in the field.

8.1.2 General properties of the harmonic modes

Because $\hat{\Theta}$ is Hermitian, its eigenvalues $(\omega/c)^2$ must be real numbers, and one can show that they must be positive. In particular, consider the following:

$$(\mathbf{H}, \mathbf{H}) \left(\frac{\omega}{c} \right)^2 = (\mathbf{H}, \hat{\Theta}\mathbf{H}) = \int d^3\mathbf{r} \frac{1}{\epsilon} |\nabla \times \mathbf{H}|^2. \quad (8.11)$$

Because we assumed earlier that $\epsilon(\mathbf{r}) > 0$ everywhere in our medium, we can conclude that the right-hand side of this equation is nonnegative. Then the operator $\hat{\Theta}$ is *positive semidefinite*, and therefore all eigenvalues ω^2 are nonnegative, and ω will be real.

Another important analogy with the case of quantum physics is that two solutions $\mathbf{H}_{1,2}$ to the eigenvalue equation 8.9 for two different frequencies ω_1 and ω_2 have to be orthogonal. To see this, consider the following:

$$\begin{aligned} \omega_1^2(\mathbf{H}_2, \mathbf{H}_1) &= c^2(\mathbf{H}_2, \hat{\Theta}\mathbf{H}_1) = c^2(\hat{\Theta}\mathbf{H}_2, \mathbf{H}_1) = \omega_2^2(\mathbf{H}_2, \mathbf{H}_1) \\ &\Rightarrow (\omega_1^2 - \omega_2^2)(\mathbf{H}_2, \mathbf{H}_1) = 0. \end{aligned} \quad (8.12)$$

That means, if $\omega_1 \neq \omega_2$, we must have $(\mathbf{H}_2, \mathbf{H}_1) = 0$, and \mathbf{H}_1 and \mathbf{H}_2 must be orthogonal.

If $\omega_1 = \omega_2$, the two modes are degenerate, and they do not have to be orthogonal. Because the frequency eigenspace is continuous, it would seem to be an extraordinary coincidence for two solutions to have the same frequency without having anything else in common. Typically, such degenerate solutions result from some sort of symmetry in the underlying material. For example, if the material is symmetric under a rotation of 120° , then the rotation of one solution by 120° will again give a solution at the same frequency. Because the operator $\hat{\Theta}$ is linear, any linear combination of such solution will again be a solution for the same frequency. However, we can also combine solutions of different frequencies, and we will again get a solution to the eigenvalue equation 8.9.

8.1.3 Electromagnetic energy and the variational principle

We know from classical mechanics and quantum physics that the solution given some Hamiltonian can be found by minimizing some functional using a variational theorem. Something equivalent is the case for solutions of equation 8.9: the smallest eigenvalue ω_0^2/c^2 , and thus the lowest-frequency mode, corresponds to the field pattern that minimizes the following functional[13]:

$$U_f(\mathbf{H}) = \frac{(\mathbf{H}, \hat{\Theta}\mathbf{H})}{(\mathbf{H}, \mathbf{H})}. \quad (8.13)$$

That means, ω_0^2/c^2 is the minimum of $U_f(\mathbf{H})$ taken over all possible field patterns \mathbf{H} . By “possible” we mean that they have to fulfill the transversality condition $\nabla \cdot \mathbf{H} = 0$. The functional U_f is also known as the *Rayleigh quotient* or the electromagnetic “energy” functional. The latter term is used to stress the similarity with quantum and classical mechanics where some physical energy is minimized by the solution.

One can show that if some function \mathbf{H} minimizes $U_f(\mathbf{H})$ for every variation $\delta\mathbf{H}$ of \mathbf{H} , then \mathbf{H} must be an eigenfunction of $\hat{\Theta}$. The eigenfunction \mathbf{H} with the lowest eigenvalue ω_0 will minimize U_f .

We can conclude another interesting fact from studying U_f . For that purpose, let us rewrite it in the following form:

$$U_f(\mathbf{H}) = \frac{(\nabla \times \mathbf{E}, \nabla \times \mathbf{E})}{(\mathbf{E}, \epsilon(\mathbf{r})\mathbf{E})} = \frac{\int d^3\mathbf{r} |\nabla \times \mathbf{E}(\mathbf{r})|^2}{\int d^3\mathbf{r} \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r})|^2}. \quad (8.14)$$

If we look more closely at the denominator in the last equation, we can see that U_f will be minimal if the electric field intensity is concentrated at those positions where we have a high dielectric constant ϵ .

8.1.4 The effect of small perturbations

In the beginning of this chapter, we made assumptions to simplify the full Maxwell equations to the simpler form of equations 8.2. One can extend the validity of that simplified approach by using perturbation theory for linear Hermitian eigenproblems[13].

To this end, let us assume we have a Hermitian operator \hat{O} , and that we alter it by some small amount $\Delta\hat{O}$. Then the resulting eigenvectors and eigenvalues of the perturbed operator can be written as a series expansion in increasing powers of the perturbation strength $\Delta\hat{O}$, and one can solve the resulting eigenequations order by order in $\Delta\hat{O}$.

In our particular case, we will be interested in small changes $\Delta\epsilon(\mathbf{r})$ to our dielectric function. Entering that into the master equation 8.6 proves to be inconvenient, and it turns out to be easier to instead use an equivalent master equation in terms of the electric field:

$$\nabla \times \nabla \times \mathbf{E}(\mathbf{r}) = \left(\frac{\omega}{c}\right)^2 \epsilon(\mathbf{r}) \mathbf{E}(\mathbf{r}). \quad (8.15)$$

As an eigenvalue equation, this was not convenient because there are operators on both sides, so that this is a generalized eigenproblem. In the context of perturbations, this proves to be useful. In particular, this allows us to get a simple expression for the frequency shift resulting from a small perturbation $\Delta\epsilon$:

$$\Delta\omega = -\frac{\omega}{2} \frac{\int d^3\mathbf{r} \Delta\epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r})|^2}{\int d^3\mathbf{r} \epsilon(\mathbf{r}) |\mathbf{E}(\mathbf{r})|^2} + \mathcal{O}(\Delta\epsilon^2). \quad (8.16)$$

In many practical cases, $|\Delta\epsilon|/\epsilon$ is 1% or smaller, which allows us to neglect the second order term of the approximation.

Let us consider an instructive example. Suppose that we have a material with a refractive index $n = \sqrt{\epsilon}$, and that this refractive index is disturbed in some regions by an amount Δn . Then the integral in the numerator of equation 8.16 has non-zero contributions only in the perturbed regions. If we write $\Delta\epsilon \approx \epsilon \cdot 2\Delta n/n$, and if we assume that $\Delta n/n$ is the same in all the perturbed regions, then we get:

$$\frac{\Delta\omega}{\omega} \approx -\frac{\Delta n}{n} \cdot \eta, \quad (8.17)$$

η is the fraction of the field intensity in the perturbed regions. That means, that the fractional change in frequency is approximately equal to the fractional change of the refractive index times the field intensity found inside the disturbed regions. The negative sign appears because an increase in the refractive index lowers the mode frequencies, and that reduces the field energy[13].

This perturbative approach can also be used to take into account absorption or a (small) gain. Losses or gain of course are not compatible with a Hermitian operator, but in the perturbation approach, we only require that the unperturbed operator $\hat{\Theta}$

be Hermitian, while the small perturbative change does not need to be. For example, a small perturbation $\Delta\epsilon = i\delta$ will lead to a small imaginary $\Delta\omega = -i\gamma/2$, where $\gamma = \omega \int |\mathbf{E}|^2 \delta / \int \epsilon |\mathbf{E}|^2$. This will lead to an exponential decay of the field in time as $e^{-\gamma t/2}$. A small gain can be implemented by using a negative value for γ , but an exponential growth will quickly stop being a small perturbation.

8.1.5 Scaling properties of the Maxwell equations

While we noted some similarities between electromagnetic waves in photonic crystals and the situation in quantum mechanics, there also is an important difference. While Planck's constant results in a fundamental length scale in quantum physics, this is not the case for electromagnetic waves in photonic crystals. The master equation is *scale invariant*. As a result, there are simple relations between electromagnetic problems that only differ by a scaling of all distances[13].

For example, assume that $\mathbf{H}(\mathbf{r})$ is an electromagnetic eigenmode for the frequency ω :

$$\nabla \times \left[\frac{1}{\epsilon(\mathbf{r})} \nabla \times \mathbf{H}(\mathbf{r}) \right] = \left(\frac{\omega}{c} \right)^2 \mathbf{H}(\mathbf{r}). \quad (8.18)$$

Let us try to find the harmonic modes for a scaled dielectric $\epsilon'(\mathbf{r})$ fulfilling $\epsilon'(\mathbf{r}) = \epsilon(\mathbf{r}/s)$ with some scaling factor s . If we change the variables in equation 8.18 such that $\mathbf{r}' = s\mathbf{r}$ and $\nabla' = \nabla/s$, we get:

$$s \nabla' \times \left[\frac{1}{\epsilon'(\mathbf{r}'/s)} s \nabla' \times \mathbf{H}(\mathbf{r}'/s) \right] = \left(\frac{\omega}{c} \right)^2 \mathbf{H}(\mathbf{r}'/s). \quad (8.19)$$

But we know that $\epsilon(\mathbf{r}'/s) = \epsilon'(\mathbf{r}')$. We can therefore rewrite equation 8.19 as:

$$\nabla' \times \left[\frac{1}{\epsilon'(\mathbf{r}')} \nabla' \times \mathbf{H}(\mathbf{r}'/s) \right] = \left(\frac{\omega}{cs} \right)^2 \mathbf{H}(\mathbf{r}'/s). \quad (8.20)$$

This is the same as our original master equation but with the scaled mode profile $\mathbf{H}'(\mathbf{r}') = \mathbf{H}(\mathbf{r}'/s)$ and a scaled frequency $\omega' = \omega/s$ [13].

A practical use is that it can be difficult to fabricate photonic crystals for short optical wavelengths. For testing purposes, one can first fabricate the photonic crystal on a larger length scale and test it using microwaves or THz waves.

8.1.6 Analytic vs numerical solutions

One particular difference between quantum mechanics and electromagnetism does not play out in our favour. In quantum mechanics, the Hamiltonian will be separable if the potential is separable. Such a factorization is typically not possible in the case of electromagnetic fields because $\hat{\Theta}$ will couple the different spatial variables. For that

reason, it is usually not possible to find an analytic solution to the master equation, and one has to make use of numeric solutions.

8.1.7 Discrete versus continuous frequency ranges

From quantum physics, we are familiar with the observation that boundary conditions will influence the eigenfrequency spectrum. If a particle is unconfined, its energy can take on any continuous value, but we get a discrete energy spectrum if a particle's motion is confined. Often, we have a combination of both cases. In photonic crystals, the situation is similar. For example, our modifications to the geometry of our optical material may be limited to a certain volume. The energies of the eigenmodes within that volume will be discrete, but the energies of modes outside that volume may take on continuous values.

8.2 Symmetries and solid-state electromagnetism

The structures in photonic crystals are often repetitive. This can be arrays of holes or layers, but it can also be repetitions of complex structures. This is similar to solid-state physics, where the unit cell of a crystal is repeated many times. As in solid-state physics, symmetries play an important role in the context of photonic crystals. In the following, we will discuss several symmetries, how to utilize them, and how such symmetries result in or affect *band gaps*.

8.2.1 Inversion symmetry

An important example of a symmetry is inversion symmetry. In order to formally define what we mean by that, let us first introduce an operator I that inverts vectors such that $I\mathbf{a} = -\mathbf{a}$ ¹. If we want to invert a vector field $\mathbf{f}(\mathbf{r})$, we need to use an operator \hat{O}_I that inverts the both the vector and its argument. That means: $\hat{O}_I \mathbf{f}(\mathbf{r}) = I\mathbf{f}(I\mathbf{r})$. If our system is symmetric under inversion, than we must have:

$$\hat{\Theta} = \hat{O}_I^{-1} \hat{\Theta} \hat{O}_I. \quad (8.21)$$

We can rewrite that as a commutator relation:

$$[\hat{O}_I, \hat{\Theta}] = 0. \quad (8.22)$$

¹ I will be a 3×3 matrix.

Because the two operators commute, if an eigenmode \mathbf{H} solves the eigenproblem 8.9 with eigenfrequency ω , then $\hat{O}_I \mathbf{H}$ will also be an eigenmode with frequency ω because:

$$\hat{\Theta}(\hat{O}_I \mathbf{H}) = \hat{O}_I(\hat{\Theta} \mathbf{H}) = \frac{\omega^2}{c^2}(\hat{O}_I \mathbf{H}). \quad (8.23)$$

Since the $\hat{O}_I \mathbf{H}$ and \mathbf{H} are eigenmodes of the same frequency ω , they can only differ by some factor α . By applying the inversion twice, one can show that we must have $\alpha = \pm 1$. Depending on the sign, we can distinguish between even or odd inversion symmetry.

8.2.2 Continuous translational symmetry

An important symmetry for photonic crystals is translational symmetry. For example, if a 1D photonic crystal consists of an array of identical holes in a waveguide with a constant relative distance a , we will have a discrete translational symmetry under a translation by a . In many cases, however, we will have continuous translational symmetries. That will be the case if the dielectric function is constant in some volume, or with respect to translation along one or more dimensions. For example, in the case of multiplayer films illustrated in Fig. 73, the dielectric function changes from layer to layer, but it will not vary in the other two dimensions.

Let us assume our system remains unchanged by a translation by a displacement \mathbf{d} , which is described by the translation operator $\hat{T}_{\mathbf{d}}$. If this operator is applied to a function $f(\mathbf{r})$, its argument is shifted by \mathbf{d} . If our system is invariant under such a translation, we will have:

$$\hat{T}_{\mathbf{d}}\epsilon(\mathbf{r}) = \epsilon(\mathbf{r} - \mathbf{d}) = \epsilon(\mathbf{r}). \quad (8.24)$$

In terms of the commutator, that means $[\hat{T}_{\mathbf{d}}, \hat{\Theta}] = 0$.

A mode of the form e^{ikz} is an eigenfunction of the translation operator along z :

$$\hat{T}_{d\hat{\mathbf{z}}}e^{ikz} = e^{ik(z-d)} = (e^{-ikd})e^{ikz}. \quad (8.25)$$

That means, it is an eigenmode with the eigenvalue e^{-ikd} . More generally, any eigenfunction of $\hat{T}_{\mathbf{d}}$ for $\mathbf{d} = d\hat{\mathbf{z}}$ must be proportional to e^{ikz} for some k .

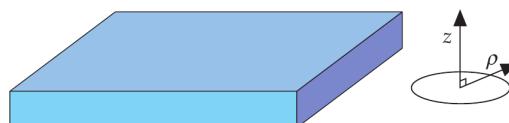


Figure 75: **An infinite plane of glass.** If the medium extends much further in the transverse direction than in the z direction, we can consider the situation as a 1D problem in the z direction. The figure is from Ref. [13].

If we have a system that is invariant under translations in *all* spatial dimensions, $\epsilon(\mathbf{r})$ must be a constant ϵ , and the eigenmodes must be *plane waves*:

$$\mathbf{H}(\mathbf{r}) = \mathbf{H}_0 e^{i\mathbf{k}\cdot\mathbf{r}}, \quad (8.26)$$

where \mathbf{H}_0 is a constant vector. Because of the transversality condition, we will get the restriction $\mathbf{k} \cdot \mathbf{H}_0 = 0$, and our plane waves have to satisfy the usual dispersion relation $\omega = c|\mathbf{k}|/\sqrt{\epsilon}$.

As another simple example, assume we have an infinite plane of glass or some other dielectric material as shown in Fig. 75. We will assume that the dielectric function is constant along the xy plane but not along the z direction, and that the medium extends much further in those dimensions than in the z direction. Let us introduce the vector $\rho = x\hat{x} + y\hat{y}$ and the in-plane wave vectors $\mathbf{k} = k_x\hat{x} + k_y\hat{y}$.

Because of the translation symmetry in the xy plane, the eigenmodes must have the form:

$$\mathbf{H}_{\mathbf{k}}(\mathbf{r}) = e^{i\mathbf{k}\cdot\rho} \mathbf{h}(z). \quad (8.27)$$

The function $\mathbf{h}(z)$ depends on \mathbf{k} and on the properties of the dielectric medium in the z direction. The transversality condition implies a restriction on \mathbf{h} : from $\nabla \cdot \mathbf{H}_{\mathbf{k}} = 0$ it follows that $\mathbf{k} \cdot \mathbf{h} = i\partial h_z / \partial z$.

While we cannot say much about $\mathbf{h}(z)$, we can classify the possible modes by their values of \mathbf{k} : for a given value of \mathbf{k} , we can line the modes up according to increasing frequencies. We use increasing values of n to enumerate the frequencies, such that every mode can be represented by a unique description (\mathbf{k}, n) . If we have multiple degenerate modes, we can introduce an additional index to distinguish them.

The index n is called the *band number*. If the spectrum is discrete for a given \mathbf{k} , we can use integer values for n , but the band number can also be a continuous variable. As n grows, so does the value of the corresponding frequency. If we make a plot, where we plot the mode frequency over the wave vector for our plane of glass, the different bands correspond to different lines that rise for increasing an increasing wavenumber. Fig. 76 shows that band strucutre. The authors of Ref. [13] computed it numerically from the master equation 8.6. In the following, we will discuss this example in a bit more detail because it also provides insight into another important phenomenon in photonic crystals: *index guiding*.

8.2.3 Index guiding

Fig. 76 shows that we have a continuous band of modes if the wavevector is sufficiently flat such that the light can travel between the two confining surfaces of the glass plate, reflecting back up and down due to total internal reflection. To describe the situation, let us split the wave vector \mathbf{k} into parts \mathbf{k}_{\parallel} and \mathbf{k}_{\perp} , that are parallel or perpendicular to the surface of the glass plane. If θ is the angle between \mathbf{k} and the normal to our glass plate, then $\mathbf{k}_{\parallel} = |\mathbf{k}|\sin\theta$. The critical angle for total internal reflection is $\theta_c =$

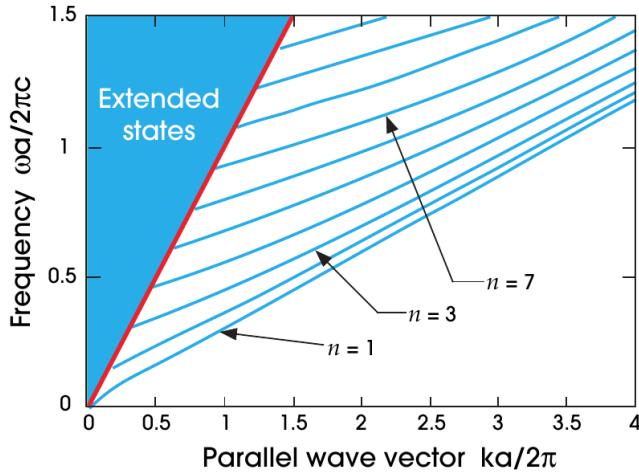


Figure 76: **Band structure for a plane of glass.** Mode frequencies for a plane of glass of thickness a and $\epsilon = 11.4$. The blue lines indicate modes that are localized in the glass. The shaded blue region represents a continuum of states that extend both in the glass and in the surrounding air. The red line represents the *light line* with $\omega = ck$. The plot only shows modes for one particular polarization, for which \mathbf{H} is perpendicular to the z and k directions. The figure is from Ref. [13].

$\sin^{-1}(n_2/n_1)$, where n_2 and n_1 are the refractive indices outside and inside the glass plane, respectively. Let a be the thickness of the glass.

For each mode crossing the glass surfaces, the frequency ω and the parallel part of the wavevector are conserved. Unconfined modes need to resemble free-space plane waves when they are far from the interfaces. They will be superpositions of plane waves fulfilling $\omega = c|\mathbf{k}| = c\sqrt{k_{\parallel}^2 + k_{\perp}^2}$ for some perpendicular wave vector component \mathbf{k}_{\perp} . For any given value of k_{\parallel} , there will be modes with every possible frequency greater than ck_{\parallel} because k_{\perp} can take any value[13]. For that reason, the spectrum of states will be continuous for frequencies beyond the line that is denoted as . The region enclosed by the *light line* $\omega = ck_{\parallel}$ in Fig. 76 is also called the *light cone*.

8.2.4 Discrete translational symmetry

As a simple example let us assume periodicity in one direction as shown in Fig. 77. The structure repeats after a basic step a in the y direction, but it is homogeneous in the x direction. a is the *lattice constant*, and $\mathbf{a} = a\hat{\mathbf{y}}$ is the *primitive lattice vector*. Because of this symmetry, we have $\epsilon(\mathbf{r}) = \epsilon(\mathbf{r} + \mathbf{R})$ if $\mathbf{R} = l\mathbf{a}$ with some integer l .

The eigenmodes of $\hat{\Theta}$ must be simultaneous eigenfunctions of the continuous translation operator along x and the discrete translation operator along y . In particular, plane

waves will be eigenfunctions:

$$\begin{aligned}\hat{T}_{d\hat{\mathbf{x}}} e^{ik_x x} &= e^{ik_x(x-d)} = (e^{-ik_x d}) e^{ik_x x} \\ \hat{T}_{\mathbf{R}} e^{ik_y y} &= e^{ik_y(y-la)} = (e^{-ik_y la}) e^{ik_y y}.\end{aligned}\quad (8.28)$$

Not all values of k_y yield different eigenvalues (frequencies). In particular, adding an integer multiple of $b = 2\pi/a$ to k_y leaves the eigenstate unchanged. The vector $\hat{\mathbf{b}} = b\hat{\mathbf{y}}$ is known as the primitive *reciprocal lattice vector*. Any linear combination of such degenerate eigenfunctions will again yield a solution with the same eigenfrequency. We can write them in the following form [13]:

$$\begin{aligned}\mathbf{H}_{k_x, k_y}(\mathbf{r}) &= e^{ik_x x} \sum_m \mathbf{c}_{k_y, m}(z) e^{i(k_y + mb)y} \\ &= e^{ik_x x} \cdot e^{ik_y y} \sum_m \mathbf{c}_{k_y, m}(z) e^{imb y} \\ &= e^{ik_x x} \cdot e^{ik_y y} \cdot \mathbf{u}_{k_y}(y, z).\end{aligned}\quad (8.29)$$

Here, we introduced the expansion coefficients $\mathbf{c}_{k_y, m}$, and $\mathbf{u}(y, z)$ is a periodic function in y such that $\mathbf{u}(y + la, z) = \mathbf{u}(y, z)$.

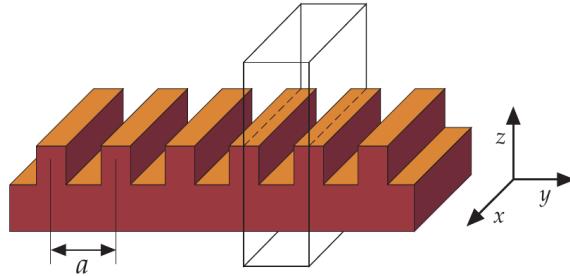


Figure 77: **An optical medium with a discrete translational symmetry in the y direction.** This example shows a structure that repeats itself after a distance a in the y direction, and it is assumed that the system has continuous translational symmetry along x . The figure is from Ref. [13].

That means, that the discrete translational symmetry along y will result in solutions \mathbf{H} that are proportional to a product of a plane wave with a y -periodic function:

$$\mathbf{H}(\dots, y, \dots) \propto e^{ik_y y} \cdot \mathbf{u}_{k_y}(y, \dots). \quad (8.30)$$

This is known as *Bloch's theorem*, which is well known from solid-state physics[13], and a state \mathbf{H} that fulfills it is known as a *Bloch state*.

This can be generalized to the three-dimensional case where the dielectric is invariant

under translations by multiples of a lattice vector \mathbf{R} . We can write any lattice vector as a combination of primitive lattice vectors: $\mathbf{R} = l\mathbf{a}_1 + m\mathbf{a}_2 + n\mathbf{a}_3$. These primitive lattice vectors will give rise to *reciprocal* lattice vectors $(\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3)$ such that $\mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}$. These vectors span a *reciprocal lattice* of wavevectors.

The corresponding Bloch states can be labelled by a *Bloch wave vector* $\mathbf{k} = k_1\mathbf{b}_1 + k_2\mathbf{b}_2 + k_3\mathbf{b}_3$. \mathbf{k} lies in the Brioullin zone, which is determined by the primitive unit cell. Each value of a wave vector \mathbf{k} inside the Brioullin zone identifies an eigenstate of $\hat{\Theta}$ with frequency $\omega(\mathbf{k})$ and an eigenvector of the following form[13]:

$$\mathbf{H}_{\mathbf{k}} = e^{i\mathbf{k}\cdot\mathbf{r}}\mathbf{u}_{\mathbf{k}}(\mathbf{r}). \quad (8.31)$$

Here, $\mathbf{u}_{\mathbf{k}}(\mathbf{r}) = \mathbf{u}_{\mathbf{k}}(\mathbf{r} + \mathbf{R})$ is a periodic function of the lattice, with lattice vectors \mathbf{R} .

8.2.5 Photonic band structures

The band structure of a photonic crystal (as illustrated in Fig. 76) determines the optical properties of a photonic crystal. There are many numeric tools for calculating the band structure. The rough idea behind these calculations is based on solving a standard eigenvalue equation that one can solve by iterative minimization techniques for each value of \mathbf{k} . This eigen equation has a slightly different form than our original master equation 8.6. The advantage of this description is that it makes it easier to take advantage of symmetries in the crystal.

To get this new equation, we insert a general Bloch vector as in equation 8.31 into our master equation 8.6. If we define a new Hermitian operator depending on \mathbf{k} :

$$\hat{\Theta}_{\mathbf{k}} \equiv (i\mathbf{k} + \boldsymbol{\nabla}) \times \frac{1}{\epsilon(\mathbf{r})}(i\mathbf{k} + \boldsymbol{\nabla}) \times, \quad (8.32)$$

then we can rewrite our master equation as:

$$\hat{\Theta}_{\mathbf{k}}\mathbf{u}_{\mathbf{k}}(\mathbf{r}) = [\omega(\mathbf{k})/c]^2\mathbf{u}_{\mathbf{k}}(\mathbf{r}). \quad (8.33)$$

The functions $\mathbf{u}_{\mathbf{k}}$ are determined by the eigenvalue problem of equation 8.33, and they have to fulfill the transversality condition $(i\mathbf{k} + \boldsymbol{\nabla}) \cdot \mathbf{u}_{\mathbf{k}} = 0$ and the periodicity condition $\mathbf{u}_{\mathbf{k}}(\mathbf{r}) = \mathbf{u}_{\mathbf{k}}(\mathbf{r} + \mathbf{R})$.

Because the solution is periodic, it will be sufficient to find the solution in a single unit cell of the photonic crystal. We expect to find an infinite set of discretely spaced frequencies, which we can then label by a band index n [13].

Just a quick example why symmetries are important: if we have a square lattice, e.g., a 2D array of holes in a medium, such a lattice often exhibits a rotational symmetry. In such cases, we do not have to take into account the full Brioullin zone because parts of it will be redundant due to the symmetries. The smallest region for which the $\omega_n(\mathbf{k})$ are not related by symmetry is called the *irreducible Brioullin zone*[13].

8.3 Multilayer films

We already mentioned multilayer films as a simple example for (1D) photonic crystals. A schematic drawing of such multilayer films can be seen in Fig. 73. Here, we will investigate this example in more detail. Similar to Bragg gratings, the interference between partially reflected waves from different layers can exhibit constructive interference for wavelengths within a certain frequency band close to the Bragg frequency. This can be used to realize *Bragg mirrors* or dielectric mirrors or optical filters. The usual approach to treat such Bragg mirrors is to sum the partially reflected waves to see under which conditions one achieves constructive interference. Here, we will instead use the methods we developed to describe photonic crystals, and we will describe the properties of multilayer films by analyzing their band structure. We will later generalize this approach to higher-dimensional photonic crystals.

Due to the periodic symmetry in the z direction and the transverse direction, the modes have to be of Bloch form:

$$\mathbf{H}_{n,k_z,\mathbf{k}_\parallel}(\mathbf{r}) = e^{i\mathbf{k}_\parallel \cdot \boldsymbol{\rho}} e^{ik_z z} \mathbf{u}_{n,k_z,\mathbf{k}_\parallel}(z). \quad (8.34)$$

Here, we introduced the periodic function $\mathbf{u}(z) = \mathbf{u}(z+R)$, where R is an integer multiple of the period a . Because our medium is homogeneous in the transverse direction, we also get a plane wave factor depending on the transverse offset $\boldsymbol{\rho}$. The wavenumber k_z can be restricted to the Brillouin zone. In particular the primitive reciprocal lattice vector is $(2\pi/a)\hat{\mathbf{z}}$, and the Brillouin zone is $-\pi/a < k_z \leq \pi/a$.

8.3.1 The physical origin of photonic band gaps

If the light travels along the z axis, perpendicular to the multilayer films, we have $\mathbf{k}_\parallel = 0$, and only k_z is important. For the moment, we can assume that k is equal to k_z . The periodicity of the multilayer film will result in a photonic bandgap. Fig. 78 shows such bandgaps for different choices of material parameters. The photonic bandgap resulting becomes more pronounced for a larger difference between the values of the dielectric constants in the different layers.

The left-hand panel of that plot shows the propagation of light in a bulk material. In this case, a is arbitrary, and the plot shows the normal dispersion relation in bulk material:

$$\omega(k) = \frac{ck}{\sqrt{\epsilon}}. \quad (8.35)$$

The slope of the straight line is reduced by a factor of $\sqrt{\epsilon}$.

Because we required earlier that k repeats itself outside the Brioullin zone, the light line folds back into the Brioullin zone instead of leaving it. This results in the pattern of lines in Fig. 78 because $k + 2\pi/a$ is replaced by k . The center plot shows a nearly

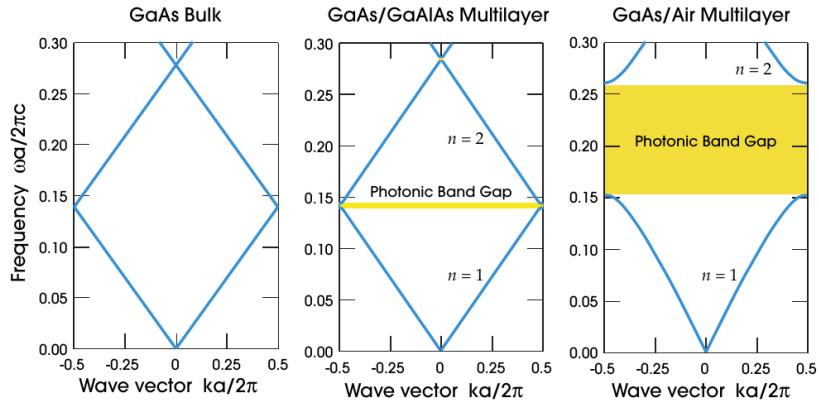


Figure 78: **Bandgaps in multilayer films.** Photonic band gap structures for on-axis propagation for three different choices of materials. (left) if all the layers consist of GaAs with the same material properties, there is no band-gap, and the dispersion relation will be as in bulk material. (middle) this shows the case if there is the dielectric function varies between values of 12 and 13 between the layers. This corresponds to a stack of GaAs and GaAlAs layers. In this case, we have a narrow band gap. (right) in the extreme case, where the dielectric function varies between values of 13 and 1, the photonic band gap is much wider. The figure is from Ref. [13].

homogeneous material because the dielectric constant does not change by much from layer to layer. Even there, we see a small bandgap appearing. That means, there are some frequencies for which there are no solutions of our eigenvalue problem defining the propagation inside our photonic crystal. This frequency gap is called a *photonic band gap*. This gap widens if we increase the difference between the dielectric constants of the different layers.

Let us now discuss what such band gaps mean and why they appear in more detail. For example, light with a frequency in the band gap will not simply disappear - it will be reflected. By placing two such photonic crystals behind each other, one could therefore form a very good cavity.

But why is there a bandgap? From Fig. 78, we can see that the bandgap between frequency bands 1 and 2 appears right at the edge of the Brillouin zone (close to $k = \pi/a$). For that k , the wavelength will be close to $2a$ (twice the crystal's period). In order to keep the symmetry of the problem, there are only two ways how the nodes of these waves can be centered on the various layers - either the node is at the center of the high- ϵ layer or at the center of the low- ϵ layer. When we discussed the variational principle in the context of photonic crystals (see section 8.1.3), we said that the energy functional would be minimized if the field were concentrated in regions of high ϵ . This

is true for low-frequency fields. For high-frequency fields, where the integrands in the numerator and the denominator in equation 8.14 have many nodes and antinodes, the numerator may become sufficiently large to counter-balance the larger denominator in high-density regions. For that reason, high-frequency fields will have a larger fraction (but not necessarily the majority) in the low- ϵ regions.

It is often the case that the low- ϵ regions are air regions (similar to the case illustrated in the right-hand panel of Fig. 78). For that reason, one often refers to the second band as the *air band* and to the first band as the *dielectric band* - even if the low- ϵ layer does not consist of air. This is analogous to the *conduction band* and the *valence band* in semiconductors.

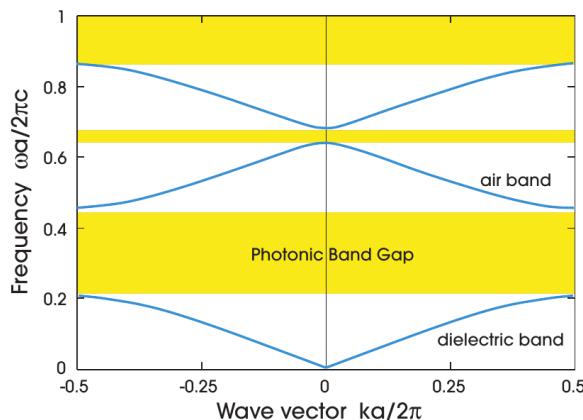


Figure 79: **Multiple bandgaps in multilayer films.** Bandgaps appear between all adjacent frequency bands. The figure is from Ref. [13].

The above argument that the frequency band appears at the edge of the Brioullin zone applies for the cases where there is only little difference between the dielectric constants of the various layers. If this difference becomes larger, the field energy of *both bands* is mostly concentrated in the high- ϵ regions. However, the distribution will be different. The lower energy band will be concentrated more on the high- ϵ zones than the higher-frequency fields. This is the origin of the band gap.

As we just saw, if the difference between the dielectric constants is small, the bandgap will appear close to the edge of the Brioullin zone. If the difference between the dielectric constants grows larger, the bandgap can also appear at the center of the Brioullin zone. Fig 79 shows the appearance of multiple bandgaps, some at the edge of the Brioullin zone, others at the center of the Brioullin zone.

8.3.2 The size of band gaps

To describe the size of bandgaps, let us introduce a dimensionless quantity that is independent of scaling: the *gap-midgap ratio*[13]. If we assume ω_m to denote the center

frequency of the band gap and $\Delta\omega$ to describe the width of the gap, then the gap-midgap ratio is $\Delta\omega/\omega_0$, and it is often expressed in percent. To present band gaps in a way that is independent of scaling, the photonic gap plots in Figures 78 and 79 are given in dimensionless units $\omega a/2\pi c$ and $ka/2\pi$.

While many of the concepts we are introducing here can be extended to 2D or 3D, there are some useful analytical results that can only be derived in 1D. For example, if the multilayer film has a *weak* difference between the dielectric constants in the periodic structure, one can make use of perturbation theory. Let us assume that one layer has a dielectric constant ϵ and the other a dielectric constant $\epsilon + \Delta\epsilon$, and that one layer has thickness d and the other $a - d$. If the contrast is weak ($\Delta\epsilon/\epsilon \ll 1$) *or* if the thickness d/a is small, then one can show that the gap-midgap ratio for the gap between the first two bands will be[13]:

$$\frac{\Delta\omega}{\omega_m} \approx \frac{\Delta\epsilon}{\pi\epsilon} \sin(\pi d/a). \quad (8.36)$$

From this we can learn two things: (1) there is a non-zero gap for any non-zero $\Delta\epsilon$, and (2) we can estimate the gap for the middle panel of Fig. 78. In that case, we have $\Delta\epsilon/\epsilon = 1/12$ and $d = a/2$. Then equation 8.36 predicts a 2.65% gap while a more precise calculation yields the predictions of a 2.55% gap[13].

For more general 1D cases, one can get an analytical prediction for arbitrary $\Delta\epsilon$, which we will not derive here. A derivation can be found in Ref. [15]. If the two layers have refractive indices $n_{1,2}$ and thicknesses $d_{1,2}$ such that $a = d_1 + d_2$, the gap for normal-incidence will be maximal if $d_1 n_1 = d_2 n_2$. For this specific case, the midgap frequency is[13]:

$$\omega_m = \frac{2\pi c(n_1 + n_2)}{4an_1n_2}. \quad (8.37)$$

If one looks at the corresponding vacuum wavelength $\lambda_m = 2\pi c/\omega_m$, one gets $\lambda_m/n_1 = 4d_1$ and $\lambda_m/n_2 = 4d_2$. That means, the individual layers all have a thickness of a quarter-wavelength. Such multilayer films are therefore known as *quarter-wave stacks*.

The gap-midgap ratio for such a quarter-wave stack is[13]:

$$\frac{\Delta\omega}{\omega_m} = \frac{4}{\pi} \sin^{-1} \left(\frac{|n_1 - n_2|}{n_1 + n_2} \right). \quad (8.38)$$

8.3.3 Evanescent modes in photonic band gaps

Similar to the tunneling effect in quantum physics, light with a frequency in the band gap is not immediately reflected at the surface, but it can enter the crystal to some extent. To be more precise, for frequencies in the bandgap no eigenmode exists with a *real* wavevector corresponding to that frequency. Instead, the wavevector will be complex, and the amplitude will decay exponentially as it travels through the photonic crystal unless it encounters a **defect** along the way. This can lead to a localized state. If

something like that occurs at the surface of the crystal, one calls the resulting localized state a **surface state**. States with frequencies in the band gap are typically not *extended* states but decay only exponentially. Such modes are called *evanescent*. For our example of a multilayer film periodic in the z direction, we can write such a mode as:

$$\mathbf{H}(\mathbf{r}) = e^{ikz}\mathbf{u}(z)e^{-\kappa z}. \quad (8.39)$$

The wave will decay on a length scale given by $1/\kappa$.

To estimate κ , one has to take a closer look at the frequency bands surrounding the band gap and take into account *time-reversal symmetry*. If we have a loss-less material, one can show that:

$$\omega_n(\mathbf{k}) = \omega_n(-\mathbf{k}), \quad (8.40)$$

but this relation actually holds for nearly all photonic crystals. A noteworthy exception are magneto-optic materials[13].

To describe evanescent fields, consider the dispersion relation $w_2(k)$ in the second frequency band and expand it in powers of k close to the edge of the Brillouin zone $k = \pi/a$. Because of the time-reversal symmetry, this expansion cannot contain odd powers of k , and we get the following to the lowest order in Δk :

$$\Delta\omega = \omega_2(k) - \omega_2\left(\frac{\pi}{a}\right) \approx \alpha \left(k - \frac{\pi}{a}\right)^2 = \alpha \Delta k^2, \quad (8.41)$$

where α depends on the curvature (the 2nd derivative) of the frequency band.

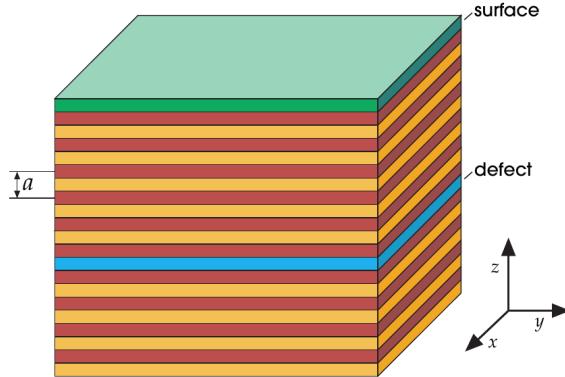


Figure 80: **Localized states in a 1D photonic crystal.** We distinguish between defect states (blue) and surface states (green). The figure is from Ref. [13].

Equation 8.41 shows that Δk will be real if $\Delta\omega \geq 0$, but it will be complex if $\Delta\omega$ is negative. That means, we get imaginary solutions for states within the bandgap and real solutions for states right above the gap. Typically, a larger gap will result in a larger κ . In our 1D situation, a quarter-wave stack will therefore have the minimum

penetration. If a photonic crystal is perfect and extends to infinity, all evanescent modes will eventually decay completely. In reality, there will be defects, and often it is desirable to design a defect on purpose such that evanescent light can excite a localized mode at that defect. In some special cases, it is also possible to excite modes on the surface of the material. These two situations are illustrated in Fig. 80.

8.3.4 Off-axis propagation

So far, we have assumed that the light travels along the z axis, perpendicular to photonic-crystal layers ($\mathbf{k}_{\parallel} = 0$). Now, we will briefly discuss how the situation changes if there are off-axis components to our wave vector. To discuss the fundamental differences between on-axis and off-axis propagation, let us first compare the two extreme cases with $\mathbf{k} = k_z \hat{\mathbf{z}}$ and $\mathbf{k} = k_y \hat{\mathbf{y}}$. There will not be any frequency bands appearing for light travelling along the y axis. However, as we will see, the dependence of ω on k will depend on the polarization of the light. This is a strong contrast to the on-axis case because there the modes for different polarizations will be degenerate due to rotation symmetry around the z axis. This symmetry is broken for off-axis propagation.

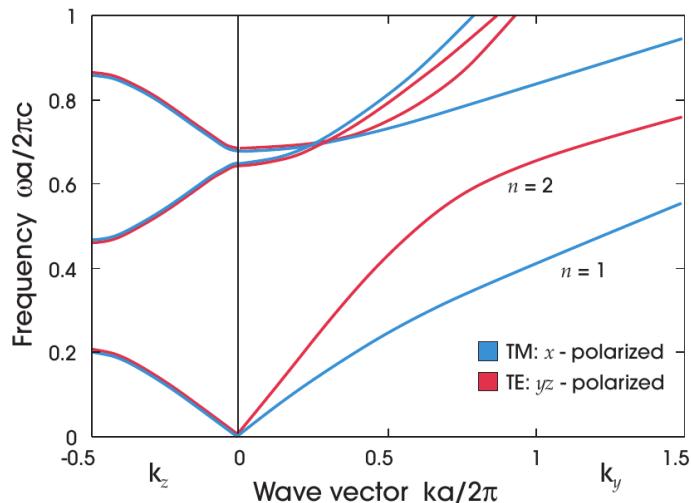


Figure 81: **Off-axis vs on-axis propagation of light in a multilayer film.** The band structure for on-axis propagation is shown to the left as a function of k_z . For off-axis propagation, the frequency bands are shown on the right as a function of k_y . For TM modes, the electric field points in the x direction. In the case of TE modes, the field vector lies in the yz plane. For on-axis propagation, these modes are degenerate. The figure is from Ref. [13].

As a result, the frequency bands will be non-degenerate for off-axis propagation, and one distinguishes between two cases: when the light is polarized in the x direction

(transverse-magnetic or TM modes) or when it is polarized in the yz plane (transverse electric or TE modes). Fig. 81 compares the different behaviour of frequency bands on the polarization for off- and on-axis propagation.

$\omega(k)$ behaves linearly for all modes for long wavelengths as $\omega \rightarrow 0$ because the electrical field will eventually have a wavelength much longer than the structure of our photonic crystal. The field will therefore experience a dielectric corresponding to the average over many periods of the crystal. This dielectric may still not be isotropic though because the average may give different values in different directions.

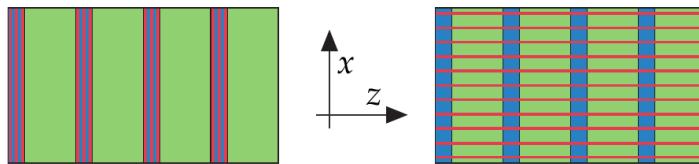


Figure 82: **Sketch of displacement field lines for long-wavelength modes.** The band structure for on-axis propagation is shown to the left as a function of k_z . For off-axis propagation, the frequency bands are shown on the right as a function of k_y . For TM modes, the electric field points in the x direction. In the case of TE modes, the field vector lies in the yz plane. For on-axis propagation, these modes are degenerate. The figure is from Ref. [13].

To get an intuitive picture of why the frequencies will be different for the different polarizations, remember that we said that low-frequency modes will have their intensity concentrated in regions with high ϵ . In this low-frequency case, we can assume that the wavelength of the light will be larger than the period of our multi-layer film. If the field is polarized along the x direction, then most of the field will concentrate in the high- ϵ regions as shown on the left in Fig. 82. If the field is polarized along the z direction, it will oscillate over many periods of the multilayer structure, the field will not be concentrated in particular regions of the crystal. Because the field will also penetrate low- ϵ regions, the frequency of such modes will be higher. This is indicated on the right in Fig. 82.

If our wavevector points in an arbitrary direction along the yz plane, one can see a continuous transition from the on-axis case to the off-axis case as we increase k_y . We can have a look at how the widths of frequency bands change as we increase k_y . The width of the bands are determined by the frequency at the zone center ($k_z = 0$) and the frequency at the edge of the zone ($k_z = \pi/a$). For increasing k_y , this difference will go to zero because $k_z/k_y \rightarrow 0$. That means we converge towards the off-axis, see Fig. 83.

As we saw earlier, when we are on the right of the light line (the red line in Fig. 83), the modes will be index guided because of total internal reflection. Of course, some of the light may still leak out into the air region, but its amplitude will decay exponentially. In this case, we can have independently guided modes in each of the high- ϵ regions.

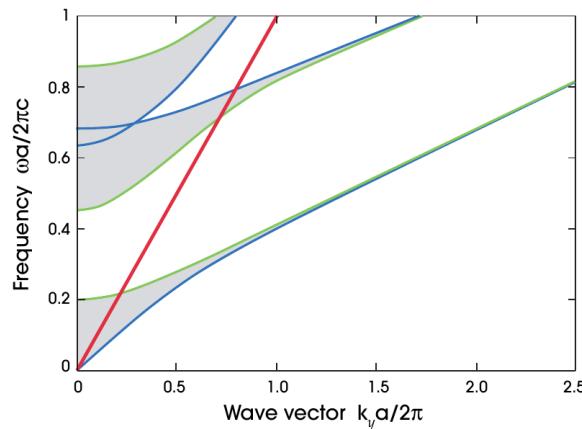


Figure 83: **x -polarized band structures for varying k_y .** The blue lines correspond to light along $(0, k_y, 0)$, and the green lines correspond to light along $(0, k_y, \pi/a)$. The light line is indicated in red, and the shaded areas indicate where the bands for an intermediate k_z would lie. The figure is from Ref. [13].

8.3.5 Localized modes at defects

Up to now, we concentrated on perfect crystals. Let us now briefly discuss the case when we encounter imperfections or *defects* in the crystal or in the periodicity of the crystal, how they affect the band structure, and how they can lead to the formation of localized states.

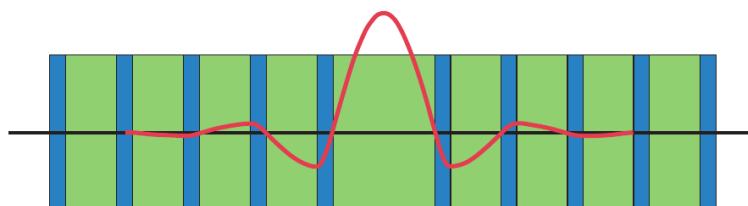


Figure 84: **A defect in a multilayer film.** The defect illustrated results from one of the low- ϵ layers having twice the normal thickness. The red line indicates the electric-field strength corresponding to the defect state associated with this structure. The figure is from Ref. [13].

Consider a stack of layers that all *except one* have the same thickness. This situation is illustrated in Fig. 84. At positions far from the defect, we would expect that the band structure will approximately look similar to the undisturbed band structure. Frequencies that lie within the band gap will not support extended modes. This is not changed

by the presence of a local defect. Because the periodicity is destroyed, we cannot associate frequencies with particular wave vectors in the Brillouin zone, but we can analyze whether particular frequencies support extended states in the rest of the crystal. This way we can divide the frequency range into regions where we have extended states and regions where we have evanescent states.

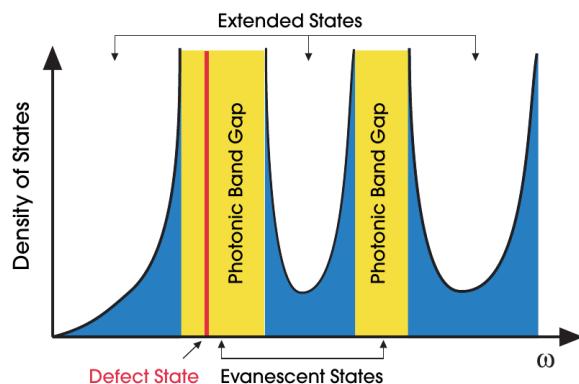


Figure 85: **Densities of extended and evanescent states.** The plot illustrates the densities of extended states outside the band gaps and evanescent states inside the band gaps. In red, a defect state is marked inside one of the band gaps. The figure is from Ref. [13].

Because of the defect, we can have a localized mode at the defect within the band gap, and we have a multilayer film to both sides of that defect that act as frequency-specific mirrors. Like in a cavity, the spectrum that is supported by this structure will be discrete with distinct and potentially narrow resonances. These modes are spatially localized between the two mirrors. The larger the thickness of the defect becomes, the longer the round-trip time of light in the cavity and the lower the associated frequency. That means, the numerator in the energy functional in equation 8.13 will become smaller, and if we minimize the functional, we can at some point get a energy value low enough to lie *inside* the band gap. The thicker the defect becomes, the more such lines will appear in the band gap. In Fig. 85, an example line is marked red. One can imagine that lines are drawn from the upper end of the band gap down inside the band gap. As a simple *application* consider the following case: we have a multilayer dielectric mirror that reflects a broad band of frequencies well. But, we introduce a defect such that one particular wavelength is allowed to pass through. In this way, one can realize what is known as a *dielectric Fabry-Perot filter*.

8.4 2D photonic crystals

After exhaustively introducing the treatment of 1D photonic crystals in the previous section, let us quickly apply this method where we can and extend it where we must to treat 2D photonic crystals. In particular, we will deal with geometries that have a discrete translational symmetry in the xy plane such that $\epsilon(\mathbf{r}) = \epsilon(\mathbf{r} + \mathbf{R})$ with $\mathbf{R} = a\hat{\mathbf{x}} + b\hat{\mathbf{y}}$. $a\hat{\mathbf{x}}$ and $b\hat{\mathbf{y}}$ are the primitive lattice vectors. If we also assume (for now) that our geometry will be homogeneous in z , then symmetry considerations will tell us that Bloch's theorem in two dimensions has to be fulfilled, and that the electromagnetic modes in this photonic crystal can be written in the form:

$$\mathbf{H}_{(n,k_z,\mathbf{k}_\parallel)}(\mathbf{r}) = e^{i\mathbf{k}_\parallel \cdot \boldsymbol{\rho}} e^{ik_z z} \mathbf{u}_{(n,k_z,\mathbf{k}_\parallel)}(\mathbf{r})(\boldsymbol{\rho}). \quad (8.42)$$

Here, $\boldsymbol{\rho}$ is the projection of \mathbf{r} on the xy plane, and $\mathbf{u}(\boldsymbol{\rho})$ is a periodic function with $\mathbf{u}(\mathbf{r}) = \mathbf{u}(\mathbf{r} + \mathbf{R})$ for all lattic vectors \mathbf{R} [13]. A photonic crystal that fulfills these conditions is illustrated in Fig. 86, where we assume a square grid of thin dielectric rods that extend to infinity in the z direction.

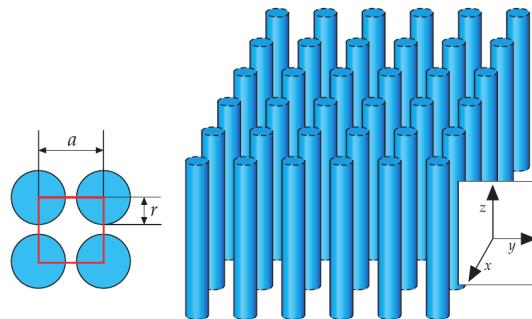


Figure 86: **2D photonic crystal consisting of dielectric rods.** A square lattice with side length a consisting of dielectric columns, each with radius r and dielectric constant ϵ . The figure is from Ref. [13].

Because the dielectric rods are assumed to be infinitely long in the z direction, we will have mirror symmetry around any plane parallel to the xy plane. This allows to separate TE modes that have \mathbf{H} perpendicular to the xy plane such a that $\mathbf{H} = H(\boldsymbol{\rho})\hat{\mathbf{z}}$ from TM modes, where $\mathbf{E} = E(\boldsymbol{\rho})\hat{\mathbf{z}}$ and $\mathbf{H} \cdot \hat{\mathbf{z}} = 0$. We already saw that the band structures can differ depending on polarization in the case of off-axis propagation in 1D multilayer films. This is also the case here. The band structures of TE and TM modes can be completely different. For example, there can be band gaps in one case while there are none in the other case. In the following, we will look at the band structures for TE and TM modes for two different kinds of 2D photonic crystals. For simplicity, we will focus on in-plane propagation ($k_z = 0$).

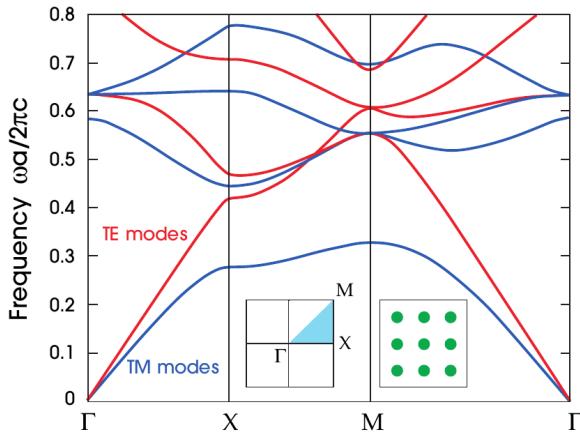


Figure 87: **Photonic band structure for a square lattice of dielectric rods.** The plots show a marked difference between the frequency bands for TM and TE modes (blue and red, respectively). It is assumed that $r = 0.2a$, and that $\epsilon = 8.9$ for the rods (alumina) and that the rods are surrounded by air. Because of the symmetry the irreducible Brioullin zone is given by a triangle between the points Γ , X and M as indicated in the inset. The figure is from Ref. [13].

Because of the square distribution of dielectric rods in the 2D array we consider here, the Brioullin zone will also have a square shape with side length $2\pi/a$. Due to the symmetries of the geometry, one can however describe all eigenmodes by concentrating on the irreducible Brioullin zone, which has the shape of a triangle. Its three corners are the points Γ , X and M , which correspond to $\mathbf{k}_{\parallel} = 0$, $\mathbf{k}_{\parallel} = \pi/a\hat{\mathbf{x}}$, and $\mathbf{k}_{\parallel} = \pi/a\hat{\mathbf{x}} + \pi/a\hat{\mathbf{y}}$, respectively[13].

For the TM modes, one gets a 31.4% gap-midgap ratio while one does not get a complete band gap for the TE modes. One can understand this significant difference in the behaviour of these two types of modes if one looks at the distribution of the field amplitudes. The field associated with the lowest TM modes (the dielectric band) is strongly concentrated in the dielectric rods. In the air band, a nodal plane (low amplitude plane) cuts through the dielectric columns. This leads to some of the displacement field amplitude to be expelled from the high- ϵ region, which goes along with higher frequency of these modes.

We know from earlier observations that low-frequency modes will be concentrated in regions of high ϵ . At the same time we know that a higher frequency mode must necessarily be orthogonal to the lower frequency mode. Therefore it must have nodes where the lower-frequency mode has anti-nodes. That means the higher-frequency mode will necessarily be less concentrated in the high- ϵ regions and more in the air-region. This is the reason why the higher-order mode must have a nodal plane passing through the

dielectric rods.

One can quantify this statement by calculating the *concentration factor*[13]:

$$\frac{\int_{\epsilon=8.9} d^3\mathbf{r}\epsilon(\mathbf{r})|\mathbf{E}(\mathbf{r})|^2}{\int d^3\mathbf{r}\epsilon(\mathbf{r})|\mathbf{E}(\mathbf{r})|^2}. \quad (8.43)$$

This expression calculates the fraction of electromagnetic energy located in the high- ϵ regions. The dielectric-band TM mode has a concentration factor of 83% while the air-band TM modes only have a concentration factor of 32%[13]. This difference in the amplitude distribution is the reason for the large photonic band gap for TM modes.

For TE modes, this difference is much less pronounced. While there are again nodal planes in the dielectric-band mode where there are maxima in the air-band mode, both modes have significant amplitudes in the air region. That means, there will be no clear distinction in their frequencies. For TE modes, one gets 23% concentration factor for the dielectric band and 9% concentration factor for the air band. The difference is therefore much less pronounced than for TM modes.

For comparison, Fig. 88 shows the frequency band for a 2D structure with a different design. Instead of a square lattice of dielectric rods, the authors of Ref. [13] described this as “dielectric veins”. I would rather say these are square columns of air in a medium with a high dielectric constant ϵ . By describing the geometry in this way, it is also straight-forward to see that the irreducible Brioullin zone must again have the same shape as in the earlier example. Because the high- ϵ distribution is significantly different, we can expect that the frequency bands will also look different. Indeed, in this case, we have a distinct band-gap for TE modes while the gaps are a bit less pronounced for the TM modes. The band gap is 18.9% for the TE modes, and there is no gap between the dielectric band and the air band for the TM modes. In this sense, our intuition that having air columns in a dielectric medium seems to be the opposite of having dielectric rods in air is correct because we see exactly the opposite behaviour in terms of the band gap.

Again, one can also see that if one looks at the actual distribution of the field amplitudes. In the case of the TM modes at position X in the irreducible Brioullin zone, the field will be concentrated mostly in the high- ϵ region in band 1 (along the vertical struts) as well as in band 2 (in the horizontal struts). Because both modes experience the same ϵ , there does not have to be a band gap. The difference in concentration for the TM modes is 89% to 77% for bands 1 and 2, respectively.

For the TE modes, on the other hand, the amplitude of the TE mode (also at point X) will mostly be concentrated in the low- ϵ regions for band 1, and it will be concentrated in the air region for band 2. This leads to a high frequency difference. The difference in concentration for the TE modes is 83% to 14% for bands 1 and 2, respectively [13].

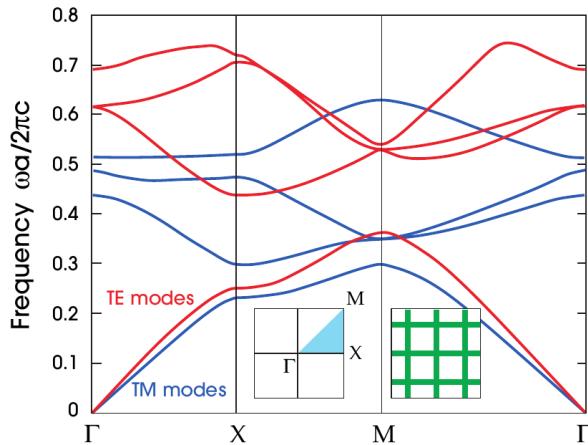


Figure 88: **Photonic band structure for a square lattice of dielectric veins.** Frequency bands for a structure of dielectric veins ($\epsilon = 8.9$) with a thickness of $0.156a$ in air. Because of the symmetry the irreducible Brillouin zone is given by a triangle between the points Γ , X and M as indicated in the inset. The figure is from Ref. [13].

8.5 A bandgap for both polarizations

We just saw that one gets a significant band gap for the TM modes in the case of dielectric rods in air, while one gets a significant band gap for TE modes if one instead has air columns embedded in a dielectric material. The question is, can we design a photonic crystal that exhibits a large bandgap for both polarizations? At first, this seems implausible because the two cases contradict each other. Isolated high- ϵ regions forced TM modes in adjacent bands to have very different concentration factors, leading to a large band gap. At the same time, the connectivity of high- ϵ regions was the key to achieve a high band gap for TE modes because the TE modes were forced to penetrate the low- ϵ regions.

The solution is to put a triangular lattice of low- ϵ columns (for example air) inside a high- ϵ material. If the radius of the columns is large enough, the intersection spots between the remaining ridges of dielectric material look themselves like localized columns of dielectric material although they are connected via a narrow squeeze between the air columns. Such a geometry is shown in Fig. 89. That such a geometry finds a good compromise between realizing isolated columns of high- ϵ material, but at the same time features connectivity between these high- ϵ regions can be seen in Fig. 90. The thin veins of material shown in this figure assure the connectivity that leads to a band gap for TE modes, while the isolated spots lead to a band gap in the TM modes.

The band structure for the particular geometry shown here is plotted in Fig. 91. The

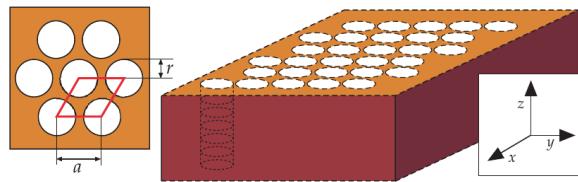


Figure 89: **Triangular array of air columns in dielectric material.** The columns have a radius r , and a distance a such that $2r$ is not much different from a . In the z direction, the columns are assumed to extend (effectively) to infinity. In this case, $r/a = 0.48$, and $\epsilon = 13$. The figure is from Ref. [13].

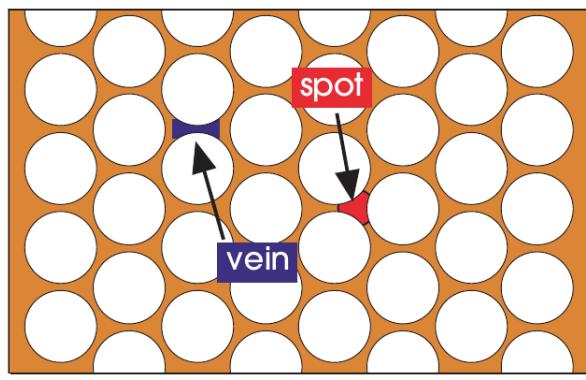


Figure 90: **Spots and veins in a triangular lattice.** Features of a triangular lattice that realize the connectivity of high- ϵ material (via veins), while at the same time realizing “isolated” dielectric columns (spots). The figure is from Ref. [13].

structure leads to an 18.6% complete photonic band gap[13].

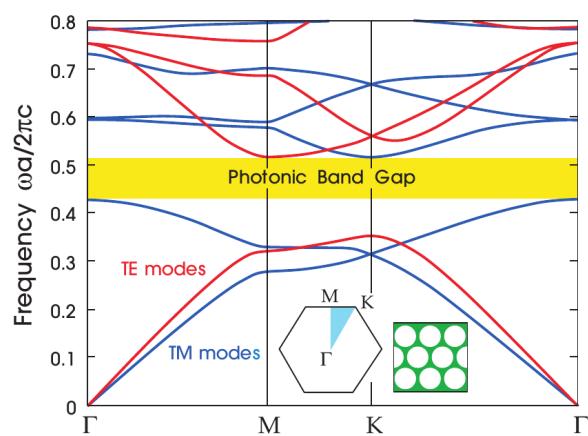


Figure 91: **Photonic band gap for both polarizations.** Frequency bands for TM and TE modes in the triangular lattice of Fig. 89. The figure is from Ref. [13].

Bibliography

- [1] Robert W. Boyd. *Nonlinear Optics*. Academic Press, Inc., USA, 3rd edition, 2008.
2, 3, 5, 6, 8, 9, 10, 11, 12, 13, 14, 17, 23, 27, 29, 46, 49, 54, 55, 56, 58, 63, 64, 67, 81, 82, 83, 84, 85, 86, 87, 88, 90, 91, 92, 94, 96, 105, 106, 111, 113, 114, 115, 118, 120, 121, 122, 123, 124, 127, 128
- [2] D. Meschede. *Optics, Light and Lasers*. Wiley-VCH, Weinheim, Germany, 2017.
19, 21, 30
- [3] Ilias Pappas, Stylianios Siskos, and Charalambos A. Active-Matrix Liquid Crystal Displays - Operation, Electronics and Analog Circuits Design. In *New Developments in Liquid Crystals*. InTech, nov 2009. 31
- [4] R. Y. Chiao, P. L. Kelley, and E. Garmire. Stimulated Four-Photon Interaction and Its Influence on Stimulated Rayleigh-Wing Scattering. *Physical Review Letters*, 17(22):1158–1161, November 1966. Publisher: American Physical Society. 50
- [5] R. Y. Chiao, E. Garmire, and C. H. Townes. Self-Trapping of Optical Beams. *Physical Review Letters*, 13(15):479–482, oct 1964. 56, 59
- [6] R. W. Hellwarth. Generation of time-reversed wave fronts by nonlinear refraction*. *Journal of the Optical Society of America*, 67(1):1, jan 1977. 63
- [7] A. Yariv and D. M. Pepper. Amplified reflection, phase conjugation, and oscillation in degenerate four-wave mixing. *Optics Letters*, 1(1):16, jul 1977. 63
- [8] D. M. Bloom and G. C. Bjorklund. Conjugate wave-front generation and image reconstruction by four-wave mixing. *Applied Physics Letters*, 31(9):592–594, nov 1977. 63
- [9] K. Okamoto. *Fundamentals of Optical Waveguides*. Elsevier Academic Press, 2006.
76, 77, 78, 79, 80, 134, 137, 140, 142, 144, 145, 146, 147
- [10] Elsa Garmire. Stimulated Brillouin Review: Invented 50 Years Ago and Applied Today. *International Journal of Optics*, 2018:1–17, dec 2018. 104, 107
- [11] Bahaa E. A. Saleh and Malvin C. Teich. *Fundamentals of Photonics*. Wiley, USA, 2nd edition, 2008. 108, 131, 132, 133

- [12] Christian W Freudiger, Wei Min, Brian G Saar, Sijia Lu, Gary R Holtom, Chengwei He, Jason C Tsai, Jing X Kang, and X Sunney Xie. Label-free biomedical imaging with high sensitivity by stimulated Raman scattering microscopy. *Science (New York, N.Y.)*, 322(5909):1857–61, dec 2008. [108](#), [109](#)
- [13] John D. Joannopoulos, Steven G. Johnson, Joshua N. Winn, and Robert D. Meade. *Photonic Crystals – Molding the Flow of Light*. Princeton University Press, Berlin, 2 edition, 2007. [151](#), [153](#), [154](#), [156](#), [157](#), [158](#), [160](#), [161](#), [162](#), [163](#), [164](#), [166](#), [167](#), [168](#), [169](#), [170](#), [171](#), [172](#), [173](#), [174](#), [175](#), [176](#), [177](#), [178](#), [179](#)
- [14] Wentao Jiang, Rishi N. Patel, Felix M. Mayor, Timothy P. McKenna, Patricio Arrangoiz-Arriola, Christopher J. Sarabalis, Jeremy D. Witmer, Raphaël Van Laer, and Amir H. Safavi-Naeini. Lithium niobate piezo-optomechanical crystals. *Optica*, 6(7):845, jul 2019. [152](#)
- [15] Pochi Yeh. *Optical Waves in Layered Media*. Wiley, 2005. [168](#)