Notes of Photonic Devices, 2021/2022

Tommaso Dezuanni

Contents

1 Introduction

A brief list of optics features that are useful for the application in the interconnection field:

- High frequency:
  - No frequency dependent loss or crosstalk;
  - Very high bandwidth
  - Well synchronized signals, always propagating at high speed
- Short wavelength:
  - Essentially no distant-dependent loss or degradation
  - Wavelenght Division Multiplexing (WDM)
- Large Photon energy:
  - Electrical isolation
  - Immunity to electromagnetic interference
  - Fundamental lower communication energy

The first optical device that we're going to study, for logical and historical reasons, is the waveguide.
The waveguides are the fundamental building blocks to design every optical system of communication, every photonic integrated circuit.
When it takes to choose a dielectric waveguide, the most important parameter is the index contrast:

$$\Delta n = \frac{n_{core} - n_{cladding}}{n_{core}}$$

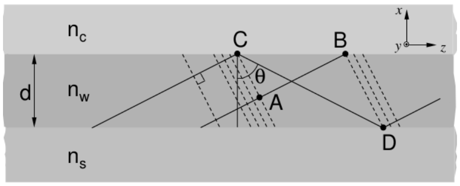
Many characteristics and properties of a waveguide depend on this parameter, or are related somehow to its varying: dimensions, fiber to waveguide coupling, bending radius, losses, directional couplers gap, birefringe and so on...
The first devices in the field of integrated optics were designed for low values of index contrast. Actually, these waveguides are still on the market today because they are reliable and ensure an excellent fibers coupling.
It is importanto to understand that the current research are aiming to merge efficiently photonics and electronics to exploit the best properties of the two options. In the end, we must mention the large, and growing, interest for silicon photonics. Indeed, since silicon is a semiconductor has a very high index contrast, but in the same time provide excellent optical properties. It is probably the best choice to work with in order to find a way to standardize photonics.

2 Waveguides

With the term *waveguide* we mean every structure, typically dielectric, that is able to confine the electromagnetic radiation in the range of light wavelengths. The most common example is the optical fiber. However, the waveguides used in integrated photonics are different because their structure is planar, not cilindrical.

Geometric optics is not a rigorous approach and we will avoid to use it, exception done for teh next paragraph, in which is shown how to derive the fundamental equations describing the electromagnetic propagation in a slab waveguide.

2.1 Slab Waveguide: introduction



Hypothesis: it extends infinitely alongside y-direction. For *symmetric* slab $n_{substrate} = n_{cladding}$, but more in general $n_{core} > n_{substrate} > n_{cladding}$. According to the laws of reflection and refraction (the rays come form medium 1 and enters in medium 2) there are two critical angles, one referring to the "core-substrate" interface and the other to the "core-cladding" interface. The electromagnetic wave can propagate if the following relation is respected:

$$\theta_{critic} = \sin^{-1}\left(\frac{n_2}{n_1}\right) \longrightarrow \theta \geq \theta_{gs} \geq \theta_{gc}$$

This is a necessary condition. However, it's not true that every wave that respects it can be called *guided.* To be guided by the structure, the waves must respect also the condition of *phase matching.*
The picture above is analyzed:

1. The distance covered from A to B is $\overline{AB} = \dfrac{(\sin^2\theta - \cos^2\theta)\cdot d}{\cos\theta}$.
2. The distance covered from C to D is $\overline{CD} = \dfrac{d}{\cos\theta}$

We assume A and C to be on the same wavefront (likewise B and D). The total phase difference between the rays must be equal to a multiple of 2π:

$$k_0 n_w(\overline{CD} - \overline{AB}) + 2\phi_c + 2\phi_s = 2\cdot N\pi$$

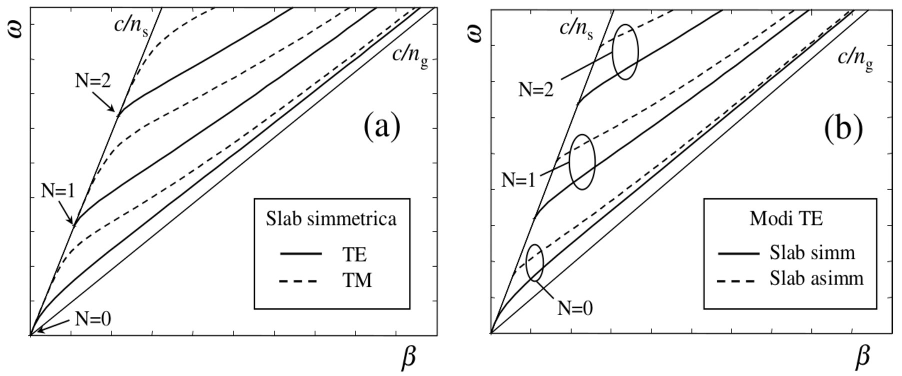
Where ϕ_c and ϕ_s are the differences of phase introduced by the reflections for the ray going from C (first reflection) to D (second reflection). Since they depend on the polarization we have different values for TM and TE modes. At the critic angle, they are nill.
Replacing the explicit expressions of the distances in the previous equation, we have:

$$k_0 n_w \cdot d\cos\theta + 2\phi_c + 2\phi_s = 2\cdot N\pi$$

It's called *eigenvalues equation* or *dispersion equation*, because for every possible λ (eigenvalue) we determine the corresponding angle, field configuration, cut-off frequency and the other parameters of interest.
In particular, we derive the definition of the propagation constant: the component of the wave vector alongside the direction of propagation:

$$\beta = \frac{\omega}{v_p} = k_0 n_w \sin\theta$$

Where v_p is the phase velocity of the wave in the core section: $c/n_w \sin\theta$.
From the equations expressed before we derive that $k_0 n_s < \beta < k_0 n_w$, an important limit on the value of the constant propagation.



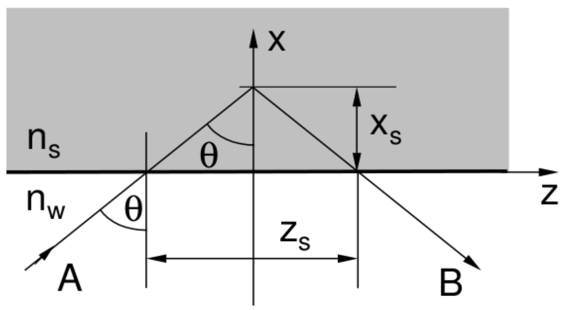
This picture shows the important relation between the propagation constant and the frequency, i.e. the phase velocity of the first three modes.
Every curve respect the following condition: $c/n_s < \omega/\beta < c/n_w$.

The cut-off frequency of every mode is determined by the intersection between the mode curve and the lower limit c/n_s. Since at the cut-off the angle of reflection is the critic one, the phase displacement is 0, thus we can compute its expression from the dispersion equation.
For symmetric structure:

$$\frac{d}{\lambda_{cut-off}} = \frac{N}{2 \cdot \sqrt{n_w^2 - n_s^2}}$$

Then, for symmetric slab the cut-off frequency is the same for TE and TM modes. Moreover the first mode is alaways guided, that is, for every possible ω.

2.2 Goos-Hänchen shift (z_s) and effective width



We suppose that the planar waves forming the total incident one have slightly different wave vectors.
Obviously we are not in the case of a critical reflection.

Amplitude incident wave:	$A(z) = e^{-\jmath(\beta-\Delta\beta)z} + e^{-\jmath(\beta+\Delta\beta)z} = 2\cos(\Delta\beta)e^{-\jmath\beta\cdot z}$
Phase displacement:	$\phi(\beta+\Delta\beta) = \phi(\beta) + \frac{d\phi}{d\beta}\Delta\beta$
Ampltiude reflected wave:	$B(z) = R \cdot A(z) = \cos(\Delta\beta(z - z_s))e^{-\jmath(\beta\cdot z - \phi)}$
"Depth of evacescence":	$x_s = \frac{z_s}{2\tan\theta}$
Effective Width:	$d_{eff} = d + x_s + x_c$

2.3 Electromagnetic theory of the guides

Reasonable hypotesis on the medium material:

- *linearity* : ϵ and μ do not depend on the field intensity
- *isotropy* : ϵ and μ do not depend on the direction of the field
- *abscence of source*: no free currents and no free charges

Hence, the Maxwell equations became:

$$\begin{cases} \nabla \times \mathcal{E} = -\jmath\omega\mu_0\mathcal{H} \\ \nabla \times \mathcal{H} = \jmath\omega\epsilon_0 n^2\mathcal{E} \\ \epsilon_0\nabla \cdot (n^2\mathcal{E}) = 0 \\ \mu_0\nabla \cdot \mathcal{H} = 0 \end{cases}$$

Where:

$$\epsilon_0 = 8.854 \cdot 10^{-12} \ \ {C^2}/{N \cdot m^2} \qquad \mu_0 = 4\pi \cdot 10{-}7 \ \ {N \cdot s^2}/{C^2}$$

Now, starting from the Maxwell's equations and exploiting some vector identities, we derive the equation of the wave, in the most general case:

$$\nabla^2 E + \nabla\left(\frac{1}{n^2}\nabla n^2 \cdot E\right) - \epsilon_0\mu_0\omega^2 n^2 E = 0$$

Obviously there is a equivalent expression also for the magnetic field.
Next, in order to simplify our problem, we separate the structure of the guide into sub-homogeneous segments. In this way, we only have to solve the wave equation in every one of them. Once we have these solutions, we finally impose the appropriate boundary conditions and the continuity of the fields at the interfaces.

Actually, the main advantage of this method is that the second term vanishes: the three cartesian components of each wave equations become scalar.

Last hypothesis: *z-invariance* $\longrightarrow$ n = n(x,y). The field distribution became:

$$\mathcal{E}(x,y,z) = \mathbf{E}(x,y) \cdot e^{-\jmath\beta \cdot z}$$

Finally, substituting this latter relation in the wave equation, we obtain:

$$\nabla_t^2\mathbf{E}(x,y) + [k_0^2 n^2(x,y) - \beta^2]\mathbf{E}(x,y) = 0 \qquad (*)$$

Where:

$$\nabla_t^2 = \left(\frac{\partial^2}{\partial^2 x}, \frac{\partial^2}{\partial^2 y}, 0\right) \qquad k_0 = \frac{2\pi}{\lambda}$$

Observations

1. The equation * give us eigenvalues (possible values of *beta*), related to eigensolutions (or eigenvectors) that correspond to the orientations of the field in the guide.
2. For a dielectric waveguide and an incoming wave with a known wavelenght there is a set of allowed solutions - these are defined as the *modes* of the guide
3. It is generally prefered to separate the electric field into longitudinal and trasverse components ($E_z(x,y)$ and $E_t(x,y)$) and solve the wave equation using only the trasverse one. Once you have E_t we can derive E_z from the Maxwell equations.
4. Trasverse components of E and H are in phase, longitudinal components are in phase. Travserse components are out of phase by $\pi/2$, with respect to the corresponding longitudinal components.

2.4 Modal Expansion Field Theory and properties of waves

The (eigen)solutions of the wave equation are of two types:

1. *Guided*: modes confined in the core region of the structure. Moreover: $E(x, y) \to 0 \quad \leftrightarrow \quad x, y \to \infty$. The guided modes are waves that propagate without changing neither their shape nor their amplitude, unless there were losses.

2. *Radiative*: modes not confined - since there is not this limit, there is not a discrete number of solution but a continuous range of possible values. The most important difference with guided modes is that, while guided are evanescent in the direction of x and y (they have pure imaginary transverse propagation constant), radiative modes substract field in waveguide (they have a perpendicular component to the power). In certain extreme cases they can have an imaginary value of β along z-direction.
In general they are excitede by discontinuites and so, at a great distance from origin, only the power associated to the guided modes remains in the guide.

Now, a really important point is discussed: the radiative modes are very similar to planar waves. However, a discontinuity does not excite only one radiative mode, indeed to have a planar way you should have a infinite extended source. It's always generated a set of them, summed in a so called *packet*, and they have different phase constant β. They interact and add up, and they form the configuration of extension over the field. In other words, we could describe the radiative modes as harmonics that contribute to the definition of the *total* field distribution.

Each radiative mode of the packet travels with its own phase velocity $v_p = \omega/\beta$. It's logical that the shape and amplitude are not mantained, and the velocity with which they change is proportional to the differnce between the phase constants.
However, if the difference is small, this change can be neglected and the packet resembles a guided mode. These leaky modes don't meet the conditions to be considered waves, and they don't have the properties of wave. We can anyway associate a phase constant to them, a specfic β equals to a weighted average of β of radiative modes that constitute it.

Any field distribution that propagates in the guide and satisfies the wave equation can be seen as a linear combination of guided and radiative modes:

$$\mathbf{E}(x, y) = \sum_m a_m \cdot e_m(x, y) + \int_0^\infty a_r(\beta) \cdot e_r(\beta, x, y) d\beta$$

They respect the *principle of orthogonality*: if we take any two generic waves the integral of overlap of transvers fields is null in any section of the guide:

$$\int_{area} (e_m \times h_n^*) \cdot ds = \delta_{m,n}$$

Where m,n are the mode indeces. Next, delta is better specified:

$$\delta_{m,n} = 0 \qquad \text{for:} m \neq n$$

$$\delta_{m,n} = 1 \qquad \text{for:} m = n$$

An important consequence: modes don't interact and don't exchange power between them (regardless to their nature).
This actually means that the total Poyinting vector is the sum of the individual mode powers:

$$P = \frac{1}{2} \sum_m a_m \cdot a_m^* \left(\int_{section} (e_m \times h_m^*) dS \right) = \frac{1}{2} \sum_m |a_m|^2$$

CASE: *What if there were also the reflected waves?*

$$\begin{cases} \mathbf{E} = \sum(a_m + b_m)e_m \\ \mathbf{H} = \sum(a_m - b_m)h_m \end{cases} \Rightarrow P = \frac{1}{2}\sum_m(|a_m|^2 - |b_m|^2)$$

CASE: *What if the were evanescent modes in the guide?*

Firtsly, evanescent means that β is imaginary. These condition leads to a phase difference of $\pi/2$ between the trasverse components of the fields. Then: $P = j|a_m|^2$. It's a reactive power not a real one.
However, if there was a counter wave, the problem becomes even more complex:

$$P = j(|a_m|^2 - |b_m|^2) + j(a_m \cdot a_m^* - b_m \cdot b_m^*)$$

While the first term is alaways imaginary, the second one can be real for some values. In that case there would be an actual trasport of power mediated by evanescent modes, pheonemena called optical tunneling. It explains losses caused by bends.

2.5 Slab Waveguide: analytical solution of the wave equation

The structure is infinitely extended in y-direction so we can consider the propagation independent from y.
Suppose we want to find solution for TE modes, that are, waves with purely transverse electric field.
From Maxwell equations:

$$E_z = 0 \wedge \text{z-invariance} \Rightarrow H_y = E_x = E_z = 0$$

Now, the wave equation is scalar and one-dimensional:

$$\frac{\partial^2 E_y(x)}{\partial x^2} + (n(x)^2 k_0^2 - \beta^2)E_y(x) = 0$$

It's important to address that $n(x)$ is the profile of the refractive index along x direction.
Once we've evaluated E_y, we derive H_z and H_x from Maxwell equations.

Suppose to have a core wide "d" and suppose that the interfaces are at $x = 0$ and at $x = d$. The acceptable solutions are those field distributions that satify the continuity at these interfaces. Moreover, if they are guided modes, they must vanish at great distance from the core.
From a mathematical point of view:

$$\begin{aligned} &E_y = E_s \exp(\gamma_s \cdot x) && \text{In the substrate, for } x < 0 \\ &E_y = E_g \cos(k_g \cdot x - \phi_s) && \text{In the core region, for } 0 < x < d \\ &E_y = E_c \exp(\gamma_c \cdot (x - d)) && \text{In the cladding, for } x > d \end{aligned}$$

We replace these solution in the wave equation and we obtain:

$$\begin{aligned} k_g^2 &= k_0^2 n_w^2 - \beta^2 \\ -\gamma_s^2 &= k_0^2 n_s^2 - \beta^2 \\ -\gamma_c^2 &= k_0^2 n_c^2 - \beta^2 \end{aligned}$$

Since the supposed solutions describes guided modes, k_g, γ_s and γ_c should be real.

$$\begin{aligned} k_0^2 n_w^2 - \beta^2 &> 0 \\ k_0^2 n_s^2 - \beta^2 &< 0 \\ k_0^2 n_c^2 - \beta^2 &< 0 \end{aligned}$$

Next, boundary conditions are imposed:

$$E_{ys} = E_{yg} \Rightarrow E_s = E_g \cos(-\phi_)$$
$$E_{yc} = E_{yg} \Rightarrow E_c = E_g \cos(k_g d - \phi_s)$$

Without showing every step, we isolate $tan(k_g d)$ removing ϕ_s. The resulting relationship is the following one:

$$\tan(k_g d) = \frac{\frac{\gamma_c}{k_g} + \frac{\gamma_s}{k_g}}{1 - \frac{\gamma_c \cdot \gamma_s}{k_g^2}}$$

It is used to determine the values of the constant β of TE propagation modes, and allows to build a scatter diagram of $\omega - \beta$.
Furthemore, relationships between the amplitudes of the electric field in the three regions are derived:

$$\frac{E_g}{E_s} = \frac{k_0}{k_g}\sqrt{n_w^2 - n_s^2}$$
$$\frac{E_g}{E_c} = \frac{k_0}{k_g}\sqrt{n_w^2 - n_c^2}$$

A similar result is obtained for TM modes:

$$\tan(k_g d) = \frac{\frac{n_w^2}{n_c^2}\frac{\gamma_c}{k_g} + \frac{n_w^2}{n_s^2}\frac{\gamma_s}{k_g}}{1 - \frac{n_w^4}{n_c^2 n_s^2}\frac{\gamma_c \gamma_s}{k_g^2}}$$

The *cut-off conditions* ($\gamma_s = 0$) are:

$$\text{For TE modes:} \quad \tan(k_g d) = \frac{\gamma_c}{k_g}$$
$$\text{For TM modes:} \quad \tan(k_g d) = \frac{n_w^2}{n_c^2}\frac{\gamma_c}{k_g}$$

In the case of a symmetric flat waveguide: $n_s = n_c$.
Hence the previous relationships becomes:

$$\text{For TE modes:} \quad \tan\left(\frac{k_g d}{2} + N\frac{\pi}{2}\right) = \left(\frac{n_w}{n_s}\right)\frac{\gamma}{k_g}$$
$$\text{For TM modes:} \quad \tan\left(\frac{k_g d}{2} + N\frac{\pi}{2}\right) = \frac{\gamma}{k_g}$$

N is the order of the mode.
Note that the cut-off condition is the same for the two cases, once we impose $\gamma = 0$.

It is evident that we msut rely on a numerical method to evaluate the eigenvalues of the dispersion relation.
However, we can still understand some properties of the propagation following a study graph.
For a flat symmetrical slab waveguide, the *normalized frequency* is defined as:

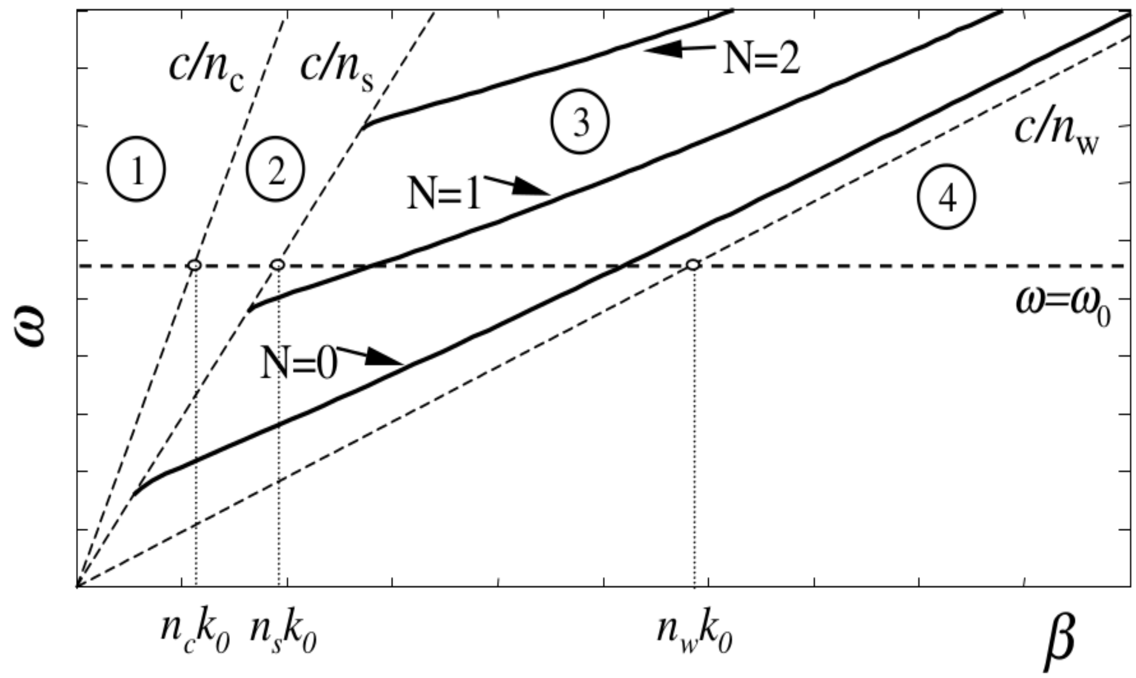
$$V = \frac{k_0 d}{2}\sqrt{n_w^2 - n_s^2}$$

Another relation in which there is V and that is also valid is:

$$V^2 = \left(\frac{\gamma d}{2}\right)^2 + \left(\frac{k_g d}{2}\right)^2$$

Furtemore we define *numerical aperture* the parameter $NA = \sqrt{n_w^2 - n_s^2}$.
We will observe that for $V < \pi/2$ we are in the monomodal regime: only the fundamental mode propagate inside the guide.

NUMERICAL METHOD

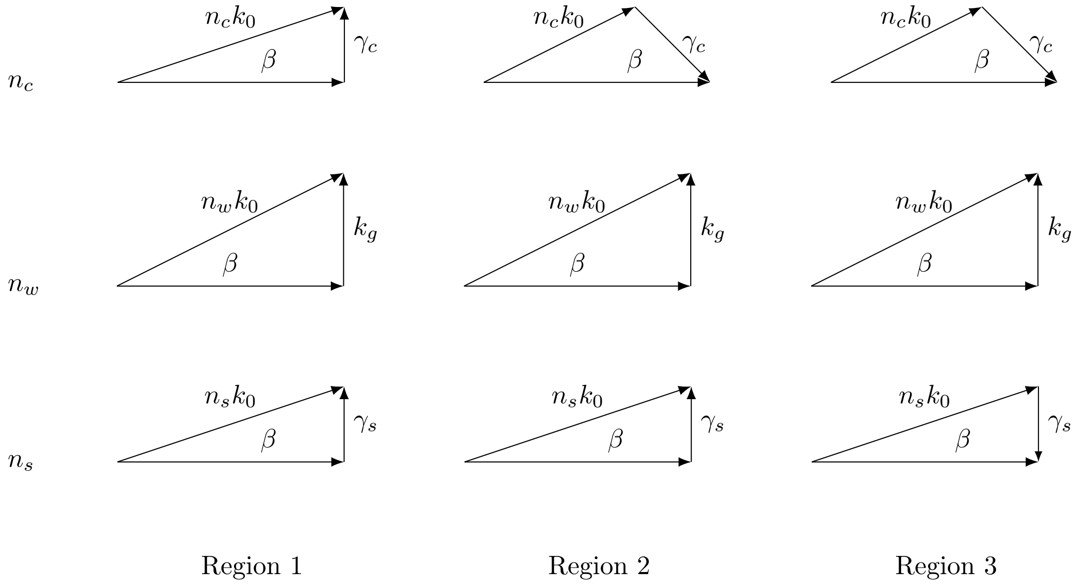


The picture above shows the relationship between the eigenvalues, i.e. the propagation constant of the modes, and the corresponding (angular) frequencies. It must be reminded that this diagram is referring to a infinite flat slab waveguide.

As we can see, once we've fixed the frequency the space is divided into 4 regions.
Keep in mind that we started analyzing this structure assuming that the field distribution, for the *guided modes*, is sinusoidal in the core and decreasing exponential in the substrate/cladding.
In order to have a distribution like this, the relationships between the parameters of the field functions tell us that k_g, γ_c and γ_s must be real.
This conclusion holds true in the third region of the diagram, whose boundaries are represented by the lines c/n_s and c/n_w.
On the horizzontal axis we can easily verify that the corresponding interval in which β varies is $(n_s k_0;\ n_w k_0)$.
As an example, the curves of the first three guided modes are drawn: note that the ω choosen is below the cut-off value of the mode of order N=2, then we can conclude that only the first two are actually propagatin in the waveguide.

Another possible approach to explain the confinement of the guided modes in the third region is the vector representation, shown in the next page.
Long story short: when γ_c or γ_s are directed as β they are real, otherwise they are imaginary.
Real γ correspond to guided modes field distribution. However, whenever thery are imaginary there is a power flow in the direction perpendicular to that of propagation (key words: radiative mode, losses).
The vectors tell us that only in the third region $n_s k_0 < \beta < n_w k_0$ is true, pointing out that usually we consider $n_c < n_s$.

Vector Representation:



Observations:

- At the same frequency high order modes have a smaller phase constant, which means that the corresponding field distribution is less confined in the core region. Indeed, the closer β is to the upper limit ($n_w k_0$), the more the propagation is limited. Moreover, considering only one, the confinement improves increasing the frequency;
- <u>Effective Index</u>: a weighted mean value of the three indeces n_s, n_c and n_w. The ray theory lead us to the following definition:

$$\beta = \frac{2\pi}{\lambda} n_{eff} = \frac{2\pi}{\lambda} n_w \cdot \sin\theta$$

- In the first region we have radiative modes either in the cladding and in the substrate. In the second region only in the substrate;
- In the fourth region we have $\beta > n_w k_0$. Observe the wave equation:

$$\frac{\partial^2 E_y(x)}{\partial x^2} = (n(x)^2 k_0^2 - \beta^2) E_y(x)$$

  Since the second term is always positive the correpsonding field distribution tends to infinite at a great distance from the core: physically speaking it makes no sense. As a matter of fact this region is proibitive.

- In physiscs, the speed at which the phase of the field advances is named *phase velocity*: $v_p = \dfrac{\omega}{\beta}$

  However, the transport of energy is associated with the *group velocity*: $v_g = \left.\dfrac{d\omega}{d\beta}\right|_{\omega=\omega_0}$

  We need to understand that these two values are different if there is dispersion in a wave packet. In this latter case, more than one wave is propagating inside the medium and each wave has its own phase constant. The group velocity is the velocity at which the paket travels, while the phase velocity becomes an "apparent" quantity.

- Recalling the effective index definition:

$$n_g(\omega) = \frac{c}{v_g(\omega)} = n_{eff}(\omega) + \omega \frac{dn_{eff}}{d\omega}$$

2.6 Real case: channel waveguide

In reality the structure are always limited in the three directions. The most simple and common case is the channel waveguide, in which the field is confined in bot directions x and y.
As for the design, we need to understand clearly that there is not an perfect waveguide: it'all about trade-off choices and the application that we aim to achieve.

1. The first main difference with the infinite slab waveguide concerns the modes, which are **never purely TE or TM** in the channel model. We call them quasi-TE or quasi-TM.

2. The properties of waveguides are **seriously limited by the available fabbrication technologies** exploited to craft the structures. They impose the most restrictive constraints on materials and dimensions. Obviously, the most critical step concerns the choice of the refractive indexes. In this regard, a very important parameter is the *refactive index contrast*, defined as

$$\Delta n = \frac{n_w - n_s}{n_w}$$

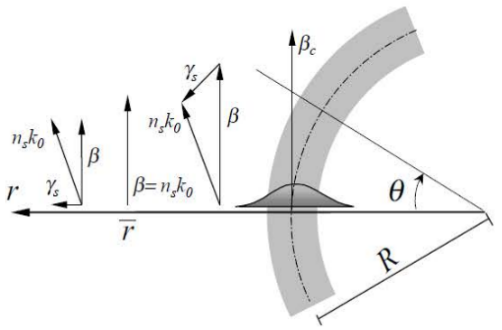
   A lot of properties vary significantly with it, so it's helpful to plot their dependency and choose the best value in according with the application required.

3. One of the desired requirement of a waveguide is the **condition of single-mode**. Set the refractive indexes, the order of magnitude for the dimensions of the channel waveguide is the same of the thickness d of an ideal slab. Typically, w and h are increased to improve the confinement of the fundamental mode.

4. However, the choice of the ratio w/h must be taken considering also another parameter, the **birefringe**. It is defined as the difference of the refractive indexes of the quasi-TE and quasi-TM mode (hereinafter called simply TE and TM), and is due to two separate contirbutions: birefringe of the material B_m and of the shape B_f.

2.7 Curved guides

The bending of waveguides is very common in photonics circuit because allow the connection of different parts. Unfortunately it could give rise to radiative modes.



The radius of curvature is constant. The phase fronts move by rotating around the point O with velocity:

$$r\frac{d\theta}{dt} = \frac{\omega}{\beta(r)}$$

Where $\beta(r)$ is the phase constant function in dependece of the distance from the center. From the previous relation, naming β_c the value in the center of the guide:

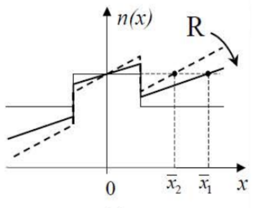
$$\beta(r) = \beta_c \frac{R}{r}$$

Now, it's evident that β decreases radially outwardly of the guide. Reminding the ω-β graph, the smaller is beta, the closer it is to the lower limit for guided modes, that is $k_0 n_c$. Over this point the propatation is in the trasverse direction.
This limit will be reached by $\beta(r)$, at some point. However, the smaller the radius of curvature, the closer the cut-off to the guide, but since here the field is higher the losses increase as well. Strictly speaking, the fundamental mode of a curve guide is therefore alsways leaky, but the losses become negligible rapidly increasing the radius of curvature.

Formal Analysis: applying a conformal transformation we can map the curve driving on a straight guide equivalent. Under condition of large radius of curvature, the refractive index can be written:

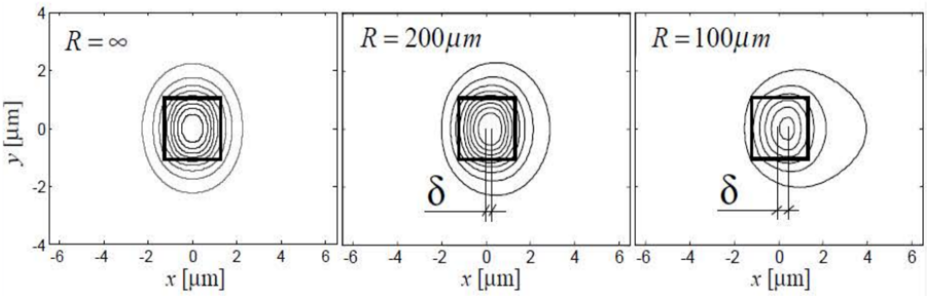
$$n(x,y) = n_0(x,y)e^{x/R} \simeq n_0(x,y)\left(1 + \frac{2x}{R}\right)$$

We set $x = r - R$.



Larger is R and less steep is the relation between $n(x)$ and n.
Paying attention to the directions and verses of the reference axis, we note that outer we are and higher is the refractive index.

After having done our transformation we can simulate and plot the guide.
For example, considering a buried waveguide ($n_w = 1.513, n_s = 1.445, w = h = 2.2$):



As we can see the effect of the curvature is a distortion with a displacement of δ of the center of gravity towards the outside of the curve.
Must be pointed out that this distorsion change the effective modal refractive index, so the phase constant of the curved guide is higher than that of the straight structure. Besides, we can also understand that this increase is due to the convention of taking as distance travelled the one referred to the centerline, but the mode is distorted outwardly and so it travels a longer distance. We can calculate the increase as follows:

$$\beta_c = \beta_0 + \frac{B}{R^2} \qquad B: \text{ dimensional parameter}$$

As for the radiative losses we can approximately calculate them as follows:

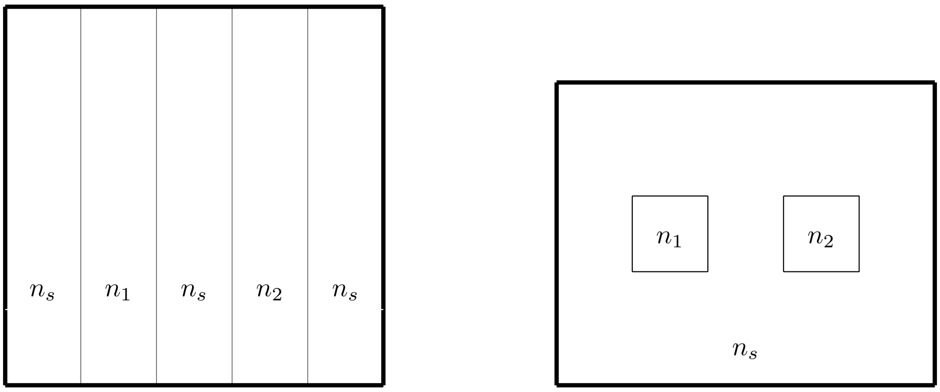
$$\alpha_r = \frac{C_1}{\sqrt{R}} e^{-C_2 \cdot R}$$

C_1, C_2 are determined by numerical simulation. In particular, C_2 is a constant which increase with the difference between the refractive indexes (of guide and cladding) or with the dimensions of the guide.

3 Directional Couplers and Dividers

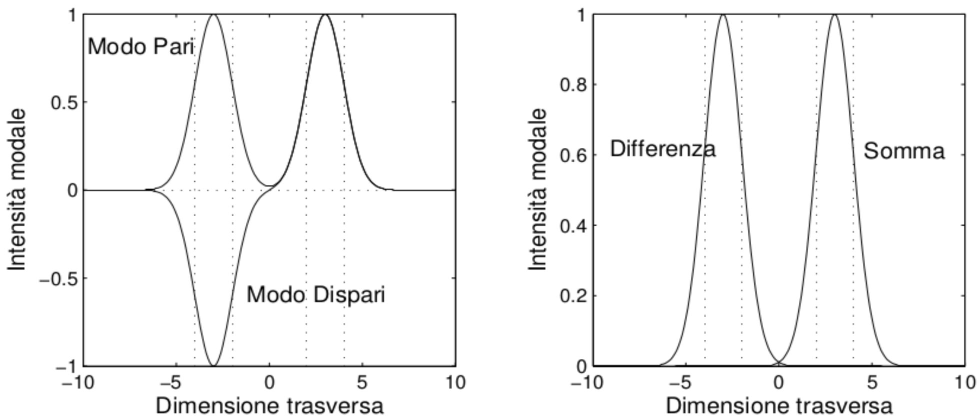
3.1 Coupled Mode Theory

This theory is applied to adiabatic perturbations, that are described as slow trasfers of power from a mode to another. Possible reflections or losses can be neglected.
Let's start with the following system: two parallel dielectric guide that are such close that their modes affect each other. We could determine the field distributions like we did in the single guide case: we solve analytically the wave equation, imposing the specific boundary conditions. Usually, structure like these are at least bi-modal.



The direction of propagation is the *z-axis*, while the section plane is *xy*.

The first two modes are called: *symmetric* (or even) and *asymmetric* (or odd). These modes ar mutually orthogonal, they do not exchange power, and they have different velocities of propagation. We name them ψ_s and ψ_a.
Moreover, since they propagate indipendently of each other they do not change "shape"!
Once we have these mdoes we can actually derive the description of the fields evolution along the guides. Unfortunately, it is not an easy task...



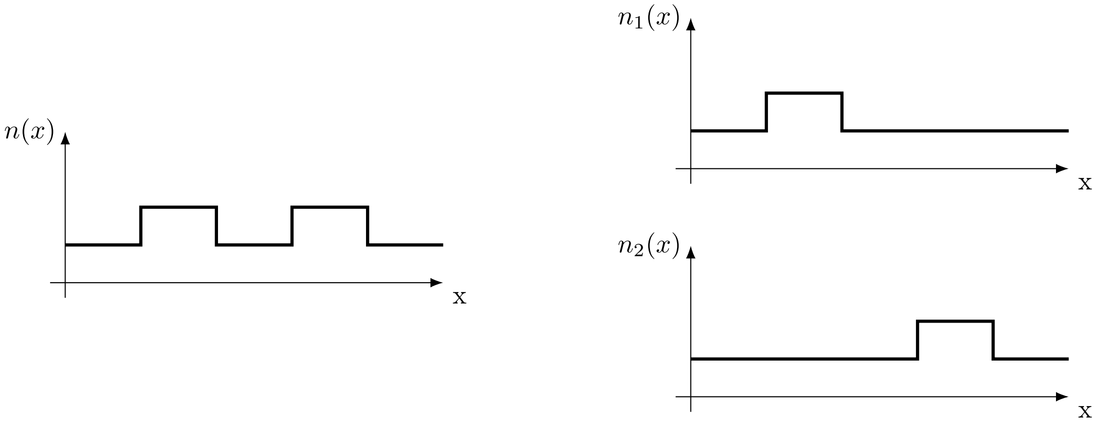
Now, if we sum them and subtract them, we would obtain two graphs that resemble the modes of the guides taken separately: ψ_1 and ψ_2. These are easier to find, but more importantly we can always derive the *exact* modes of the structure with a proper combination of them. The farther are the guides, the better is this approximation. Since the exact modes have different phase velocities these mode change shape while propagating.
We called ψ_1 and ψ_2 *coupled modes*, because they are the result of two different guides close enough to allow the coupling of their modes.
We can define them as follows:

$$\psi_1 \simeq \frac{\psi_s + \psi_a}{2} \qquad \psi_2 \simeq \frac{\psi_s - \psi_a}{2}$$

The formal derivation of the *coupled modes equations* starts with the wave equation of the coupled structure:

$$\nabla_t^2 \Psi + \frac{\partial^2 \Psi}{\partial z^2} + k_0^2 n^2(x,y)\Psi = 0$$

∇_t is the gradient along the trasverse plane.
Note that the refractive index does not depend on z because of the z-invariant hypothesis. Moreover, we have the following schemes for the refractive indexes:



In order to use the coupled modes theory we consider two guides that are not strongly coupled. In this way, the supposed solution becomes a linear combination of the coupled modes:

$$\Psi(x,y,z) = A(z)\psi_1(x,y)e^{-\jmath\beta_1 z} + B(z)\psi_2(x,y)e^{-\jmath\beta_2 z}$$

We have two complex amplitudes because the "shape" of these approximated modes change with z.
Next, we replace the assumed solution in the equation of the wave, and we obtain:

$$k_0^2 \Delta n_1^2 A\psi_1 + k_0^2 \Delta n_2^2 B\psi_2 e^{j\Delta\beta z} - 2\jmath\beta_1\psi_1\frac{dA}{dz} - 2\jmath\beta_2\psi_2\frac{dB}{dz}e^{j\Delta\beta z} = 0$$

Observations on how we obtained this result:

- We define:
$$\Delta n_1^2 = n^2(x,y) - n_1^2(x,y) \qquad \Delta n_2^2 = n^2(x,y) - n_2^2(x,y) \qquad \Delta\beta = \beta_1 - \beta_2$$
- Since the guides are weakly coupled, $A(z)$ and $B(z)$ vary very slow:
$$\frac{\partial^2 A(z)}{\partial z^2} \to 0 \qquad \frac{\partial^2 B(z)}{\partial z^2} \to 0$$

Now, a very important step: to find the coupled equations we need to project the previous equation on the two modes. We do this multiplying the equation to ψ_1^* one time and to ψ_2^* the other.

$$\begin{cases} \dfrac{dA}{dz} = -\jmath\kappa_{11}A(z) - \jmath\kappa_{12}B(z)e^{j\Delta z} \\ \dfrac{dB}{dz} = -\jmath\kappa_{22}B(z) - \jmath\kappa_{21}A(z)e^{-j\Delta z} \end{cases} \Leftrightarrow \quad \kappa_{ij} = \frac{k_0^2}{2\beta_i} \cdot \frac{\int\int_{-\infty}^{+\infty} \psi_j \Delta n_j^2 \psi_i^* dxdy}{\int\int_{-\infty}^{+\infty} \psi_i \psi_i^* dxdy}$$

However, this is not the form that we prefer.
The actual expressions are derived after having done the following substitution:

$$\boxed{\begin{cases} a(z) = A(z)e^{-\jmath\beta_1 z} \\ b(z) = B(z)e^{-\jmath\beta_2 z} \end{cases} \Leftrightarrow \begin{cases} \dfrac{da}{dz} = -\jmath(\beta_1 + \kappa_{11})a(z) - \jmath\kappa_{12}b(z) \\ \dfrac{db}{dz} = -\jmath(\beta_2 + \kappa_{22})b(z) - \jmath\kappa_{21}a(z) \end{cases}}$$

The physical meanings of the coefficients are:

- κ_{ii}: phase constant corrections due to the presence of the second guide. They are usually very small;
- κ_{ij}: constants that quantify the coupling between the guides. They are equal if the guides are identical and they go to zero increasing the distance.

It must be reminded, another time, that ψ_1 and ψ_2 are not solution of the structure, they exchange power and they are coupled.
The exact solutions of the guides are the modes ψ_a and ψ_s, that have different phase velocities and that don't exchange power. We can write the total solution as follows:

$$\Psi = A\psi_a e^{-j\beta_a z} + B\psi_s e^{-j\beta_s z}$$

The solution.
We suppose to have $a(z) = a_s e^{-j\beta z}$ and $b(z) = a_a e^{-j\beta z}$, particular solutions. The important detail is that these waves don't change shape or amplitude.
The system of the coupled equations becomes:

$$\begin{cases} a_s(\beta - \beta_1) - \kappa_{12} a_a = 0 \\ \\ a_a(\beta - \beta_2) - \kappa_{21} a_s = 0 \end{cases}$$

The system leads to notrivial solutions only if the determinant is zero, that occurs if β is equal to the eigenvalues.

$$\boxed{\beta_{a,s} = \frac{\beta_1 + \beta_2}{2} \pm \sqrt{\frac{(\beta_1 - \beta_2)^2}{4} + \kappa_{12}\kappa_{21}}}$$

As we know the eigenvalues determine the eigenvectors, that are the fields configurations in the structure.
The general solutions of the coupled equations are:

$$\boxed{\begin{aligned} a(z) &= a_s e^{-j\beta_s z} + a_a e^{-j\beta_a z} \\ \\ b(z) &= \frac{\beta_s - \beta_1}{\kappa_{12}} a_s e^{-j\beta_s z} + \frac{\beta_a - \beta_2}{\kappa_{21}} a_a e^{-j\beta_a z} \end{aligned}}$$

Keep in mind that a_a and a_s are the amplitudes of the exact modes, related to the exact solutions of the coupled structure. Instead, $a(z)$ and $b(z)$ are the amplitudes of the fields in the guides.

The matrix configuration.
We call I_1 and I_2 the input field complex amplitudes.

$$\begin{bmatrix} O_1 \\ O_2 \end{bmatrix} = T_C \begin{bmatrix} I_1 \\ I_2 \end{bmatrix}$$

T_C is known as transmission matrix.
From the solution obtained before, we derive its explicit expression:

$$T_C = e^{-j\bar{\beta}z} \begin{bmatrix} \cos\delta z - jRsin\delta z & -jS\sin\delta z \\ \\ -jS\sin\delta z & \cos\delta z + jRsin\delta z \end{bmatrix}$$

In the following there are the definitions of the parameters and some properties.

$$R = \frac{\Delta\beta}{2\delta}$$

$$S = \frac{\kappa}{\delta}$$

$$\Delta\beta = \beta_1 - \beta_2$$

$$\bar{\beta} = \frac{\beta_1 + \beta_2}{2}$$

$$\kappa^2 = \kappa_{12} \cdot \kappa_{21}$$

$$\delta = \sqrt{\frac{\Delta\beta^2}{4} + \kappa^2}$$

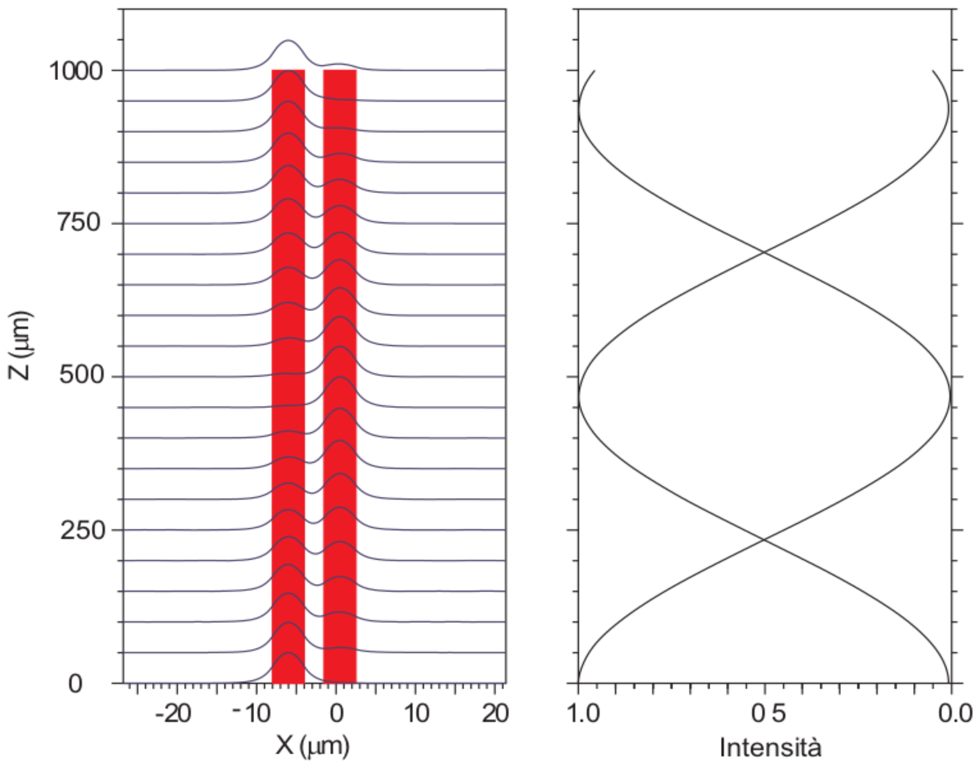
- R and S gauges the symmetry of the structure. If $\beta_1 = \beta_2$, then $R = 0, S = 1$.
- In this course the losses are to be considered negligible:

$$R^2 + S^2 = 1 \quad \wedge \quad det(T_C) = 1$$

- If $z = 0$, the field is unitary in the first guide and nill in the second. Along z:

$$\begin{cases} |a(z)|^2 = 1 - \frac{\kappa^2}{\delta^2}\sin^2(\delta z) \\ |b(z)|^2 = \frac{\kappa^2}{\delta^2}\sin^2(\delta z) \end{cases}$$

The plot for $\Delta\beta = 0$:



Another time: the total field is give by the superposition of the exact modes, which are the symmetric and asymmetric ones. They have different phase velocities, so the shape of the total field varyies along the guide. This plot can be also described like an exchange of power between the guides of the modes, ψ_1 and ψ_2. This exchange occurs with a periodicity, in general:

$$L_c = \frac{\pi}{2\delta} = \frac{\pi}{\beta_1 - \beta_2}$$

It is defined as the minimal length at which there is the maximum power exchange, that is:

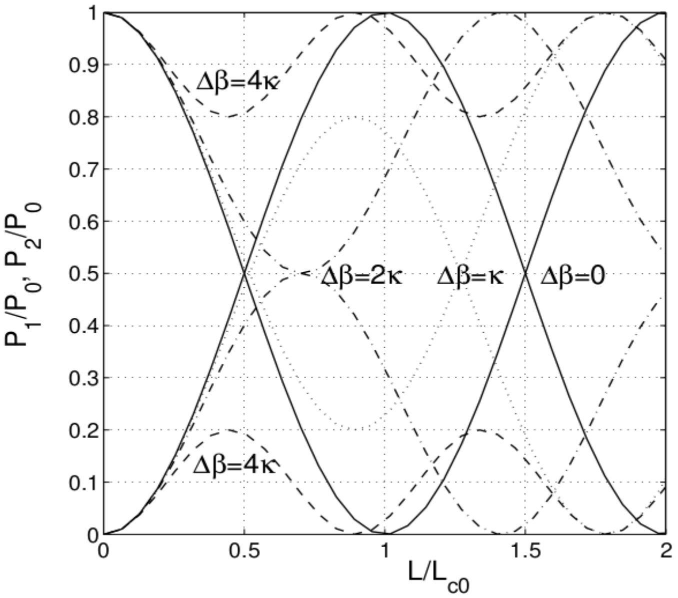
$$|b|^2_{max} = \frac{\kappa^2}{\delta^2}$$

However, if $\Delta\beta = 0$ the power exchange is total and it happens after a distance equals to $L_{c0} = \frac{\pi}{2\kappa}$.

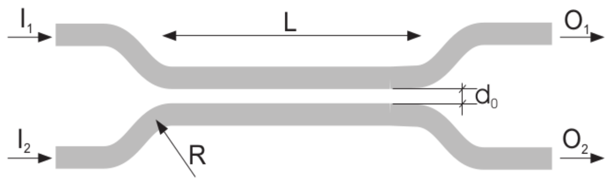
We end up showing the particular T_c for this case, that is called more oftenly *synchronous coupling*:

$$T_c \begin{bmatrix} \cos(\kappa L) & -\jmath \sin(\kappa L) \\ -\jmath \sin(\kappa L) & \cos(\kappa L) \end{bmatrix} \Rightarrow \begin{cases} P_1(z) = P_0 \cos^2(\kappa z) \\ P_2(z) = P_0 \sin^2(\kappa z) \end{cases}$$

The relationship between $\Delta\beta$ and z is shown in the following picture:



3.2 Directional Coupler: a general overview



The directional coupler is a fundamental device in integrated optics. It's a 2-ports device that makes possible to divide or merge the light, and we will learn how to describe it with a matrix relationship between the inputs and ouputs.
The guides are coupled for a length L only. The transition is designed to avoid radiative losses.
The main parameter is the *coupling ratio*:

$$K = \frac{P_2}{P_1 + P_2}$$

But, supposing that there are no losses and that power flows out from one guide only, we have that:

$$K = \frac{P_2}{P_0}$$

However, in general we call P_{bar} the power that comes out the guide the wave enters at the input. Instead, P_{cross} is the power flowing out the coupled guide.
The critical step is choosing κ and $\Delta\beta$: if κ is large the guides are close but the coupler is very sensitive to construction tollerances. On the other hand, if κ is small, the coupler is less sensitive but it has greater dimensions.
We set them by trial, without neglecting technological aspects. Anyway, in general the smallest sensitivity to the parameters κ, $\Delta\beta$ and L is obtained if:

$$L = \frac{\pi}{2\kappa}\sqrt{K} \qquad \Delta\beta = 2\kappa\sqrt{\frac{1}{K} - 1}$$

Moreover, it's possible to compute also the lenght of the curved parts:

$$L_t = \int_0^{\infty} e^{-\gamma(d(z)-d_0)} dz = \frac{1}{2}\sqrt{\frac{\pi R}{\gamma}}$$

This result is obtained considering the radius of the transitions constant.

3.2.1 "-3dB" Coupler

Like the title suggests, this directional coupler divides equally the input light.
In the simple case of $\Delta\beta = 0$, we derive the condition $\kappa L = \pi/4$. Since the transfer function is periodical, an alternative solution is $\kappa L = (2N+1)(\pi/4)$. However, even if it's right mathematically, the technological realization is different.
The transmission matrix for a synchronous coupling:

$$T_c = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -j \\ -j & 1 \end{bmatrix}$$

Considering the graph $\Delta\beta$ versus L_c/L_{c0} it's possible to obtain the division in equal halfs until the limit value $\Delta\beta = 2\kappa$: $\kappa L = \pi/(2\sqrt{2})$:

$$T_c = -\frac{\jmath}{2\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

3.2.2 WDM coupler

The properties of the directional coupler depend on parameters κ and $\Delta\beta$, which both vary with the frequency.
While $\Delta\beta$ depends on λ^{-1}, we can consider κ varying with λ: the superposition integral depends on λ^2 for a lot of different guides.
Consequently, K increases with λ, but $\Delta\beta$ can be choosen to make K insensitive to the wavelength for large bandwidths: this particular value would be found through numerical simulations.
Actually, for large κL the directional coupler becomes very sensitive to λ. We can exploit this relation to design coupler that can filter signals: they are called WDM.
We suppose to enter in one guide with two different wavelengths λ_1, and λ_2. We want to obtain at the ouput this following configuration:

$$\begin{cases} P_{cross}(\lambda_1, L) = \sin^2(\kappa_1 L) = \sin^2(\kappa_0\lambda_1 L) = 1 \\ P_{bar}(\lambda_2, L) = \cos^2(\kappa_2 L) = \cos^2(\kappa_0\lambda_2 L) = 1 \end{cases}$$

In this way the signal is completely ($\Delta\beta = 0$) coupled to λ_1, that ends in the bar. Instead, λ_2 wil be discriminated and diverted into cross guide.

$$\begin{cases} \kappa_0\lambda_1 L = \pi\left(N_2 \pm \frac{1}{2}\right) & \textit{cross} \\ \kappa_0\lambda_2 L = \pi N_1 & \textit{bar} \end{cases}$$

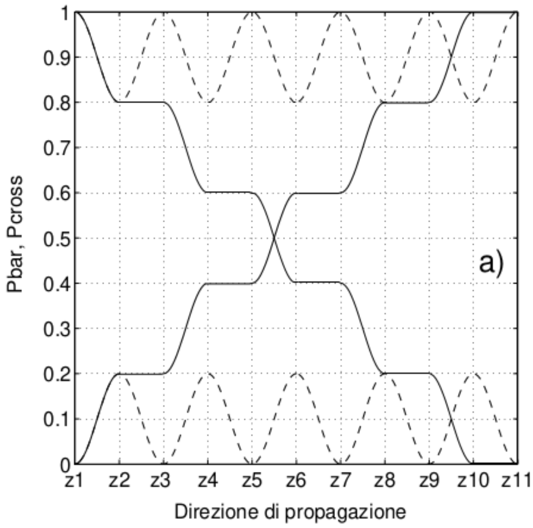
We use the sign + when $\lambda_1 > \lambda_2$, - in the opposite case. Moreover $N_1 = N_2 = N$.

$$\kappa_0 L = \frac{\pi}{2} \cdot \frac{1}{\lambda_1 - \lambda_2} \qquad\qquad N = \frac{1}{2} \cdot \frac{\lambda_2}{|\lambda_1 - \lambda_2|}$$

In the end, we derive $\kappa_0 L$ to have a certain wavelength in a specific output guide.
The problem is that in general, λ_1 and λ_2 cannot satisfy the equation to compute N: since N is a integer number, the operation makes sense only for values of λ that are far away in the spectrum.

3.3 Coupling assisted by grating

In this paragraph we will study how to guarantee the maximum power transfer for an asynchronous coupling. We recall that for two guides with different phase constants the maximum energy transfer that is achievable is $|b(z)|^2_{max} = \kappa^2/\delta^2$.
Anyway, do observe the following graph:



The dashed line represents the power in the guides. As we can see, if we could stop the coupling at z_2 and let it start again at z_3, the power will contiuosly move from a guide to the other. At the end the guides will have exchanged all the power, as it was a synchronous directional coupler.
This periodical interruption is realizable varying the spacing within the guides, or the refractive index profile.

For a formal computation, we do the following hypothesis:

1. the guides are weakly coupled: $a(0) = a_0 \simeq a(z)$;
2. there are no retrofelctions;
3. $\Delta\beta$ is small;
4. $b(0) = 0$;
5. κ_{22} is negligible.

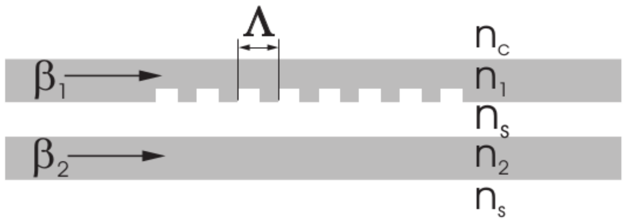
We can write:

$$b(z) = a(0) \int_0^z \jmath\kappa_{21}(\xi) e^{\jmath\Delta\beta\xi} d\xi$$

As we can see, it's a Fourier transformation.
The coupling is actually possible if $\kappa_{21}(z)$ is periodical, with period equals to or multiple of $\Delta\beta$. Obviously changing the profile we should change also $\kappa(z)$.
For any possible case, anyway, there is a condition that must be respected, called *phase matching*:

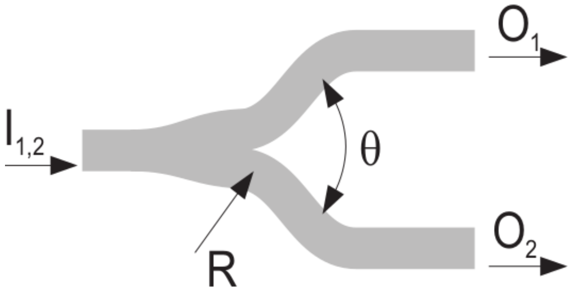
$$(\beta_1 - \beta_2) \cdot \Lambda = 2\pi$$

Where Λ is the lattice parameter ("passo reticolare").



3.4 Bifurcations

Like suggested by the title, these devices consists in a division of a single guide into two different ones.
As for the Directional Couplers an exact analysis of the structure is really difficult, but we can rely on the coupled mode theory.
The functional principle: the original guide is bimodal (at least) and they are the fundamental modes of the output guides. If the central guide is monomodal, one of the output mode is radiative.



A symmetric bifurcation is called Y-branch and it divides the light equally in halfs.
Usually, if the input guide is monomodal, in order to make it bimodal we insert a taper to double the dimensions. The transition zones must be as gentle as possible to avoid the excitation of radiative modes.

Let's analyze two possible cases:

a) do consider the excitation of the fundamental mode of the central guide, with an amplitude a_1. Whent it reaches the bifurcation zone only the symmetric mode arises. It is then divided into the fundamental modes of the guides, with equal ampltiudes:

$$b_1 = b_2 = \frac{a_1}{\sqrt{2}}$$

b) If we enter from one of the previusly labeled output guides with the relative fundamental mode, both symmetric and asymetric modes are excited. These are the same of the central guide and they have the same amplitude. However, since we have take the central guide as monomodal, the asymmetric one is radiated away.

Next, we derive the expression of the transmission matrix, T_Y.
We call a_1 and a_2 the amplitudes of the *fields* at the input, and b_1 and b_2 those of the fileds at the ouput.

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = T_Y \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Where:

$$T_Y = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

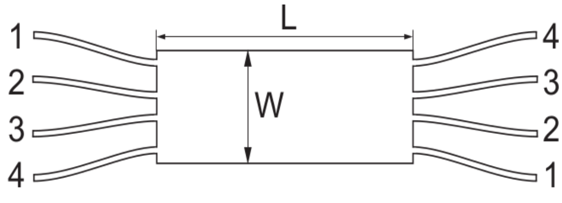
If we had complex amplitudes $a_{1,2} = a \cdot \exp(\pm \jmath\varphi)$:

$$b_1 = \frac{2}{\sqrt{2}} a \cos\varphi$$

$$b_2 = \jmath \frac{2}{\sqrt{2}} \sin\varphi$$

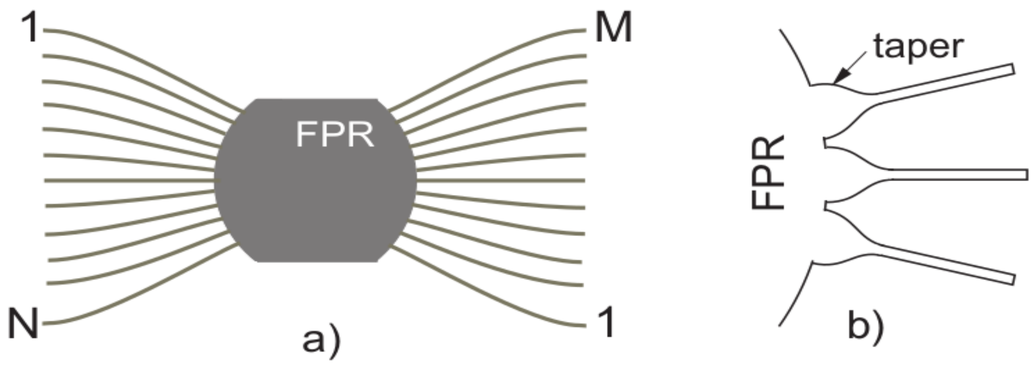
3.5 MMIC: Mulitmode interference Coupler

We have N guides at intput, a slab waveguid whose dimensions are Width and Length, and finally P guides at the ouput.
The input filed can be described as the linear combination fo the slab modes.



Transverse coordinate:	x
Number of modes:	M
Mode implicit expression:	$\psi(x)$
Complex amplitude:	C_m
Constant of propagation:	β_m
Field in generic z:	$\Psi(x,z) = \sum_{m=0}^{M-1} C_m \psi_m(x) \exp(-\jmath\beta_m z)$

3.5.1 Star Couplers



- N,M: numbers of input/output guides. They are arranged radially.
- The FPR, free propagation region, is guided only in the vertical plane. It is limited by two arcs of circumference with radius R, whose centers lay on the oppsite arc. This latter requirement let us describe the distribution of the field in the ouput array as the Fourier transformation of the field distribution in the input guides.
- The dimensions of the star coupler are such that we can consider the output guides in the far-field region (useful to apply some simplifictions).
- The angular difference between two adjacent guides is α. The spatial distance is $a = \alpha \cdot R$. The total extension of the arc, that is the angular extension of the input/output arrays, is θ_M, hence $M \cdot a = \theta_M \cdot R$.
- The width of each guide is d.

We suppose that in the FPR the optic field propagates like a gaussian beam.
Moreover, we suppose also that the power of the guide p is equally divided into the M guides, determing the follwing tranfer function between two generic guides "p" and "q":

$$T(p,q) = \frac{\eta}{\sqrt{M}} e^{\jmath\phi_{pq}}$$

where η is the efficiency. Exploiting the formula describing gaussian beams, to reach an approximated uniformity in the ouput guides, the following relations is obtained:

$$R = \frac{M \cdot a \cdot d \cdot \pi n_{eff}}{2\lambda}$$

4 Filters

The transfer function of a generic filter can be expressed as a Fourier series:

$$H(f) = \sum_{m=0}^{N} c_m e^{\jmath\phi_m} e^{-2(2m\pi fT)}$$

where N is the order of the filter and $c_m e^{\jmath\phi_m}$ are the coefficients. In the classic optics this operation is realized with interferometers.

4.1 Spectral characteristics

The filtering, in general, is obtained inducing an interference between two or more waves delayed one another. They must have the same polarization, the same frequency and they must be temporally coherent.
Considering two waves that travels along two differet paths, their relative phase is:

$$\Delta\phi = \frac{2\pi}{\lambda}(n_{eff,1}(\lambda)L_1 - n_{eff,2}(\lambda)L_2)$$

This difference is equal to multiples of 2π everytime that $f = f_m$, so it's periodical:

$$\Delta\phi(f_m) = \frac{2\pi f_m}{c}\Delta L_{eff}(f_m) = 2m\pi \quad \Rightarrow \quad \Delta L_{eff}(f_m) = \frac{m \cdot c}{f_m}$$

Keep in mind that the transfer function mantains the same periodicity of $\Delta\phi$, and this periodicity is named *FSR*, or *free spectral range*. We compute it referring to the central frequency, f_0, between f_m and f_{m+1}.

$$\Delta L_{eff}\begin{pmatrix} f_{m+1} \\ f_m \end{pmatrix} = \Delta L_{eff}(f_0) \pm \frac{FSR}{2}\frac{d\Delta L_{eff}}{df}\bigg|_{f_0}$$

Now we can isolate FSR:

$$FSR = \frac{c}{\Delta L_{eff}(f_0) + f_0 \left.\frac{\Delta L_{eff}}{df}\right|_{f_0}} = \frac{c}{n_g \Delta L}$$

Where we supposed that $n_{eff,1} = n_{eff,2}$. Besides, n_g is commonly called the group index, defined as follows:

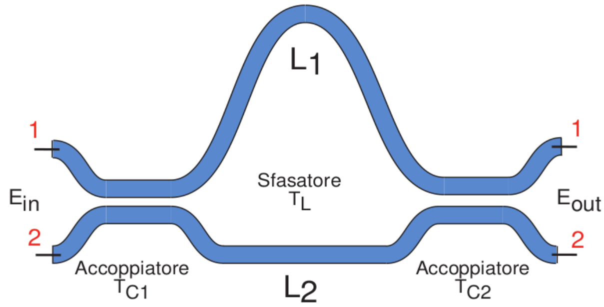
$$n_g = n_{eff}(f_0) + f_0 \left.\frac{dn_{eff}}{df}\right|_{f_0} = n_{eff}(\lambda_0) + \lambda_0 \left.\frac{dn_{eff}}{d\lambda}\right|_{\lambda_0}$$

It can be useful to point out that the FSR is the inverse of the delay between the waves.
In the end, just for the sake of completeness:

$$\Delta\lambda_{FSR} \approx \frac{\lambda^2}{n_g \Delta L}$$

4.2 Mach-Zehnder filter

It's the easiest filter in integrated optics. The transfer function is sinusoidal, that is not an excellent function to filter WDM channels, but it's in fact the actual basis to realize more efficient filters.



The filter consists of three elementar blocks: two couplers (usually -3dB) and one phasor shifter in between. To find the total transfer function we consider the blocks separately, and in the following we suppose that $T_{C1} = T_{C2} = T_C$.
Explicitly:

$$T_C = \begin{bmatrix} \cos(\kappa L_c) & -\jmath\sin(\kappa L_c) \\ -\jmath\sin(\kappa L_c) & \cos(\kappa L_c) \end{bmatrix} \qquad T_L = \begin{bmatrix} \exp(-\jmath\varphi_{L1}) & 0 \\ 0 & \exp(-\jmath\varphi_{L2}) \end{bmatrix} = \exp(-\jmath\varphi_{L1}) \begin{bmatrix} 1 & 0 \\ 0 & \exp(\jmath\Delta\varphi_L) \end{bmatrix}$$

Where φ_{Li} is the phase shift along the guide i of the shifter. They are computable doing the integral of the propagation vectors over the lengths of the guides:

$$\varphi_{L1,2} = \int_0^{L_{1,2}} \beta_{1,2}(l)dl$$

Neglecting the dependency of β on the curvature:

$$\Delta\varphi_L = \varphi_{L1} - \varphi_{L2} \simeq \beta_1 L_1 - \beta_2 L_2 = \frac{2\pi}{\lambda}(n_{eff}L_1 - n_{eff2}L_2) = \frac{2\pi}{\lambda}\Delta L_{eff}$$

So we see that the phase shift is due to different lengths, different refractive indexes, or both.

The general input-output relation is:

$$\begin{bmatrix} E_{out,1} \\ E_{out,2} \end{bmatrix} = T_C \cdot T_L \cdot T_C \cdot \begin{bmatrix} E_{in,1} \\ E_{in,2} \end{bmatrix} \equiv T_{MZ} \cdot \begin{bmatrix} E_{in,1} \\ E_{in,2} \end{bmatrix} = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} \begin{bmatrix} E_{in,1} \\ E_{in,2} \end{bmatrix}$$

Hence:

$$\begin{cases} P_{bar} = |T_{11}|^2 = \cos^2\left(\frac{\Delta\varphi_L}{2}\right)\cos^2(2\kappa L_c) + \sin^2\left(\frac{\Delta\varphi_L}{2}\right) \\ \\ P_{cross} = |T_{21}|^2 = \cos^2\left(\frac{\Delta\varphi_L}{2}\right)\sin^2(2\kappa L_c) \end{cases}$$

In the case of ideal directional couplers, with $\kappa L_c = \pi/4$, we have:

$$\begin{cases} P_{bar} = \sin^2\left(\frac{\Delta\varphi_L}{2}\right) \\ \\ P_{cross} = \cos^2\left(\frac{\Delta\varphi_L}{2}\right) \end{cases}$$

Obviosluy, under condition of lossless propagation, $P_{bar} + P_{cros} = 1$.
The condition $P_{bar} = 0 \wedge P_{cross} = 1$ is obtained when $\Delta\varphi_L = 0$ (or $2N\pi$, in general).
Even if the Mach-Zehnder transfer function resembles that of a classic directional coupler, it is not proportional to κL, but to ΔL_{eff}. This is the reason why the Mach-Zehnder is way more controllable. Moreover, the function can be modified after the realization, changing suitably the n_{eff} thanks to different techniques: we could change the temperature of the guide, or through electro-optic operations, or also with a UV trimming.

4.2.1 Mach-Zehnder Design: an example

We want to separate two channels: one centered at $\lambda_1 = 1550$ nm and the other one at $\lambda_2 = 1550.8$ nm.
We suppose to realize the filter with glass on silicon technology and employing guides with $n_{eff} = 1.46$ and $n_g = 1.51$.
We must determine ΔL.

Since FSR must be equal to twice the spectral spacing:

$$\Delta\lambda = 0.8 \ \ nm \qquad \Leftrightarrow \qquad \Delta f = \frac{c\Delta\lambda}{\lambda^2} = 99.84 \ \ GHz$$

We obtain:

$$\Delta L_{eff} = 1502.34 \ \ \mu m \Rightarrow \Delta L = 994.93 \ \ \mu m$$

However, as it is exaplained for WDM couplers, the latter condition is not enough to guarantee that $P_{bar}(\lambda_1) = 1$ and $P_{cross}(\lambda_2) = 1$. In other words, these two conditions cannot be satisfied in the same time. For example, if we impose $P_{cross}(\lambda_1) = 1$, we have:

$$\frac{\pi n_{eff} \Delta L}{\lambda_1} = N\pi \Rightarrow N = 937.16$$

"N" will be approximated to the closest integer value, $N = 937$.
Hence $\Delta L = 994.76$ and the *cross-talk* (the signal at λ_2 flowing in the cross guide) is -25.6 dB, a value that is acceptable or not depending on the case.

In general it is more important to study the sensibility to errors in the realization of the ΔL_{eff}.
An error in balance does not alter the shape of the tranfer function, but it leads to a translation. Indeed, it's sufficient that ΔL_{eff} changes by one wavelength to have the response traslated by a FSR.
For istance, to guarantee a traslation $\delta f \leq B$, where B is the bandwidth for which the cross talk is not higher than X_t, it must be imposed that

$$\delta f \leq B/2 = FSR\sqrt{X_t}/\pi \quad \rightarrow \quad \delta \Delta L_{eff} \leq \frac{\lambda}{\pi}\sqrt{X_t}$$

4.3 Multi-state Mach-Zehnder

a) **Cascade Mach-Zehnder**
If we connect more directional couplers in succession, the total transfer function is the product of the transfer functions of the single stages.
If we have N stages the project requires 2N couplers.
The simplest design is realized with the same stage repeated. In this way the transfer function is simply $P_{bar} = \sin^2(\Delta\varphi/2)$. The selectivity and the rejection out-of-bandwidth are high, but the shape of the bandwidth is not flat.
To increase the rejection, we could realize the filter with growing imbalances in power of 2. This technology is exploited to realize filters that exctract WDM channels.

As for MUX or DEMUX we need $\log_2 N$ stages in order to merge N channels, that are joined 2 by 2. This type of device it's a good choice for a small numbers of channels (from 4 to 8) and when there are not required strict spectral characteristics, i.e. when the channels are not too close. Actually, to improve the characteristics each stage can be realized with cascade Mach-Zehnder.
The stages employ 3dB couplers and they are realized to center properly the output signals around the wavelenghts of interest, thus they can differs one from each other. The transfer function of a generic filter of order N and for the n_{th} path:

$$H_n = \frac{1}{N}\sum_{k=0}^{N-1} \exp\left(\jmath 2\pi \frac{nk}{N}\right) \exp\left(-\jmath \beta k \Delta L\right)$$

b) **Lattice Mach-Zehnder filter**
These devices are realized connecting both outputs to the inputs of the next stage. It is a structure way more flexible but with better properties.
The main advantages consist of small losses in the passband and N+1 couplers for a N-order filter.
The phase shifter are all equal, then we can express the transfer function with Fourier series:

$$H(\lambda) = \sum_{k=0}^{N_1} c_k \exp(-\jmath \beta k \Delta L)$$

Actually, these filters are designed relying on iterative techniques imposing an objective function for amplitude and phase. To have all the power in the cross guide for the desired frequencies we must impose $\beta \Delta L = M\pi/2$. Thus, the whole structure can be seen as a big directional coupler long as much as the

sum of the stages lengths.
It results:

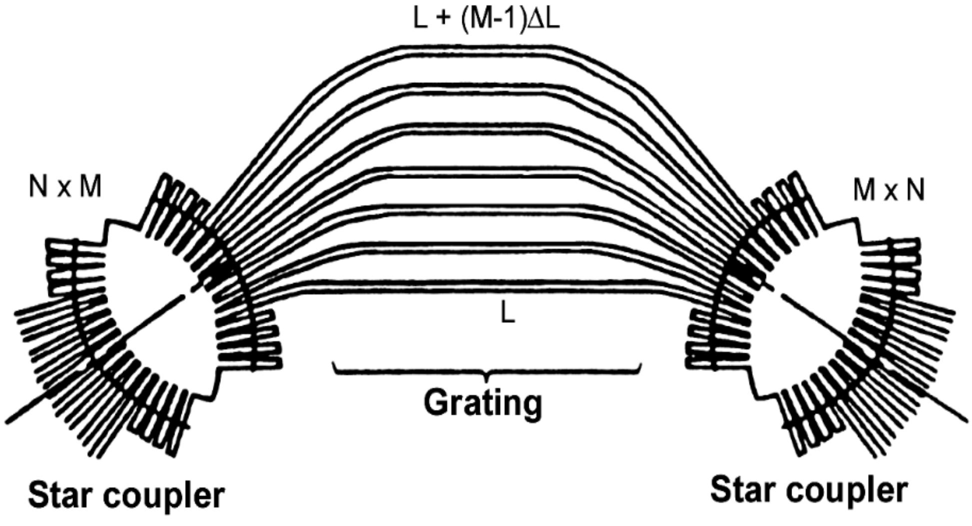
$$\sum_{i=0}^{N-1} \cos^{-1}(\sqrt{K_i}) = \frac{\pi}{2}$$

The main applications of these filters are to realize FIR or function of add/dropp, and equalize amplitude and phase on large bandwidths.

4.4 Arrayed waveguide gratings

To generilize, an AWG consists of two MMI (or star coupler), the first has $N \times M$ guides while the second one has $M \times N$ guides. The two devices are joined through the array of the M-guides, which have increasing length. The difference in optical path of two adjacent array guides is ΔL. If the first has a value of L_0 the last one is long $L_0 + \Delta L \cdot (M-1)$.

The operting principle: the signal that reaches the first star coupler incoming from the N_{th} guide is divided equally in the M guides of the array. In particular *is the power of the signal that is divided*!
In the second star coupler the waves recombine.
Consider what happens for the generic central guide "s". We use "p" to indicate the guide in the first star coupler the wave is coming fro, and "q" for the guide of the second in which the wave is collected.



$$E_{pq} = E_{in} \frac{1}{\sqrt{M}} \frac{1}{\sqrt{M}} e^{j\phi_{psq}} \qquad \phi_{psq} = \phi_{ps} + \phi_s + \phi_{sq}$$

Where ϕ_{psq} is the total phase shift caused by the three devices (first star coupler, guide s, second star coupler), shown explicitly in the follwing.

$$\begin{cases} \phi_{ps} = \frac{2\pi}{\lambda} n_{eff} \cdot R \cdot (1 - ps\alpha^2) \\ \phi_s = \frac{2\pi}{\lambda} \int_0^{L_s} n_{eff}(l) dl \simeq \frac{2\pi}{\lambda} n_{eff}(L_s - L_1) = \frac{2\pi}{\lambda} n_{eff} \cdot s \cdot \Delta L \\ \phi_{sq} = \frac{2\pi}{\lambda} n_{eff} \cdot R \cdot (1 - sq\alpha^2) \end{cases}$$

We've assumed that the star coupler are identical and we called n_{eff} the index of both guides and FPR (Free Propagation Region). The approximation made is valid for very large radius of curvature, as previoulsy shown for Mach-Zehnder shifters.
Now, the interesting parameter is the phase shift between two adjacent guides, *s* and *s-1*.

$$\Delta\phi_{pq} = \phi_{psq} - \phi_{p(s-1)q} = \frac{2\pi n_{eff}}{\lambda}[\Delta L - R(p + q\alpha^2)]$$

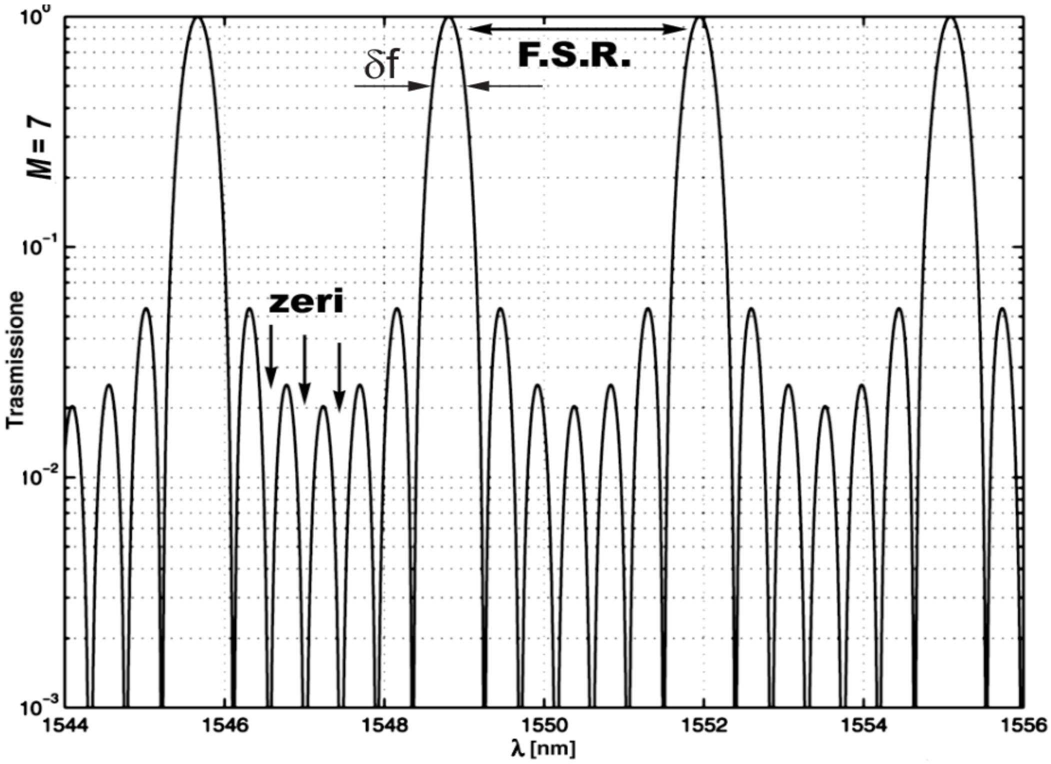
Long story short, the power transfer function between the port p and q is:

$$P_{pq} = \left| \sum_{s=0}^{M-1} E_{in} \frac{1}{M} e^{\jmath \phi_{psq}} \right|^2 = \frac{P_{in}}{M^2} \left| \sum_{s=0}^{M-1} e^{\jmath s \Delta\phi_{pq}} \right|^2$$

But computing the series, we obtain in the end:

$$T(p,q) = \frac{1}{M^2} \frac{\sin^2\left(M \cdot \frac{\Delta\phi_{pq}}{2}\right)}{\sin^2\left(\cdot \frac{\Delta\phi_{pq}}{2}\right)}$$

For M=7:



4.5 Channels sepration, FSR and Applications

The transfer function is sstrongly dependent on M, which represents the number of the guides in the central array.
This function is periodical with respect to $\Delta\phi_{pq}$. In each period there are one peak corresponding to a main lobe, also called primary, and $M-2$ secondary lobes separeted by $M-1$ zeros.
For the values of $\Delta\phi_{pq}$ for which peaks are determined the transfer of power from the port p to the port q is complete. Mathematically:

$$\frac{\Delta\phi_{pq}}{2} = Q \cdot \pi$$

where Q is onviously an integer number.
Exploiting the expression derived in the previous subsection:

$$\lambda_{pq} = \frac{n_{eff}\Delta L}{Q} - n_{eff} \cdot R(p+q)\frac{\alpha^2}{Q}$$

Or:

$$\lambda_{pq} = \lambda_0 - (p+q)\Delta\lambda$$

If we define:

$$\lambda_0 = \frac{n_{eff} \cdot \Delta L}{Q} \qquad \Delta\lambda = \frac{n_g R \alpha^2}{Q} = \frac{n_g \alpha^2}{Q \cdot R}$$

where n_g is the group index of the guides.
Recallin what we have said about star coupler, to obtain an approximated uniformity in the ouput guides we have to optimize the ratio d/a. If N is the number of channels, and FSR is equal to $N\Delta\lambda$, we find that

$$\frac{d}{a} \leq \frac{2N}{\pi M}$$

Similarly to what we have said about Mach-Zehnder interferometer, the periodicity of the trasnfer function is equal to the FSR:

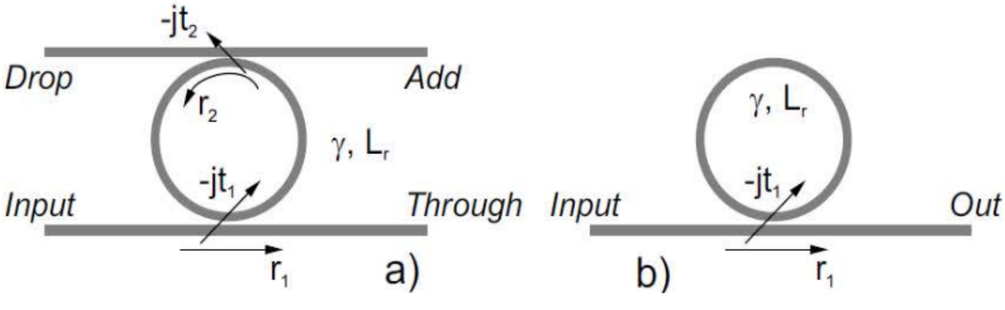
$$FSR = \frac{c}{n_g \Delta L} = \frac{c \cdot n_{eff}}{\lambda_0 \cdot Q \cdot n_g}$$

The FSR deos not depend on the parameter M, but is dependent on optic path difference.
When we design a filter like this one, we need to choose a good compromise between the available bandwith and the croostalk. Moreover, AWG are known to be very sensitive to temperature, polarization and technology process tollerances.

First of all, AWG can be employed to realize DEMUX/MUX. In the first case the input port is only one and the signal contains different wavelengths. They are suitably spaced one another. Ideally each wavelength will be directed to a different output guide, respecting the equation of λ_{pq}. Changing the input guide also the distribution of ouput channels will change.
For the second case, from each input guide enters a signal at a different wavelength. However, the values are such that all the waves end in the same channels, i.e. the same ouput guide.
Another possible application of the AWG is the router. The most promising employment of these WGR is the realization of Optical ADD/DROP.

4.6 Resonators Filters

The Fabry-Perot cavity consists of two mirrors placed one in front of another with an active medium in between. It's the most famous interferometer in the classic optics, but the difficulties inherent in making mirrors with low losses for integrated optics lead to a different design.
The cavity is substituted with a guide closed in a loop and coupled to the outer guides with a directional coupler. These devices are called resonant rings, their dimensions are in the order of the wavelength and they belong to the category of IIR.



4.6.1 The Ring Filter

The model *a* is a filter. If the signal incoming from the *input port* has the same frequency of the resonant signal of the ring, it exits from the *drop port*. Otherwise, it goes on and reaches the *through* port. The fourth port is used to add the resonant signal to the through ouput. The trasfer function is calculcated by placing in cascade the blocks of the various elementary matrices.

For istance, in the hypothesis of two different couplers, ring length L_r and round-trip losses of γ, the realation H_t (through) between input and output (through) is:

$$H_t = \frac{r_1 - \gamma r_2 e^{-\jmath\beta L_r}}{1 - \gamma r_1 r_2 e^{-\jmath\beta L_r}}$$

And similarly we obtain:

$$H_d = -\frac{t_1 t_2 \sqrt{\gamma} e^{-\jmath\beta L_r/2}}{1 - \gamma r_1 r_2 e^{-\jmath\beta L_r}}$$

- The resonance occurs when $\beta L_r = 2m\pi$, that is, when light make a whole number of wavelengths in the ring. "m" is known as full order of resonance. At resonance, if $r_1 = \gamma r_2$ the through port is isolated because the signal is completely diverted in the drop port.
- Periodicity of the transfer function is $FSR = {}^{c}/_{(n_g L_r)}$.
- In the case of identical couplers, with $t_i^2 = K$, and lossless condition, the rejection (the ratio between H_d in resonance and anti-resonance) is

  $$ER = \frac{(K-2)^2}{K^2}$$

  It gauges the cross-talk intoduced. The smaller the K and the smaller the cross-talk, although the bandwidth narrows.
- The bandwidth is:

  $$B = \frac{FSR}{\pi} \cdot \frac{K}{\sqrt{1-K}}$$

- Other important parameters are the quality factor and the finesse.

  $$Q = f_m/B \qquad \mathcal{F} = FSR/B = Q/m$$

  Finesse is used more and it indicates the filter selectivity and the number of rond trip that light do in the ring. Moreover: the greater the $\mathcal{F}$, the smaller the K, the greater the effect of losses. However, also the group delay and the insertion loss are proportional to the finess.

4.6.2 The Ring shifter

The *b* diagram represents a ring shifter, that is equivalent to an all-pass filter. The phase is non-linear with the frequency:

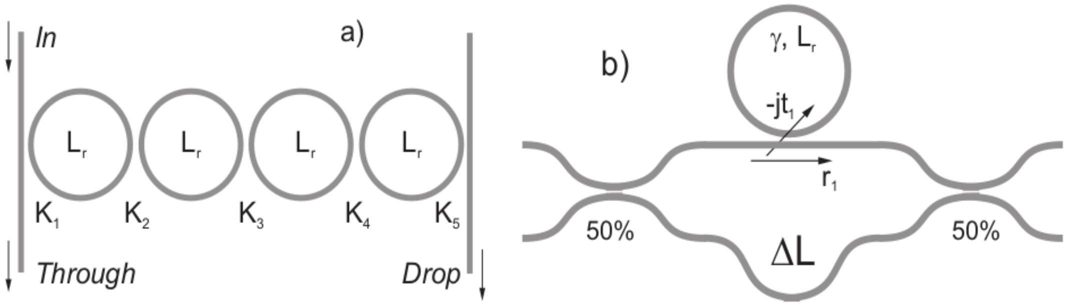
$$\Phi(\omega) = \tan^{-1}\left(\frac{K \cdot \sin(\beta L_r)}{2r - (1+r^2)\cos(\beta L_r)}\right)$$

In general there are two main applications: if we want to use it as a phase-shifter we need low losses, i.e. γ is small; otherwise we could employ the ring as a modulator. Especially if $r = \gamma$ we are in the condition known as "critic coupling": all the light is dissipated by the ring, letting the output isolated.
At the resonance, $r_2 = 1$ and $\beta L_r = 2m\pi$. Hence we have:

$$\tau_g = T\left(\frac{\gamma}{\gamma - r} + \frac{\gamma r}{1 - \gamma r}\right) \xrightarrow{\gamma=1} \frac{1+r}{1-r}$$

T is the time that it takes for the light to complete a round trip.
τ_g can be larger than T, so the rings can be used also to realize optic delay lines.

4.6.3 Rings resonator filters



The expamples shown in the pciture are two different devices: a pass-band filter with directly coupled rings and an interleaver realized with a Mach-Zehnder loaded with a ring shifter.
These are possible solution to achieve a larger integration for optic circuit.

5 Magneto-optic components

5.1 The optical attivity and the Faraday effect

What does it mean the term *optical activity*?
First of all, it is a shorter way to say that the material in which the light is going to pass through can change the polarization of the wave, without affecting the other properties.
More specifically, the dielectric tensor has elements outside the main diagonal that lead to the coupling between the transverse field components. A material like this one is commonly known as anisotropic. For istance and simplicity:

$$\epsilon = \begin{Vmatrix} \epsilon_\perp & \jmath\delta\epsilon & 0 \\ -\jmath\delta\epsilon & \epsilon_\perp & 0 \\ 0 & 0 & \epsilon_\parallel \end{Vmatrix}$$

If we put this tensor in the wave equation we obtain two scalar equations (the parallel component doesn't couple with the others), shown in matrix form:

$$\begin{bmatrix} -\beta^2 + \omega^2\mu_0\epsilon_0\epsilon_\perp & \jmath\omega^2\mu_0\epsilon_0\epsilon_\parallel \\ -\jmath\omega^2\mu_0\epsilon_0\epsilon_\parallel & -\beta^2 + \omega^2\mu_0\epsilon_0\epsilon_\perp \end{bmatrix} \cdot \begin{bmatrix} E_x \\ E_y \end{bmatrix} = 0$$

The eigenvectors and eigenvalues are respectively the fields and their constant propagation. They define which waves can propagate in the medium. This equation allows non-trivial solutions if the determinant of the matrix is zero, hence we have:

$$\textit{Eigenvalues} \qquad \beta_{R,L} = \tfrac{2\pi}{\lambda}\sqrt{\epsilon_\perp \pm \delta\epsilon}$$

$$\textit{Eigenvectors} \qquad E_x = \pm\jmath E_y$$

The second equation tell us that we have found waves with right/left circular polarization.

$$\text{Right circular} \qquad n = \sqrt{\epsilon_\perp + \delta\epsilon}$$

$$\text{Left circular} \qquad n = \sqrt{\epsilon_\perp - \delta\epsilon}$$

Now we define the birefringe of the medium, assuming that $\delta\epsilon \ll \epsilon_\perp$:

$$n_R - n_L \simeq \frac{\delta\epsilon}{\sqrt{\epsilon_\perp}}$$

Therefore, an optical active medium is a medium with circular birefringe that can rotate the polarization plane by an angle θ, called also power optical rotation:

$$\theta = \frac{1}{2}(\beta_R - beta_L)L = \frac{\pi\delta\epsilon}{\lambda\sqrt{\epsilon_\perp}}L$$

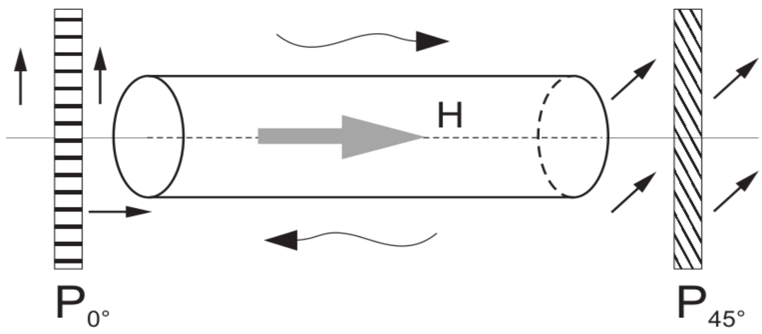
Some materials can be activated with a magnetic field directed along the direction of propagation of the wave. In this case we're talking about the Faraday effect and the component $\delta\epsilon$ depends linearly to the longitudinal component of H:

$$\theta = VBL$$

where V is the Verdet constant, $B = \mu_0 H$ is the magnetic inductance vector and L is the lenght travelled in the medium by the wave.

5.2 The insulator

It is a device that allows the propagation in one direction but not in the oppisite one. To understand the working principle do consider the following system:



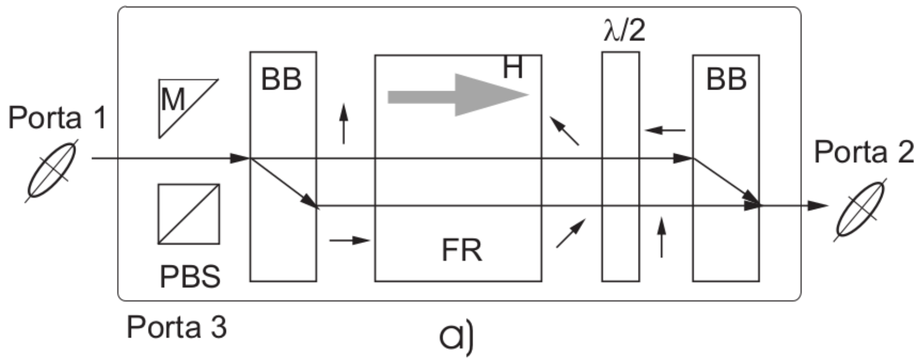
- the incident wave in vertically polarized and is approaching the magneto-optic material in the same direction of the magnetic field applied, H
- the length of the magneto-optical material is such that the polarization plane at the ouput is rotated by 45°
- the right polarizer does not counter the travel of the beam

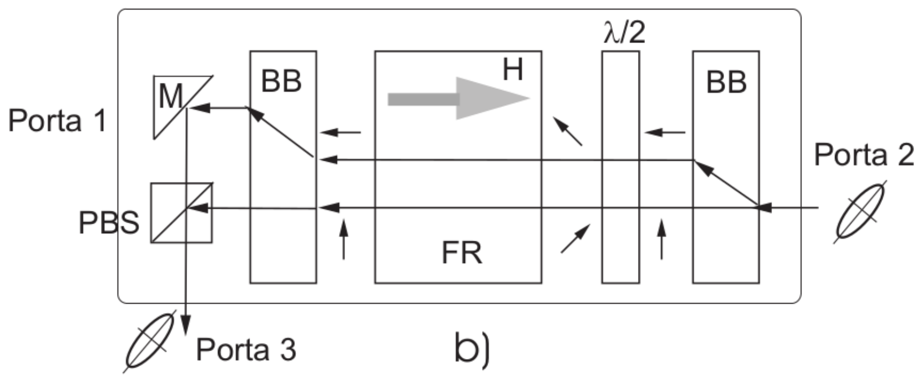
Now, what happens if the light is moving in the opposite direction?

- the wave is polarized by the first polarizer, the plane rotates by 45°
- the active medium has an opposite effect with respect to the previous case because the sign of H is the opposite, so the sign of the element outside the diagonal changes
- the wave met the left polarizer (the vertical one) with a plane that is now horizzontal: the light does not pass!!

5.3 The circulator

The circulators are passive devices with three or more ports that allow the signal to propagate from one port to the other in one direction. For istance, suppose to have three ports:





a) the light enters from the port 1 and it is splitted by the birefringence block, BB: the vertical and horizzontal state of polarization, or the ordinary and extrordinary ray, are divided. Netx, the Faraday rotator and a $\lambda/2$ plate give rise to a rotation of the polarization plane by 90°. Basically they exchange their polarization, so at the second BB the rays are reunited into a single beam that exits the port 2.

b) the light enters from the port 2. The effect of the second BB is the same as in the previous case, but the effect of the FR rotator and the plate is opposite! Therefore, when the rays reach the first BB, they are separated but they present the original polarization: so the BB separates them further. Now, the extraordinary beam is reflected by a mirror M and it is combined with the ordinary one by a polarizing beam splitter toward the port 3.

5.3.1 Application of the circulators

They are widely oused to realize devices used in the WDM systems, to make measuremants of reflectometry type, in bidiectional optical systems and in some optical amplifiers. They are tyipically used:

1. to recover a refelcted signal,
2. to achieve add-drop and demux functions,
3. to compensate the dispersion caused by the optical fiber.

6 Electro-optics

6.1 Hints on anisotropic materials

Eletric Field:	$\mathbf{E}$
Dielectric susceptibility:	χ
Polarization vector:	$\mathbf{P} = \epsilon_0 \chi \mathbf{E} = \epsilon_0(\epsilon_r - 1)\mathbf{E}$
Dielectric Displacement:	$\mathbf{D}$
Dielectric Material:	$\mathbf{D}(\mathbf{E}) = \epsilon_0 \mathbf{E} + \mathbf{P}(\mathbf{E}) = \epsilon_0 \epsilon_r \mathbf{E}$

In general ϵ_r is a scalar for isotropic materials, otherwise it is a tensor, because it changes with respect to the direction that is observed. For the sake of completeness:

$$\begin{aligned} D_x &= \epsilon_0[\epsilon_{xx}E_x + \epsilon_{xy}E_y + \epsilon_{xz}E_z] \\ D_y &= \epsilon_0[\epsilon_{yx}E_x + \epsilon_{yy}E_y + \epsilon_{yz}E_z] \\ D_z &= \epsilon_0[\epsilon_{zx}E_x + \epsilon_{zy}E_y + \epsilon_{zz}E_z] \end{aligned}$$

However, we assume to have used a system of reference that let us to simplify the matrix: the element outside diagonal are zeroed, the remaining ones are called principal dielectric constants:

$$\epsilon_r = \begin{bmatrix} \epsilon_x & 0 & 0 \\ 0 & \epsilon_y & 0 \\ 0 & 0 & \epsilon_z \end{bmatrix}$$

Now, if $\epsilon_x = \epsilon_y = \epsilon_z$ the material is isotropic, while is anisotropic if they are not equal. In particular, we make a difference between *uniaxal* anisotropic ($\epsilon_x = \epsilon_y \neq \epsilon_z$, like the $LiNbO_3$) and *biaxial* anisotropic ($\epsilon_x \neq \epsilon_y \neq \epsilon_z$). At optical frequencies, the index of refranction along the generic direction is given by $n_i = \sqrt{\epsilon_i}$. For uniaxial material we have two indexes to compute:

$n_0 = \sqrt{\epsilon_x} = \sqrt{\epsilon_y}$	ordinary index
$n_e = \sqrt{\epsilon_z}$	extraordinary index

If $n_e > n_0$ we call the material positive uniaxial, negative if $n_e < n_0$.

6.1.1 The ellipsoid of indeces

It is derived by the expression of the electrical energy density in the material, doing a simple normalization:

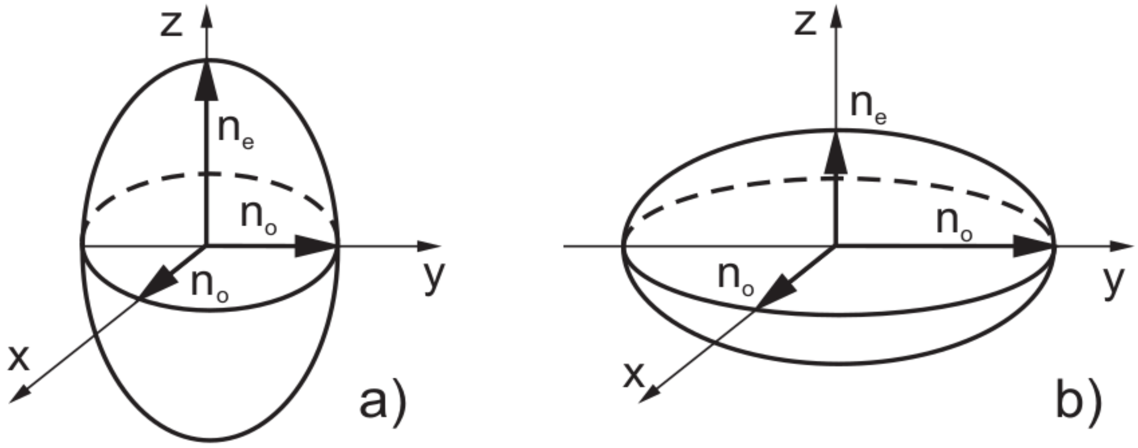
1. Energy density:

$$W_E = \frac{1}{2}\mathbf{E} \cdot \mathbf{D} = \frac{1}{2\epsilon_0}\left(\frac{D_x^2}{\epsilon_x} + \frac{D_y^2}{\epsilon_y} + \frac{D_z^2}{\epsilon_z}\right)$$

2. Normalization replacing $\mathbf{D}/\sqrt{2\epsilon_0 W_E}$ with the direction vector $\vec{r}$:

$$\frac{x^2}{n_x^2} + \frac{y^2}{n_y^2} + \frac{z^2}{n_z^2} = 1$$

We obtain:



The case *a* represents a positive uniaxial crystal, while *b* instead represents a negative one.
Once we have decided the direction of propagation of the wave, we consider the plane perpendicular that passes through the center of the ellipsoid. This intersection gives as result an ellipse whose major and minor semi-axes are the eigenstates that give the directions of the two othogonal polarizations that propagate undisturbed in the medium. Note that the *optical axis* does not necessarily coincide with the direction of propagation. The optical axis represents the axis of rotational symmetry for the ellipsoid, along which we have a constant refractive index.

6.2 The elctro-optic effect

The electro-optic effect is defined as the dependance of the dielectric properties of one material on the electric field E induced in its interior.

- The response of the material is described with the dielectric susceptibility, developed in series as:

$$\chi(E) = \chi(0) + \left.\frac{\partial\chi(E)}{\partial E}\right|_{E=0} E + \frac{1}{2}\left.\frac{\partial^2\chi(E)}{\partial E^2}\right|_{E=0} E^2 + \cdots = \chi^{(1)} + \chi^{(2)}E + \chi^{(3)}E^2 + \cdots$$

Now:

$\chi^{(1)}$	Linear susceptibility (or first order). If the field E is much less than the Coulomb attraction field of the material, the medium responds linearly
$\chi^{(2)}$	Linear electro-optic effect, also known as Pockels effect. It causes the dependancy of the refractive index on the field.
$\chi^{(3)}$	Quadratic electro-optic effect, or Kerr effect. It causes the dependendancy on the intensity of the field.

The critical step is to remember that the susceptibility is in fact a tensor of range (n+1) having 3^{n+1} components $\chi^n_{ij\ldots}$.

To better understand this latter property, note that the linear susceptibility has three components per direction. Besides, it is not multiplied by the field.
Otherwise, when we consider the second order term we have the product with the field E_k, where $k = x, y, z$. Indeed, there is a biunivocal relation between χ and E_k, so we add the letter k to relate the

coefficient to the field component. The number of components for the Pockels coefficient increases: 9 (3 per direction) times 3 (number of components of the field): $\chi^{(2)}_{ijk}$.

Now in general the scusceptibility is a complex quantity that depends on the frequency: we have a real part and an imaginary one, linked by the Kramers-Kronig relations. Thus, refractive index variations are correlated to the variation of the material absorption index.
This coexstince could be a problem, but in many electro-optic materials one of them prevails. Indeed the material that we will consider is the $LiNbO_3$, in which the type of electro-optic effect is the electro-refraction: the change in absorption can be ignored.

- The application of the field E causes a re-distribution of the electric charges and a deformation of the crystal structure, resulting in a change in the size and orientation of the ellipsoid of the indeces and a variation in the dielectric tensor.
The generic expression of the ellipsoid:
$$\sum_{m,j=1}^{3} \frac{x_m \cdot x_j}{n^2_{mj}(E)} = 1$$
We are interested into the variation of n_{mj}, but it is more convenient to determine the variation $\Delta(1/n^2)_{mj}$.
In this way, we define the electro-optic coefficients as follows:
$$\Delta\left(\frac{1}{n^2}\right)_{mj} = r_{mjk}E_k + s_{mjkl}E_kE_l$$
Ther terms of higher order are neglected.
Next we exploit the properties of symmetry of the dielectric tensor and we introduce a compact matrix notation. For istance: $mj = 11$, that is the first element of the matrix and the component xx, becomes the element $i = 1$.
In the hypothesis that the applied field is much less than the one in the atoms, the Pockels effect prevails. This linear approximation leads to the following expression:
$$\Delta\left(\frac{1}{n^2}\right)_i = \sum_{k=1}^{3} r_{ik}E_k \qquad i = 1\cdots 6$$
This tensor gauges how much the elctric field modify the refractive indeces and the orientation of the principal axis of the ellipsoid.
However, with a little approximation:
$$\Delta\left(\frac{1}{n^2}_i\right) = \frac{1}{[n_i + \Delta n_i]^2} - \frac{1}{n_i^2} \cong -\frac{2\Delta n_i}{n_i^2[n_i 2\Delta n_i]} \cong -\frac{2\Delta n_i}{n_i^3}$$
And finally:
$$\Delta n_i = -\frac{n_i^3}{2}\sum_{k=1}^{3} r_{ik}E_k, \qquad i = 1\cdots 6$$

6.3 $LiNbO_3$

The Lithium Niobate is a uniaxial negative crystal, birifrangent and non-centrosymmetric. Low index. Assumed z as optical axis (or c), the refractive index along x and y is the ordinary one (n_0), while the one along z is the extraordinary index (n_e). For this particular material, the matrix of the six coefficients obtained in the previous chapter can be represented as follows:

$$\Delta n_k = -\frac{n^3}{2}\begin{bmatrix} -r_{22}E_y + r_{13}E_z & -r_{22}E_x & r_{51}E_x \\ -r_{22}E_x & r_{22}E_y + r_{13}E_z & r_{51}E_y \\ r_{51}E_x & r_{51}E_y & r_{33}E_z \end{bmatrix}$$

It's important to remember that $r_{33} = 30.8 \cdot 10^{-12} m/V$, that is the highest coefficient, and n is n_e or n_o depending on which direction we're observing.

Now, it's important to understand that the variation Δn_k is given by the linear combination of the corresponding row (or column, the matrix is symmetric).
For istance, if the field applied is directed as x and z, the component y is zero and the expressions of the variations are:

$$\Delta n_x = -\frac{n_0^3}{2} r_{13} E_z + \frac{n_0^3}{2} r_{22} E_x - \frac{n_0^3}{2} r_{51} E_x$$

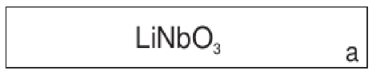
$$\Delta n_y = \frac{n_0^3}{2} r_{22} E_x - \frac{n_0^3}{2} r_{13} E_z$$

$$\Delta n_z = -\frac{n_e^3}{2} r_{51} E_x - \frac{n_e^3}{2} r_{33} E_z$$

First of all, because of E_x there is a rotation of the axis of the ellipsoid.
Secondly, since the indeces change and now $n_x \neq n_y$ the crystal became biaxial.
In general, once that we decided the optical axis we are interested into the transverse directions. So, it's useful to choose x (or y) as optical axis to exploit r_{33} in the z direction, it allows larger change of index.

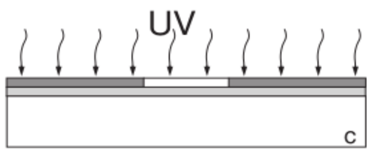
The are different processes to realize waveguides with $LiNbO_3$, but the technology that we study is the titanium diffusion.
There are seven main steps:

1. Preparation and cleaning of the substrate:

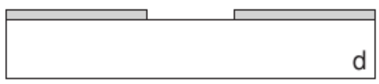


2. Photoresist deposition:

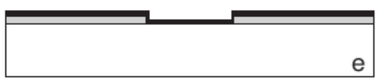
3. Masking and exposure:



4. Removing of the exposed photoresist:

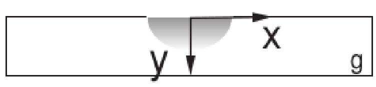


5. Titanium deposition:



6. Removing of the remaining photoresist:

7. Titanium diffusion:



6.4 Electrodes

Note that we can induce a refractive index variation Δn that exploit the high value of r_{33} if the optical axis is aligned with the modulating field and the state of polarization. To avoid rotations we have to direct the substrate of $LiNbO_3$ such that the optical field enters in parallel with the axis of the crystal. In this way we would end up obtaining a pure phase modulation.
For istance, if the electric field applied, i.e. the *modulating* field, has only one component, z, then

$$\Delta n = -\left(n_e^3/2\right) r_{33} \cdot E_m$$

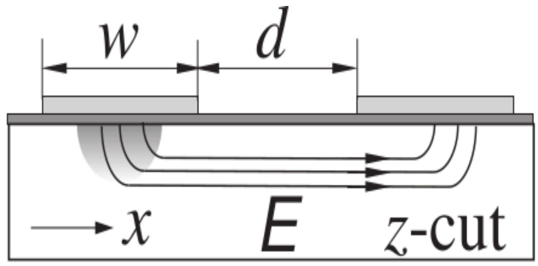
Since the field can be approximated as V/d, in order to work with small voltages (at high frequencies) we need a small distance, d. This is the reason why the electrodes are placed on the same surface: to reduce the distance, the thickness of the substrate is usually around 500, 800 nm.
At this point, the corresponding phase variation is:

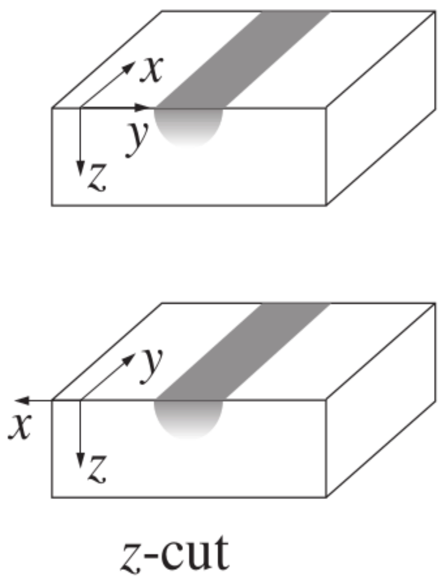
$$\Delta\varphi = \frac{2\pi}{\lambda}\Delta n \cdot L = -\frac{2\pi}{\lambda} \cdot \frac{1}{2} n_e^3 r_{33} \cdot E_m \cdot L$$

where L is the length.

Actually, the crtical step is choosing their position in the design.

Suppose to have a wafer of $LiNbO_3$, cut along the x-y plane (z-cut), with a TE mode. The usual configurations are shown the following:





The closer the electrodes, the higher the intensity of the field. However they must be placed in order to have electrical lines of force directed as z in the region of the guide where the titanium diffused.
More precisely, the indeces modulation depends on the distribution of the field E and of the optical field ψ, so it's actually convenient defining a superposition integral:

$$\Gamma = \frac{d}{V}\int E \cdot |\psi|^2 d\sigma$$

This parameter represents the tradeoff and the problem consists of optimize Γ in according with the application and the electronic constraints.
Now we can re-write the phase variation definition:

$$\Delta\varphi = -\frac{\pi}{\lambda} n_e^3 r_{33} \Gamma \cdot \frac{V}{d} \cdot L$$

How the electrodes are realized?

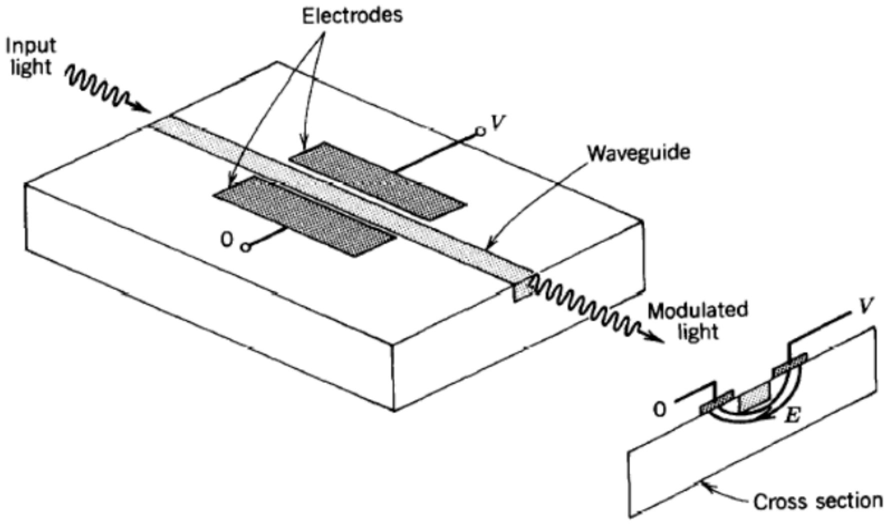
First of all, in order to isolate them a thin strate of SiO_2 is deposited on the guide through *sputtering*. With a process of masking, exposure and *liftoff*, the useless SiO_2 is removed. Above this insulator it is placed a thin strate of cromium to improve the adhesion to the glass or the crystal. The electrodes are finally made of gold, deposited by sputtering as well, and they are made thicker through galvanic grows.

6.4.1 Phase Modulation

A phase modulator is obtained by means of a guide and properly placed electrodes. For future pourpose it's important to know that, in order to have a phase shift of π, we have to isolate V from the last relation found in the previous chapter. It is commonly called half-wave voltage or V_π:

$$V_\pi = \frac{\lambda \cdot d}{n_e^3 r_{33} \Gamma L}$$

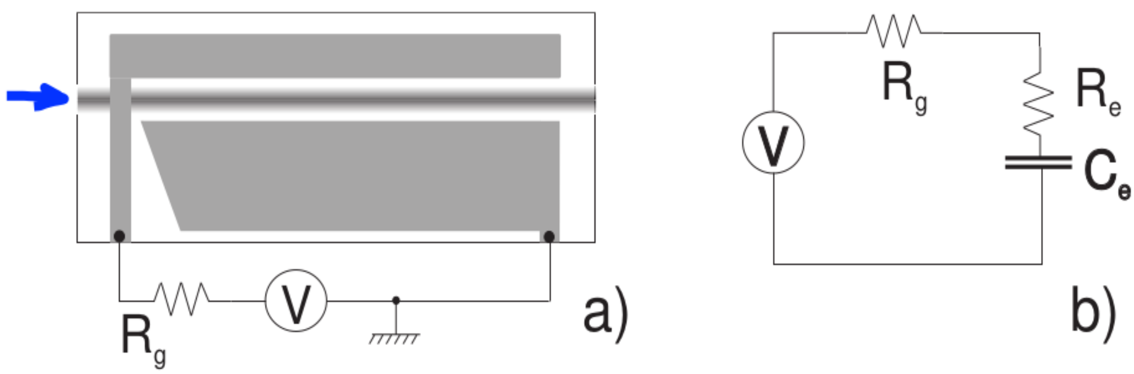
The generic 3D desgin of an integrated phase modultor, with a x-cut wafer of $LiNbO_3$, is the following:



6.4.2 Type of electrodes

The shape of electrodes can affect the electro-optic efficiency, the response of the modulator in frequency and the quality of the modulated signal. Therefore, the bandwidth is limited either by electrodes and driving circuits; since the response of the electro-optic effect is in the order of the picosecond.
There are two types of electrodes: lumped or distributed (labeled also as *travelling wave*).
In both cases the effective relative dielectric constant seen by the modulating field is about the algebric average between that of air and that of the substrate ($LiNbO_3$: $\epsilon_{r,eff} = (\epsilon_{rs} + 1/2) \simeq 18$).

1- Lumped Electrodes: electrodes are capacitances!



The figure a represents the actual design of the phase modulator with its driving circuit. The figure b instead represents the equivalent circuit.
The electrodes are placed to maximize the electro-optic effect, trying to contain the total capacity and hence the time constant of the driver.
The lumped electrodes assumption makes sense whenever it's verified the following condition:

$$L \ll \frac{c}{2 f_m \sqrt{\epsilon_{r,eff}}}$$

where f_m is the maximum modulating frequency and the factor 2 relates L to the half-wavelength.
Just for curiosity: for the $LiNbO_3$, it is $f_m L \ll 3.3 \;\; GHz \cdot cm$.
The most strictly constraints are due to the driving circuit. Actually, the maximum modulating frequency coincides with the cut-off frequency of the circuit, and it defines also the bandwidth:

$$f_m = \frac{1}{2\pi (R_g + R_e) C_e}$$

R_e is the resistance of the electrodes, computable as follows:

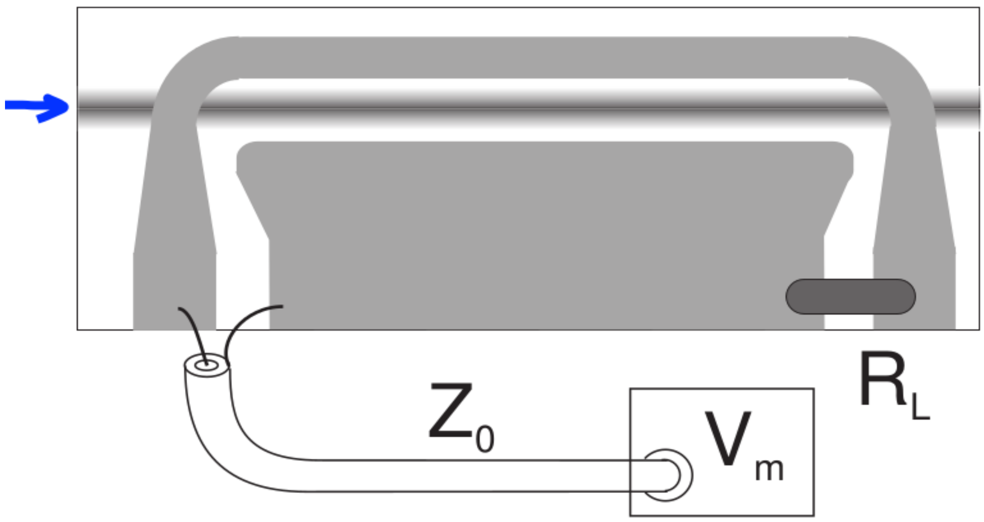
$$R_e = \frac{\rho_{Au} L}{A_e}$$

It is much less than the internal resistance of the generator, so it can be neglected.
Suppose that:

$$C_e = \epsilon_{r,eff} \frac{L \cdot w}{d}$$

If d is too short, part of the electric field rises directly between the electrodes and the lines of force are too curvy to satisfy the alignment requirements. On the other hand, decreasing d makes V_π smaller and C_e higher. I have smaller voltages but narrower bandwidth.
In general, lumped electrodes are used up to few hundred Mega Hertz. The opposite reasoning can be done observing what happens if L changes.

2- Travelling wave electrodes: electrodes are transmission lines!
If L is not way smaller than the wavelength we modelize the electodes as a transmission line. They are characterized by an impedance Z_0, effective dielectric constant $\epsilon_{r,eff}$ and losses α.
In order to avoid retroflections and stationary waves, the electrodes must be matched either to the generator resistances and to the load resistance.



$$Z_0 = \sqrt{\frac{\mathcal{R} + \jmath\omega\mathcal{L}}{\jmath\omega C_e}} \qquad v_m = \frac{c}{\sqrt{\epsilon_{r,eff}}} = \frac{1}{\sqrt{\mathcal{L}\mathcal{C}}} \qquad \beta = \frac{2\pi}{\lambda}\sqrt{\epsilon_{r,eff}}$$

{The characteristic impedance, the velocity of the wave and the propagation constant}

These devices are usually made with coaxial cable and micro-striplines.

The dimensioning of the electrodes must respect optical aspects as musch as radiofrequency constraints. This process is quite delicate, because every geometrical or physiscal parameter affects Z_0.
There are two main configurations for the transmission line: coplanar and micro-stripline.
With equal ratio $^d/_w$, the coplanar structure has lower impedance. However with the micro-stripline the matching at $Z_0 = 50\Omega$ is actually achievable with a good electro-optical efficency.

With travelling wave electrodes the constraint is no more the time constant of the driving circuit. To begin with, one issue that arises concerns *the difference in velocity between the modulating wave and the optical signal.* Indeed, for modulators made with $LiNbO_3$, $v_m = c/\sqrt{\epsilon_{r,eff}} \approx c/4,4$, while $v_o = c/n_e \approx c/2,2$: the electrical signal propagates slower than the optical. Hence, the wave undergoes an electro-optical effect which depends on the position because it's not constant along the line. Consequently, the optical wave that is emitted is not a correct copy of the modulating signal, but it has a longer lifespan and it is distorted. Besides, the effect is lower because it's temporally distributed.
Let's assumed matched electrodes, lossless lines and a sinusoidal modualting electrical signal. The wafer is z-cut. In other words, the voltage along the elctrode is:

$$V_m(x,t) = V_0 \sin\left(\frac{2\pi f_m}{c}\sqrt{\epsilon_{r,eff}} \cdot x - 2\pi f_m \cdot t\right)$$

When the optic wavefront enters in the guide, at the instant t_0, it's modulated by this value:

$$V_m(x,t_0) = V_0 \sin\left(\frac{2\pi f_m}{c}(\sqrt{\epsilon_{r,eff}} - n_{eff}) \cdot x - 2\pi f_m \cdot t_0\right)$$

The phase shift due to the velocity mismatch is:

$$\Delta\phi(t) = \Delta\beta_0 L \frac{\sin(\pi f_m/f_0)}{\pi f_m/f_0} \sin\left(2\pi f_m t - \pi\frac{f_m}{f_0}\right)$$

where

$$\Delta\beta_0 L = \frac{2\pi}{\lambda}\frac{n^3 r_{33}}{2}\frac{V_0}{d}\Gamma L$$

and f_0 is the frequency at which the modualtion is zeroed:

$$f_0 = \frac{c}{L\left(\sqrt{\epsilon_{r,eff}} - n_{eff}\right)}$$

For the travelling wave the bandwidth is defined as the frequency at which the phase shift is reduced by a factor 2, $B = 2f_0/\pi$. Nowadays (2021-2022), on the market, there are modulators in $LiNbO_3$ that work up to 40 GHz of bandwidth.

Now, let's consider *how losses affect the electro-optical effect.*
The finite resistivity of the metals causes the *skin effect*: the electric field does not reach the interior part and the current remains concentrated on the surface of the electrodes. The depth of penetration (or skin thickness) at which the field is reducted by a factor e is labeled δ:

$$\delta = \sqrt{\frac{\rho}{\pi\mu_0 f}}$$

This peculiar behaviour makes Z_0 dependent to the frequency.
For the quasi-static approximation we have that:

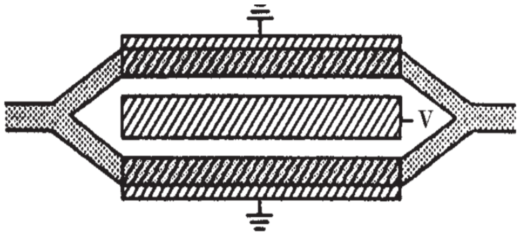
$$\alpha = \frac{\mathcal{R}}{2Z_0} = \cdots = \alpha_0\sqrt{f}$$

The losses can limit the modulator bandwidth because the modulating signal is attenuated along the line. The phase shift is:

$$\Delta\phi = \int_0^L \Delta\beta(z)dz = \Delta\beta_0 L\left[\frac{1-\exp(-\alpha L)}{\alpha L}\right]$$

6.5 Devices with $LiNbO_3$

6.5.1 Mach-Zhender interferometric modulator



This device is realized with two -3 dB couplers and a pair of uncoupled monomodal lines. However, in the picture is shown a simplified version of this design: we have two ports instead of fours, thank to the use of Y bifurcations. It resembles a balanced Mach-Zhender. Obviously, in both cases the obtained transfer function depends on the phase shift between the optical fields induced by the electro-optical effect.
Also for the electrodes there are two options:

1. the driving voltage is supplied by two electrodes placed on the sides of a guide only,
2. three electrodes are used (coplanar design). The central electrode is fed with the radfiofrequency signal while the other two are grounded (the roles can be inverted). In this way the voltages are halved, but we need to point out that with this design the matching is not excellent: tapers are required. This solution is called push-pull. Indeed, observing the electrodes we can easily understand that the induced phase variations are opposite in the guides (the verse of the voltage drops are opposite). Hence, at the output, the total phase shift is $2\Delta\phi$.

The total transfer function for a configuration with two electrodes is calculated as the product of the matrices:

$$I_0 = \cos^2\left(\frac{\Delta\beta L}{2}\right)$$

where

$$\Delta\beta = \frac{2\pi}{\lambda}\Delta n = \frac{2\pi}{\lambda}\left(\frac{1}{2}n^3 r \frac{V}{d}\Gamma\right)$$

In the absence of an electric field the input wave is divided into two equal components which propagate each in one of the guides. The components would eventually recombine in phase at the output to form the same signal which enters in the first port. Unfortunately, it's impossible to technically realize two guides perfectly identical, so the material must be biased to work without errors.
Observing the transfer function the the status off of the modulator can be also obtained for voltages applied to the electrodes that cause a phase shift of $\Delta\beta L = \pi$. For a push-pull design, this voltage is $V_\pi/2$.
Moreover, since the driving voltages are dissipitaed by the loads, if the voltages are halved the power are reduced by a factor 4. The push-pull design is much less expensive in terms of driving power.

In general, there is a residual phase variation if we don't use a push-pull configuration. Resorting on the equations describing the behaviour of a Mach-Zhender interferometer, if we calculate the field at the ouput we get:

$$\frac{1 + e^{-\jmath\Delta\phi}}{2} = e^{-\jmath\frac{\Delta\phi}{2}}\frac{e^{+\jmath\frac{\Delta\phi}{2}} + e^{-\jmath\frac{\Delta\phi}{2}}}{2} = e^{-\jmath\frac{\Delta\phi}{2}} \cdot \cos\left(\frac{\Delta\phi(t)}{2}\right)$$

As we can easily understand the modulator change the signal both in intensity and in phase. This residual phase modualtion is named *chirp*. Generally the chirp is counter-balanced setting properly electrodes and voltages, or usign a push-pull configuration. Moreover, to obtain an even more efficient system we could work with four electrodes in order to have two indipendet modulators. This obviously means that we need two driving circuits which work indipendently, that is quite expensive but it allows a more strict control of the component.

7 Fibers Gratings

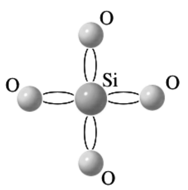
The gratings are a class of devices that enable the power transfer between two modes of the same guide. Indeed, the working principle relies on the fact that modes of the unperturbed guide loss their orthogonality. To make it happen, in the guide there is a periodic lengthwise variation of the effective index, induced by the photorefractive effect or alternatively with geometrical corrugations. The coupled modes can be wheather copropagants or counter-propagating, even of different orders. Since the disruption is usually small, we can apply the coupled mode theory already seen for the Directional Couplers.
The grating that cause the coupling between counter-propagating modes are called Bragg gratings, or reflection gratings, and permitt to realize selective refectors in frequency.

7.1 The photorefractive effect

It's possible to change the physical properties of materials by exposing them to light radiation. More specifically, if the optical fiber is expososed to sufficiently intense ultraviolet radiation, its refractive index could be permanently altered. The first time this phenomenon has ben observed the device was a fiber made with germanium-doped silica.
In this section the two main models will be briefly treated in order to understand the relation between the variations of refractive index and that of chemical-physical properties.

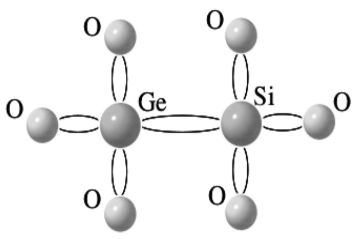
7.1.1 Color center model

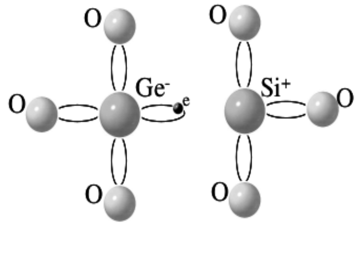
The exposure to ultraviolet radiation can actually change the absorption spectrum of the material, which is linked to the variation of the refractive index. This is justified by the Kramers-Kronig relations.
The atomic configuration in a germanium-doped silica is the following:



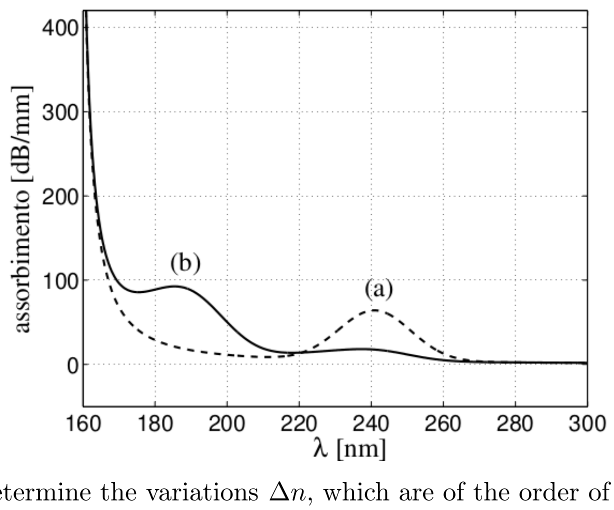
Since the Silicon and Germanium are elements of the same group in the periodic table, they both create boundings with four atoms of oxygen. After the doping, some atoms of Si are replaced with some atoms of Ge. However, the defects can be of two types: GeO and Gee$^-$.

In the first case, the germanium (or silicon) forms three boundigns with the Oxygen and the last one with the Silicon. In the second case, the germanium steals an electron from a nearby atom of silicon.





Because of these defects, the absorption specturm shows a peak (a *band*) around 240 nm. If we expose the fiber to ultraviolet radiation, the boundings of defects GeO can be broken and defects Gee$^-$ are generated instead. As result, the peak changes position and progressively shifts from 240 nm to 190 nm. Besides, more peaks appear at longer wavelengths, but we ignore them.



From the Kramers-Kronig relations we can determine the variations Δn, which are of the order of 10^{-4} at the

wavelengths of interest for otpical communications.
However, higher variations have been measured experimentally (10^{-3}) in the core, proving that this model is not sufficient to understand (conceptually) the photorefractive effect.

7.1.2 Compaction model

The matrix of SiO_2 undergoes a process of densification after being exposed to ultraviolet radiation. This process leads to periodic corrugations in the core of the preform optical fibers. Moreover this process is not irreversible because the struct can return to its original form through *annealing*, that consist of heating up the fiber up to 1000°C.
The germanium facilitates this mechanism givng to the structure less rigidness and making more likely phenomena of microstructural reorganization.
Just for the sake of completeness, the mathematical relation modelling the variation in density and polarizzability is named Lorentz-Lorentz relation.

7.2 Photosensitivity techniques

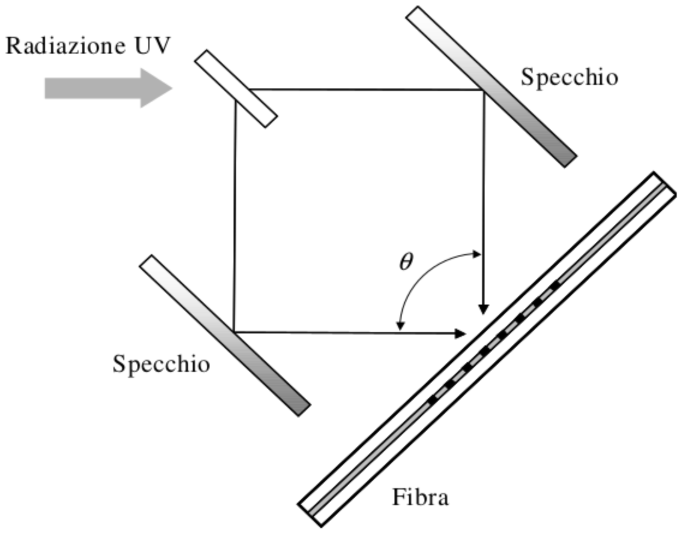
Since the presence of dopings increases the propagation losses, but the optical fibers are optimized to have losses as low as possible, the photosensitivity is very small. It's possible to increase it with an higher Ge doping, with hydrogen diffusion in the fiber or with boron/phosphrus doping.

7.3 Writing techniques of the fiber gratings

We have two types of technique: holographic (interferometric) and non-interferometric writing. The first category is preferred when the lattice step is small. The exposure pattern must be stable and the source of radiation must be both spatially and temporally coherent. On the other hand, the techniques of the second category are choosen when the lattice step is longer than $1\mu m$. In this case the fiber is directly exposed to the UV radiation properly modulated in amplitude in the direction of the core.

7.3.1 Interferometric methods

- **Volumetric interferometer**



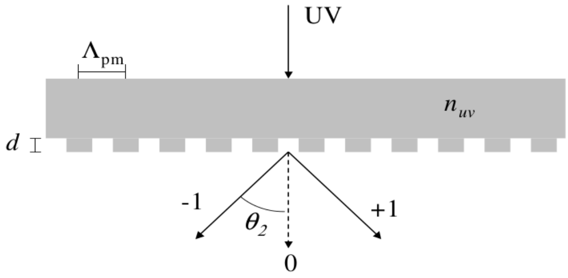
If the source is far enough, the wavefront is assumed to be plane. Under this condition the lattice step is:

$$\Lambda = \frac{\lambda_{uv}}{2n_{uv}\sin(\theta/2)}$$

where λ_{uv} is the wavelength of the source, n_{uv} is the corrsponding refractive index for the SiO2, and θ si the angle between the rays.

First of all, the lattice step can be changed varying the angle θ, that is a simple and quick solution. However, the mechanical vibrations of the mirrors can affect the quality of the gratings if the time under exposure is not short.
Concerning the source, the ultraviolet lasers are not sufficiently coherent, hence are usually replaced with argon ion laser duplicated in frequency.

- **Phase mask interferometer**



A phase mask is a quartz plate that is transparent to the ultraviolet radiation and that has periodic monodimensional corrugations obtained through photo-litographic processes. A light beam (λ_{uv}) incides on the plate with angle θ_1 and it is diffracted at the angle θ_2, given by:

$$n_{uv} \sin\theta_1 = \sin\theta_2 + m \cdot \frac{\lambda_{uv}}{\Lambda_{pm}}$$

where m is an integer that determines the order of diffraction, and Λ_{pm} is the period of the corrugation.

Usually the angle of incidence is 0°. In this way the most of the diffracted wave is contained between the orders ±1. To cancel out the order 0, the following condition must be respected:

$$d(n_{uv} - 1) = \frac{\lambda_{uv}}{2}$$

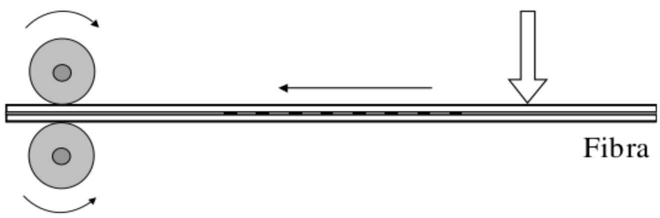
The interferance between diffracted rays gives rise to the periodic pattern of the grating:

$$\Lambda = \frac{\lambda_{uv}}{2\sin(\theta/2)} = \frac{\Lambda_{pm}}{2}$$

Differently fromt the volumetric holography it's much less flexible because the period is fixed by the mask. However, it's a technique that allow a serial fabrication of gratings with a elevate reproducibility of the parameters.

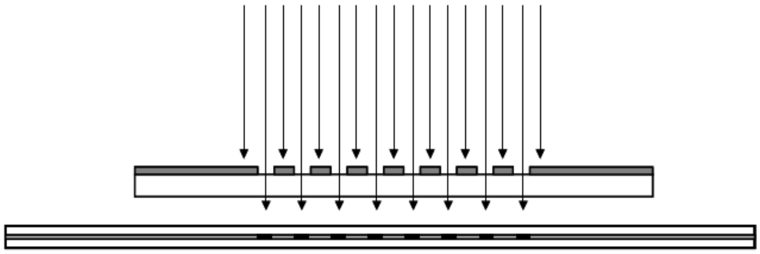
7.3.2 Non-interferometric methods

- **Point-point writing**



The modulation of the refractive index is obtained by moving rigidly the fiber with respect to the transverse UV radiation. The main advantge is the extreme flexibility, inasmuch we can obtain also very complex structure varying dinamically either the mechanical or optical parameters of the writing support.
We can realize gratings with a variable index modulazion profile (apodized gratings) or with variable period.
It is generally use to realize long period grating: $\Lambda > 10\mu m$.

- **Amplitude mask**



A phase mask is a silica plate equipped with a periodic structure on one of its surface. It's usually made of cromium or nickel because they are opacque to the Uv radiation. If the period of the mask is way larger than the wavelength of the source, and if the distance between mask and fiber is sufficiently small, the diffraction can be ignored: the variation of the refractive index follows the transparency profile of the mask.
This technique is less flexible than the point-to-point writing. It is actually more repeatable.

7.4 Coupled mode theory for gratings

First thing to recall: the generic distribution of a field inside a guide can be represent as a linear combination of a discrete number of guided modes and an eventual continuos spectrum of radiative modes.
Second thing to recall: if the guide remain unperturbed, all modes are orthogonal. Inhomogeneities causes power losses or reciprocal couplings between guided modes and thus the power is transfered from one mode to the other.
For simplicity do consider two possible guided modes of the unperturbed structure, ψ_1 and ψ_2. The total field $\Psi(x, y, z)$ is :

$$\Psi(x, y, z) \approx A(z)\psi_1(x, y)e^{-\jmath\beta_1 z} + B(z)\psi_2(x, y)e^{-\jmath\beta_2 z}$$

where $\beta_{1,2}$ are the propagation constants of the modes.
The amplitudes are functions of z, otherwise they would be solutions of a unperturbed guide.
In this case the index profile can be represented as follows:

$$n_p(x, y, z) = n(x, y) + dn(x, y, z)$$

where the first term is the index of the unperturbed guide, and the second take into account the perturbation.
In the guiding structure with the grating, the propagation is described by the wave equation:

$$\nabla_t^2 \mathbf{\Psi} + \frac{\partial^2 \mathbf{\Psi}}{\partial z^2} + k_0^2 n_p^2(x, y, z)\mathbf{\Psi} = 0$$

Differently from the case of the directional couplers, the index profile depends on z.
We assume that:

1. $\mathbf{\Psi}$ is a solution of the equation. Rememeber that modes $\psi_{1,2}$ are solution of the wave equation for the unperturbed guide;
2. the coupling is weak : $d^2A/dz^2 = 0$, $d^2B/dz^2 = 0$;

hence we find:

$$k_0^2 \Delta n^2 A\psi_1 + k_0^2 \Delta n^2 B\psi_2 e^{\jmath(\beta_1-\beta_2)z} - 2\jmath\beta_1 \frac{dA}{dz} - 2\jmath\beta_2\psi_2 \frac{dB}{dz} e^{\jmath(\beta_1-\beta_2)z} = 0$$

where $\Delta n^2 = n_p^2(x, y, z) - n^2(x, y)$, for small perturbations is aproximately equal to $\Delta n^2 \approx 2n(x, y)dn(x, y)$.

As done for directional coupler we multiply the last equation by ψ_1^* and integrating all over the transverse section we obtain that:

$$\begin{cases} \dfrac{dA}{dz} = -\jmath c_{11}A(z) - \jmath c_{12}B(z)e^{\jmath(\beta_1-\beta_2)z} \\ \\ \dfrac{dB}{dz} = -\jmath c_{21}B(z) - \jmath c_{22}A(z)e^{-\jmath(\beta_1-\beta_2)z} \end{cases}$$

We suppose that $dn(x,y,z) = \delta n(x,y)g(z)$. The coefficients are defined as follows:

$$c_{ij} = \frac{k_0^2}{\beta} g(z) \frac{\iint_{-\infty}^{\infty} \psi_j n(x,y)\delta n(x,y)\psi_i^* dxdy}{\iint_{-\infty}^{\infty} \psi_i \psi_i^* dxdy} = \tilde{\kappa}_{ij}(x,y) \cdot g(z)$$

In particular, if the refractive index of perturbation has a periodicity Λ, the function $g(z)$ can be expressed with a Fourier series in the form:

$$g(z) = \sum_m g_m e^{\jmath m \frac{2\pi}{\Lambda} z}$$

Putting these expressions all togheter, we obtain:

$$A(z) = A(0) - \jmath \tilde{\kappa}_{11} \sum_m g_m \int_0^z e^{\jmath m \frac{2\pi}{\Lambda} \nu} A(\nu) d\nu - \jmath \tilde{\kappa}_{12} \sum_m g_m \int_0^z e^{\jmath \left(m \frac{2\pi}{\Lambda} - (\beta_1 - \beta_2)\right)\nu} B(\nu) d\nu$$

Since the function A(z) and B(z) vary slowly with z, we can consider all the contributes nill, except for those whose complex exponentials have exponent equals to zero. Hence, in the first integral survives only the contribute for $m = 0$, that is related to the continous component of the idex perturbation dn. The second integral is not zero whether the following relation is satisfied:

$$\beta_1 - \beta_2 = m \frac{2\pi}{\Lambda}$$

It is commonly known as phase matching condition or synchronism condition.

7.5 Excitation and diagram

Assuming to operate in the single-mode, the only two modes that can be actually excited are the fundamental mode propagatin towards +z, and the counter propgating mode, directed as -z.
In this case, if the fundamental propagation constant is $\beta_1 = \beta$, the counter propagation constant is $\beta_2 = -\beta$. The phase matching condition becomes:

$$\beta_1 - \beta_2 = 2\beta = m \frac{2\pi}{\Lambda} \Longrightarrow \beta = m \frac{\pi}{\Lambda}$$

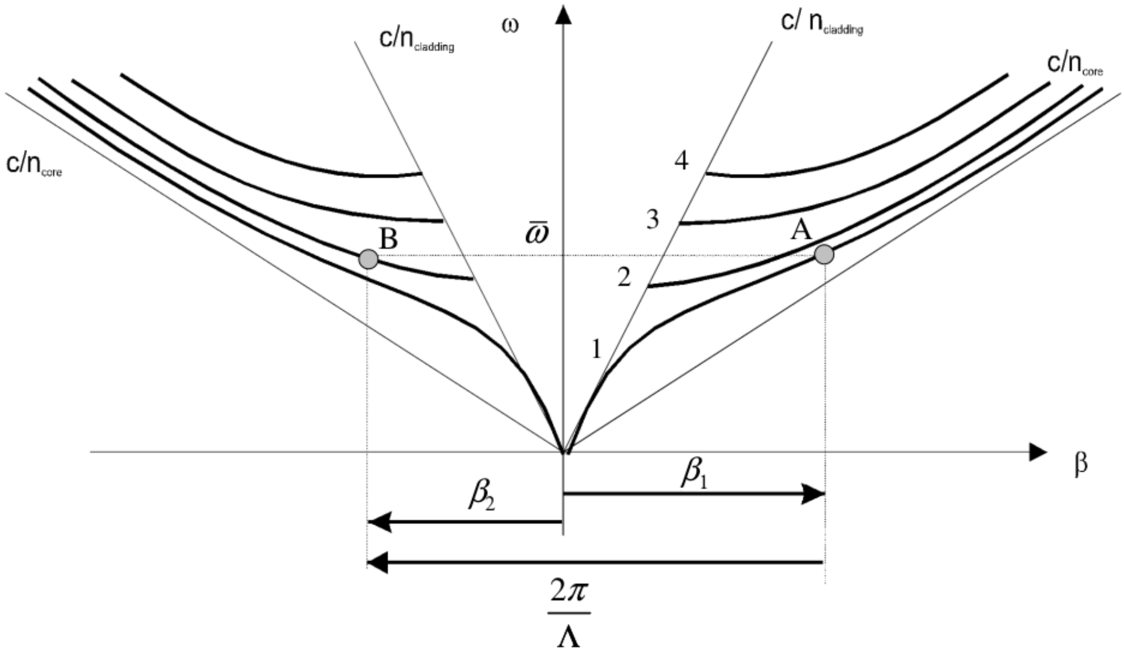
If Λ is given, the wavelength which satisfy the sychronism condition is the following:

$$\bar{\lambda} = \frac{2\pi}{\beta} n_{eff} = \frac{2 n_{eff} \Lambda}{m}$$

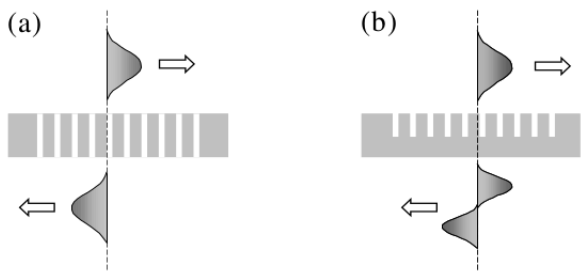
For multimode fibers the counter-propagatin coupling can occur also between two modes of different orders, provided that the generical phase matching condition is satifsied:

$$|\beta_1| + |\beta_2| = m \frac{2\pi}{\Lambda}$$

The prevailing contribution is usually related to the fundamental harmonic, defined $m = 1$.



- in this example, at the fixed value of omega ($\bar{\omega}$) the point A belongs to the fundamental propagating mode while the point B belongs to the counter-propagating mode of the first higher order. It is important to remember that this diagram refers to the modes that are solutions of the unperturbed guide.
- the couplings occurs if the lattice has a discontinuity. The exact couplings between the modes in the diagram occurs if:
  1. the disconuity is such that the modes are actually excited;
  2. the period Λ is the one related to their propagation constants;
  3. the perturbation $\delta n(x, y)$ is odd. Otherwise, since the counter-propagating mode has an odd field distribution on the transverse section, the counter porpagating mode wouldnt' be excited. Indeed, an even perturbation causes the coupling between modes with same symmetry (a), while an odd perturbation causes the coupling between modes that are anti-symmetric (b).



The gratings that allow the couplings with counter-propagating modes are called *Bragg gratings*.

7.6 Uniform Bragg Gratings

Let's assume that the perturbation along z is periodical. In this way the system of equations for the computation of the amplitudes (at page 47) can be solved in a closed form. We recall that:

$$dn(x, y, z) = \delta n(x, y) \cdot g(z) = \delta n(x, y) \cdot \sum_m g_m e^{\jmath m \frac{2\pi}{\Lambda} z}$$

Next, relying on the superposition theorem we can consider every harmonic separately and deduce the behaviour of the grating summing up all the contributions.

Formally:

$$dn(x,y,z) = \delta n(x,y)\left[g_0 + v\cos\left(\frac{2\pi}{\Lambda}\right)z\right]$$

We call theese gratings uniform because δn and Λ do not depend on z.

The system of equation becomes:

$$\begin{cases} \frac{dA}{dz} = -\jmath\kappa_{11}A - \jmath\kappa_{12}Be^{\jmath\Delta\beta z} \\ \frac{dB}{dz} = -\jmath\kappa_{21}B - \jmath\kappa_{22}Ae^{-\jmath\Delta\beta z} \end{cases}$$

where $\kappa_{11} = g_0 \cdot \tilde{\kappa_{11}}$ and $\kappa_{12} = v \cdot \tilde{\kappa_{12}}/2$.
Moreover $\Delta\beta = \beta_1 - \beta_2 - 2\pi/\Lambda$.

Let's take the case of a counter-propagating coupling between two modes of the same order. For simplicity, ψ_1 and ψ_2 have the same form and so $\beta_1 = -\beta_2 = \beta$. In this way, we obtain:

$$\kappa_{11} = -\kappa_{22} = \tfrac{2\pi}{\lambda}g_0\bar{\delta n}_{eff}$$

$$\kappa_{12} = -\kappa_{21} = \kappa = \tfrac{\pi}{\lambda}v\bar{\delta n}_{eff}$$

where we call $g_0\bar{\delta n}_{eff}$ the average perturbation of the refractive index of the guide, and $v\bar{\delta n}_{eff}$ the modulation depth.
Next, we do the following substitution:

$$a(z) = A(z)e^{-\jmath\left(\beta_1 - \frac{\pi}{\Lambda}\right)z}$$

$$a(z) = B(z)e^{-\jmath\left(\beta_2 + \frac{\pi}{\Lambda}\right)z}$$

So we find:

$$\begin{cases} \frac{da}{dz} = -\jmath\sigma a(z) - \jmath\kappa b(z) \\ \frac{db}{dz} = \jmath\sigma b(z) + \jmath\kappa a(z) \end{cases}$$

where we have that:

$$\sigma = \kappa_{11} + \beta - \frac{\pi}{\Lambda}$$

Inspi<red by the similarity with the directional coupler we want to determine the transfer function between complex amplitudes at input and the complex amplitudes at the ouput.
For a grating of length L, we have:

$$T_G = \begin{bmatrix} T_{G_{11}} & T_{G_{12}} \\ T_{G_{21}} & T_{G_{22}} \end{bmatrix} = \begin{bmatrix} \cosh\delta L - \jmath R\sinh\delta L & -\jmath S\sinh\delta L \\ \jmath S\sinh\delta L & \cosh\delta L + \jmath R\sinh\delta L \end{bmatrix}$$

With the following definitions:

$$R = \sigma/\delta \qquad S = \kappa/\delta \qquad \delta = \sqrt{\kappa^2 - \sigma^2}$$

Without losses $S^2 - R^2 = 1$ and $\det(T_G) = 1$. For simplicity for now on the losses will be neglected.

In the case of $I_1 = 1$ and $O_2 = 0$ in the grating there is a *direct* wave whose amplitude at the output $O_1 = 1/T_{G_{22}}$. It decreases with the increase of L. In the same time the amplitude at the input $I_2 = T_{G_{21}}/T_{G_{22}}$, that is related to the reflected wave, increase with L.
If the grating is suficiently long, it is possible to reach the condition for which the wave is totally reflected. The grating resembles a distributed mirror.
In general the power reflectivity of the grating is defined as the ration at the input port between the intensity of the reflected wave and that of the incident wave:

$$R = \left|\frac{I_2}{I_1}\right|^2 = \left|\frac{T_{G_{21}}}{T_{G_{11}}}\right|^2 = \frac{\sinh^2\delta L}{\cosh^2\delta L - \left(\frac{R}{S}\right)^2}$$

The reflectivity is maximum whenever $\sigma = 0$, that occurs when the synchronism condition is satisfied. In this way:

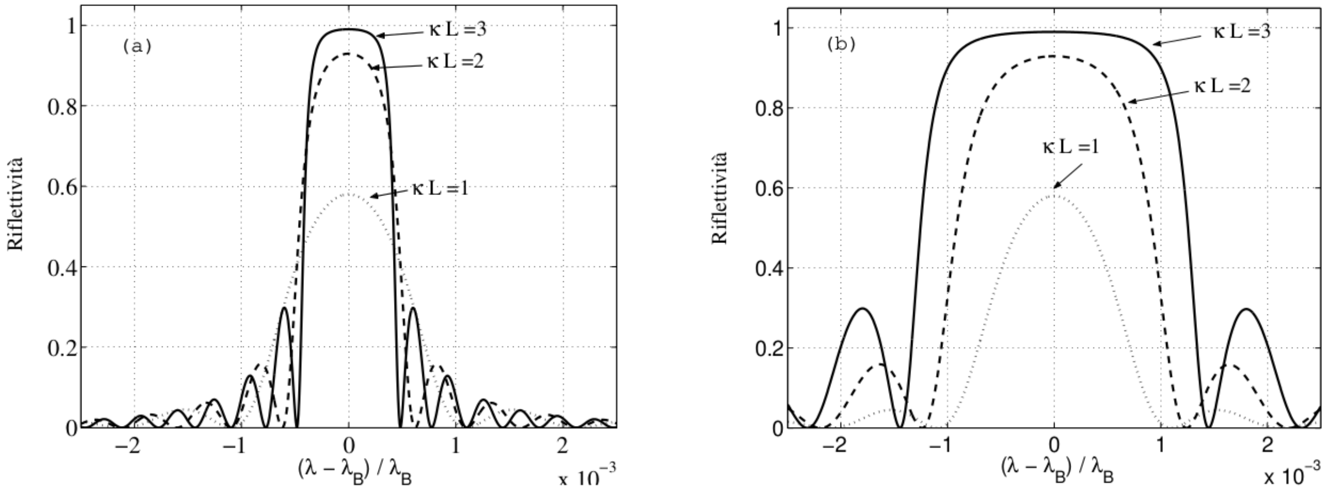
$$R_M = \tanh^2 \kappa L$$

The wavelength corresponding to R_M is named Bragg wavelength λ_B. Its value is derived subtituing in the expression of σ the expression of κ_{11} and β:

$$\lambda_B = 2\Lambda n_{eff} \left(1 + g_0 \frac{\delta n_{eff}}{n_{eff}}\right)$$

In the end, note that if $g_0 = 0$ the Bragg wavelength is reduced to $\lambda_B = 2\Lambda n_{eff}$, that is the same result deduced from the phase-matching condition.

7.6.1 Spectral characteristics



a) The maximum refelctivity R_M increases with the length of the grating and for values of $\kappa L > 3$, it is basically unitary. However, the reflectivity is high only around λ_B, that is, untill the parameter σ is small. Note that σ gauges the deviation from the synchronism condition.
For high values of σ ($\sigma \gg \kappa$):

$$T_G = \begin{bmatrix} e^{-\jmath\sigma L} & 0 \\ 0 & e^{\jmath\sigma L} \end{bmatrix}$$

That is the matrix for a simple line of transmission.

b) If we increase the modulation depth, with equal κL, the reflectivity has the same peak but the reflectivity bandwidth has increased!
A possible measurement of the bandwidth is the distance in wavelengths between the first zero of the reflectivity spectrum.

$$\frac{\Delta\Lambda}{\lambda_B} = \frac{v\bar{\delta n}_{eff}}{n_{eff}} \sqrt{1 + \left(\frac{\lambda_B}{v\bar{\delta n}_{eff} L}\right)^2}$$

Since the badwidth is proportional to the modulation depth, we need a high depth to have a large bandwidth.

Weak grating	$v\bar{\delta n}_{eff} L \ll \lambda_B$	$\Delta\lambda \approx \frac{\lambda_B^2}{n_{eff} L}$
Strong grating	$v\bar{\delta n}_{eff} L \gg \lambda_B$	$\Delta\lambda \approx \frac{v\bar{\delta n}_{eff}}{n_{eff}} \lambda_B$

Secondly, looking at the pictures we can observe secondary side lobes. These are due to the finite length of the grating, whose input and ouput interfaces are like small but sharp discontinuities behaving like mirrors in a Fabry-Perot cavity.
The wavelengths at which the grating reflectivity spectrum is nill are the resonant values of the structure. They are characterized by a unitary transmission, provided that losses are zero. The spacing between reflectivity zeros is $\Delta\lambda_z \simeq \lambda^2/(2nL)$, equals to the FSR of a Fabry-Perot cavity of length L.

7.6.2 Group delay

The grating can introduce a group delay with the reflection, computable as follows:

$$\tau_g = \frac{d\phi}{d\omega} = \frac{\lambda^2}{2\pi c}\frac{d\phi}{d\lambda}$$

where ϕ is the phase of the reflection coefficient.
For a uniform grating the group delay is symmetric, like the reflectivity, with respect to λ_B.
In the chirped grating is possible to obtain non-symmetric diagrams, required by some applications.
It is possible also to define a group length:

$$L_p = \frac{c\tau_g}{n_{eff}} = \frac{r_M \lambda_b}{2\pi v \delta n_{eff}}$$

This parameter is an index of the penetration of the field in the grating, and it defines the point in which the reflection is assumed to be concentrated.
Keep in mind that the group delay gauges the time inteval between the enter of the incident wave in the grating and the exit of the reflected wave. Hence, when the reflectivity is maximum, the reflected wave returns quickly. On the other hand, when reflectivity is zero, the transmission is unitary and the group delay measures how long the transmitted wave takes to reach the output.

7.7 Non uniform gratings

The unifrom gratings allow the analysis with simple equations in closed form. However, there are many applications that needs devices performing efficiently out of bandwitdh. In this case we exploit non-uniform gratings.

$$dn(x,y,z) = \delta n(x,y)\left[g_0(z) + v(z)\cos\left(\frac{2\pi}{\Lambda}z + \phi(z)\right)\right]$$

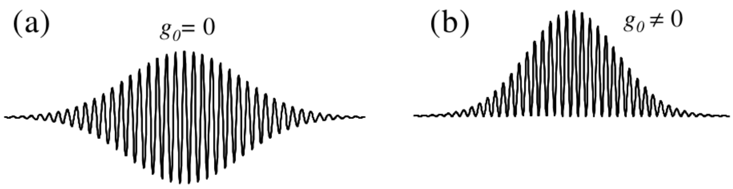
Differently from the unifrom case, the visibility v and the average value of the perturbation g_0 are depending on z. Besides, there is an additional phase term $\phi(z)$, which represents the modualtion along z of the grating period.

In general we have two cases:

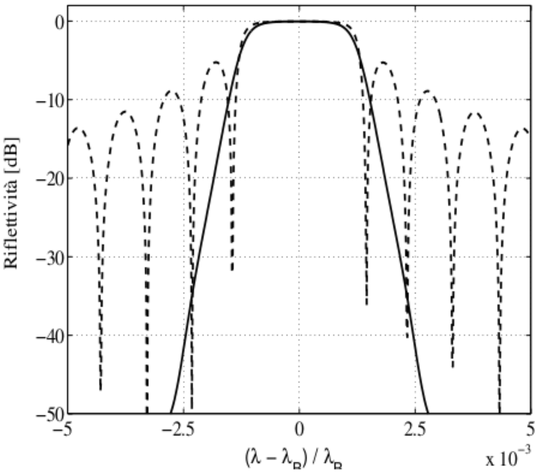
- the *apodized gratings* are used to minimized the secondary lobes of the reflectivity;
- the *chirped gratings* are used to compensate the chromatic dispersion.

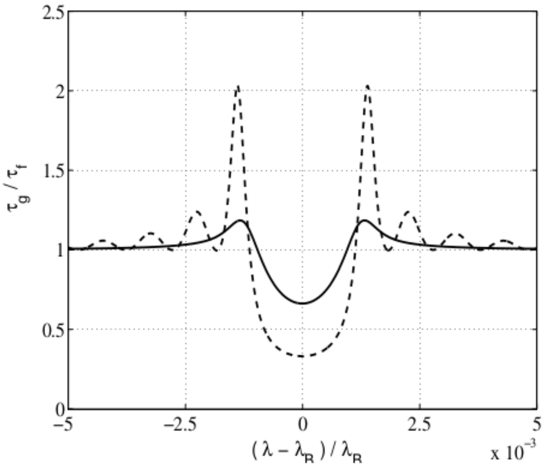
7.7.1 Apodized gratings

The impedance "jump" that leads to the secondary lobes it's related to the impedance matching between the grating and the input/output guides. For this reason, $v(z)$ is designed to decrease gradually from the center towards the exterior regions of the grating.
For istance, the profile can be gaussian or raised cosine.

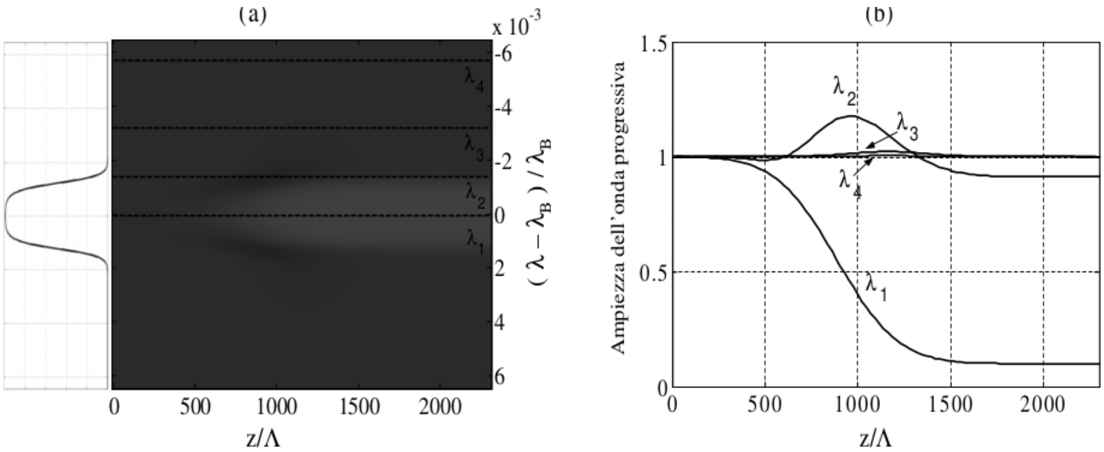


With a gaussian profile we obtain diagrams that are similar to the following ones:





The dotted lines represent the diagram for the uniform grating. As we can see, the secondary lobes are very small in decibel. Moreover, the group delay undergoes a reduction. In the uniform case the resonance is caused by the discontinuities at the ends, but in this case the matchinf is better so the peaks are lower. Viceversa, when the R is at its maximum, the delay is higher with respect to the uniform case. This happens because the the maximum of R depends on κ which depends on $v\overline{\delta n_{eff}}$, thus we have a weak reflective grating at the beginning of the fiber and the wave penetrates more.
Look at the following field diagrams to see what happens for different wavelengths:



7.7.2 Chirped gratings

This category of grating does not have a constant period Λ. The chirp function $\phi(z)$ represents a degree of freedom that can be optimized to satisfy specific applications.
We only consider gratings with a linear chirp function, usually called *Dispersion Compensating Grating*, or DCG. Note that these devices can only compensate the first order dispersion. To compensate higher order dispersion non-linear function are required.

If we define Λ_L and Λ_S the maximum and minimum period, the generic function of the period is:

$$\Lambda(z) = \Lambda(0) \pm C \cdot z$$

Where:

$$\Lambda_0 = \frac{\Lambda_L + \Lambda_S}{2} \qquad C = \frac{\Lambda_L - \Lambda_S}{L} \qquad z \in [-L/2; +L/2]$$

Since the period varies, different wavelengths satisfy the synchronism condition in differet points. Thus, the Bragg wavelength is a function of z. Assuming $g_0 = 0$: $\lambda_B(z) = 2n_{eff}\Lambda(z)$.
In first approximation, the bandwidth is the spectral spacing between the maximum and minimum wavelengths that satisfy the previous definition:

$$\Delta\lambda = 2n_{eff}(\Lambda_L - \Lambda_S) = 2n_{eff}CL$$

The bandwidth grows with C and L. However, once that δn is fixed the grating cannot be arbitrarily elongated. Indeed, to have the generic wavelength reflected it's necessary that the grating contains a sufficient number of periods that satisfy the synchronism condition at that wavelength. Hence $\kappa\Delta L$, where ΔL is a distance along which the period can be considered constant, must be large enough.

For a positively chirped grating, the group delay is:

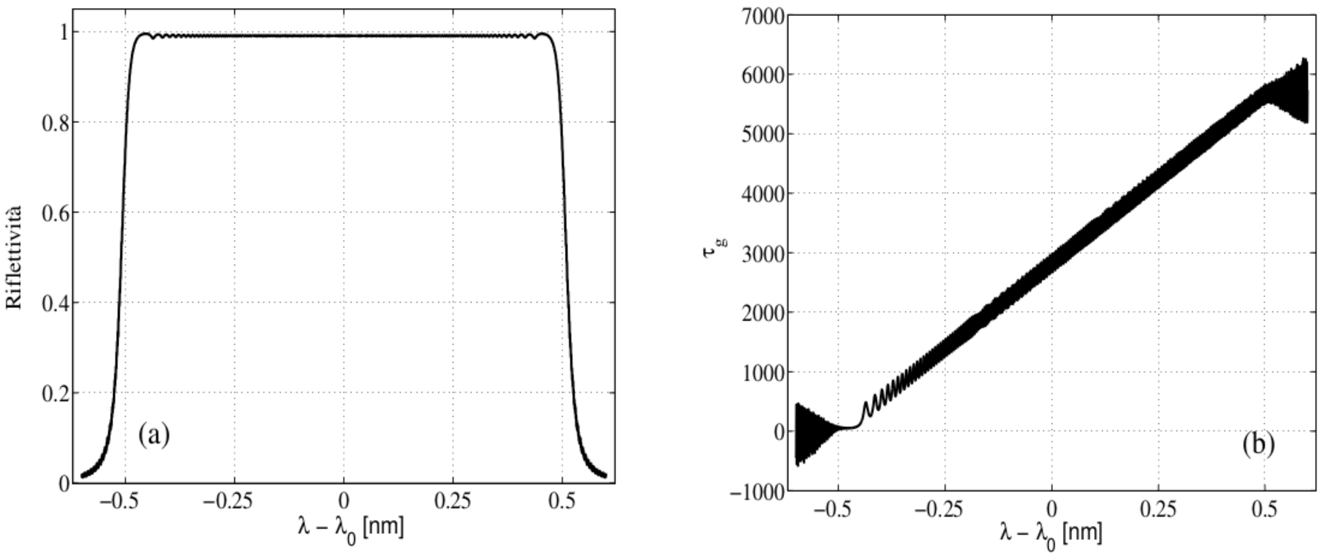
$$\tau_g(\lambda) = \frac{\lambda - \lambda_S}{c \cdot C}$$

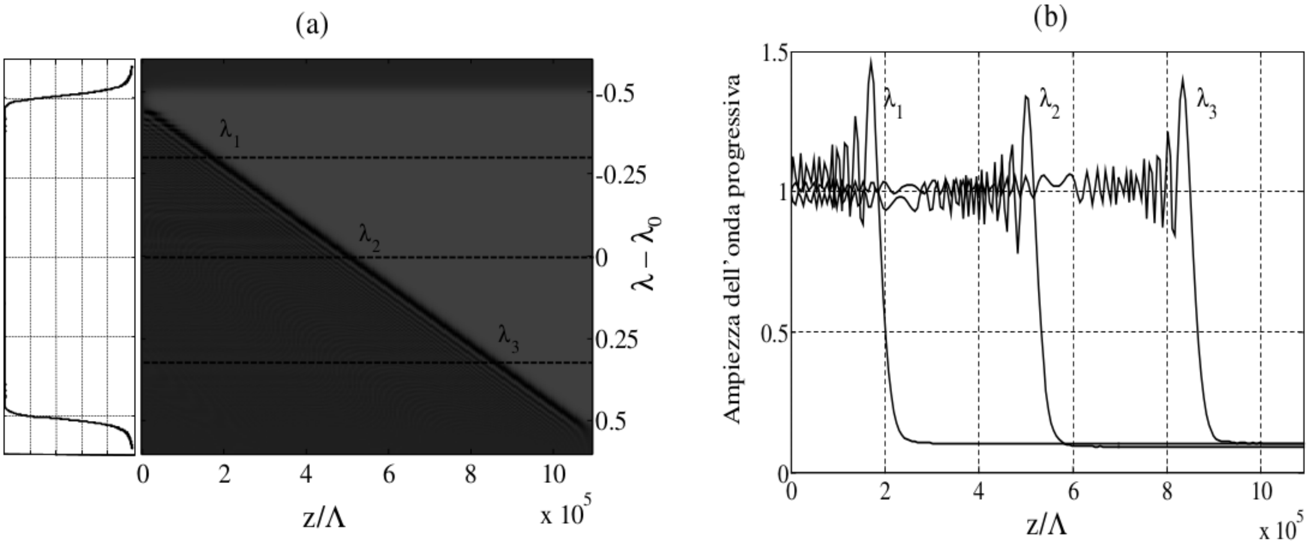
where the chromatic dispersion is defined as:

$$D_G = \frac{d\tau_g}{d\lambda} = \frac{1}{c \cdot C}$$

Once we have fixed D_G, i.e. C is fixed as well, the bandwidth increases with L.

For a chirped grating long 564nm, with $\delta n = 5 \cdot 10^{-5}$ and $C = 5.86 \cdot 10^{-10}$:





The oscillations in the group delay depend on the wavelenght: as shown in the group delay diagram the frequency of the osccillations increases linearly (for a positive chirped).
Moreover, the field appears perturbated because of the cavity effect, which causes also the oscillations in the group delay.
These oscillations cannot be ignored because they can deteriorate the devices based on chirped gratings. To avoid this problem, we rely on the apodization of the infex profile in order to minimize the disconinuity at the input. In this way, the oscillations of the group delay are reduced, but unfortunately also the operational bandwith of the device. An even more improved design is derived from an assymetric apodization: since the discontinuity at the ouput does not affect as the input one, the apodization concern only the input port of the gratings.

As it is mentioned in the introduction of this subsection, the chirped gratings are used to compensate the chromatic dispersion. This phenomenon is due to the different velocities with which different spectral components propagate along the guidance (for istance, an optical fiber).
We can represent formlly the chromatic dispersion extending the constant propagation around the central frequency of the spectrum:

$$\beta(\omega) = n(\omega)\frac{\omega}{c} = \beta_0 + \beta_1(\omega - \omega_0) + \frac{\beta_2}{2}(\omega - \omega_0)^2 + \cdots$$

where

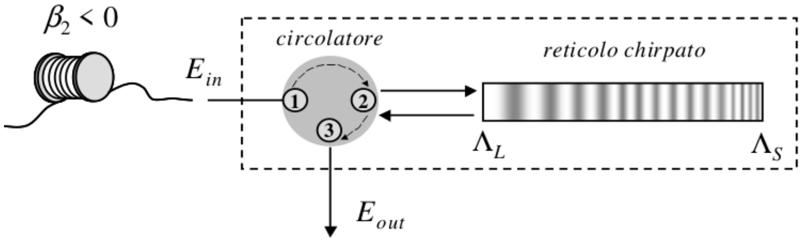
$$\beta_m = \left.\frac{d^m \beta_m}{d\omega^m}\right|_{\omega=\omega_0}$$

The coefficient β_1 is related to the well-known group velocity as follows:

$$v_g = \frac{d\omega}{d\beta} = \frac{1}{\beta_1}$$

If the group velocity (which gauges the velocity of the envelope) does not depend on ω all the higher coefficients are zero and the signal can propagate undisturbed.
On the other hand the temporal distribution of the signal is altered. In parituclar, the prevailing contribution is give by the β_2, that has the physical meaning of a delay and that varies linearly with the frequency.
If $\beta_2 < 0$ the low frequencies are delayed, viceversa if $\beta_2 > 0$.

To counter-part the chromatic dispersion is exploited the linear relation between the group delay and the wavelength. For this porpouse the chirped grting is used with a circulator, as shown in the following picture:



The circulator ha the single duty to separate the compensated signal E_{out} from the signal coming from the fiber. In general, the dispersion introduced by the fiber is compensated by the negative chirped grating. Formally the toal dispersion caused by the fiber is

$$D_F = \frac{d\tau_g}{d\lambda} = \frac{d\tau_g}{d\omega}\frac{d\omega}{d\lambda} = -\frac{2\pi c \beta_2 L_F}{\lambda^2}$$

where L_F is the total length of the fiber.
Next, imposing that $D_F + D_G = 0$ the value of C is derived:

$$C = \frac{\lambda^2}{2\pi c \beta_2 L_F}$$

The behaviour of the signal can be generally understood considering the following picture, representing an example for a sequence of three gaussian pulses: