Introduction

An Electro-magnetic field configuration inside a dielectric slab is considered a *mode* if it satisfies the following two requirements:

- being a wave inside the *core*, it impinges on each of the interfaces towards the *cover* with an incidence angle $\theta \geq \theta^\ell$, so always experiencing a **total internal reflection**: this allows the field to keep being confined in the core;
- it is **self-consistent**: after two consecutive reflections, its phase is incremented by a multiple of 2π. This guarantees that the field keeps a uniform magnitude along the propagation direction. If this direction is z, the z-dependence of the field expression should only be $e^{-jk_z z}$: it can immediately be verified that, this way, the value of z does not alter the magnitude of the field, being $|e^{-jk_z z}| = 1$, $\forall z \in \mathbb{R}$.

Dielectric slab waveguide

Planar dielectric waveguides come in various flavours, according to the available materials and the circuits where are embedded.

They are always composed, however, by a number n of layers. The most important of them is the *guiding* layer: the Electro-magnetic field must be bounded as much as possible inside it. It has the highest refractive index, n_1. The other layers have lower refractive indexes.

Symmetrical slab waveguide

The simplest kind of dielectric slab is symmetrical with respect to the (y, z)-plane, which has equation $x = 0$. It is composed by only two materials: a *cover* (sometimes alternatively called *cladding*, following the fiber optics terminology) with refractive index n_2, and a *core*, with refractive index $n_1 > n_2$.

Note that this structure has an infinite extension along the y-axis. The z-axis enters the page, according to the right-hand rule. The permettivity $\varepsilon = \varepsilon(x)$ is not a constant through the *vertical* direction x. Its variation shows how the two types of materials fill the whole three-dimensional space:

$$\varepsilon(x) = \begin{cases} \varepsilon_2, & x > a \\ \varepsilon_1, & a \leq x \leq -a \\ \varepsilon_2, & x < -a \end{cases} \quad = \quad \begin{cases} \varepsilon_1, & |x| \leq a \\ \varepsilon_2, & |x| > a \end{cases} \tag{1}$$

Correspondingly, the refractive index $n = n(x) = \sqrt{\varepsilon_r(x)} = \sqrt{\varepsilon(x)/\varepsilon_0}$ is:

$$n(x) = \begin{cases} n_1, & |x| \leq a \\ n_2, & |x| > a \end{cases} \tag{2}$$

This structure can be trivially realized by a slab of a specific material as **core** with some $n_1 > 1$, and the air itself as cover, with $n_2 = 1$.

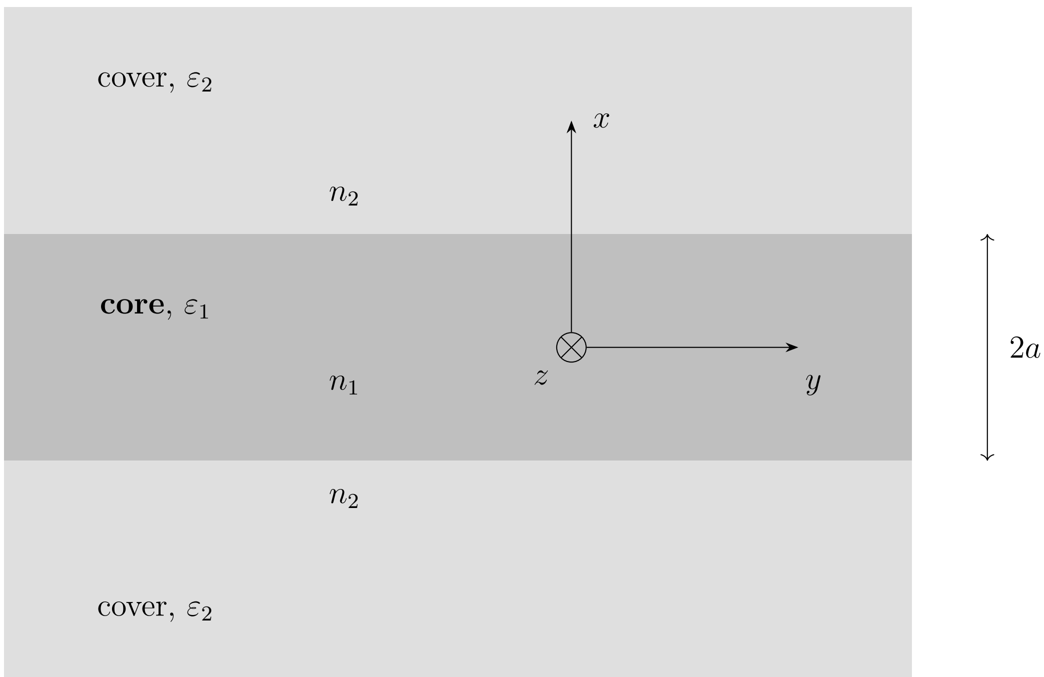

Figure 1: Structure of a symmetrical dielectric slab.

Such a waveguide is uniform with respect to both the y and the z direction: along each of them, fields are allowed to be *uniform* or *propagating*.

The objective of a waveguide is the *propagation* of a signal along a specific direction: then, uniform fields along both y and z would not meet this requirement[1]. A field which is (a) uniform along one of these directions, and (b) propagating along the other, will instead be obtained. The choice (a) significantly reduces the complexity of the problem: the whole structure will appear as bi-dimensional, with an eased computation[2]. The choice (b) will allow for the propagation of a signal[3].

The following hypothesis will be applied:

1. $\dfrac{\partial}{\partial y} = 0$ for all the field quantities. The field components can then be represented as $E(x, y, z) = E(x, z)$ or $H(x, y, z) = H(x, z)$. The y direction is chosen as the uniform direction.

2. The field components have separable dependencies with respect to the remaining variables x and z, such that $E(x, z) = e(x)f(z)$, and $f(z)$ should represent a *propagator*: $f(z) = e^{-jk_z z}$, $k_z \in \mathbb{R}$. This is the only dependency on z allowing for a propagating field. The z direction is therefore chosen as the propagation direction.

3. $n_1 > n_2$.

4. Absence of field sources[4].

5. Both the core and the cover media are linear and isotropic[5].

---

[1] Uniform fields along the whole (y, z) plane would be appropriate for a capacitor, not a waveguide.

[2] While such a dielectric slab is unrealisable (like the parallel-plate waveguide), it shows the major features of this whole family of waveguides in a simple way.

[3] Propagation can obviously take place also along a generic direction with both a y and a z component, but it wouldn't ease the computations. The coordinate reference system can always be chosen such that the direction of propagation is one of them, without loosing generality.

[4] The fields will be related to a region of space far away from their sources: here, it is not important the generation of the fields, but only the fields themselves, in a steady-state configuration.

[5] If the medium is linear, $\mathbf{D} = [\varepsilon]\mathbf{E}$, being $[\varepsilon]$ a tensor. If the medium is *also* isotropic, then $\mathbf{D} = \varepsilon\mathbf{E}$, being ε a scalar quantity.

The reference Maxwell equations to be solved are then:

$$\begin{cases} \nabla \cdot \mathbf{D} = 0 \\ \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{E} = -j\omega\mu\mathbf{H} \\ \nabla \times \mathbf{H} = j\omega\varepsilon\mathbf{E} \end{cases} \tag{3}$$

From the first equation $\nabla \cdot \mathbf{D} = 0$, using linearity and isotropy of the media,

$$\begin{aligned} &\nabla \cdot (\varepsilon\mathbf{E}) = 0 \\ &\varepsilon\nabla \cdot \mathbf{E} = 0 \\ &\nabla \cdot \mathbf{E} = 0 \end{aligned} \tag{4}$$

being always $\varepsilon \neq 0$[6]. Last equation in (4) lets some useful expressions to be obtained, expanding the ∇ operator:

$$\frac{\partial}{\partial x}E_x(x,z) + \frac{\partial}{\partial y}E_y(x,z) + \frac{\partial}{\partial z}E_z(x,z) = 0 \tag{5}$$

Remembering the hypotheses 1 and 2,

$$\begin{aligned} &\frac{\partial}{\partial y}E_y(x,z) = 0 \\ &\frac{\partial}{\partial z} = -jk_z \end{aligned} \tag{6}$$

The operator of partial derivative with respect to z simply becomes a scalar multiplication to $-jk_z$, due to the form of $f(z) = e^{-jk_z z}$ which is common to all the field components. So, (5) becomes

$$\begin{aligned} &\frac{\partial}{\partial x}E_x(x,z) - jk_z E_z(x,z) = 0 \\ &\frac{\partial}{\partial x}E_x(x,z) = jk_z E_z(x,z) \end{aligned} \tag{7}$$

Similarly, from the 2nd Maxwell equation[7]:

[6] The permettivity is a property of the material: its lowest value is about vacuum, $\varepsilon_0 \simeq 8.8541878 \cdot 10^{-12}$ F/m.

[7] Also μ is a property of the medium and it is always not zero.

$$
\begin{aligned}
&\nabla \cdot (\mu \mathbf{H}) = 0 \\
&\nabla \cdot \mathbf{H} = 0 \\
&\frac{\partial}{\partial x} H_x(x,z) = j k_z H_z(x,z)
\end{aligned} \tag{8}
$$

The 3rd Maxwell equations expands as follows:

$$
\nabla \times \mathbf{E} = \det \begin{vmatrix} \mathbf{u}_x & \mathbf{u}_y & \mathbf{u}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ E_x(x,z) & E_y(x,z) & E_z(x,z) \end{vmatrix} =
$$

$$
\begin{aligned}
&= \mathbf{u}_x \left(\frac{\partial}{\partial y} E_z(x,z) - \frac{\partial}{\partial z} E_y(x,z) \right) + \mathbf{u}_y \left(\frac{\partial}{\partial z} E_x(x,z) - \frac{\partial}{\partial x} E_z(x,z) \right) + \\
&+ \mathbf{u}_z \left(\frac{\partial}{\partial x} E_y(x,z) - \frac{\partial}{\partial y} E_x(x,z) \right) = -j\omega\mu\mathbf{H}
\end{aligned} \tag{9}
$$

Remembering that all the members with a partial derivatives with respect to y are zero, equation (9) becomes:

$$
\begin{aligned}
&-\mathbf{u}_x \frac{\partial}{\partial z} E_y(x,z) + \mathbf{u}_y \left(\frac{\partial}{\partial z} E_x(x,z) - \frac{\partial}{\partial x} E_z(x,z) \right) + \mathbf{u}_z \frac{\partial}{\partial x} E_y(x,z) = \\
&= -j\omega\mu\mathbf{H}
\end{aligned} \tag{10}
$$

Being (10) a vector equation, it represents three single equations, along the three directions $\mathbf{u}_x$, $\mathbf{u}_y$, $\mathbf{u}_z$ of the three-dimensional space. They are, respectively:

$$
\begin{gathered}
-\frac{\partial}{\partial z}E_y(x,z) = -j\omega\mu H_x(x,z) \\
\frac{\partial}{\partial z}E_y(x,z) = j\omega\mu H_x(x,z) \\
-jk_zE_y(x,z) = j\omega\mu H_x(x,z) \\
\boxed{k_z\, E_y(x,z) = -\omega\mu\, H_x(x,z)}
\end{gathered}
\tag{11}
$$

$$
\begin{gathered}
\frac{\partial}{\partial z}E_x(x,z) - \frac{\partial}{\partial x}E_z(x,z) = -j\omega\mu H_y(x,z) \\
-jk_zE_x(x,z) - \frac{\partial}{\partial x}E_z(x,z) = -j\omega\mu H_y(x,z) \\
\boxed{jk_z\, E_x(x,z) + \frac{\partial}{\partial x}\, E_z(x,z) = j\omega\mu\, H_y(x,y)}
\end{gathered}
\tag{12}
$$

$$
\boxed{\frac{\partial}{\partial x}\, E_y(x,z) = -j\omega\mu\, H_z(x,z)}
\tag{13}
$$

Applying the same procedure for the 4th Maxwell equation,

$$
\begin{gathered}
\nabla \times \mathbf{H} = \det \begin{vmatrix} \mathbf{u}_x & \mathbf{u}_y & \mathbf{u}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ H_x(x,z) & H_y(x,z) & H_z(x,z) \end{vmatrix} = \\
= \mathbf{u}_x\left(\frac{\partial}{\partial y}H_z(x,z) - \frac{\partial}{\partial z}H_y(x,z)\right) + \mathbf{u}_y\left(\frac{\partial}{\partial z}H_x(x,z) - \frac{\partial}{\partial x}H_z(x,z)\right) + \\
+\mathbf{u}_z\left(\frac{\partial}{\partial x}H_y(x,z) - \frac{\partial}{\partial y}H_x(x,z)\right) = j\omega\varepsilon\mathbf{E}
\end{gathered}
\tag{14}
$$

$$
\boxed{k_z\, H_y(x,z) = \omega\varepsilon\, E_x(x,z)}
\tag{15}
$$

$$\boxed{jk_z H_x(x,z) + \frac{\partial}{\partial x} H_z(x,z) = -j\omega\varepsilon E_y(x,y)} \tag{16}$$

$$\boxed{\frac{\partial}{\partial x} H_y(x,z) = j\omega\varepsilon E_z(x,z)} \tag{17}$$

The relations (11), (12), (13), (15), (16), (17) have been obtained only with the very general and simple hypotheses 1-5.

After inspection of the field components involved in each of the 6 equations, it is possible to identify two *completely independent* sets of equations:

- One set is composed by (11), (13) and (16) and it involves *only* the field components

$$E_y(x,z),\ H_x(x,z),\ H_z(x,z) \tag{18}$$

- Equations (12), (15) and (17) represent another set, with the field components

$$H_y(x,z),\ E_x(x,z),\ E_z(x,z) \tag{19}$$

The value of the components $E_y(x,z)$, $H_x(x,z)$ and $H_z(x,z)$ does not depend on the remaining $H_y(x,z)$, $E_x(x,z)$ and $E_z(x,z)$: therefore, a field with only the first three components can exist with or without the other ones, and vice-versa. These two separate sets are *independent from each other*.

The curl operator is now applied to both sides of the 3rd and 4th Maxwell equations:

$$\begin{aligned} \nabla \times \nabla \times \mathbf{E} &= -j\omega\mu\nabla \times \mathbf{H} \\ \nabla \times \nabla \times \mathbf{H} &= j\omega\varepsilon\nabla \times \mathbf{E} \end{aligned} \tag{20}$$

Given a vector field $\mathbf{A}$, the following relation holds:

$$\nabla \times \nabla \times \mathbf{A} = \nabla\left(\nabla \cdot \mathbf{A}\right) - \nabla^2\mathbf{A} \tag{21}$$

Recalling that $\nabla \cdot \mathbf{E} = 0$ and $\nabla \cdot \mathbf{H} = 0$, equations (20) become:

$$
\begin{aligned}
&\nabla \times \nabla \times \mathbf{E} = -\nabla^2 \mathbf{E} = -j\omega\mu \nabla \times \mathbf{H} \\
&\nabla \times \nabla \times \mathbf{H} = -\nabla^2 \mathbf{H} = j\omega\varepsilon \nabla \times \mathbf{E} \\
&\nabla^2 \mathbf{E} = j\omega\mu \nabla \times \mathbf{H} \\
&\nabla^2 \mathbf{H} = -j\omega\varepsilon \nabla \times \mathbf{E}
\end{aligned} \tag{22}
$$

$\nabla \times \mathbf{E}$ in the Right Hand Side of the last equation is given by the 3rd Maxwell equation, as well as $\nabla \times \mathbf{H}$ is given by the 4th one. They can be substituted:

$$
\begin{aligned}
&\nabla^2 \mathbf{E} = j\omega\mu \left(j\omega\varepsilon \mathbf{E} \right) \\
&\nabla^2 \mathbf{H} = -j\omega\varepsilon \left(-j\omega\mu \mathbf{H} \right) \\
&\nabla^2 \mathbf{E} = -\omega^2 \mu\varepsilon \mathbf{E} \\
&\nabla^2 \mathbf{H} = -\omega^2 \mu\varepsilon \mathbf{H}
\end{aligned} \tag{23}
$$

$$
\begin{aligned}
&\nabla^2 \mathbf{E} + \omega^2 \mu\varepsilon \mathbf{E} = 0 \\
&\nabla^2 \mathbf{H} + \omega^2 \mu\varepsilon \mathbf{H} = 0
\end{aligned} \tag{24}
$$

Two Helmoltz equations, which are *wave equations*, have been obtained: one for the Electric field and one for the Magnetic field. This **allows the existence of waves** in the space represented in Figure 1.

Equations (24) are actually a synthesis of 6 scalar equations, one for each component of the Electric and Magnetic fields:

$$
\begin{aligned}
&\nabla^2 E_x(x,z) + \omega^2 \mu\varepsilon E_x(x,z) = 0 \\
&\nabla^2 E_y(x,z) + \omega^2 \mu\varepsilon E_y(x,z) = 0 \\
&\nabla^2 E_z(x,z) + \omega^2 \mu\varepsilon E_z(x,z) = 0 \\
&\nabla^2 H_x(x,z) + \omega^2 \mu\varepsilon H_x(x,z) = 0 \\
&\nabla^2 H_y(x,z) + \omega^2 \mu\varepsilon H_y(x,z) = 0 \\
&\nabla^2 H_z(x,z) + \omega^2 \mu\varepsilon H_z(x,z) = 0
\end{aligned} \tag{25}
$$

TE modes

Let's now assume that only the components of the first set, $E_y(x,z)$, $H_x(x,z)$ and $H_z(x,z)$, are not zero. Therefore, $E_x(x,z) = E_z(x,z) = 0$. Only one of the three Electric field scalar equations from (25) is not a zero identity:

$$\begin{aligned}
&\nabla^2 E_y(x,z) + \omega^2\mu\varepsilon E_y(x,z) = 0\\
&\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) E_y(x,z) + \omega^2\mu\varepsilon E_y(x,z) = 0\\
&\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right) E_y(x,z) + \omega^2\mu\varepsilon E_y(x,z) = 0
\end{aligned} \tag{26}$$

Remembering hypothesis 2,

$$E_y(x,z) = e(x)f(z) = e(x)e^{-jk_z z} \tag{27}$$

Note that each field component of $\mathbf{E}$ is represented by a *scalar* function $g(x,y,z)$. The space variables x, y, z, that this scalar function depends on, have absolutely no relation to the direction of the component being evaluated. In *this* particular problem, for the reasons already explained, E_y depends on x and z, but in general, it is an ordinary scalar function defined in the $\mathbb{R}^3$ space: $g(x,y,z)$.

$$\begin{aligned}
&\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right) e(x)e^{-jk_z z} + \omega^2\mu\varepsilon e(x)e^{-jk_z z} = 0\\
&\left[\frac{\partial^2}{\partial x^2}e(x)\right] e^{-jk_z z} - k_z^2 e(x)e^{-jk_z z} + \omega^2\mu\varepsilon e(x)e^{-jk_z z} = 0
\end{aligned} \tag{28}$$

The complex exponential $e^{-jk_z z}$ is such that $e^{-jk_z z} \neq 0$, $\forall z \in \mathbb{C}$. So, being a common factor, it can be erased:

$$\begin{aligned}
&\frac{\partial^2}{\partial x^2}e(x) - k_z^2 e(x) + \omega^2\mu\varepsilon e(x) = 0\\
&\frac{\partial^2}{\partial x^2}e(x) + (k^2 - k_z^2)e(x) = 0
\end{aligned} \tag{29}$$

The wave vector is, in general:

$$\mathbf{k} = k_x\mathbf{u}_x + k_y\mathbf{u}_y + k_z\mathbf{u}_z \tag{30}$$

Hypotheses 1 implies that $k_y = 0$, because all the fields are uniform along y. As regards the remaining quantities:

$$\begin{aligned} &k^2 = |\mathbf{k}|^2 = \omega^2\mu\varepsilon \\ &k^2 = k_x^2 + k_z^2 \\ &k^2 - k_z^2 = k_x^2 \end{aligned} \tag{31}$$

and then

$$\boxed{\frac{\mathrm{d}^2}{\mathrm{d}x^2}e(x) + k_x^2 e(x) = 0} \tag{32}$$

The symbol of partial derivative has been replaced by d, being x the only variable that $e(x)$ depends on. This is the equation that will determine the shape of all the fields components, starting from E_y that is being investigated now.

The space is filled with different materials along x. Their parameters will be $\mu = \mu_0$ and, according to (1), $\varepsilon = \varepsilon(x)$. If so, then

$$k^2(x) = \omega^2\mu_0\varepsilon(x) = k_x^2(x) + k_z^2 \tag{33}$$

k, as well as its x component k_x, will assume different values in the core and the cover. The structure is instead uniform along z and so k_z remains constant.

Given that, equation (32) actually represents *two* equations:

$$\begin{cases} \dfrac{\mathrm{d}^2}{\mathrm{d}x^2}e(x) + k_{x_1}^2 e(x) = 0, & |x| > a \\ \dfrac{\mathrm{d}^2}{\mathrm{d}x^2}e(x) + k_{x_2}^2 e(x) = 0, & |x| \le a \end{cases} \tag{34}$$

These are the only constraints that $e(x)$ is forced to have in the two regions of space.

They must be solved separately, because of the different k_x, but both are of the form (32), which admits several solutions. The unknown function $e(x)$ is allowed to be:

1. $Ae^{rx} + Be^{-rx}$, $r \in \mathbb{R}$: an increasing and/or decreasing exponential with respect to x;

2. $Ae^{jhx} + Be^{-jhx}$, $r = jh$, $h \in \mathbb{R}$: a complex exponential;

3. $A\sin(rx) + B\cos(rx)$, $r \in \mathbb{R}$: a sinusoid.

For each equation in (34), $e(x)$ is mathematically allowed to assume one of these three shapes, indifferently. Remembering that $e(x)$ represents the x dependency of an Electric field, these solutions provide different kinds of fields:

1. represents a field which exponentially increases, or which undergoes an exponential attenuation, along x.

2. provides a field which is propagating along x.

3. corresponds to the field of a standing wave along x.

A custom field with $\partial/\partial y = 0$ and $f(z) = e^{-jk_z z}$ has already been chosen. Proceeding in the same way, a shape for $e(x)$ is chosen, too.

Note, in fact, that this is not a full analysis of all the available solutions of the equations (34). This is rather an attempt to find if, *between* those solutions, a field which meets the desired requirements can be found.

Such a field should be able to propagate a signal inside the core along z and to confine as much as possible this signal inside the core, being negligible outside. The first feature is already granted by the condition $f(z) = e^{-jk_z z}$. Then, the following solutions are chosen for $e(x)$:

1. a sinusoidal shape inside the core, a stationary wave along x for $|x| \leq a$, representing a field which is confined inside this region;

2. an exponential decreasing shape in the cover, when $|x| > a$: while getting away from the core, the field should vanish as rapidly as possible.

This is not the consequence of a physical constraint upon this problem, but an arbitrary choice, aimed at seeking a desired field between the available ones.

Considering only the half-space $x \geq 0$ for easiness:

$$e(x) = \begin{cases} A\cos(k_{x_1}x) + B\sin(k_{x_1}x), & 0 \leq x \leq a \\ Ce^{-jk_{x_2}(x-a)} + De^{jk_{x_2}(x-a)}, & x > a \end{cases} \tag{35}$$

Note that k_{x_1} is a real quantity (as specified in the sinusoidal, standing wave solution of equations (34)).

The only way condition 2 is satisfied is $D = 0$ and k_{x_2} negative and purely imaginary:

$$\begin{aligned}
&k_{x_2} = -j\gamma \\
&-jk_{x_2} = j^2\gamma = -\gamma \\
&\gamma = |k_{x_2}| \\
&k_{x_2} = -j|k_{x_2}| \\
&-jk_{x_2} = -|k_{x_2}|
\end{aligned} \tag{36}$$

and from (35)

$$e(x) = Ce^{-|k_{x_2}|(x-a)}, \quad x > a \tag{37}$$

Having $D \neq 0$ would give rise to a term $De^{|k_{x_2}|(x-a)}$, which is exponentially increasing while taking away from the core (x growing above a): it would be incompatible with the above requirements and it would be physically unattainable.

The set of field components (18) has the property that the entire electric field is transverse to the direction of propagation z. For this reason, this field configuration gives rise to a Transverse Electric, TE, mode. Because of the first equation in (35), such a mode can be symmetrical or anti-symmetrical with respect to the plane $x = 0$.

Even TE modes

In the system of equations (35), $B = 0$ creates a symmetrical mode with respect to the plane (y, z) or $x = 0$, with $e(x)$ reaching its maximum on the plane itself. The function $e(x)$ is defined in the whole space as:

$$e(x) = \begin{cases} A\cos(k_{x_1}x), & |x| \leq a \\ A'e^{-|k_{x_2}|(|x|-a)}, & |x| > a \end{cases} \tag{38}$$

Note that $|x| \leq a$ is the synthesis of $-a \leq x \leq a$ and the second line is the synthesis[8] of

[8]This is a double expansion. The absolute value notation $|x|$ represents x if $x \geq 0$ and

$$\begin{cases} A'e^{-|k_{x2}|(x-a)}, & x > a \\ A'e^{|k_{x2}|(x+a)}, & x < -a \end{cases} \tag{39}$$

The function $e(x)$ is an *even* function, because it is symmetrical with respect to x. The TE mode represented by this $e(x)$ is therefore called *even TE mode*.

Such a field must satisfy the boundary conditions at the interface between the two materials, represented by the plane $x = a$ (and also the plane $x = -a$). In particular, the *tangent electric field* must be continuous across this interface.

$$\begin{gathered} A\cos(k_{x_1}x)\big|_{x=a} = A'e^{-|k_{x2}|(x-a)}\Big|_{x=a} \\ A' = A\cos(k_{x_1}a) \end{gathered} \tag{40}$$

This is a constraint on the coefficient A', whose value is determined when A is given. A will remain a parameter: it represents the field amplitude and it depends on the *source* strength. When the source and its intensity are given, a numerical value can be assigned to A.

The *tangent magnetic field* must satisfy the same continuity condition at the same interface. Between the (18) magnetic field components, only $H_z(x, y)$ is tangent to the $x = a$ plane.

Given $e(x)$, the field component in the set (18) that can be immediately determined is $E_y(x, z)$, according to (27). In order to obtain $H_z(x, z)$, instead, equation (13) (which is derived from the expansion of the curl equation (9)) must be recalled.

$$H_z(x, z) = -\frac{1}{j\omega\mu}\frac{\partial}{\partial x}E_y(x, z) \tag{41}$$

$E_y(x, z)$ has different definitions according to the region of space being considered, and $H_z(x, z)$ will, too.

$-x$ if $x < 0$. So, $|x| > a$ gives rise to two branches, one with $x > a$ and one with $x < -a$. In the first branch, where the x is certainly positive, $|x| - a$ in the exponent simply becomes $x - a$. In the second branch, where x is certainly negative, $|x| - a$ is $-x - a$.

$$
\begin{aligned}
H_z^{(1)}(x,z) &= -\frac{1}{j\omega\mu}\frac{\partial}{\partial x}E_y^{(1)}(x,z) \\
H_z^{(2)}(x,z) &= -\frac{1}{j\omega\mu}\frac{\partial}{\partial x}E_y^{(2)}(x,z)
\end{aligned}
\tag{42}
$$

where region 1 is $|x| \leq a$ and region 2 is $|x| > a$.

$$
\begin{aligned}
H_z^{(1)}(x,z)\Big|_{x=a} &= H_z^{(2)}(x,z)\Big|_{x=a} \\
-\frac{1}{j\omega\mu}\left[\frac{\partial}{\partial x}E_y^{(1)}(x,z)\right]_{x=a} &= -\frac{1}{j\omega\mu}\left[\frac{\partial}{\partial x}E_y^{(2)}(x,z)\right]_{x=a} \\
\left[\frac{\partial}{\partial x}E_y^{(1)}(x,z)\right]_{x=a} &= \left[\frac{\partial}{\partial x}E_y^{(2)}(x,z)\right]_{x=a}
\end{aligned}
\tag{43}
$$

The single derivatives have the following values:

$$
\begin{aligned}
\left[\frac{\partial}{\partial x}E_y^{(1)}(x,z)\right]_{x=a} &= -\, k_{x_1}A\sin(k_{x_1}x)\big|_{x=a} = -k_{x_1}A\sin(k_{x_1}a) \\
\left[\frac{\partial}{\partial x}E_y^{(2)}(x,z)\right]_{x=a} &= -\, |k_{x_2}|A'e^{-|k_{x_2}|(x-a)}\Big|_{x=a} = -|k_{x_2}|A'
\end{aligned}
\tag{44}
$$

Applying (43),

$$
\begin{aligned}
-k_{x_1}A\sin(k_{x_1}a) &= -|k_{x_2}|A' \\
A' &= \frac{k_{x_1}}{|k_{x_2}|}A\sin(k_{x_1}a)
\end{aligned}
\tag{45}
$$

Both equations (40) and (45) are related to A' and they must be verified at the same time. Then,

$$
\begin{aligned}
A\cos(k_{x_1}a) &= \frac{k_{x_1}}{|k_{x_2}|}A\sin(k_{x_1}a) \\
&\boxed{|k_{x_2}| = k_{x_1}\tan(k_{x_1}a)}
\end{aligned}
\tag{46}
$$

This first relation for the – so far – unknowns k_{x_1} and $|k_{x_2}|$ is called *characteristic equation*. It must be verified in order for the desired fields to exist, given that also the hypotheses 1-5 are already satisfied.

When having two unknowns, two equations are required to obtain their values.

The relation (31) will assume two different forms, one for each medium of the dielectric waveguide. In the medium 1, representing the *core*, by simple substitution of the appropriate quantities, (31) becomes:

$$\begin{gathered} k^2 = k_x^2 + k_z^2 \\ \boxed{k_1^2 = k_{x_1}^2 + k_z^2} \end{gathered} \tag{47}$$

In the medium 2, representing the *cover*:

$$\begin{gathered} k^2 = k_x^2 + k_z^2 \\ k_2^2 = k_{x_2}^2 + k_z^2 \end{gathered} \tag{48}$$

but from (36) it is known that k_{x_2} is a purely imaginary, negative quantity.

$$\begin{aligned} &k_{x_2} = -j\gamma \\ &\gamma = |k_{x_2}| \\ &k_{x_2} = -j|k_{x_2}| \\ &k_{x_2}^2 = j^2|k_{x_2}|^2 = -|k_{x_2}|^2 \end{aligned} \tag{49}$$

Equation (48) can then be rewritten in a more explicit form as

$$\boxed{k_2^2 = k_z^2 - |k_{x_2}|^2} \tag{50}$$

Note that k_z is the same quantity in both the media, so in both equations (47) and (50). Subtracting side-by-side equation (50) from equation (47):

$$\begin{aligned} &k_1^2 - k_2^2 = k_{x_1}^2 + |k_{x_2}|^2 \\ &\omega^2\mu_0\varepsilon_0(\varepsilon_{r_1} - \varepsilon_{r_2}) = k_{x_1}^2 + |k_{x_2}|^2 \\ &\boxed{k_{x_1}^2 + |k_{x_2}|^2 = k_0^2(n_1^2 - n_2^2)} \end{aligned} \tag{51}$$

The variables k_{x_1} and $|k_{x_2}|$ must satisfy at the same time the equations (46) and (51): together, they are called the *dispersion equations* of the TE even mode. The system composed by them can be solved to determine the unknown values of these two variables. This system contains a trascendent equation and a closed form for its solution is **not available**: therefore, a graphical method is recommended to find it. The field defined in (38) will then be fully determined, becoming a valid solution of Maxwell's equations for the present problem.

Odd TE modes

When $A = 0$ in the system of equations (35), an anti-symmetrical mode with respect to the plane (y, z) or $x = 0$ is obtained. The function $e(x)$ is:

$$e(x) = \begin{cases} B' e^{-|k_{x_2}|(x-a)}, & x > a \\ B \sin(k_{x_1} x), & |x| \leq a \\ -B' e^{|k_{x_2}|(x+a)}, & x < -a \end{cases} \tag{52}$$

This $e(x)$ function is an *odd* function, because it is anti-symmetrical with respect to x. It represents an *odd TE mode*.

The procedure to fully determine $e(x)$ is the same as in the case of the even TE modes. The electric field tangent to the plane $x = a$ must be continuous: so, as in (40),

$$\begin{gathered} B \sin(k_{x_1} x)\big|_{x=a} = B' e^{-|k_{x_2}|(x-a)}\Big|_{x=a} \\ B' = B \sin(k_{x_1} a) \end{gathered} \tag{53}$$

Note that the sign of B' in the first line of (52) is determined by the value of the sine in $x = a$. In the last line of (52), the coefficient of the exponential function will have an opposite sign.

Using again equations (41) and (42), referring only to the region of space $x \geq 0$ to avoid ambiguity,

$$\begin{gathered} \left[\frac{\partial}{\partial x} E_y^{(1)}(x, z)\right]_{x=a} = k_{x_1} B \cos(k_{x_1} a) \\ \left[\frac{\partial}{\partial x} E_y^{(2)}(x, z)\right]_{x=a} = -|k_{x_2}| B' \end{gathered} \tag{54}$$

Then, according to (43),

$$
\begin{aligned}
&k_{x_1} B\cos(k_{x_1}a) = -|k_{x_2}|B' \\
&B' = -\frac{k_{x_1}}{|k_{x_2}|} B\cos(k_{x_1}a)
\end{aligned}
\tag{55}
$$

It follows that (as in (46)):

$$
\begin{gathered}
B\sin(k_{x_1}a) = -\frac{k_{x_1}}{|k_{x_2}|} B\cos(k_{x_1}a) \\
\boxed{|k_{x_2}| = -k_{x_1}\cot(k_{x_1}a)}
\end{gathered}
\tag{56}
$$

The system composed by this equation, along with the unvaried equation (51), can be solved to obtain the unknown values k_{x_1} and $|k_{x_2}|$, using a graphical method as already pointed out. The field defined in (52) will be fully determined for the TE odd modes, as well.

TM modes

Alternatevely to the first set (18) of components, the second one (19) can be used. If only $H_y(x,z)$, $E_x(x,z)$ and $E_z(x,z)$, are not zero, again, only one of the three Magnetic field scalar equations from (25) is not a zero identity:

$$
\begin{aligned}
&\nabla^2 H_y(x,z) + \omega^2\mu\varepsilon H_y(x,z) = 0 \\
&\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right) H_y(x,z) + \omega^2\mu\varepsilon H_y(x,z) = 0
\end{aligned}
\tag{57}
$$

Hypothesis 2 can also be applied to this single Magnetic field component:

$$
H_y(x,z) = h(x)f(z) = h(x)e^{-jk_z z} \tag{58}
$$

With the same considerations made for the TE modes, the following couple of equation can be obtained:

$$\begin{cases} \dfrac{\mathrm{d}^2}{\mathrm{d}x^2}h(x) + k_{x_1}^2 h(x) = 0, & |x| > a \\ \dfrac{\mathrm{d}^2}{\mathrm{d}x^2}h(x) + k_{x_2}^2 h(x) = 0, & |x| \leq a \end{cases} \tag{59}$$

remembering that the materials are the same as before, so only ε is varying, and μ remains constant. The same solution, with $D = 0$, is chosen for the same reasons.

$$h(x) = \begin{cases} A\cos(k_{x_1}x) + B\sin(k_{x_1}x), & 0 \leq x \leq a \\ Ce^{-jk_{x_2}(x-a)}, & x > a \end{cases} \tag{60}$$

The Magnetic field, being a single component, is entirely transverse to the direction of propagation z: a Transverse Magnetic, TM, mode will be considered. It can be symmetrical or anti-symmetrical with respect to the plane $x = 0$.

Even TM modes

The term $B = 0$ in (57) originates a symmetrical TM mode with respect to the plane (y, z). The function $h(x)$, which reaches its maximum on the plane itself, is defined as follows:

$$h(x) = \begin{cases} A\cos(k_{x_1}x), & |x| \leq a \\ A'e^{-|k_{x_2}|(|x|-a)}, & |x| > a \end{cases} \tag{61}$$

This is an even function, and it represents an even TM mode.

The fields must again satisfy the boundary conditions at the interface between the two materials. The *tangent magnetic field* must be continuous:

$$\begin{gathered} A\cos(k_{x_1}x)\big|_{x=a} = A'e^{-|k_{x_2}|(x-a)}\Big|_{x=a} \\ A' = A\cos(k_{x_1}a) \end{gathered} \tag{62}$$

which is identical to (40) (the fields are, too).

Then, the continuity requirement of the *electric magnetic field* concerns $E_z(x, y)$: this is the only electric field component in (19) to be tangent to the $x = a$ plane.

The function $h(x)$ is directly related to $H_y(x,z)$, according to (58). $E_z(x,z)$ can be obtained, instead, from equation (17) (which is derived from the expansion of the curl equation (14)).

$$E_z(x,z) = \frac{1}{j\omega\varepsilon}\frac{\partial}{\partial x}H_y(x,z) \tag{63}$$

It has different definitions according to the region of space being considered:

$$\begin{aligned} E_z^{(1)}(x,z) &= \frac{1}{j\omega\varepsilon_1}\frac{\partial}{\partial x}H_y^{(1)}(x,z) \\ E_z^{(2)}(x,z) &= \frac{1}{j\omega\varepsilon_2}\frac{\partial}{\partial x}H_y^{(2)}(x,z) \end{aligned} \tag{64}$$

where region 1 is $|x| \leq a$ and region 2 is $|x| > a$. Differently from (42), here the permettivity explicitly shows the change in the materials between the two regions.

This field component must be continuous across the $x = a$ plane:

$$\begin{aligned} E_z^{(1)}(x,z)\Big|_{x=a} &= E_z^{(2)}(x,z)\Big|_{x=a} \\ \frac{1}{j\omega\varepsilon_1}\left[\frac{\partial}{\partial x}E_y^{(1)}(x,z)\right]_{x=a} &= \frac{1}{j\omega\varepsilon_2}\left[\frac{\partial}{\partial x}E_y^{(2)}(x,z)\right]_{x=a} \\ \frac{1}{\varepsilon_1}\left[\frac{\partial}{\partial x}E_y^{(1)}(x,z)\right]_{x=a} &= \frac{1}{\varepsilon_2}\left[\frac{\partial}{\partial x}E_y^{(2)}(x,z)\right]_{x=a} \end{aligned} \tag{65}$$

Note that the permittivities don't erase like the permeabilities μ_0 in (43). The single derivatives have the following values:

$$\begin{aligned} \left[\frac{\partial}{\partial x}H_y^{(1)}(x,z)\right]_{x=a} &= -\, k_{x_1} A \sin(k_{x_1}x)\big|_{x=a} = -k_{x_1} A \sin(k_{x_1}a) \\ \left[\frac{\partial}{\partial x}H_y^{(2)}(x,z)\right]_{x=a} &= -\, |k_{x_2}| A' e^{-|k_{x_2}|(x-a)}\Big|_{x=a} = -|k_{x_2}| A' \end{aligned} \tag{66}$$

Applying (65),

$$
\begin{gathered}
-\frac{1}{\varepsilon_1} k_{x_1} A \sin(k_{x_1} a) = -\frac{1}{\varepsilon_2} |k_{x_2}| A' \\
A' = \frac{\varepsilon_2}{\varepsilon_1} \frac{k_{x_1}}{|k_{x_2}|} A \sin(k_{x_1} a)
\end{gathered}
\tag{67}
$$

Using both equations (62) and (67)

$$
\begin{gathered}
A \cos(k_{x_1} a) = \frac{\varepsilon_2}{\varepsilon_1} \frac{k_{x_1}}{|k_{x_2}|} A \sin(k_{x_1} a) \\
\boxed{|k_{x_2}| = \frac{n_2^2}{n_1^2} k_{x_1} \tan(k_{x_1} a)}
\end{gathered}
\tag{68}
$$

A quick comparison between this relation and its correspondent (46) shows a new term, that was previously not present: n_2^2/n_1^2. The more the refractive indexes will be close to each other, the more this ratio will be close to unity.

The unknown values k_{x_1} and $|k_{x_2}|$ can be obtained as graphical solutions of the system composed by equations (68) and (51). The TM even modes will be represented by the field fully defined in (57).

Odd TM modes

A zero A coefficient in the system of equations (60) will provide an anti-symmetrical mode with respect to the plane (y, z) or $x = 0$. The function $h(x)$ is, exactly as in equation (52):

$$
h(x) = \begin{cases} B' e^{-|k_{x_2}|(x-a)}, & x > a \\ B \sin(k_{x_1} x), & |x| \leq a \\ -B' e^{|k_{x_2}|(x+a)}, & x < -a \end{cases}
\tag{69}
$$

An odd $h(x)$ function is obtained: it is anti-symmetrical with respect to x and represents an *odd TM mode*.

The usual procedure is applied. Continuity of the magnetic field tangent to the plane $x = a$ provides (as in equation (53)):

$$
B' = B \sin(k_{x_1} a) \tag{70}
$$

Equations (63) and (64) (referring only to the region of space $x \geq 0$ to avoid ambiguity) provide

$$\begin{aligned} \left[\frac{\partial}{\partial x} H_y^{(1)}(x,z)\right]_{x=a} &= k_{x_1} B \cos(k_{x_1} a) \\ \left[\frac{\partial}{\partial x} H_y^{(2)}(x,z)\right]_{x=a} &= -|k_{x_2}| B' \end{aligned} \tag{71}$$

From (65),

$$\begin{aligned} \frac{1}{\varepsilon_1} k_{x_1} B \cos(k_{x_1} a) &= -\frac{1}{\varepsilon_2} |k_{x_2}| B' \\ B' &= -\frac{\varepsilon_2}{\varepsilon_1} \frac{k_{x_1}}{|k_{x_2}|} B \cos(k_{x_1} a) \end{aligned} \tag{72}$$

It follows that (as in (68)):

$$\begin{aligned} B \sin(k_{x_1} a) &= -\frac{\varepsilon_2}{\varepsilon_1} \frac{k_{x_1}}{|k_{x_2}|} B \cos(k_{x_1} a) \\ &\boxed{|k_{x_2}| = -\frac{n_2^2}{n_1^2} k_{x_1} \cot(k_{x_1} a)} \end{aligned} \tag{73}$$

As for the even TM modes, the only difference between (73) and (56) is the n_2^2/n_1^2 term.

This equation, along with (51), composes a system that can be graphically solved to obtain the unknown values k_{x_1} and $|k_{x_2}|$. The field for the odd TM modes, defined in (69), will then be fully determined.

Normalized equations

For each of the TE and TM modes, a system of two equations must be solved in order to fully characterize the Electro-magnetic field which represents a solution of the Maxwell's equations. These couples of equations (as anticipated) are called *dispersion equations* of the modes. They are:

- equations (46) and (51) for TE even modes;

$$
\begin{cases}
|k_{x_2}| = k_{x_1} \tan(k_{x_1} a) \\
k_{x_1}^2 + |k_{x_2}|^2 = k_0^2 (n_1^2 - n_2^2)
\end{cases}
\tag{74}
$$

- equations (56) and (51) for TE odd modes;

$$
\begin{cases}
|k_{x_2}| = -k_{x_1} \cot(k_{x_1} a) \\
k_{x_1}^2 + |k_{x_2}|^2 = k_0^2 (n_1^2 - n_2^2)
\end{cases}
\tag{75}
$$

- equations (68) and (51) for TM even modes;

$$
\begin{cases}
|k_{x_2}| = \dfrac{n_2^2}{n_1^2} k_{x_1} \tan(k_{x_1} a) \\
k_{x_1}^2 + |k_{x_2}|^2 = k_0^2 (n_1^2 - n_2^2)
\end{cases}
\tag{76}
$$

- equations (73) and (51) for TM odd modes.

$$
\begin{cases}
|k_{x_2}| = -\dfrac{n_2^2}{n_1^2} k_{x_1} \cot(k_{x_1} a) \\
k_{x_1}^2 + |k_{x_2}|^2 = k_0^2 (n_1^2 - n_2^2)
\end{cases}
\tag{77}
$$

These systems, however, are not solved in the form they have been presented so far.

Both sides of equations (46), (56), (68) and (73) are multiplied by a and both sides of equation (51) are multiplied by a^2. For example, the system (74) becomes:

$$
\begin{cases}
a|k_{x_2}| = a k_{x_1} \tan(a k_{x_1}) \\
a^2 k_{x_1}^2 + a^2 |k_{x_2}|^2 = a^2 k_0^2 (n_1^2 - n_2^2)
\end{cases}
\tag{78}
$$

Note that each member like ak_{x_1} and $a|k_{x_2}|$ is a dimensionless quantity and that the tangent has now ak_{x_1} both as argument and as multiplying factor.

Let

$$
\begin{aligned}
u &= ak_{x_1} \\
w &= a|k_{x_2}| \\
v^2 &= a^2 k_0^2 (n_1^2 - n_2^2)
\end{aligned}
\tag{79}
$$

So, system (74) can also be written in a more essential, straightforward manner:

$$
\boxed{\begin{cases} w = u \tan(u) \\ v^2 = u^2 + w^2 \end{cases}}
\tag{80}
$$

This provides a standard set of equations related to the TE even modes. They do not depend on the actual thickness a of the core, so they are always the same for all the problems. This is the standard, normalized form for the dispersion equations.

Each equation in the system (80) represents a curve: if and where the two curves intersect, the system has a solution (u, w). It may have zero, one or more solutions, corresponding to zero, one or more TE even modes available for the dielectric slab.

Because the system (80) is standard for each TE even modes problem, it is worth mentioning some features of the curves it generates.

These curves will be evaluated *only in the 1st quadrant*, where $u \geq 0, w \geq 0$. The half-plane $w < 0$ (composed by the 3rd and 4th quadrants) is not relevant, being $w = a|k_{x_2}|$. a is a thickness, and it is a real, positive quantity due to physical reasons. $|k_{x_2}|$ is the modulus of a complex number, which is a real, non-negative quantity. So, $w \geq 0$.

As pointed out in the solution (35) of equations (34), k_{x_1} is a real quantity too. Given that, the 2nd quadrant, where $w \geq 0$ and $u < 0$, would provide redundant information with respect to the 1st quadrant. In fact, let's consider a given solution $(u^{(0)}, w^{(0)})$ found in the 1st quadrant. Its abscissa value $u^{(0)}$ corresponds to $ak_{x_1}^{(0)}$. Considering $-k_{x_1}^{(0)}$ in the 2nd quadrant would provide the *same* TE and TM even modes, because

$$
\cos(k_{x_1}^{(0)} x) = \cos(-k_{x_1}^{(0)} x)
$$

(the cosine is an even function). As regards the TE and TM odd modes, $-k_{x_1}^{(0)}$ would provide modes which are symmetrical with respect to the (y, z) plane[9] to the modes provided by $k_{x_1}^{(0)}$, in fact

$$\sin(-k_{x_1}^{(0)} x) = -\sin(k_{x_1}^{(0)} x)$$

(being sine an odd function). With a symmetrical dielectric slab, this would not be a significant change; it is actually not a new field configuration; it is like observing the same solution from an opposite point of view. Moreover, the modulus of the field would be the same in both cases.

So, the graphical solution of system (80) will be only found for $u \geq 0, w \geq 0$.

The function $u \tan(u)$ resembles the well-known $\tan(u)$, especially as regards the following features: both intersect the origin $(0, 0)$, both diverge as $u \to \pi/2 + m\pi$ $(m \in \mathbb{Z})$ and both are not defined for $u = \pi/2 + m\pi$. However, $u \tan(u)$ is not a periodic function, due to the multiplying factor u; it is also an *even* function, being the product of two odd functions. As shown in Figure 2, $u \tan(u)$ raises slowly for small u (the u multiplication lowers the overall value of $u \tan(u)$ with respect to $\tan(u)$); for greater u instead, it diverges more rapidly (above $\tan(u)$, due to the same multiplication by u).

This function must intersect the circumference or radius v, defined with the equation $v^2 = u^2 + w^2$. The only value in (80) that depends on the specific dielectric slab problem is v itself: **this value** will determine the *number* and the *position* of the intersections with the other curve. Respectively, these features are related to the number of available modes and to the quality of the modes, in particular how much they are bounded to the core, because a higher $|k_{x_2}|$ will imply a higher bound.

The designer must only specify four parameters:

- the core half-thickness a;
- the vacuum wavelength of the signal λ;
- the refractive indexes of the two materials n_1 and n_2.

being

[9] In the space of the dielectric slab.

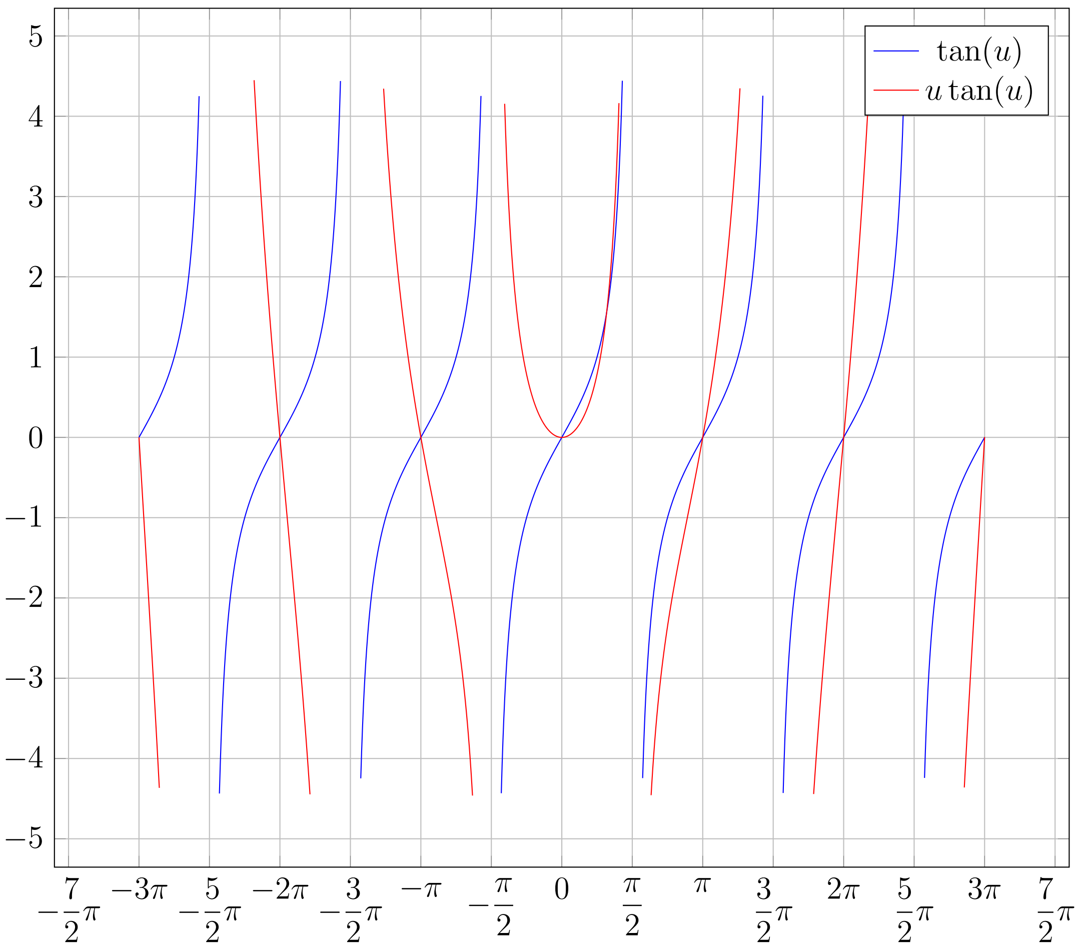

Figure 2: Comparison between $\tan(u)$ and $u \tan(u)$.

$$v^2 = a^2 k_0^2 (n_1^2 - n_2^2) = a^2 \left(\frac{2\pi}{\lambda} \right)^2 (n_1^2 - n_2^2) \tag{81}$$

Alternatively to λ, the knowledge of ω is sufficient, because

$$k_0^2 = \omega^2 \mu_0 \varepsilon_0 \tag{82}$$

but the use of λ is more common.

With this information, v is known and a solution to system (80) can be looked for. Each of these four fundamental parameters (a, λ, n_1 and n_2) determines the radius of the circumference.

With the equations (80), at least one intersection is theoretically always possible, so *one* TE even mode should always be available: this makes the TE even mode the *fundamental mode* of dielectric symmetrical slabs.

$$\boxed{\begin{cases} w = -u \cot(u) \\ v^2 = u^2 + w^2 \end{cases}} \tag{83}$$

is the system of equations correspondent to (75), for the TE odd modes.

$$\boxed{\begin{cases} w = \dfrac{n_2^2}{n_1^2} u \tan(u) \\ v^2 = u^2 + w^2 \end{cases}} \tag{84}$$

corresponds to (76), as regards the TM even modes.

$$\boxed{\begin{cases} w = -\dfrac{n_2^2}{n_1^2} u \cot(u) \\ v^2 = u^2 + w^2 \end{cases}} \tag{85}$$

is derived from (77) for the TM odd modes.

Note that while the first equation of systems (80) and (83) do not depend on the custom quantities n_1 and n_2, the first equations of systems (84) and (85) do, instead.

Given the numerical solution obtained by graphical methods for u and w, the actual values of k_{x_1} and $|k_{x_2}|$ can be obtained dividing the results by a.

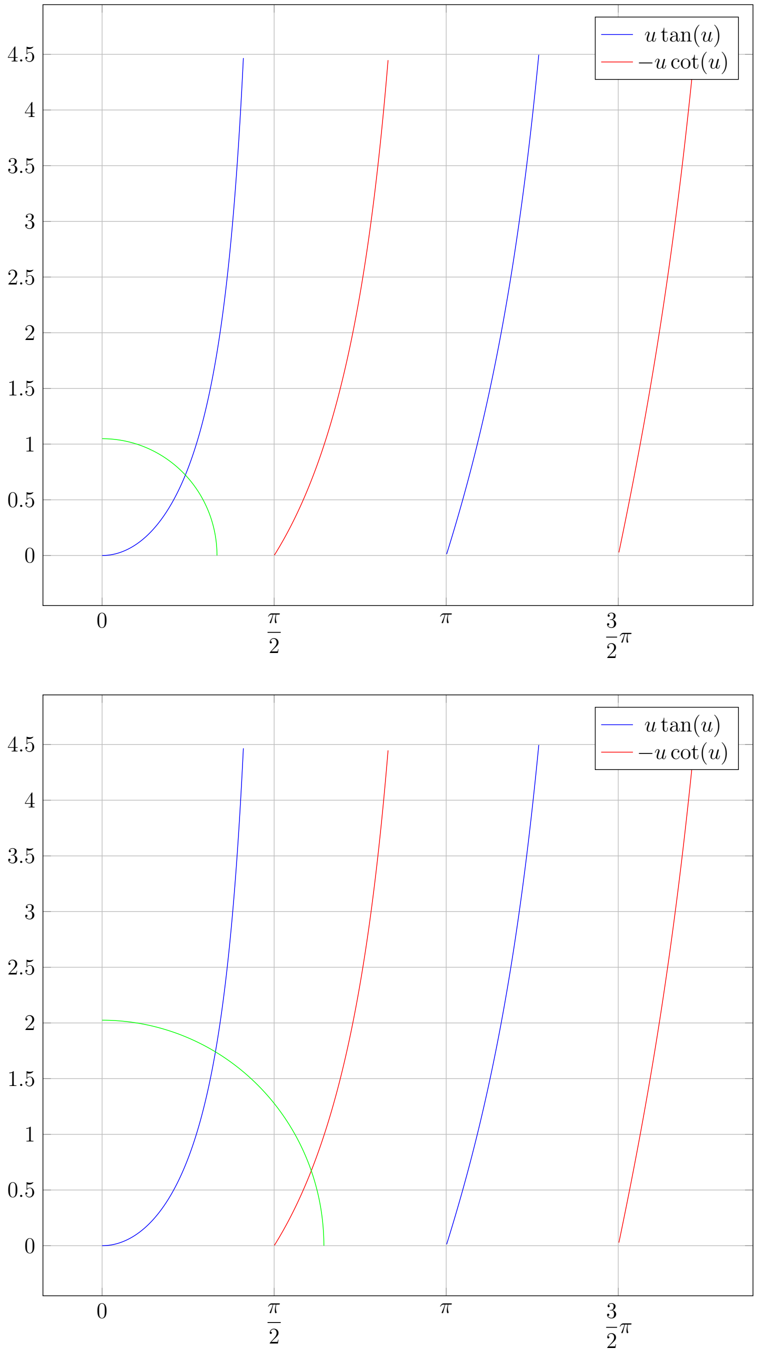

Figure 3: Variable number of available modes according to the values of the radius v: in the first plot, only one intersection occurs; in the second plot, where the circumference has a greater v, two intersections are generated.

Assuming that the physical parameters of the guide are known (a, n_1, n_2), if the *frequency* is fixed, Figure 3 shows that a *finite set of solutions* is available, that is: a **finite number of modes**. It is the same for TM modes.

This is different from parallel plate or rectangular metallic waveguides, for example, where an *infinite*, *numerable* set of solutions could be found. This kind of dielectric waveguides provides only a finite number of modes.

Figure 3 shows another fundamental feature of these solutions: considering always the same mode, for example the first TE even mode, if the frequency increases it becomes **more confined**. In fact, raising the frequency leads to an increment of the radius v of the circumference, then the intersection with the tangent (or cotangent) branch is higher, and the correspondent $|k_{x_2}|$ is greater. The higher its value, the stronger the field confinement.

Dielectric cutoff condition

There is a non-obvious relation between the frequency ω of a mode and its angle of incidence θ_c: the knowledge of the former quantity is equivalent to the knowledge of the latter.

From equation (47):

$$k_z = \sqrt{k_1^2 - k_{x_1}^2} \tag{86}$$

From equation (50):

$$k_z = \sqrt{k_2^2 + |k_{x_2}|^2} \tag{87}$$

These are *equivalent expressions* to determine k_z, the propagation constant of the mode, which (as specified in the initial hypothesis 2) is a real quantity: $k_z \in \mathbb{R}$.

As already observed in the solution (35) of equations (34), k_{x_1} is a real quantity, so $k_{x_1}^2$ is a real, non-negative quantity. Given that, equation (86) implies that

$$k_z \leq k_1 \tag{88}$$

Being k_z in any case a real quantity, of course it should be $k_{x_1} \leq k_1$ in (86) and therefore (88) is actually $0 \leq k_z \leq k_1$, but this full expression is not necessary, as it will be immediately shown.

Also $|k_{x_2}|$ is a real, non-negative quantity: therefore, from equation (87),

$$k_z \geq k_2 \tag{89}$$

Equations (86) and (87) are equivalent, alternative definitions of k_z and must be verified at the same time; so, also the inequalities (88) and (89) must be verified at the same time:

$$\begin{cases} k_z \leq k_1 \\ k_z \geq k_2 \end{cases} \tag{90}$$

It follows that, for each mode:

$$\boxed{k_2 \leq k_z \leq k_1} \tag{91}$$

This provides a *range* where k_z is allowed to span. It can not be a value outside this interval.

Inequality (91) can be written in an alternative form, knowing that $k_1 = n_1 k_0$ and $k_2 = n_2 k_0$. Dividing all members by k_0:

$$\begin{gathered} \frac{k_2}{k_0} \leq \frac{k_z}{k_0} \leq \frac{k_1}{k_0} \\ \boxed{n_2 \leq n_{\text{eff}} \leq n_1} \end{gathered} \tag{92}$$

where n_{eff} is the effective refractive index and its knowledge is equivalent to the knowledge of k_z.

By inspection of Figure 3, considering a single branch of $u \tan(u)$ or $-u \cot(u)$ and so a *single* mode, the quantity that varies the most is $|k_{x_2}|$. In fact, while k_{x_1} will at least vary in a $\pi/2$ interval starting from its initial value, $|k_{x_2}|$ can range from 0 to (theoretically) infinity. As already mentioned, this quantity determines the confinement of the mode.

For low frequencies, such that the circumference does not intersect the branch of the considered mode, the mode is not confined and not available for propagation *inside* the core.

Increasing the frequency, the *cutoff* condition is verified when the circumference *reaches* the considered branch. At this point, $|k_{x_2}|$ reaches the value 0 and the confinement of the mode is barely possible[10]. The mode **activates** and $k_z = k_2$ assumes its inferior limit value.

[10] This is the limit condition for the confinement: when $|k_{x_2}| = 0$, the mode is not confined, but for (even little) higher values, it is, because the exponent of $e^{|k_{x_2}|(x-a)}$ becomes not zero.

For higher frequencies, overcome the *cutoff* condition, the mode is confined and fully available for propagation: the higher the frequency, the higher $|k_{x_2}|$, the stronger the confinement. In the limit $|k_{x_2}| \to \infty$, even if inequality (89) and equation (87) could induce to the conclusion that also $k_z \to \infty$, it should be remembered that also inequality (87) must be verified: therefore, in the best achievable confinement[11], $k_z \to k_1$.

Snell's law

Let the three-dimensional space be divided in two regions: for $x \geq 0$, filled by a medium with refractive index n_2, and for $x < 0$, filled by a medium with refractive index $n_1 > n_2$. The plane $x = 0$ (orthogonal to the paper, in Figure 4) represents the interface between these two media.

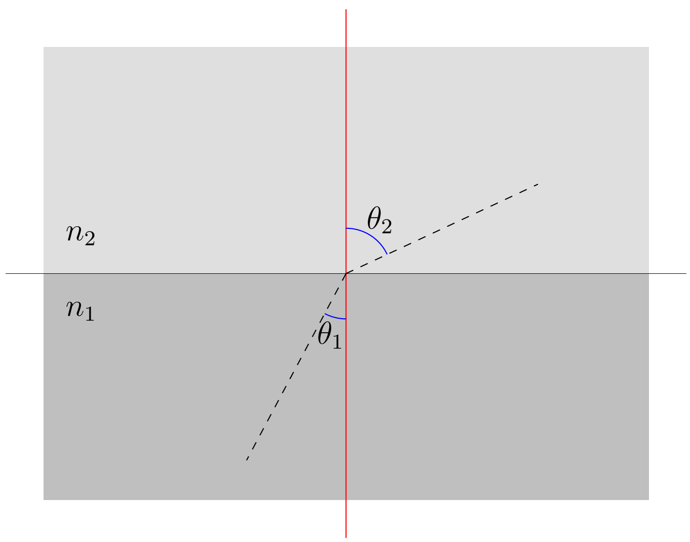

Figure 4: Refraction of a plane wave impinging on the interface between two media.

In Figure 4, the red line is normal (orthogonal) to the interface. Usually, n_1 and n_2 are two small quantities, close to each other and close to 1. For example, for light at $\lambda = 600$ nm: water (at 20 °C) has $n = 1.33$, crown glass (which is used in optics) has $n = 1.52$.

[11]This is actually unpractical, though, because it would require a huge circumference radius, which would activate several modes more than the chosen one, with undesirable effects.

Let a plane wave impinge on the interface. The direction of propagation of the wave (that is, the direction of the wavevector) forms an angle θ_1 with the direction of the normal line. After the plane wave crossed the interface, it will assume in general a *new* direction, forming an angle $\theta_2 \neq \theta_1$ with the normal line.

Snell's law regulates the relation between these two angles:

$$\boxed{n_1 \sin\theta_1 = n_2 \sin\theta_2} \tag{93}$$

The range of acceptable values for θ_1 is $[0, \pi/2]$. As regards the symmetrical range $[-\pi/2, 0]$, the same considerations apply and it would be redundant to consider it. For angles $\theta \geq \pi/2$ or $\theta \leq -\pi/2$, the initial plane wave would not be in the medium n_1, so this law would be reversed and applied with n_2 as the initial medium.

If $\theta_1 \in [0, \pi/2]$, $\sin\theta_1 \in [0, 1]$, and the higher the angle, the higher its sine.

Let initially be θ_1 almost 0, and then let this value raise. (93) imposes a relation between the **sines** of these angles:

$$\sin\theta_2 = \frac{n_1}{n_2}\sin\theta_1 \tag{94}$$

In particular, the sine of θ_2 grows as much as n_1/n_2 times the sine of θ_1: if $n_1 > n_2$, $\sin\theta_2$ is greater[12] than $\sin\theta_1$. So, also the *angle* θ_2 that is generated will be greater than θ_1 for *each* test value of θ_1.

While θ_1 increases, it will reach a value $\theta_1^\ell < \pi/2$ for which the corresponding θ_2^ℓ will equal $\pi/2$: the plane wave is bended so much that **it is no more able to exit the interface** in $x = 0$. Therefore, the wave follows a direction which is *parallel* to that interface and **does not reach** the medium 2. This is a limit condition and the correspondent angle θ_1^ℓ

$$\sin\theta_1^\ell = \frac{n_2}{n_1}\sin\theta_2 = \frac{n_2}{n_1} \tag{95}$$

is called **limit angle** or critical angle. If $\theta_1 \geq \theta_1^\ell$, the incident plane wave will be confined inside the medium 1, because no real value of θ_2 would satisfy Snell's law (93). In fact, as more explicitly shown by (94), $\sin\theta_2$ should

[12] Not far greater, remembering the considerations made on the n values, but anyway higher than 1.

assume values greater than 1, which is not possible for a sine function of a real angle[13].

When dealing with Snell's law, only refraction is often considered. Anyway, it is not the only phenomenon which occurs: each plane wave impinging on an interface will undergo both **refraction** and **reflection**. Reflection is well considered in textbooks[14] in the simple case when $\theta_1 = 0$, so for a wave normally incident on an interface plane.

As regards the power carried by the plane wave, part of it crosses the interface, is refracted and can proceed in the medium 2: its direction of propagation is bended at the interface according to Snell's law; part of it, instead, is reflected, and it goes back in the medium 1 with the same angle of incidence θ_1, but with opposite direction. A minimalist example of this joint action, referred to the same plane wave and the same geometry of Figure 4, is shown in Figure 5.

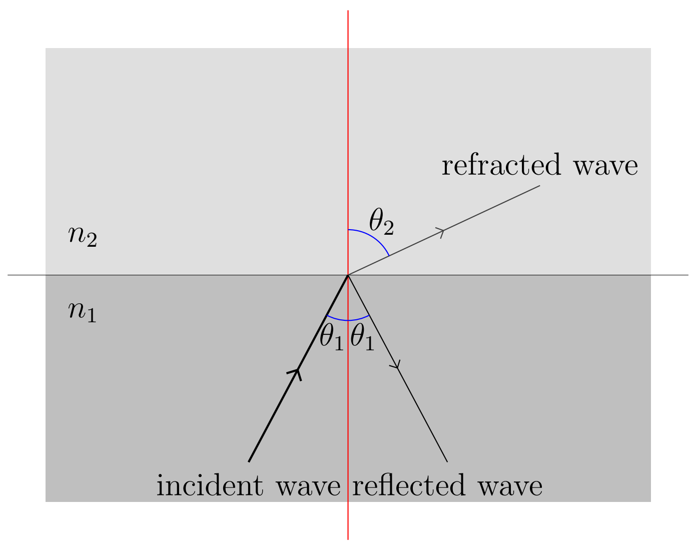

Figure 5: Complete description of the plane wave behaviour at the interface between two media.

If $\theta_1 > \theta_1^\ell$, ideally *all* the incident power is reflected back in the medium 1: **there is no refraction, but only reflection**. It is worth observing

[13]It is possible when angles have a purely imaginary value. So, there is still a solution of the type $\theta_2 = j\alpha$, $\alpha \in \mathbb{R}$, for equation (94), but it has no physical meaning.

[14]For example, D. M. Pozar, *Microwave Engineering*, Wiley.

in fact that the θ_1 angle of the reflected wave is not a *new* value of the refraction angle, because that angle *does not exist* any more and refraction *does not occur*. The "second" θ_1 angle is the result of a *different phenomenon*: reflection. It has always occurred: however, when refraction occurs too, reflection is often negligible, because most of the power of the incident wave crosses the interface.

On the contrary, when refraction *can not* occur, due to an angle of incidence which exceeds the limit-angle θ_1^ℓ, reflection becomes significant. Ideally *all* the power carried by the incident wave is **reflected back** at the interface: there is no transmission of power to the medium 2 and a full reflection occurs.

A waveguide includes not just one, but *two* interfaces between the medium 1 and the medium 2: one above the medium 1, and one below it, according to Figure 1. Assuming that the plane wave has been generated in the medium 1 in an appropriate way, if it is impinging on the upper interface with an angle $\theta_1 > \theta_1^\ell$, after the first full reflection it will meet the second interface, where the same condition applies, and so on.

This way, *all the power* carried by the plane wave *remains confined* in the medium 1, that is in the **core**.

This behaviour makes possible the use of dielectric waveguides to carry signals.

Figure 6 shows the path followed by a plain wave in a waveguide, with generic wavevector $\mathbf{k}_1$ having an angle of incidence $\theta_1 > \theta_1^\ell$. Note that, in order to explicitly show the direction of propagation z, a different perspective is assumed with respect to Figure 1, which has been rotated by 90° in the plane of the core, towards the observer.

Figure 6 also explicitly represents equations (47) and (86). The wavevector components can be put in relation with the angle θ_1, too. θ_1 is the angle between the direction of $\mathbf{k}_1$ and the direction which is normal to the dielectric interface (red line), as well as (by geometric considerations) between $\mathbf{k}_1$ and the vertical dashed line parallel to k_{x_1}. Consequently,

$$k_z = k_1 \sin\theta_1 \tag{96}$$

and also

$$k_{x_1} = k_1 \cos\theta_1 \tag{97}$$

Equation (96) can alternatively be written as

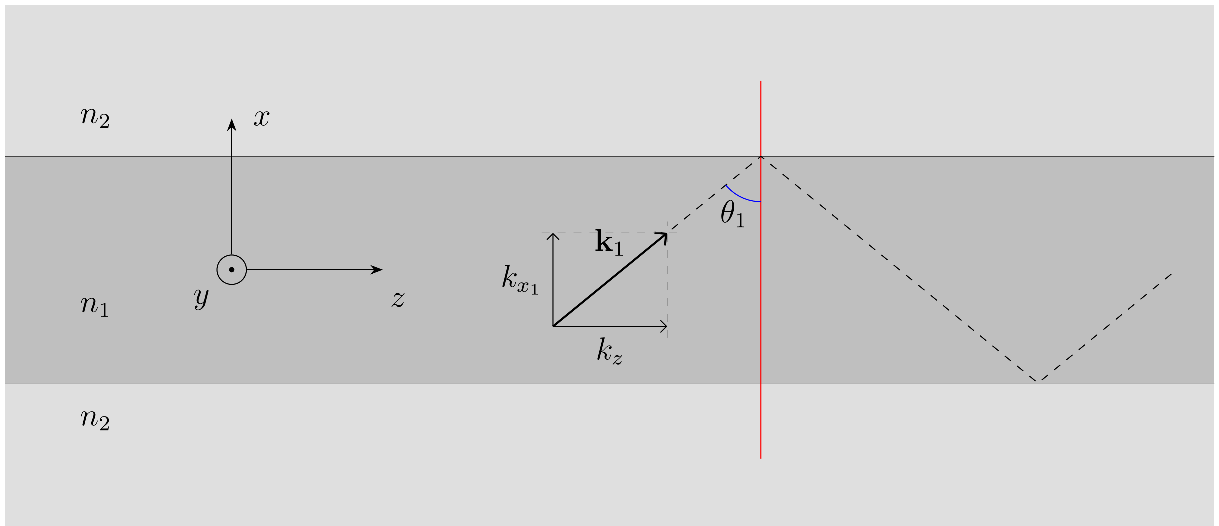

Figure 6: Path of a plane wave inside the waveguide core when the total internal reflection condition is satisfied.

$$k_z = n_1 k_0 \sin\theta_1 = n_1 \omega \sqrt{\mu_0 \varepsilon_0} \sin\theta_1 \tag{98}$$

This proves that there *is* a relation between the angle of incidence θ_1 and the value of k_z.

Moreover, being

$$0 \leq \sin\theta_1 \leq 1 \tag{99}$$

necessarily will also be from (98)

$$0 \leq k_z \leq n_1 k_0 \tag{100}$$

Also,

$$\theta_1 \geq \theta_1^\ell \tag{101}$$

that can be rewritten (applying the sine function to both members and remembering (95)) as

$$\sin\theta_1 \geq \frac{n_2}{n_1} \tag{102}$$

So (again from (98)) it will be

$$k_z \geq n_1 k_0 \cdot \frac{n_2}{n_1} \tag{103}$$

that is

$$k_z \geq n_2 k_0 \tag{104}$$

Because both the conditions (100) and (104) on k_z must be verified at the same time,

$$\begin{cases} 0 \leq k_z \leq n_1 k_0 \\ k_z \geq n_2 k_0 \end{cases} \Rightarrow \begin{cases} 0 \leq k_z \leq k_1 \\ k_z \geq k_2 \end{cases} \Rightarrow \boxed{k_2 \leq k_z \leq k_1} \tag{105}$$

These are the same conditions in (90) and (91), obtained here through an alternative approach leading to the same results.

As already stated, the cutoff condition for a mode occurs when the value of $|k_{x_2}|$ equals 0. Referring to Figure 3 and considering only the TE even modes as a first example, the $w = u\tan(u)$ branches will satisfy this condition when they cross the u-axis. It is the abscissa axis, with equation $w = 0$: $|k_{x_2}| = 0$ is equivalent to $w = a|k_{x_2}| = 0$.

$$\begin{aligned} &u\tan(u) = 0 \\ &u = 0 \vee \tan(u) = 0 \end{aligned} \tag{106}$$

The trivial solution $u = 0$, which implies $k_{x_1} = 0$, is usually not considered: it would correspond to a field which is not confined and, inside the core, would not create a stationary wave, having a zero k_{x_1}. Also $\tan(u)$ reaches 0 for $u = 0$. All the other zeros of the $\tan(u)$ in the 1st quadrant are suitable:

$$\begin{aligned} &u = n\pi,\ \ n = 1, 2, \ldots \\ &ak_{x_1} = n\pi \end{aligned} \tag{107}$$

Each $u\tan(u)$ branch corresponds to a *mode*. It is activated – as anticipated – when the circumference $u^2 + w^2$ reaches a radius v such that it is able to intersect the specific $u\tan(u)$. The first intersection occurs when the mode is at cutoff, on the u-axis. Here,

$$w = 0 \Rightarrow \begin{cases} u\tan(u) = 0,\ \ u = n\pi \\ v^2 = u^2 + w^2 = u^2 \end{cases} \Rightarrow v = u \tag{108}$$

When the cutoff condition for a specific mode is met, the circumference radius v assumes the same value as the abscissa u: being v related to the frequency f of the actual sinusoidal signal inside the core, the value assumed by f will be called *cutoff frequency* of the mode. Its value can be easily obtained.

The circumference equation is

$$v^2 = u^2 + w^2 \tag{109}$$

At cutoff, $w = 0$. Remembering equation (79),

$$v^2 = u^2 + w^2 \Rightarrow a^2 k_0^2 (n_1^2 - n_2^2) = u^2 \tag{110}$$

Using (82), remembering that $\omega = 2\pi f$ and applying the square root to both the sides:

$$
\begin{aligned}
&a^2\omega^2\mu_0\varepsilon_0(n_1^2 - n_2^2) = u^2 \\
&a^2 4\pi^2 f^2 \mu_0\varepsilon_0(n_1^2 - n_2^2) = u^2 \\
&2a\pi f\sqrt{\mu_0\varepsilon_0}\sqrt{n_1^2 - n_2^2} = u
\end{aligned}
\tag{111}
$$

Remembering that $\sqrt{\mu_0\varepsilon_0} = 1/c$ and equation (107),

$$
\begin{aligned}
&\frac{2a\pi f}{c}\sqrt{n_1^2 - n_2^2} = u \\
&\frac{2a\pi f}{c}\sqrt{n_1^2 - n_2^2} = n\pi \\
&\boxed{f = f_c = \frac{nc}{2a\sqrt{n_1^2 - n_2^2}}}
\end{aligned}
\tag{112}
$$

According to the acceptable values $1, 2, \ldots$ of n, these are the cutoff frequencies f_c of the TE even modes of the dielectric slab waveguide. Sometimes, it is more convenient to refer to the wavelengths λ_c. From equation (112), remembering that $\lambda = c/f$,

$$
\boxed{\lambda_c = \frac{2a\sqrt{n_1^2 - n_2^2}}{n}}
\tag{113}
$$

Equation (98) shows the relation between the propagation constant k_z and the angle of incidence θ_1. Moreover:

$$
\begin{aligned}
&k_z = n_1 k_0 \sin\theta_1 \\
&n_2 k_0 = n_1 k_0 \sin\theta_1 \\
&\sin\theta_1 = \frac{n_2}{n_1}
\end{aligned}
\tag{114}
$$

When the frequency of a sinusoidal signal inside the waveguide is such that a specific mode is at cutoff, the propagation constant of this mode k_z reaches its **lowest value** $k_2 = n_2 k_0$ (this was already pointed out in the observations after the inequality (91)): not only, but its angle of incidence θ_1 reaches its **lowest limit** θ_1^ℓ.

Therefore, also the *frequency* of the signal is related to the *angle* of incidence: even if apparently there should be no relation between these two quantities, they turn out to be strictly bound. If the frequency is forced to be at its cutoff value $f = f_c$, the angle of incidence is forced to be the limit angle $\theta_1 = \theta_1^\ell$. This is the **only acceptable value** for this angle, when the mode is in that condition.

This behaviour is imposed by the equations that have been used to characterize this physical system and its structure.

It has also been shown that, when the mode reaches the highest achievable confinement[15], k_z reaches the limit value of $k_1 = n_1 k_0$. Then,

$$\begin{aligned} k_z &= n_1 k_0 \sin\theta_1 \\ n_1 k_0 &= n_1 k_0 \sin\theta_1 \\ \sin\theta_1 &= 1 \\ \theta_1 &= \frac{\pi}{2} \end{aligned} \tag{115}$$

The angle of incidence is forced to assume its highest value, $\pi/2$ (if θ_1 exceeded $\pi/2$, the wave would not be any more inside the core).

In all the intermediate situations, when the frequency f of the sinusoidal signal is above the cutoff and the mode is active, the angle of incidence for a specific mode will be forced to assume a *single* value between θ_1^ℓ and $\pi/2$, depending on f:

$$\boxed{\theta_1^\ell \leq \theta_1 \leq \frac{\pi}{2}} \tag{116}$$

From another perspective, recalling Figure 3, a frequency value f of the sinusoidal signal will cause a *finite* number M of intersection between the $v^2 = u^2 + w^2$ circumference and the $u\tan(u)$ branches: this is the finite number of modes that are *active* (above cutoff) at f. Correspondingly, this frequency allows for the same, finite number M of *angles of incidence*, **one for each active mode**, and no more.

[15]Because the frequency of the sinusoidal signal inside the waveguide is high enough to reach this condition in one of the $u\tan(u)$ branches, which corresponds to the specific mode.