Introduction

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

- It is a valid Maxwell's equations solution in the three-dimensional space[1], and it satisfies the boundary conditions imposed by the dielectric slab structure.

- Being a wave inside the *core*, it impinges on each of the interfaces towards the *cover* with an incidence angle equal to or greater than the limit-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 $|e^{-jk_z z}| = 1$, $\forall z \in \mathbb{R}$.

Oblique incidence of a plane wave on a surface

An Electro-magnetic wave is generally able to cross the surface between two different dielectric materials. Several phenomena are related to this event: reflection, transmission, refraction. Their analysis shows how and when a couple of different dielectric materials constitutes a waveguide.

The most trivial example of a wave which encounters the surface separating two different dielectric materials involves a plane wave and its wavevector $\mathbf{k}$, when it is perfectly orthogonal to the surface (normal incidence).

A more complex – but more general – case occurs when the wavevector $\mathbf{k}$ has a different angle and its incidence is therefore not normal to the surface. Let *all* the three-dimensional half space $z < 0$ be filled with a dielectric medium whose permittivity is ε_1; let the remaining space $z \geq 0$ be filled

[1] Unlike the metallic waveguides, this solution must be valid not only inside the guide, but also outside. Dielectric waveguides are *open* waveguides: there is no hard shield between their inside and their outside.

with a medium whose permittivity is $\varepsilon_2 < \varepsilon_1$. The plane separating the two dielectric media is $z = 0$.

The plane containing the vector normal to the surface $z = 0$ and the wavevector $\mathbf{k}$ is called *incidence plane*. In this particular problem, the incidence plane is itself orthogonal to the separation plane $z = 0$. With the chosen system of coordinates, shown in Figure 1, this coincides with the (x, z) plane.

The Electric field can lie on this plane (parallel polarization), or it can be orthogonal to this plane (perpendicular polarization), or it can be written as a linear combination of these two cases. All the essential observations can be made already with the first case only, and it will be the only one shown here. The following approach is the same as in David M. Pozar. *Microwave engineering*. Wiley, Hoboken (NJ), 3rd edition, 2005 (where also the simpler *normal incidence* case and the perpendicular polarization case are shown).

An Electro-magnetic field $\mathbf{E}_i, \mathbf{H}_i$ is directed from medium 1 towards the separation surface $z = 0$ and it constitutes a *plane wave*: this implies that the wavefronts are *planes* of *infinite extension*. A finite portion of them is depicted in Figure 2. Wavevector $\mathbf{k}_1$ is always normal to these planes. It should be recalled in fact that, when dealing with plane waves, the triad of vectors $\mathbf{E}_i, \mathbf{H}_i, \mathbf{k}_1$ follow the right-hand rule: they are such that $\mathbf{E}_i \times \mathbf{H}_i$ is directed along $\mathbf{k}_1$.

The same considerations apply for the field $\mathbf{E}_t, \mathbf{H}_t$ in medium 2 with wavevector $\mathbf{k}_2$, and for the reflected field $\mathbf{E}_r, \mathbf{H}_r$ in medium 1, whose wavevector has the same magnitude, but a different direction, with respect to the initial wavevector $\mathbf{k}_1$.

In this problem, the incident Electro-magnetic field $\mathbf{E}_i, \mathbf{H}_i$ is a known quantity, as well as its wave vector $\mathbf{k}_1$ and so also the angle of incidence θ_i. This field in Figure 1 is evaluated at a general position $\mathbf{r}$, along its path towards the interface plane.

The vector $\mathbf{r}$ is used to identify a specific point in space with (in this case) coordinates (x, y, z):

$$\mathbf{r} = x\mathbf{u}_x + y\mathbf{u}_y + z\mathbf{u}_z \tag{1}$$

where $\mathbf{u}_x, \mathbf{u}_y, \mathbf{u}_z$ are the unit vectors along the three-dimensional axes directions. To determine the dependencies of the propagators of the wave with respect to the variables x, y, z:

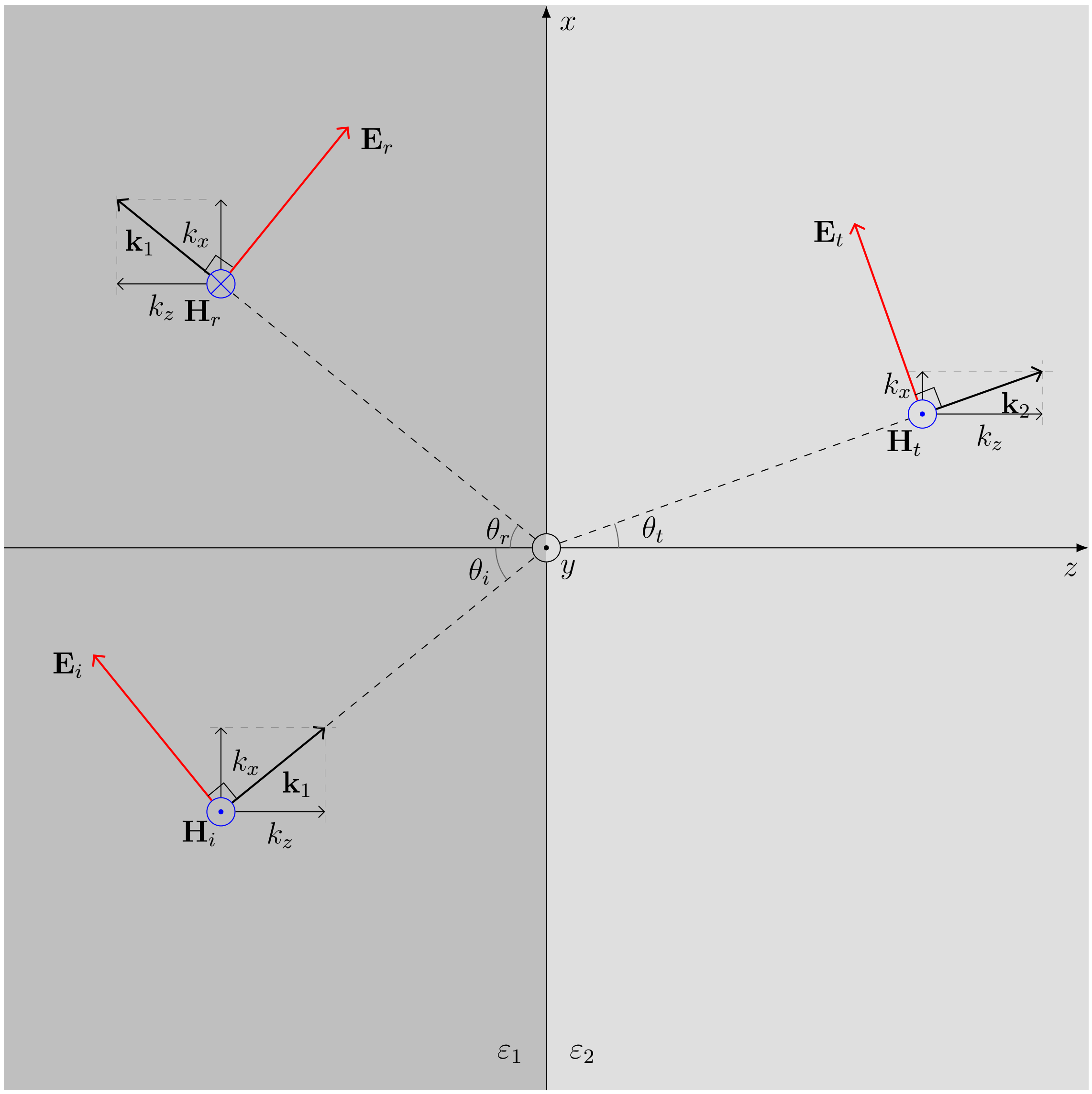


Figure 1: Incidence of a plane wave on the surface between two different dielectric media.

$$
\begin{aligned}
\mathbf{k}_1 \cdot \mathbf{r} &= k_x^{(1)} x + k_y^{(1)} y + k_z^{(1)} z \\
\mathbf{k}_2 \cdot \mathbf{r} &= k_x^{(2)} x + k_y^{(2)} y + k_z^{(2)} z
\end{aligned} \tag{2}
$$

Being (x, z) the incidence plane, the wave does not propagate along y: the related components $k_y^{(1)}$ and $k_y^{(2)}$ of the wavevectors $\mathbf{k}_1$ and $\mathbf{k}_2$ are 0. Observing Figure 1 and the value of the angle of incidence θ_i, upon geometric considerations the following identities can be derived observing the right triangles involving $\mathbf{k}_1$ and its components $k_x^{(1)}, k_z^{(1)}$:

$$
\begin{aligned}
k_x^{(1)} &= k_1 \sin\theta_i \\
k_z^{(1)} &= k_1 \cos\theta_i
\end{aligned} \tag{3}
$$

where $k_1 = |\mathbf{k}_1|$. So, the first equation in (2) becomes:

$$
\mathbf{k}_1 \cdot \mathbf{r} = k_1 x \sin\theta_i + k_1 z \cos\theta_i \tag{4}
$$

The same computations can be applied to the wavevector in the medium 2:

$$
\begin{aligned}
k_x^{(2)} &= k_2 \sin\theta_t \\
k_z^{(2)} &= k_2 \cos\theta_t
\end{aligned} \tag{5}
$$

with $k_2 = |\mathbf{k}_2|$ and the second equation in (2) is:

$$
\mathbf{k}_2 \cdot \mathbf{r} = k_2 x \sin\theta_t + k_2 z \cos\theta_t \tag{6}
$$

Evaluated in position $\mathbf{r}$, the incident Electric and Magnetic fields are:

$$
\begin{aligned}
\mathbf{E}_i(\mathbf{r}) &= E_0 \cos\theta_i e^{-jk_1 x \sin\theta_i} e^{-jk_1 z \cos\theta_i} \mathbf{u}_x - E_0 \sin\theta_i e^{-jk_1 x \sin\theta_i} e^{-jk_1 z \cos\theta_i} \mathbf{u}_z \\
\mathbf{H}_i(\mathbf{r}) &= \frac{E_0}{\eta_1} e^{-jk_1 x \sin\theta_i} e^{-jk_1 z \cos\theta_i} \mathbf{u}_y
\end{aligned} \tag{7}
$$

The interface represents a discontinuity in the path of the incident field $\mathbf{E}_i, \mathbf{H}_i$. As already pointed out, these fields are known, as well as the angle of incidence θ_i.

In the most general case, it may happen that *a fraction* of the incident field is reflected back in the medium 1 and *another fraction* of the incident field is able to go through the medium 2: moreover, both their paths can have angles (with respect to the z axis) which *differ* from θ_i.

Considering the Electric field only (then, the Magnetic field can be derived from its value), in general, if E_0 is the magnitude of the incident field, ΓE_0 conventionally represents the fraction of field which is reflected back to the medium 1; TE_0 represents the portion of field that crosses the interface. Both the reflection and the transmission *can* lead to phase shifts, so Γ and T are in general complex quantities.

Remembering that $\mathbf{E}_r, \mathbf{H}_r$ propagates backwards in the medium 1, and that they must still follow the right-hand rule together with the new wavevector $\mathbf{k}_1$, the reflected Electro-magnetic field can be written as:

$$\begin{aligned}
\mathbf{E}_r(\mathbf{r}) &= E_0\Gamma\cos\theta_r e^{-jk_1x\sin\theta_r}e^{-jk_1z\cos\theta_r}\mathbf{u}_x + E_0\Gamma\sin\theta_r e^{-jk_1x\sin\theta_r}e^{-jk_1z\cos\theta_r}\mathbf{u}_z \\
\mathbf{H}_r(\mathbf{r}) &= -\frac{E_0\Gamma}{\eta_1}e^{-jk_1x\sin\theta_r}e^{-jk_1z\cos\theta_r}\mathbf{u}_y
\end{aligned} \tag{8}$$

Unlike the incident wave, the $\mathbf{E}$ field has a positive x component and $\mathbf{H}$ is entering the paper in Figure 1. The orientation of $\mathbf{E}_r, \mathbf{H}_r$ and $\mathbf{k}_1$ is more explicitly shown is Figures 2, 3 and 4.

The portion of Electro-magnetic field which can cross the interface and propagate inside the medium 2 can be expressed as:

$$\begin{aligned}
\mathbf{E}_t(\mathbf{r}) &= E_0T\cos\theta_t e^{-jk_2x\sin\theta_t}e^{-jk_2z\cos\theta_t}\mathbf{u}_x - E_0T\sin\theta_t e^{-jk_2x\sin\theta_t}e^{-jk_2z\cos\theta_t}\mathbf{u}_z \\
\mathbf{H}_t(\mathbf{r}) &= \frac{E_0T}{\eta_2}e^{-jk_2x\sin\theta_t}e^{-jk_2z\cos\theta_t}\mathbf{u}_y
\end{aligned} \tag{9}$$

The magnitude of the wavevector $\mathbf{k}_1$ depends on the medium 1 and on the frequency ω of the field:

$$|\mathbf{k}_1| = k_1 = \omega\sqrt{\mu_1\varepsilon_1} = \omega n_1\sqrt{\mu_0\varepsilon_0} \tag{10}$$

Where it has been assumed that $n_1 = \sqrt{\varepsilon_{r_1}}$ is the refractive index of medium 1 and $\mu_1 = \mu_0$ is the same as vacuum. Also, the wave impedance in medium 1 is:

$$\eta_1 = \sqrt{\frac{\mu_0}{\varepsilon_1}} = \frac{1}{n_1}\eta_0 \tag{11}$$

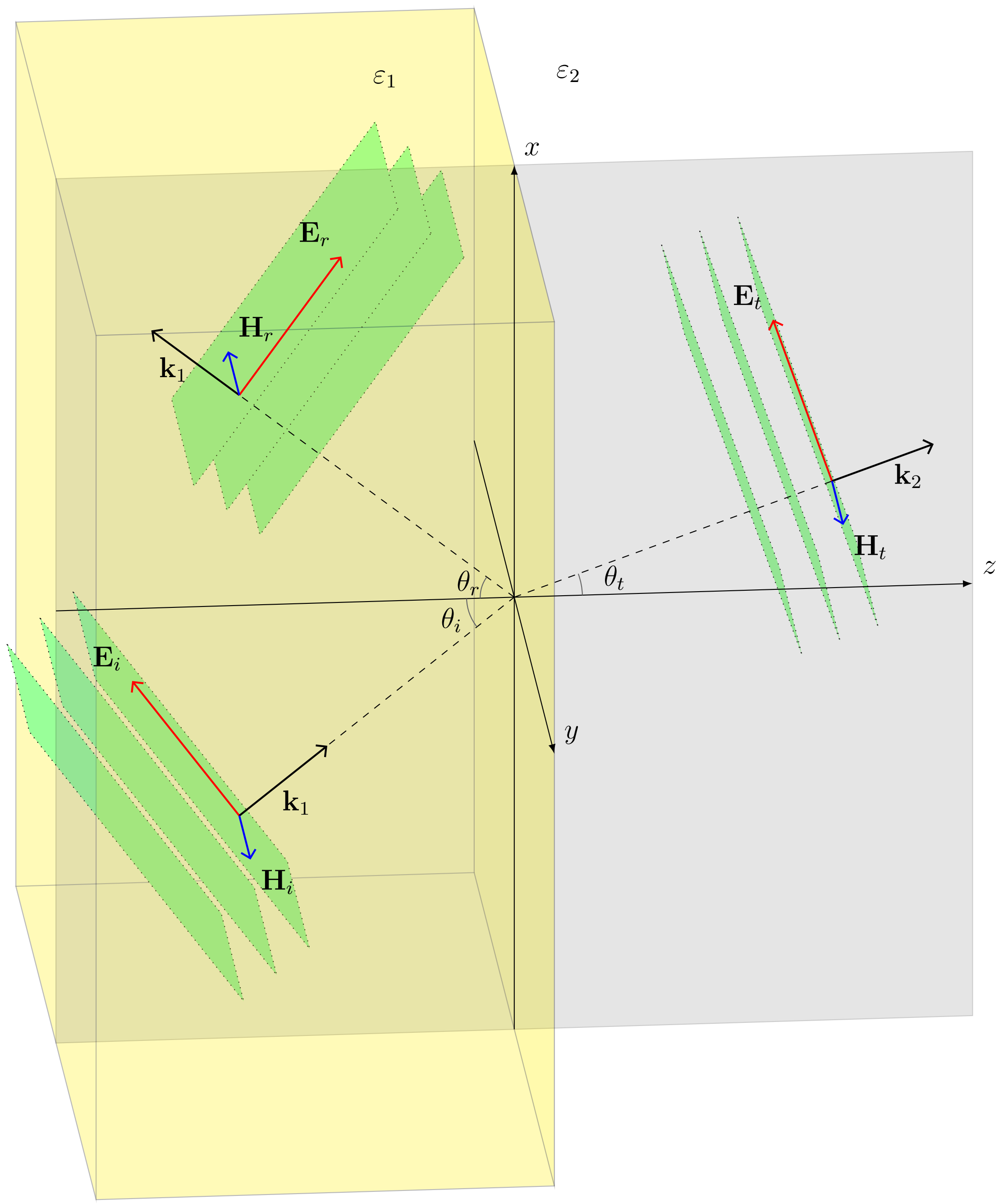


Figure 2: Incidence of a plane wave on the surface between two different dielectric media, with the perspective of Figure 1, in three-dimensional space. The (x, z) plane, which is the plane of incidence, has been highlighted in gray. The wavefronts of plane waves are depicted as green planes. The half-space $z < 0$, which hosts the medium 1, is filled with a light yellow color.

As regards medium 2:

$$|\mathbf{k}_2| = k_2 = \omega\sqrt{\mu_2\varepsilon_2} = \omega n_2\sqrt{\mu_0\varepsilon_0} \tag{12}$$

assuming $n_2 = \sqrt{\varepsilon_{r_2}}$ and $\mu_2 = \mu_0$; also

$$\eta_2 = \sqrt{\frac{\mu_0}{\varepsilon_2}} = \frac{1}{n_2}\eta_0 \tag{13}$$

Observing the field expressions (7), (8) and (9), and the quantities they involve, the unknown values are $\Gamma, T, \theta_r, \theta_t$. Some equations will be needed to determine them.

Across the interface between two dielectric materials, the tangential Electric and Magnetic field must be continuous.

The field x and y components are subjected to reflection, refraction *and* they must satisfy a continuity equation. The field z components are only subjected to reflection and refraction.

Only the x-component of the field $\mathbf{E}$ is tangent to the $z = 0$ plane; on the other hand, all the field $\mathbf{H}$ is tangent to the same plane, being always and only along y. Then, the following two continuity conditions must be verified:

$$\begin{aligned} &[\mathbf{E}_i(\mathbf{r})]_x + [\mathbf{E}_r(\mathbf{r})]_x = [\mathbf{E}_t(\mathbf{r})]_x \\ &\mathbf{H}_i(\mathbf{r}) + \mathbf{H}_r(\mathbf{r}) = \mathbf{H}_t(\mathbf{r}) \end{aligned} \tag{14}$$

for all the points $\mathbf{r}$ belonging to the (x, y) plane, which has equation $z = 0$.

Summing the x-components of the $\mathbf{E}$ fields in (7) and (8), and equalling the result to the x component in (9), remembering that $z = 0$:

$$\begin{aligned} &E_0\cos\theta_i e^{-jk_1x\sin\theta_i} + E_0\Gamma\cos\theta_r e^{-jk_1x\sin\theta_r} = E_0T\cos\theta_t e^{-jk_2x\sin\theta_t} \\ &\cos\theta_i e^{-jk_1x\sin\theta_i} + \Gamma\cos\theta_r e^{-jk_1x\sin\theta_r} = T\cos\theta_t e^{-jk_2x\sin\theta_t} \end{aligned} \tag{15}$$

Summing the Magnetic fields in (7) and (8), and equalling the result to the one in (9), remembering that $z = 0$:

$$\begin{aligned} &\frac{E_0}{\eta_1}e^{-jk_1x\sin\theta_i} - \frac{E_0\Gamma}{\eta_1}e^{-jk_1x\sin\theta_r} = \frac{E_0T}{\eta_2}e^{-jk_2x\sin\theta_t} \\ &\frac{1}{\eta_1}e^{-jk_1x\sin\theta_i} - \frac{\Gamma}{\eta_1}e^{-jk_1x\sin\theta_r} = \frac{T}{\eta_2}e^{-jk_2x\sin\theta_t} \end{aligned} \tag{16}$$

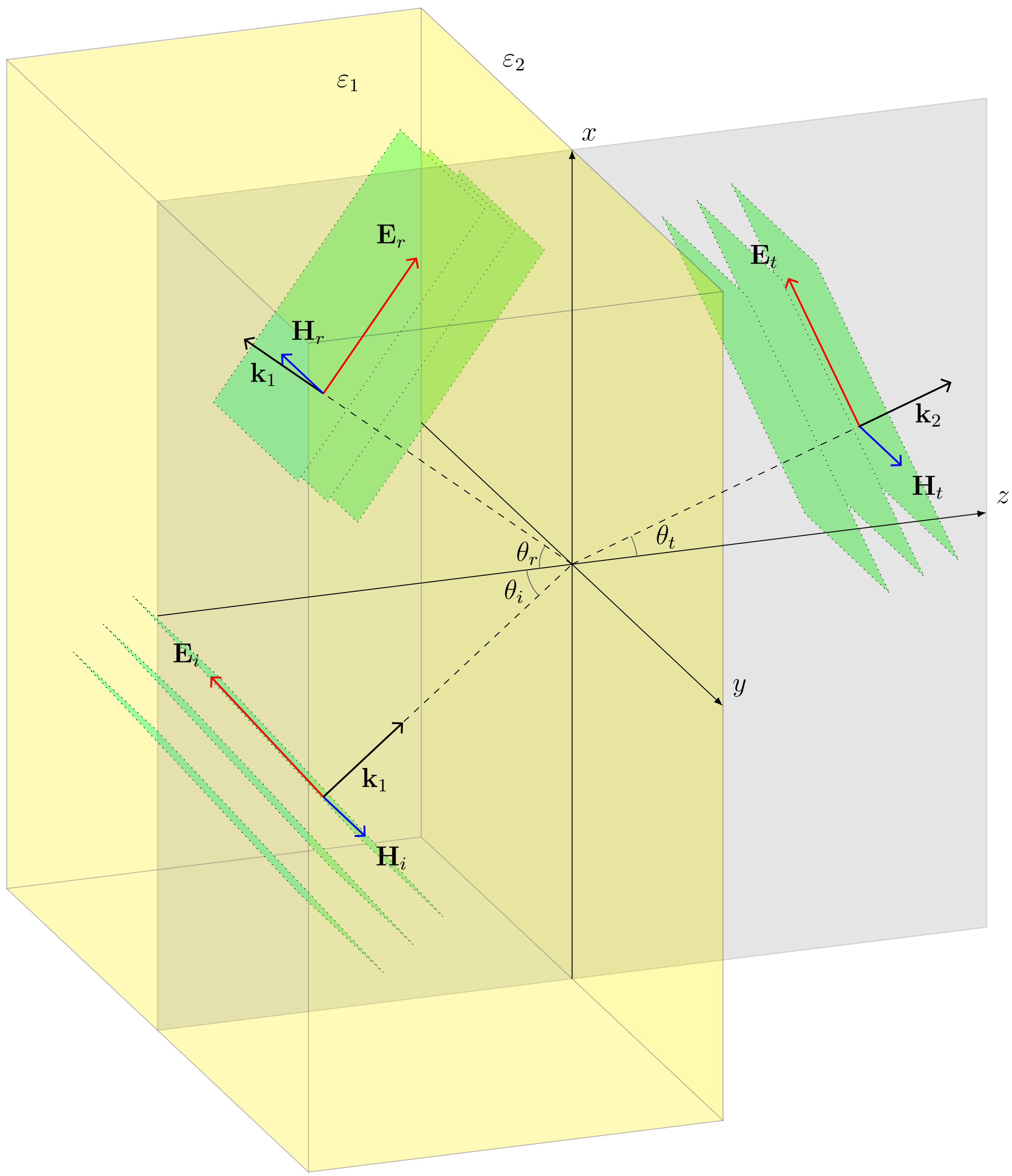


Figure 3: Incidence of a plane wave on the surface between two different dielectric media, in three-dimensional space, seen from medium 1 side and perspective.

In order to determine the values of the 4 unknowns $\Gamma, T, \theta_r, \theta_t$, more than two equations are required. However, (15) and (16) must hold at the same time for *every* real value of x: therefore, each of them is not a single equation, but rather a *system of infinite equations*, one for each value of x. Of course, only some of these equations will be significant and linearly independent.

Evaluating (15) and (16) as $x = 0$:

$$\cos\theta_i + \Gamma\cos\theta_r = T\cos\theta_t \tag{17}$$

$$\frac{1}{\eta_1} - \frac{\Gamma}{\eta_1} = \frac{T}{\eta_2} \tag{18}$$

Substituting (17) in the Right Hand Side of equation (15):

$$\begin{aligned}
\cos\theta_i e^{-jk_1 x\sin\theta_i} + \Gamma\cos\theta_r e^{-jk_1 x\sin\theta_r} &= (\cos\theta_i + \Gamma\cos\theta_r)\, e^{-jk_2 x\sin\theta_t} \\
\cos\theta_i e^{-jk_1 x\sin\theta_i} + \Gamma\cos\theta_r e^{-jk_1 x\sin\theta_r} &= \cos\theta_i e^{-jk_2 x\sin\theta_t} + \Gamma\cos\theta_r e^{-jk_2 x\sin\theta_t}
\end{aligned} \tag{19}$$

By side-by-side comparison, if (19) and (15) must be verified at the same time (the latter for $x = 0$, the former for any other $x \neq 0$),

$$\begin{cases} e^{-jk_1 x\sin\theta_i} = e^{-jk_2 x\sin\theta_t} \\ \\ e^{-jk_1 x\sin\theta_r} = e^{-jk_2 x\sin\theta_t} \end{cases} \Rightarrow \begin{cases} k_1\sin\theta_i = k_2\sin\theta_t \\ k_1\sin\theta_r = k_2\sin\theta_t \end{cases} \tag{20}$$

that is

$$\begin{cases} n_1\sin\theta_i = n_2\sin\theta_t \\ n_1\sin\theta_r = n_2\sin\theta_t \end{cases} \tag{21}$$

This leads to two conclusions:

1. The first equation in system (21) is **Snell's law** of refraction.

2. Observing that the equations in (21) have the same Right Hand Sides, it follows that

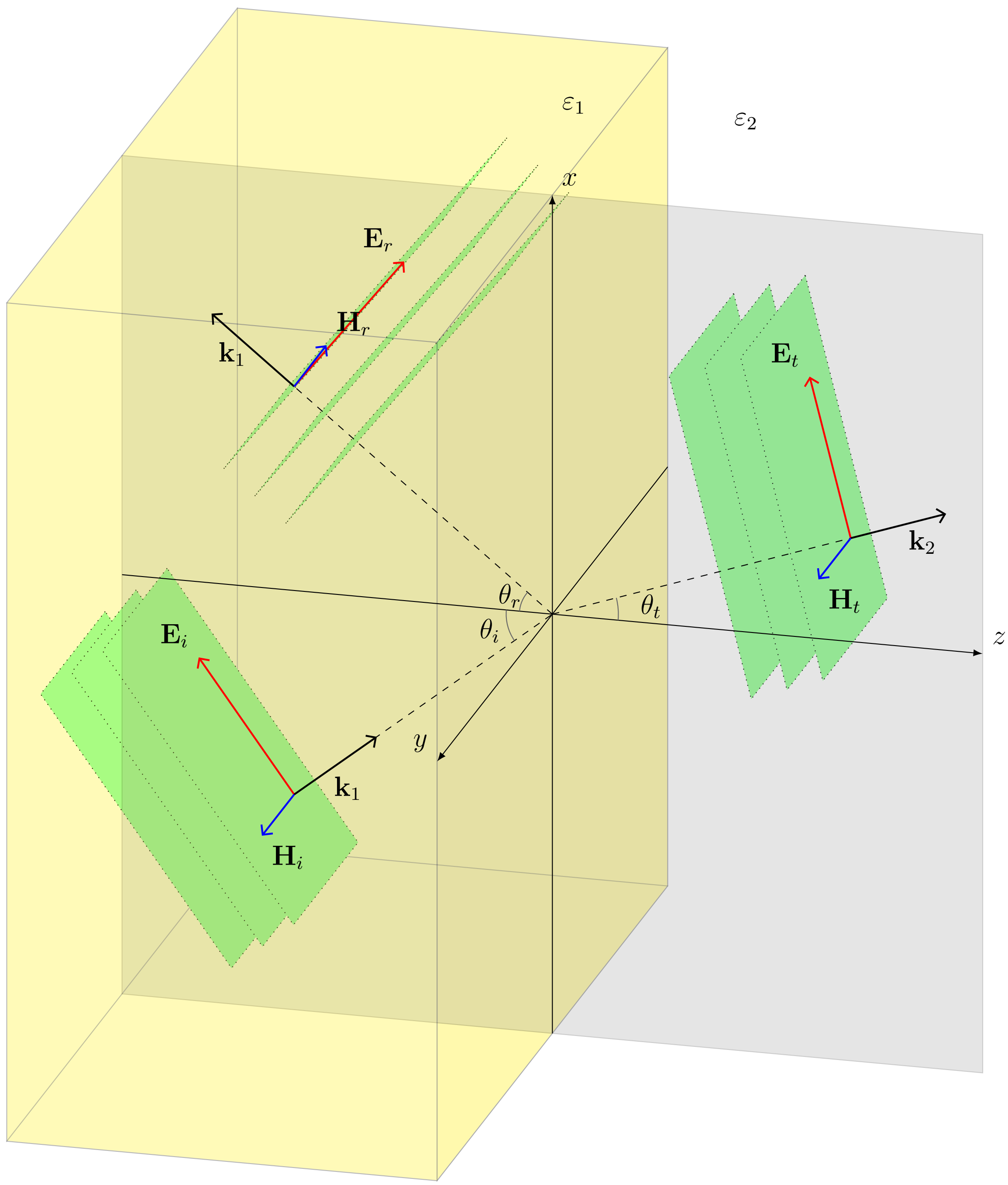


Figure 4: Incidence of a plane wave on the surface between two different dielectric media, in three-dimensional space, seen from medium 2 side and perspective.

$$n_1 \sin\theta_i = n_1 \sin\theta_r \rightarrow \theta_i = \theta_r \tag{22}$$

Being $0 \leq \theta_{i,r} \leq \pi/2$, the equality of the sines of the two angles also implies the equality of the angles themselves.

Therefore, the angle of the reflected wavevector[2] is **equal** to the angle of incidence.

These conditions together are called **Snell's laws of refraction and reflection**. They have been obtained by imposing that the tangent Electric and Magnetic fields do not vary while crossing the interface between the two media.

The condition (15) is not a single equation: it is a much harder constraint, involving *every* value of x. Its consequences determine the features of both the reflected and the transmitted wave across the surface.

The constraint of Snell's law is not only about the angles and refractive indexes. Multiplying by k_0 both sides of first equation in (21), the following equality is also obtained:

$$\begin{aligned} n_1 k_0 \sin\theta_i &= n_2 k_0 \sin\theta_t \\ k_1 \sin\theta_i &= k_2 \sin\theta_t \end{aligned} \tag{23}$$

A comparison with relations (3) and (5) shows that (23) imposes:

$$\begin{aligned} &k_x^{(1)} = k_1 \sin\theta_i = k_x^{(2)} = k_2 \sin\theta_t \\ &k_x^{(1)} = k_x^{(2)} \end{aligned} \tag{24}$$

Despite the crossing of the interface in $z = 0$ and despite the medium that hosts the wave changes from 1 to 2, the x component of the wavevector **does not vary**. The wave is forced to keep the same velocity along x regardless of the medium change. The presence of an interface and the subsequent reflections and/or refractions do not affect the longitudinal propagation of the wave. This is a not obvious and remarkable result. With the system of coordinates that will be used fo dielectric waveguides, the role of the current k_x will be represented by k_z: having this same value in the medium 1 and medium 2 leads to the equations that define a mode.

[2] With respect to the z axis (orthogonal to the interface plane $z = 0$).

Two out of the four unknowns have been determined so far: θ_r and θ_t.

Recalling (17) and (18), they constitute a system of two equations as regards the unknowns Γ and T:

$$\begin{cases} \dfrac{\cos\theta_i + \Gamma\cos\theta_r}{\cos\theta_t} = T \\ \dfrac{\eta_2}{\eta_1}(1-\Gamma) = T \end{cases} \tag{25}$$

Equating the Left Hand Sides of both the equations in (25):

$$\begin{aligned} &\frac{\cos\theta_i + \Gamma\cos\theta_r}{\cos\theta_t} = \frac{\eta_2}{\eta_1}(1-\Gamma) \\ &\frac{\eta_2}{\eta_1} - \frac{\eta_2}{\eta_1}\Gamma = \frac{\cos\theta_i}{\cos\theta_t} + \Gamma\frac{\cos\theta_r}{\cos\theta_t} \\ &\Gamma\left(\frac{\cos\theta_r}{\cos\theta_t} + \frac{\eta_2}{\eta_1}\right) = \frac{\eta_2}{\eta_1} - \frac{\cos\theta_i}{\cos\theta_t} \\ &\Gamma\left(\frac{\eta_1\cos\theta_r + \eta_2\cos\theta_t}{\eta_1\cos\theta_t}\right) = \frac{\eta_2\cos\theta_t - \eta_1\cos\theta_i}{\eta_1\cos\theta_t} \\ &\Gamma = \frac{\eta_2\cos\theta_t - \eta_1\cos\theta_i}{\eta_1\cos\theta_r + \eta_2\cos\theta_t} \end{aligned} \tag{26}$$

Observing that $\theta_r = \theta_i$:

$$\boxed{\Gamma = \frac{\eta_2\cos\theta_t - \eta_1\cos\theta_i}{\eta_2\cos\theta_t + \eta_1\cos\theta_i}} \tag{27}$$

Substituting back the value (26) of Γ in the second equation of system (25):

$$\begin{aligned} &\frac{\eta_2}{\eta_1}\left(1 - \frac{\eta_2\cos\theta_t - \eta_1\cos\theta_i}{\eta_2\cos\theta_t + \eta_1\cos\theta_i}\right) = T \\ &\frac{\eta_2}{\eta_1}\left(\frac{\eta_2\cos\theta_t + \eta_1\cos\theta_i - \eta_2\cos\theta_t + \eta_1\cos\theta_i}{\eta_2\cos\theta_t + \eta_1\cos\theta_i}\right) = T \end{aligned} \tag{28}$$

$$\boxed{T = \frac{2\eta_2\cos\theta_i}{\eta_2\cos\theta_t + \eta_1\cos\theta_i}} \tag{29}$$

Note that, in general, Γ and/or T can be complex, due to phase shifts which may be introduced during the reflection and refraction processes. Expressions (27) and (29) turn out to be a *generalization* of the already mentioned elementary case of normal incidence. The relations for that case can be re-obtained by equalling θ_i, θ_r and θ_t to zero (27) and (29):

$$\begin{aligned} \Gamma &= \frac{\eta_2 - \eta_1}{\eta_2 + \eta_1} \\ T &= \frac{2\eta_2}{\eta_2 + \eta_1} \end{aligned} \tag{30}$$

Also, the relation

$$1 + \Gamma = T \tag{31}$$

holds. When instead $\theta_i \neq 0$, the more general (27) and (29) must be used instead of the above relations (30): moreover, (18) must replace (31).

All the previously mentioned unknowns $\Gamma, T, \theta_r, \theta_t$ have been determined. It is now useful to consider some extreme cases.

The reflection coefficient Γ vanishes if the numerator of (27) is 0. If this condition can occur, the correspondent value of the incidence angle θ_i must be determined. First, the numerator should be rewritten only in terms of θ_i, as it is the only desired variable. Recalling Snell's law, first equation in (21):

$$\begin{aligned} & n_1 \sin\theta_i = n_2 \sin\theta_t \\ & \sin\theta_t = \frac{n_1}{n_2} \sin\theta_i \\ & \cos\theta_t = \sqrt{1 - \sin^2\theta_t} = \sqrt{1 - \frac{n_1^2}{n_2^2}\sin^2\theta_i} \end{aligned} \tag{32}$$

Rewriting the numerator of (27) equal to 0, substituting (32) and remembering that $\cos\alpha = \sqrt{1 - \sin^2\alpha}$:

$$
\begin{aligned}
&\eta_2 \cos\theta_t - \eta_1 \cos\theta_i = 0 \\
&\sqrt{\frac{\mu_0}{\varepsilon_0}} \frac{1}{\sqrt{\varepsilon_{r_2}}} \cos\theta_t - \sqrt{\frac{\mu_0}{\varepsilon_0}} \frac{1}{\sqrt{\varepsilon_{r_1}}} \cos\theta_i \\
&\frac{1}{\sqrt{\varepsilon_{r_2}}} \cos\theta_t - \frac{1}{\sqrt{\varepsilon_{r_1}}} \cos\theta_i = 0
\end{aligned}
\tag{33}
$$

$$
\begin{aligned}
&\frac{1}{n_2} \cos\theta_t - \frac{1}{n_1} \cos\theta_i = 0 \\
&n_1 \cos\theta_t - n_2 \cos\theta_i = 0
\end{aligned}
\tag{34}
$$

$$
\begin{aligned}
&n_1 \sqrt{1 - \frac{n_1^2}{n_2^2} \sin^2\theta_i} = n_2 \sqrt{1 - \sin^2\theta_i} \\
&n_1^2 \left(1 - \frac{n_1^2}{n_2^2} \sin^2\theta_i\right) = n_2^2 (1 - \sin^2\theta_i) \\
&n_1^2 - n_1^2 \frac{n_1^2}{n_2^2} \sin^2\theta_i = n_2^2 - n_2^2 \sin^2\theta_i
\end{aligned}
\tag{35}
$$

The $\sin\theta_i$ term must now be isolated:

$$
\begin{aligned}
n_1^2 - n_2^2 &= \left(n_1^2 \frac{n_1^2}{n_2^2} - n_2^2\right) \sin^2\theta_i \\
(n_1 + n_2)(n_1 - n_2) &= \left(\frac{n_1^4 - n_2^4}{n_2^2}\right) \sin^2\theta_i
\end{aligned}
\tag{36}
$$

$$
\begin{aligned}
(n_1 + n_2)(n_1 - n_2) &= \frac{(n_1^2 - n_2^2)(n_1^2 + n_2^2)}{n_2^2} \sin^2\theta_i \\
(n_1 + n_2)(n_1 - n_2) &= \frac{(n_1 + n_2)(n_1 - n_2)(n_1^2 + n_2^2)}{n_2^2} \sin^2\theta_i
\end{aligned}
\tag{37}
$$

$$
\begin{aligned}
1 &= \left(\frac{n_1^2 + n_2^2}{n_2^2}\right) \sin^2\theta_i \\
\frac{n_2^2}{n_1^2 + n_2^2} &= \sin^2\theta_i
\end{aligned}
\tag{38}
$$

The involved values of θ_i are between 0 and $\pi/2$, so the sine is always positive and its square root can be taken without ambiguity. The final result can be provided in several alternative ways:

$$\sin\theta_i = \sqrt{\frac{n_2^2}{n_1^2 + n_2^2}} \tag{39}$$

Multiplying numerator and denominator inside the square root by ε_0:

$$\begin{aligned}
&\sin\theta_i = \sqrt{\frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2}} \\
&\theta_i = \arcsin\left(\sqrt{\frac{\varepsilon_2}{\varepsilon_1 + \varepsilon_2}}\right) \\
&\theta_i = \arcsin\left(\frac{1}{\sqrt{\dfrac{\varepsilon_1 + \varepsilon_2}{\varepsilon_2}}}\right) = \arcsin\left(\frac{1}{\sqrt{1 + \dfrac{\varepsilon_1}{\varepsilon_2}}}\right) = \theta_B
\end{aligned} \tag{40}$$

As specified in the beginning, here the whole Electric field is contained in the incidence plane (parallel polarization, or TM wave). In this problem, a value for the angle θ_i exists such that the reflection coefficient vanishes: a wave impinging on the interface with this angle will experience *total transmission* from medium 1 to medium 2 and no reflection. This specific value of θ_i is called **Brewster's angle** and it is usually referred to as θ_B. Its value depends on the ratio $\varepsilon_1/\varepsilon_2$ between the dielectric constants of the two media.

Brewster's angle is a significant value, which causes also other features to be observed in the wave. However, they will not be considered here, as they are not related to the dielectric waveguides.

When the Electric field is orthogonal to the incidence plane (perpendicular polarization, or TE wave), it can be verified that there is *no value* for the angle θ_i that vanishes the reflection coefficient, so in that case Brewster's angle does not exist.

Dielectric slab waveguides

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 waveguides

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 permittivity $\varepsilon = \varepsilon(x)$ is not a constant through the *vertical* direction x. Its variation shows how the two types of materials fill the whole three-dimensional space:

$$\varepsilon(x) = \begin{cases} \varepsilon_2, & x > a \\ \varepsilon_1, & a \leq x \leq -a \\ \varepsilon_2, & x < -a \end{cases} \quad = \quad \begin{cases} \varepsilon_1, & |x| \leq a \\ \varepsilon_2, & |x| > a \end{cases} \tag{41}$$

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

$$n(x) = \begin{cases} n_1, & |x| \leq a \\ n_2, & |x| > a \end{cases} \tag{42}$$

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 $n_2 = 1$.

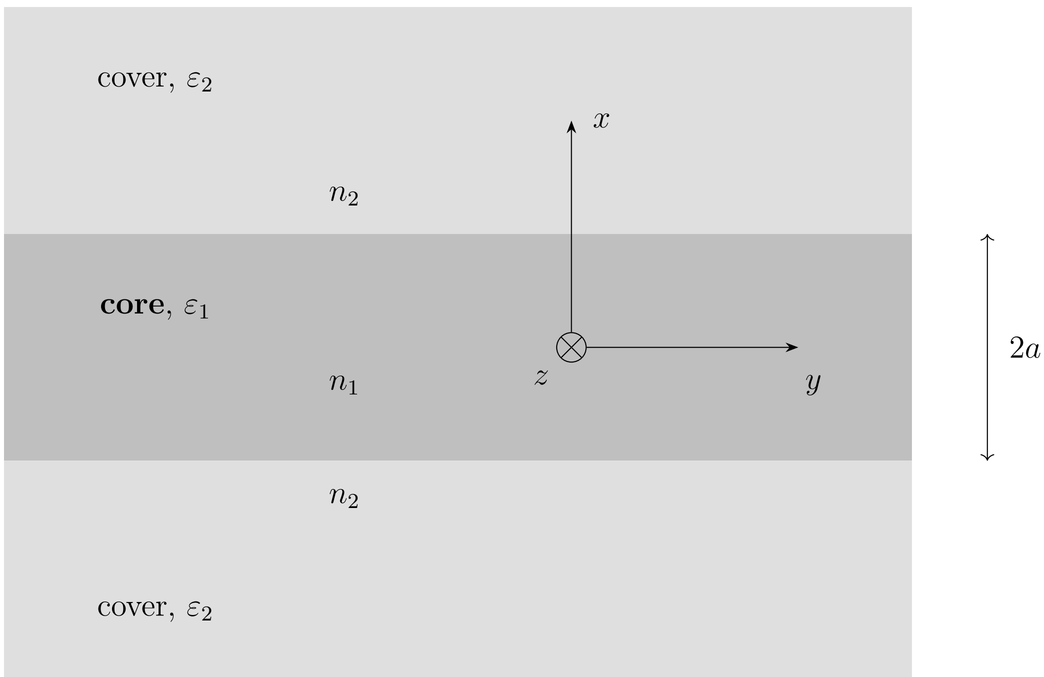


Figure 5: Structure of a symmetrical dielectric slab.

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[3]. 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[4]. The choice (b) will allow for the propagation of a signal[5].

The following hypotheses 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[6].

5. Both the core and the cover media are linear, isotropic[7] and homogeneous[8] media.

---

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

[4]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.

[5]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.

[6]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.

[7]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.

[8]Composed by a uniform material.

The reference Maxwell equations to be solved are then:

$$\begin{cases} \nabla \cdot \mathbf{D} = 0 \\ \nabla \cdot \mathbf{B} = 0 \\ \nabla \times \mathbf{E} = -j\omega\mu\mathbf{H} \\ \nabla \times \mathbf{H} = j\omega\varepsilon\mathbf{E} \end{cases} \tag{43}$$

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

$$\begin{aligned} &\nabla \cdot (\varepsilon\mathbf{E}) = 0 \\ &\varepsilon\nabla \cdot \mathbf{E} = 0 \\ &\nabla \cdot \mathbf{E} = 0 \end{aligned} \tag{44}$$

being always $\varepsilon \neq 0$[9]. Last equation in (44) lets some useful expressions to be obtained, expanding the ∇ operator:

$$\frac{\partial}{\partial x}E_x(x,z) + \frac{\partial}{\partial y}E_y(x,z) + \frac{\partial}{\partial z}E_z(x,z) = 0 \tag{45}$$

Remembering the hypotheses 1 and 2,

$$\begin{aligned} &\frac{\partial}{\partial y}E_y(x,z) = 0 \\ &\frac{\partial}{\partial z} = -jk_z \end{aligned} \tag{46}$$

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, (45) becomes

$$\begin{aligned} &\frac{\partial}{\partial x}E_x(x,z) - jk_z E_z(x,z) = 0 \\ &\frac{\partial}{\partial x}E_x(x,z) = jk_z E_z(x,z) \end{aligned} \tag{47}$$

Similarly, from the 2nd Maxwell equation[10]:

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

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

$$
\begin{aligned}
&\nabla \cdot (\mu \mathbf{H}) = 0 \\
&\nabla \cdot \mathbf{H} = 0 \\
&\frac{\partial}{\partial x} H_x(x,z) = j k_z H_z(x,z)
\end{aligned} \tag{48}
$$

The 3rd Maxwell equations expands as follows:

$$
\begin{aligned}
&\nabla \times \mathbf{E} = \det \begin{vmatrix} \mathbf{u}_x & \mathbf{u}_y & \mathbf{u}_z \\ \dfrac{\partial}{\partial x} & \dfrac{\partial}{\partial y} & \dfrac{\partial}{\partial z} \\ E_x(x,z) & E_y(x,z) & E_z(x,z) \end{vmatrix} = \\
&= \mathbf{u}_x \left(\frac{\partial}{\partial y} E_z(x,z) - \frac{\partial}{\partial z} E_y(x,z) \right) + \mathbf{u}_y \left(\frac{\partial}{\partial z} E_x(x,z) - \frac{\partial}{\partial x} E_z(x,z) \right) + \\
&+ \mathbf{u}_z \left(\frac{\partial}{\partial x} E_y(x,z) - \frac{\partial}{\partial y} E_x(x,z) \right) = -j\omega\mu\mathbf{H}
\end{aligned} \tag{49}
$$

Remembering that all the members with a partial derivatives with respect to y are zero, equation (49) becomes:

$$
\begin{aligned}
&-\mathbf{u}_x \frac{\partial}{\partial z} E_y(x,z) + \mathbf{u}_y \left(\frac{\partial}{\partial z} E_x(x,z) - \frac{\partial}{\partial x} E_z(x,z) \right) + \mathbf{u}_z \frac{\partial}{\partial x} E_y(x,z) = \\
&= -j\omega\mu\mathbf{H}
\end{aligned} \tag{50}
$$

Being (50) 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:

$$
\begin{aligned}
-\frac{\partial}{\partial z}E_y(x,z) &= -j\omega\mu H_x(x,z) \\
\frac{\partial}{\partial z}E_y(x,z) &= j\omega\mu H_x(x,z) \\
-jk_z E_y(x,z) &= j\omega\mu H_x(x,z) \\
&\boxed{k_z\, E_y(x,z) = -\omega\mu\, H_x(x,z)}
\end{aligned}
\tag{51}
$$

$$
\begin{aligned}
\frac{\partial}{\partial z}E_x(x,z) - \frac{\partial}{\partial x}E_z(x,z) &= -j\omega\mu H_y(x,z) \\
-jk_z E_x(x,z) - \frac{\partial}{\partial x}E_z(x,z) &= -j\omega\mu H_y(x,z) \\
&\boxed{jk_z\, E_x(x,z) + \frac{\partial}{\partial x}\, E_z(x,z) = j\omega\mu\, H_y(x,y)}
\end{aligned}
\tag{52}
$$

$$
\boxed{\frac{\partial}{\partial x}\, E_y(x,z) = -j\omega\mu\, H_z(x,z)}
\tag{53}
$$

Applying the same procedure for the 4th Maxwell equation,

$$
\begin{aligned}
&\nabla \times \mathbf{H} = \det \begin{vmatrix} \mathbf{u}_x & \mathbf{u}_y & \mathbf{u}_z \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ H_x(x,z) & H_y(x,z) & H_z(x,z) \end{vmatrix} = \\
&= \mathbf{u}_x \left(\frac{\partial}{\partial y}H_z(x,z) - \frac{\partial}{\partial z}H_y(x,z) \right) + \mathbf{u}_y \left(\frac{\partial}{\partial z}H_x(x,z) - \frac{\partial}{\partial x}H_z(x,z) \right) + \\
&+ \mathbf{u}_z \left(\frac{\partial}{\partial x}H_y(x,z) - \frac{\partial}{\partial y}H_x(x,z) \right) = j\omega\varepsilon\mathbf{E}
\end{aligned}
\tag{54}
$$

$$
\boxed{k_z\, H_y(x,z) = \omega\varepsilon\, E_x(x,z)}
\tag{55}
$$

$$\boxed{jk_z H_x(x,z) + \frac{\partial}{\partial x} H_z(x,z) = -j\omega\varepsilon E_y(x,y)} \tag{56}$$

$$\boxed{\frac{\partial}{\partial x} H_y(x,z) = j\omega\varepsilon E_z(x,z)} \tag{57}$$

The relations (51), (52), (53), (55), (56), (57) 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 (51), (53) and (56) and it involves *only* the field components

$$E_y(x,z),\ H_x(x,z),\ H_z(x,z) \tag{58}$$

- Equations (52), (55) and (57) represent another set, with the field components

$$H_y(x,z),\ E_x(x,z),\ E_z(x,z) \tag{59}$$

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:

$$\begin{array}{l} \nabla \times \nabla \times \mathbf{E} = -j\omega\mu\nabla \times \mathbf{H} \\ \nabla \times \nabla \times \mathbf{H} = j\omega\varepsilon\nabla \times \mathbf{E} \end{array} \tag{60}$$

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

$$\nabla \times \nabla \times \mathbf{A} = \nabla\left(\nabla \cdot \mathbf{A}\right) - \nabla^2\mathbf{A} \tag{61}$$

Recalling that $\nabla \cdot \mathbf{E} = 0$ and $\nabla \cdot \mathbf{H} = 0$, equations (60) become:

$$\begin{aligned}
&\nabla \times \nabla \times \mathbf{E} = -\nabla^2 \mathbf{E} = -j\omega\mu \nabla \times \mathbf{H} \\
&\nabla \times \nabla \times \mathbf{H} = -\nabla^2 \mathbf{H} = j\omega\varepsilon \nabla \times \mathbf{E} \\
&\nabla^2 \mathbf{E} = j\omega\mu \nabla \times \mathbf{H} \\
&\nabla^2 \mathbf{H} = -j\omega\varepsilon \nabla \times \mathbf{E}
\end{aligned} \tag{62}$$

$\nabla \times \mathbf{E}$ 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:

$$\begin{aligned}
&\nabla^2 \mathbf{E} = j\omega\mu \left(j\omega\varepsilon \mathbf{E} \right) \\
&\nabla^2 \mathbf{H} = -j\omega\varepsilon \left(-j\omega\mu \mathbf{H} \right) \\
&\nabla^2 \mathbf{E} = -\omega^2 \mu\varepsilon \mathbf{E} \\
&\nabla^2 \mathbf{H} = -\omega^2 \mu\varepsilon \mathbf{H}
\end{aligned} \tag{63}$$

$$\begin{aligned}
&\nabla^2 \mathbf{E} + \omega^2 \mu\varepsilon \mathbf{E} = 0 \\
&\nabla^2 \mathbf{H} + \omega^2 \mu\varepsilon \mathbf{H} = 0
\end{aligned} \tag{64}$$

Two Helmholtz 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 5.

Equations (64) are actually a synthesis of 6 scalar equations, one for each component of the Electric and Magnetic fields:

$$\begin{aligned}
&\nabla^2 E_x(x,z) + \omega^2 \mu\varepsilon E_x(x,z) = 0 \\
&\nabla^2 E_y(x,z) + \omega^2 \mu\varepsilon E_y(x,z) = 0 \\
&\nabla^2 E_z(x,z) + \omega^2 \mu\varepsilon E_z(x,z) = 0 \\
&\nabla^2 H_x(x,z) + \omega^2 \mu\varepsilon H_x(x,z) = 0 \\
&\nabla^2 H_y(x,z) + \omega^2 \mu\varepsilon H_y(x,z) = 0 \\
&\nabla^2 H_z(x,z) + \omega^2 \mu\varepsilon H_z(x,z) = 0
\end{aligned} \tag{65}$$

Note that, as regards the path followed from (60) to (64), only the last two initial hypotheses have been actually applied (equations (65), where the functions are explicitly not depending on y, also consider the uniformity along y). The fact that *all* the field components satisfy a wave equation is therefore a general result, which applies not only to dielectric waveguides, but to *any* source-free, linear, isotropic, homogeneous medium.

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 (65) is not a zero identity:

$$\begin{aligned}
&\nabla^2 E_y(x,z) + \omega^2 \mu\varepsilon E_y(x,z) = 0 \\
&\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}\right) E_y(x,z) + \omega^2 \mu\varepsilon E_y(x,z) = 0 \\
&\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right) E_y(x,z) + \omega^2 \mu\varepsilon E_y(x,z) = 0
\end{aligned} \tag{66}$$

Remembering hypothesis 2,

$$E_y(x,z) = e(x)f(z) = e(x)e^{-jk_z z} \tag{67}$$

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: $g(x,y,z)$.

$$\begin{aligned}
&\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right) e(x)e^{-jk_z z} + \omega^2 \mu\varepsilon e(x)e^{-jk_z z} = 0 \\
&\left[\frac{\partial^2}{\partial x^2} e(x)\right] e^{-jk_z z} - k_z^2 e(x)e^{-jk_z z} + \omega^2 \mu\varepsilon e(x)e^{-jk_z z} = 0
\end{aligned} \tag{68}$$

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:

$$
\begin{aligned}
&\frac{\partial^2}{\partial x^2}e(x) - k_z^2 e(x) + \omega^2 \mu\varepsilon e(x) = 0 \\
&\frac{\partial^2}{\partial x^2}e(x) + (k^2 - k_z^2)e(x) = 0
\end{aligned}
\tag{69}
$$

The wave vector is, in general:

$$
\mathbf{k} = k_x \mathbf{u}_x + k_y \mathbf{u}_y + k_z \mathbf{u}_z \tag{70}
$$

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

$$
\begin{aligned}
&k^2 = |\mathbf{k}|^2 = \omega^2 \mu\varepsilon \\
&k^2 = k_x^2 + k_z^2 \\
&k^2 - k_z^2 = k_x^2
\end{aligned}
\tag{71}
$$

and then

$$
\boxed{\frac{\mathrm{d}^2}{\mathrm{d}x^2}e(x) + k_x^2 e(x) = 0} \tag{72}
$$

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 (41), $\varepsilon = \varepsilon(x)$. If so, then

$$
k^2(x) = \omega^2 \mu_0 \varepsilon(x) = k_x^2(x) + k_z^2 \tag{73}
$$

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 (72) actually represents *two* equations:

$$
\begin{cases}
\dfrac{\mathrm{d}^2}{\mathrm{d}x^2}e(x) + k_{x_1}^2 e(x) = 0, & |x| \le a \\[2ex]
\dfrac{\mathrm{d}^2}{\mathrm{d}x^2}e(x) + k_{x_2}^2 e(x) = 0, & |x| > a
\end{cases}
\tag{74}
$$

First equation is related to region 1 and second equation to region 2. They are both Helmholtz equations: 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 (72), which admits several solutions, as pointed out in Section A specific case of Appendix A subset of Second order differential equations.

The unknown functions $e(x)$ in region 1 and region 2 are both allowed to be:

1. an increasing and/or decreasing exponential with respect to x;

2. a complex exponential;

3. a sinusoid.

For each equation in (74), $e(x)$ 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.

k_{x_1} and k_{x_2} will be chosen to *realize* the type of field that is desirable for this waveguide structure.

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 (74). 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. In order to accomplish this, k_{x_1} must be *real*.

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. According to the mentioned Appendix, this can be realized when the coefficient of $e(x)$ in (74) is *negative*: consequently, k_{x_2} must be *purely imaginary*.

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:

$$e(x) = \begin{cases} A\cos(k_{x_1}x) + B\sin(k_{x_1}x), & 0 \leq x \leq a \\ Ce^{-jk_{x_2}(x-a)} + De^{jk_{x_2}(x-a)}, & x > a \end{cases} \tag{75}$$

As anticipated, k_{x_1} is a real quantity. Moreover, the only way condition 2 is satisfied is $D = 0$ and k_{x_2} purely imaginary:

$$\begin{aligned} & k_{x_2} = -j\gamma \\ & -jk_{x_2} = j^2\gamma = -\gamma \\ & \gamma = |k_{x_2}| \\ & k_{x_2} = -j|k_{x_2}| \\ & -jk_{x_2} = -|k_{x_2}| \end{aligned} \tag{76}$$

and from (75)

$$e(x) = Ce^{-|k_{x_2}|(x-a)}, \quad x > a \tag{77}$$

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 (58) 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 (75), such a mode can be symmetrical or anti-symmetrical with respect to the plane $x = 0$.

Even TE modes

In the system of equations (75), $B = 0$ 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:

$$e(x) = \begin{cases} A\cos(k_{x_1}x), & |x| \leq a \\ A'e^{-|k_{x_2}|(|x|-a)}, & |x| > a \end{cases} \tag{78}$$

Note that $|x| \leq a$ is the synthesis of $-a \leq x \leq a$ and the second line is the synthesis[11] of

$$\begin{cases} A'e^{-|k_{x_2}|(x-a)}, & x > a \\ A'e^{|k_{x_2}|(x+a)}, & x < -a \end{cases} \tag{79}$$

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.

$$\begin{gathered} A\cos(k_{x_1}x)\big|_{x=a} = A'e^{-|k_{x_2}|(x-a)}\big|_{x=a} \\ A' = A\cos(k_{x_1}a) \end{gathered} \tag{80}$$

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.

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

The *tangent magnetic field* must satisfy the same continuity condition at the same interface. Between the (58) magnetic field components, only $H_z(x, y)$ is tangent to the $x = a$ plane.

Given $e(x)$, the field component in the set (58) that can be immediately determined is $E_y(x, z)$, according to (67). In order to obtain $H_z(x, z)$, instead, equation (53) (which is derived from the expansion of the curl equation (49)) must be recalled.

$$H_z(x, z) = -\frac{1}{j\omega\mu}\frac{\partial}{\partial x}E_y(x, z) \tag{81}$$

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

$$\begin{aligned} H_z^{(1)}(x, z) &= -\frac{1}{j\omega\mu}\frac{\partial}{\partial x}E_y^{(1)}(x, z) \\ H_z^{(2)}(x, z) &= -\frac{1}{j\omega\mu}\frac{\partial}{\partial x}E_y^{(2)}(x, z) \end{aligned} \tag{82}$$

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

$$\begin{aligned} H_z^{(1)}(x, z)\Big|_{x=a} &= H_z^{(2)}(x, z)\Big|_{x=a} \\ -\frac{1}{j\omega\mu}\left[\frac{\partial}{\partial x}E_y^{(1)}(x, z)\right]_{x=a} &= -\frac{1}{j\omega\mu}\left[\frac{\partial}{\partial x}E_y^{(2)}(x, z)\right]_{x=a} \\ \left[\frac{\partial}{\partial x}E_y^{(1)}(x, z)\right]_{x=a} &= \left[\frac{\partial}{\partial x}E_y^{(2)}(x, z)\right]_{x=a} \end{aligned} \tag{83}$$

The single derivatives have the following values:

$$\begin{aligned} \left[\frac{\partial}{\partial x}E_y^{(1)}(x, z)\right]_{x=a} &= -\, k_{x_1} A \sin(k_{x_1} x)\big|_{x=a} = -k_{x_1} A \sin(k_{x_1} a) \\ \left[\frac{\partial}{\partial x}E_y^{(2)}(x, z)\right]_{x=a} &= -\, |k_{x_2}| A' e^{-|k_{x_2}|(x-a)}\Big|_{x=a} = -|k_{x_2}| A' \end{aligned} \tag{84}$$

Applying (83),

$$
\begin{gathered}
-k_{x_1} A \sin(k_{x_1} a) = -|k_{x_2}| A' \\
A' = \frac{k_{x_1}}{|k_{x_2}|} A \sin(k_{x_1} a)
\end{gathered}
\tag{85}
$$

Both equations (80) and (85) are related to A' and they must be verified at the same time. Then,

$$
\begin{gathered}
A \cos(k_{x_1} a) = \frac{k_{x_1}}{|k_{x_2}|} A \sin(k_{x_1} a) \\
\boxed{|k_{x_2}| = k_{x_1} \tan(k_{x_1} a)}
\end{gathered}
\tag{86}
$$

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 (71) 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, (71) becomes:

$$
\begin{gathered}
k^2 = k_x^2 + k_z^2 \\
\boxed{k_1^2 = k_{x_1}^2 + k_z^2}
\end{gathered}
\tag{87}
$$

In the medium 2, representing the *cover*:

$$
\begin{gathered}
k^2 = k_x^2 + k_z^2 \\
k_2^2 = k_{x_2}^2 + k_z^2
\end{gathered}
\tag{88}
$$

but from (76) it is known that k_{x_2} is a purely imaginary, negative quantity.

$$
\begin{aligned}
&k_{x_2} = -j\gamma \\
&\gamma = |k_{x_2}| \\
&k_{x_2} = -j|k_{x_2}| \\
&k_{x_2}^2 = j^2|k_{x_2}|^2 = -|k_{x_2}|^2
\end{aligned}
\tag{89}
$$

Equation (88) can then be rewritten in a more explicit form as

$$\boxed{k_2^2 = k_z^2 - |k_{x_2}|^2} \tag{90}$$

Note that k_z is the same quantity in both the media, so in both equations (87) and (90). This peculiarity has already been described in Section Oblique incidence of a plane wave on a surface. Subtracting side-by-side equation (90) from equation (87):

$$\begin{aligned} &k_1^2 - k_2^2 = k_{x_1}^2 + |k_{x_2}|^2 \\ &\omega^2 \mu_0 \varepsilon_0 (\varepsilon_{r_1} - \varepsilon_{r_2}) = k_{x_1}^2 + |k_{x_2}|^2 \\ &\boxed{k_{x_1}^2 + |k_{x_2}|^2 = k_0^2 (n_1^2 - n_2^2)} \end{aligned} \tag{91}$$

The variables k_{x_1} and $|k_{x_2}|$ must satisfy at the same time the equations (86) and (91): 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 (78) 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 (75), an anti-symmetrical mode with respect to the plane (y, z) or $x = 0$ is obtained. The function $e(x)$ is:

$$e(x) = \begin{cases} B' e^{-|k_{x_2}|(x-a)}, & x > a \\ B \sin(k_{x_1} x), & |x| \leq a \\ -B' e^{|k_{x_2}|(x+a)}, & x < -a \end{cases} \tag{92}$$

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 (80),

$$\begin{gathered} B \sin(k_{x_1} x)\big|_{x=a} = B' e^{-|k_{x_2}|(x-a)}\Big|_{x=a} \\ B' = B \sin(k_{x_1} a) \end{gathered} \tag{93}$$

Note that the sign of B' in the first line of (92) is determined by the value of the sine in $x = a$. In the last line of (92), the coefficient of the exponential function will have an opposite sign.

Using again equations (81) and (82), referring only to the region of space $x \geq 0$ to avoid ambiguity,

$$\begin{aligned} \left[\frac{\partial}{\partial x} E_y^{(1)}(x,z)\right]_{x=a} &= k_{x_1} B \cos(k_{x_1} a) \\ \left[\frac{\partial}{\partial x} E_y^{(2)}(x,z)\right]_{x=a} &= -|k_{x_2}| B' \end{aligned} \tag{94}$$

Then, according to (83),

$$\begin{aligned} k_{x_1} B \cos(k_{x_1} a) &= -|k_{x_2}| B' \\ B' &= -\frac{k_{x_1}}{|k_{x_2}|} B \cos(k_{x_1} a) \end{aligned} \tag{95}$$

It follows that (as in (86)):

$$\begin{aligned} B \sin(k_{x_1} a) &= -\frac{k_{x_1}}{|k_{x_2}|} B \cos(k_{x_1} a) \\ &\boxed{|k_{x_2}| = -k_{x_1} \cot(k_{x_1} a)} \end{aligned} \tag{96}$$

The system composed by this equation, along with the unvaried equation (91), 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 (92) will be fully determined for the TE odd modes, as well.

TM modes

Alternatevely to the first set (58) of components, the second one (59) 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 (65) is not a zero identity:

$$
\begin{aligned}
&\nabla^2 H_y(x,z) + \omega^2 \mu\varepsilon H_y(x,z) = 0 \\
&\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2}\right) H_y(x,z) + \omega^2 \mu\varepsilon H_y(x,z) = 0
\end{aligned}
\tag{97}
$$

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

$$
H_y(x,z) = h(x) f(z) = h(x) e^{-jk_z z} \tag{98}
$$

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

$$
\begin{cases}
\dfrac{\mathrm{d}^2}{\mathrm{d}x^2} h(x) + k_{x_1}^2 h(x) = 0, & |x| > a \\
\dfrac{\mathrm{d}^2}{\mathrm{d}x^2} h(x) + k_{x_2}^2 h(x) = 0, & |x| \leq a
\end{cases}
\tag{99}
$$

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.

$$
h(x) = \begin{cases}
A\cos(k_{x_1} x) + B\sin(k_{x_1} x), & 0 \leq x \leq a \\
Ce^{-jk_{x_2}(x-a)}, & x > a
\end{cases}
\tag{100}
$$

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 (97) 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:

$$
h(x) = \begin{cases}
A\cos(k_{x_1} x), & |x| \leq a \\
A' e^{-|k_{x_2}|(|x|-a)}, & |x| > a
\end{cases}
\tag{101}
$$

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:

$$
\begin{aligned}
A\cos(k_{x_1}x)\big|_{x=a} &= A' e^{-|k_{x_2}|(x-a)}\Big|_{x=a} \\
A' &= A\cos(k_{x_1}a)
\end{aligned}
\tag{102}
$$

which is identical to (80) (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 (59) to be tangent to the $x = a$ plane.

The function $h(x)$ is directly related to $H_y(x, z)$, according to (98). $E_z(x, z)$ can be obtained, instead, from equation (57) (which is derived from the expansion of the curl equation (54)).

$$
E_z(x, z) = \frac{1}{j\omega\varepsilon}\frac{\partial}{\partial x}H_y(x, z) \tag{103}
$$

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

$$
\begin{aligned}
E_z^{(1)}(x, z) &= \frac{1}{j\omega\varepsilon_1}\frac{\partial}{\partial x}H_y^{(1)}(x, z) \\
E_z^{(2)}(x, z) &= \frac{1}{j\omega\varepsilon_2}\frac{\partial}{\partial x}H_y^{(2)}(x, z)
\end{aligned}
\tag{104}
$$

where region 1 is $|x| \leq a$ and region 2 is $|x| > a$. Differently from (82), here the permittivity explicitly shows the change in the materials between the two regions.

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

$$
\begin{aligned}
E_z^{(1)}(x, z)\Big|_{x=a} &= E_z^{(2)}(x, z)\Big|_{x=a} \\
\frac{1}{j\omega\varepsilon_1}\left[\frac{\partial}{\partial x}E_y^{(1)}(x, z)\right]_{x=a} &= \frac{1}{j\omega\varepsilon_2}\left[\frac{\partial}{\partial x}E_y^{(2)}(x, z)\right]_{x=a} \\
\frac{1}{\varepsilon_1}\left[\frac{\partial}{\partial x}E_y^{(1)}(x, z)\right]_{x=a} &= \frac{1}{\varepsilon_2}\left[\frac{\partial}{\partial x}E_y^{(2)}(x, z)\right]_{x=a}
\end{aligned}
\tag{105}
$$

Note that the permittivities don't erase like the permeabilities μ_0 in (83). The single derivatives have the following values:

$$
\begin{aligned}
\left[\frac{\partial}{\partial x}H_y^{(1)}(x,z)\right]_{x=a} &= -\,k_{x_1}A\sin(k_{x_1}x)|_{x=a} = -k_{x_1}A\sin(k_{x_1}a) \\
\left[\frac{\partial}{\partial x}H_y^{(2)}(x,z)\right]_{x=a} &= -\,|k_{x_2}|A'e^{-|k_{x_2}|(x-a)}\Big|_{x=a} = -|k_{x_2}|A'
\end{aligned}
\tag{106}
$$

Applying (105),

$$
\begin{aligned}
-\frac{1}{\varepsilon_1}k_{x_1}A\sin(k_{x_1}a) &= -\frac{1}{\varepsilon_2}|k_{x_2}|A' \\
A' &= \frac{\varepsilon_2}{\varepsilon_1}\frac{k_{x_1}}{|k_{x_2}|}A\sin(k_{x_1}a)
\end{aligned}
\tag{107}
$$

Using both equations (102) and (107)

$$
\begin{aligned}
A\cos(k_{x_1}a) &= \frac{\varepsilon_2}{\varepsilon_1}\frac{k_{x_1}}{|k_{x_2}|}A\sin(k_{x_1}a) \\
&\boxed{|k_{x_2}| = \frac{n_2^2}{n_1^2}k_{x_1}\tan(k_{x_1}a)}
\end{aligned}
\tag{108}
$$

A quick comparison between this relation and its correspondent (86) 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 (108) and (91). The TM even modes will be represented by the field fully defined in (97).

Odd TM modes

A zero A coefficient in the system of equations (100) 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 (92):

$$h(x) = \begin{cases} B' e^{-|k_{x_2}|(x-a)}, & x > a \\ B \sin(k_{x_1} x), & |x| \leq a \\ -B' e^{|k_{x_2}|(x+a)}, & x < -a \end{cases} \tag{109}$$

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 (93)):

$$B' = B \sin(k_{x_1} a) \tag{110}$$

Equations (103) and (104) (referring only to the region of space $x \geq 0$ to avoid ambiguity) provide

$$\begin{gathered} \left[\frac{\partial}{\partial x} H_y^{(1)}(x,z)\right]_{x=a} = k_{x_1} B \cos(k_{x_1} a) \\ \left[\frac{\partial}{\partial x} H_y^{(2)}(x,z)\right]_{x=a} = -|k_{x_2}| B' \end{gathered} \tag{111}$$

From (105),

$$\begin{gathered} \frac{1}{\varepsilon_1} k_{x_1} B \cos(k_{x_1} a) = -\frac{1}{\varepsilon_2} |k_{x_2}| B' \\ B' = -\frac{\varepsilon_2}{\varepsilon_1} \frac{k_{x_1}}{|k_{x_2}|} B \cos(k_{x_1} a) \end{gathered} \tag{112}$$

It follows that (as in (108)):

$$\begin{gathered} B \sin(k_{x_1} a) = -\frac{\varepsilon_2}{\varepsilon_1} \frac{k_{x_1}}{|k_{x_2}|} B \cos(k_{x_1} a) \\ \boxed{|k_{x_2}| = -\frac{n_2^2}{n_1^2} k_{x_1} \cot(k_{x_1} a)} \end{gathered} \tag{113}$$

As for the even TM modes, the only difference between (113) and (96) is the n_2^2/n_1^2 term.

This equation, along with (91), 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 (109), 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 (86) and (91) for TE even modes;

$$\begin{cases} |k_{x_2}| = k_{x_1} \tan(k_{x_1} a) \\ k_{x_1}^2 + |k_{x_2}|^2 = k_0^2(n_1^2 - n_2^2) \end{cases} \tag{114}$$

- equations (96) and (91) for TE odd modes;

$$\begin{cases} |k_{x_2}| = -k_{x_1} \cot(k_{x_1} a) \\ k_{x_1}^2 + |k_{x_2}|^2 = k_0^2(n_1^2 - n_2^2) \end{cases} \tag{115}$$

- equations (108) and (91) for TM even modes;

$$\begin{cases} |k_{x_2}| = \dfrac{n_2^2}{n_1^2} k_{x_1} \tan(k_{x_1} a) \\ k_{x_1}^2 + |k_{x_2}|^2 = k_0^2(n_1^2 - n_2^2) \end{cases} \tag{116}$$

- equations (113) and (91) for TM odd modes.

$$\begin{cases} |k_{x_2}| = -\dfrac{n_2^2}{n_1^2} k_{x_1} \cot(k_{x_1} a) \\ k_{x_1}^2 + |k_{x_2}|^2 = k_0^2(n_1^2 - n_2^2) \end{cases} \tag{117}$$

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

Both sides of equations (86), (96), (108) and (113) are multiplied by a and both sides of equation (91) are multiplied by a^2. For example, the system (114) becomes:

$$\begin{cases} a|k_{x_2}| = ak_{x_1}\tan(ak_{x_1}) \\ a^2k_{x_1}^2 + a^2|k_{x_2}|^2 = a^2k_0^2(n_1^2 - n_2^2) \end{cases} \tag{118}$$

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

$$\begin{aligned} u &= ak_{x_1} \\ w &= a|k_{x_2}| \\ v^2 &= a^2k_0^2(n_1^2 - n_2^2) \end{aligned} \tag{119}$$

So, system (114) can also be written in a more essential, straightforward manner:

$$\boxed{\begin{cases} w = u\tan(u) \\ v^2 = u^2 + w^2 \end{cases}} \tag{120}$$

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 (120) 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 (120) 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 (75) of equations (74), 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[12] 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 (120) 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 6, $u\tan(u)$ 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 (120) 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

[12] In the space of the dielectric slab.

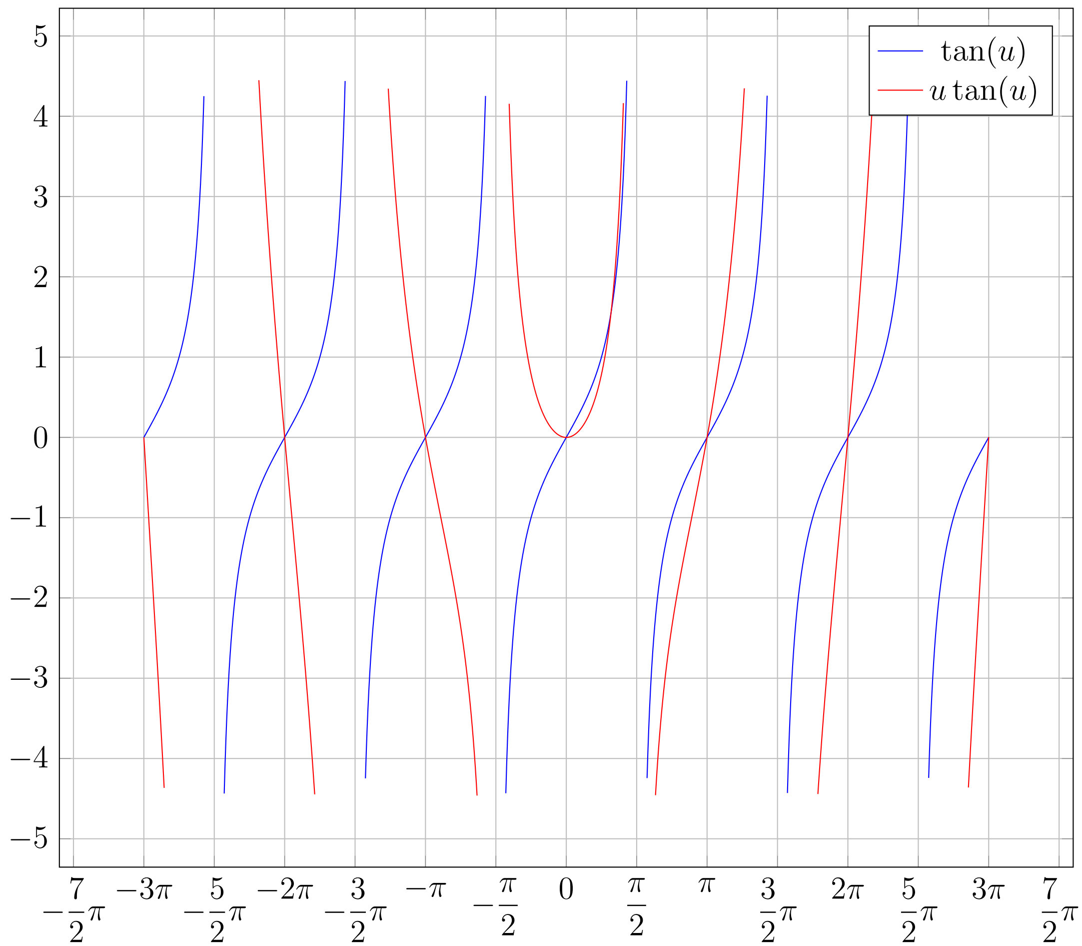


Figure 6: Comparison between $\tan(u)$ and $u\tan(u)$.

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

$$v^2 = a^2 k_0^2 (n_1^2 - n_2^2) = a^2 \left(\frac{2\pi}{\lambda}\right)^2 (n_1^2 - n_2^2) \tag{121}$$

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

$$k_0^2 = \omega^2 \mu_0 \varepsilon_0 \tag{122}$$

but the use of λ is more common.

With this information, v is known and a solution to system (120) can be looked for. Each of these four fundamental parameters (a, λ, n_1 and n_2) determines the radius of the circumference.

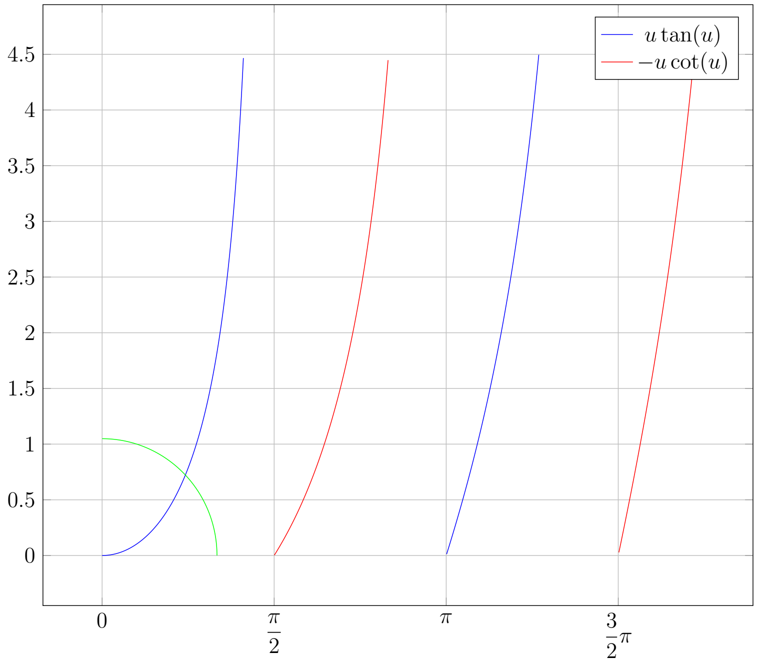
With the equations (120), at least one intersection is theoretically always possible, so *one* TE even mode should always be available: this makes the TE even mode the *fundamental mode* of dielectric symmetrical slabs.

$$\boxed{\begin{cases} w = -u\cot(u) \\ v^2 = u^2 + w^2 \end{cases}} \tag{123}$$

is the system of equations correspondent to (115), for the TE odd modes.

$$\boxed{\begin{cases} w = \dfrac{n_2^2}{n_1^2} u\tan(u) \\ v^2 = u^2 + w^2 \end{cases}} \tag{124}$$

corresponds to (116), as regards the TM even modes.



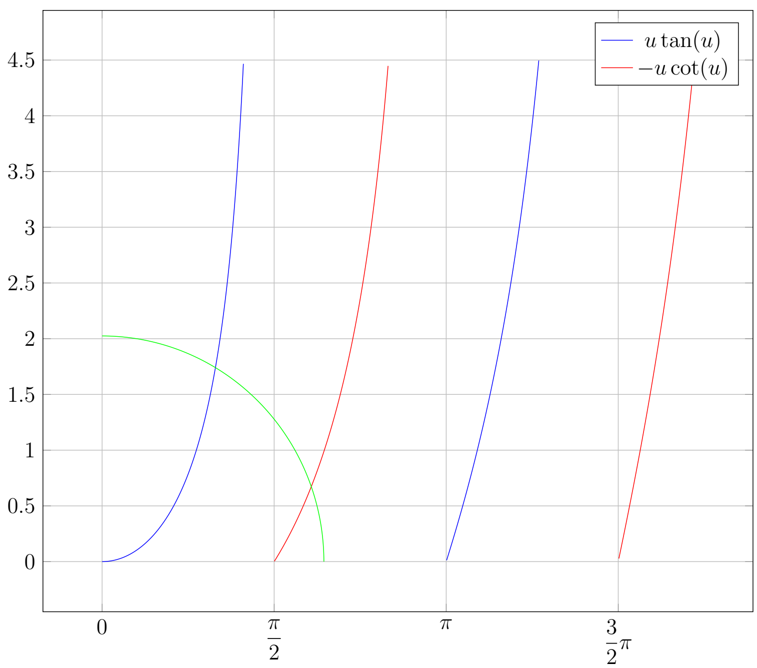


Figure 7: Variable number of available modes according to the values of the radius v: in the first plot, only one intersection occurs; in the second plot, where the circumference has a greater v, two intersections are generated.

$$\boxed{\begin{cases} w = -\dfrac{n_2^2}{n_1^2} u \cot(u) \\ v^2 = u^2 + w^2 \end{cases}} \tag{125}$$

is derived from (117) for the TM odd modes.

Note that while the first equation of systems (120) and (123) do not depend on the custom quantities n_1 and n_2, the first equations of systems (124) and (125) 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.

Assuming that the physical parameters of the guide are known (a, n_1, n_2), if the *frequency* is fixed, Figure 7 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 7 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 (87):

$$k_z = \sqrt{k_1^2 - k_{x_1}^2} \tag{126}$$

From equation (90):

$$k_z = \sqrt{k_2^2 + |k_{x_2}|^2} \tag{127}$$

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 (75) of equations (74), k_{x_1} is a real quantity, so $k_{x_1}^2$ is a real, non-negative quantity. Given that, equation (126) implies that

$$k_z \leq k_1 \tag{128}$$

Being k_z in any case a real quantity, of course it should be $k_{x_1} \leq k_1$ in (126) and therefore (128) 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 (127),

$$k_z \geq k_2 \tag{129}$$

Equations (126) and (127) are equivalent, alternative definitions of k_z and must be verified at the same time; so, also the inequalities (128) and (129) must be verified at the same time:

$$\begin{cases} k_z \leq k_1 \\ k_z \geq k_2 \end{cases} \tag{130}$$

It follows that, for each mode:

$$\boxed{k_2 \leq k_z \leq k_1} \tag{131}$$

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

Inequality (131) 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:

$$\begin{gathered} \frac{k_2}{k_0} \leq \frac{k_z}{k_0} \leq \frac{k_1}{k_0} \\ \boxed{n_2 \leq n_{\text{eff}} \leq n_1} \end{gathered} \tag{132}$$

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

By inspection of Figure 7, 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[13]. The mode **activates** and $k_z = k_2$ assumes its inferior limit value.

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 (129) and equation (127) could induce to the conclusion that also $k_z \to \infty$, it should be remembered that also inequality (127) must be verified: therefore, in the best achievable confinement[14], $k_z \to k_1$.

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 8) represents the interface between these two media.

In Figure 8, 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 $n = 1.52$.

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.

---

[13] 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.

[14] 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.

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

$$\boxed{n_1 \sin\theta_1 = n_2 \sin\theta_2} \tag{133}$$

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. (133) imposes a relation between the **sines** of these angles:

$$\sin\theta_2 = \frac{n_1}{n_2}\sin\theta_1 \tag{134}$$

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[15] 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

$$\sin\theta_1^\ell = \frac{n_2}{n_1}\sin\theta_2 = \frac{n_2}{n_1} \tag{135}$$

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 (133). In fact, as more explicitly shown by (134), $\sin\theta_2$ should assume values greater than 1, which is not possible for a sine function of a real angle[16].

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

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

[16] 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 (134), but it has no physical meaning.

well considered in textbooks[17] 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 8, is shown in Figure 9.

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 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 5. 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 10 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

[17] For example, David M. Pozar. *Microwave engineering*. Wiley, Hoboken (NJ), 3rd edition, 2005.

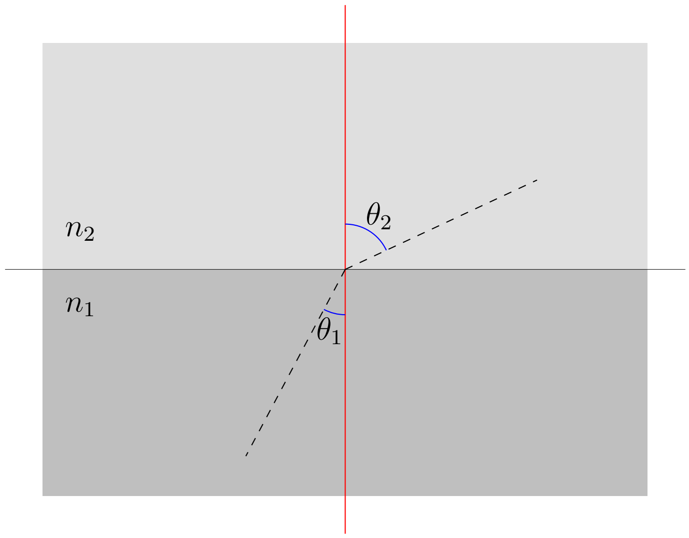


Figure 8: Refraction of a plane wave impinging on the interface between two media.

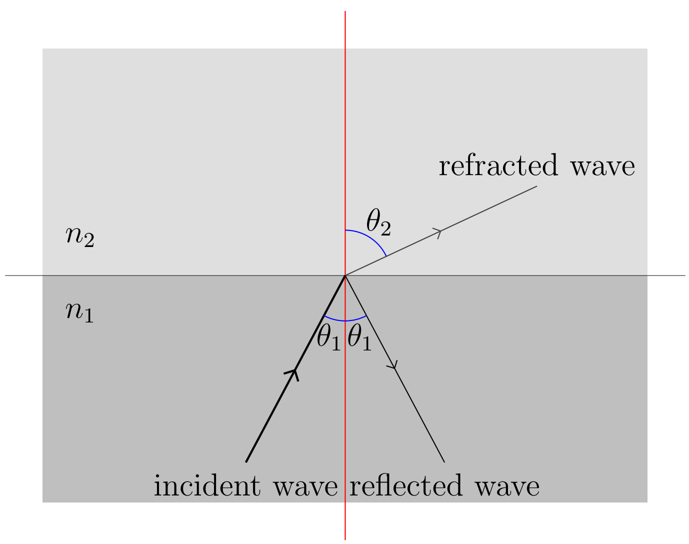


Figure 9: Complete description of the plane wave behaviour at the interface between two media.

is assumed with respect to Figure 5, which has been rotated by 90° in the plane of the core, towards the observer.

Figure 10 also explicitly represents equations (87) and (126). 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,

$$k_z = k_1 \sin\theta_1 \tag{136}$$

and also

$$k_{x_1} = k_1 \cos\theta_1 \tag{137}$$

Equation (136) can alternatively be written as

$$k_z = n_1 k_0 \sin\theta_1 = n_1 \omega \sqrt{\mu_0 \varepsilon_0} \sin\theta_1 \tag{138}$$

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

Moreover, being

$$0 \leq \sin\theta_1 \leq 1 \tag{139}$$

necessarily will also be from (138)

$$0 \leq k_z \leq n_1 k_0 \tag{140}$$

Also,

$$\theta_1 \geq \theta_1^\ell \tag{141}$$

that can be rewritten (applying the sine function to both members and remembering (135)) as

$$\sin\theta_1 \geq \frac{n_2}{n_1} \tag{142}$$

So (again from (138)) it will be

$$k_z \geq n_1 k_0 \cdot \frac{n_2}{n_1} \tag{143}$$

that is

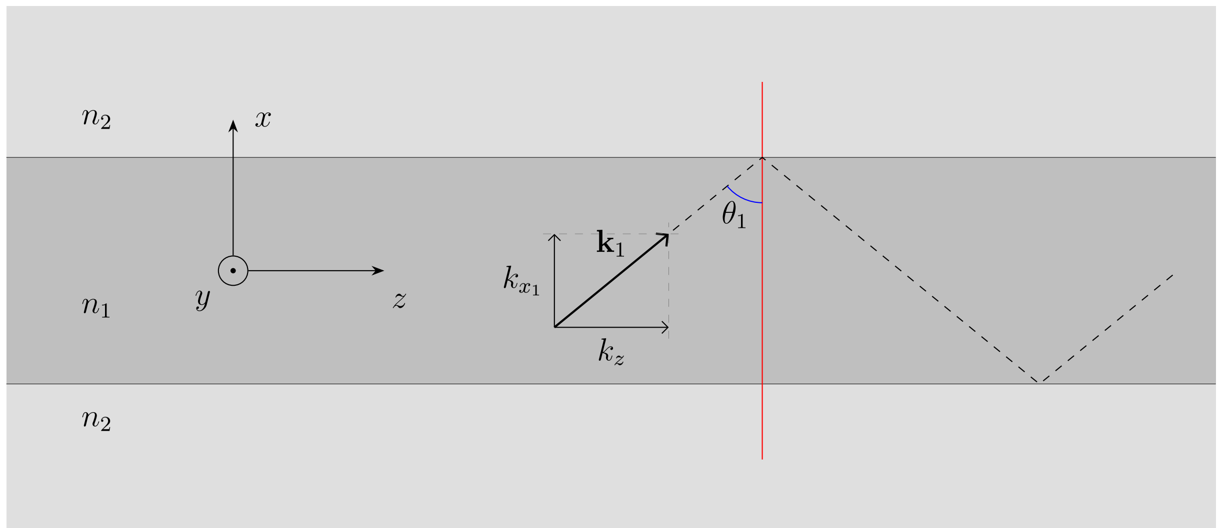


Figure 10: Path of a plane wave inside the waveguide core when the total internal reflection condition is satisfied.

$$k_z \geq n_2 k_0 \tag{144}$$

Because both the conditions (140) and (144) on k_z must be verified at the same time,

$$\begin{cases} 0 \leq k_z \leq n_1 k_0 \\ k_z \geq n_2 k_0 \end{cases} \Rightarrow \begin{cases} 0 \leq k_z \leq k_1 \\ k_z \geq k_2 \end{cases} \Rightarrow \boxed{k_2 \leq k_z \leq k_1} \tag{145}$$

These are the same conditions in (130) and (131), 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 7 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 $w = a|k_{x_2}| = 0$.

$$\begin{aligned} &u\tan(u) = 0 \\ &u = 0 \vee \tan(u) = 0 \end{aligned} \tag{146}$$

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:

$$\begin{aligned} &u = n\pi,\ n = 1, 2, \ldots \\ &ak_{x_1} = n\pi \end{aligned} \tag{147}$$

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,

$$w = 0 \Rightarrow \begin{cases} u\tan(u) = 0,\ u = n\pi \\ v^2 = u^2 + w^2 = u^2 \end{cases} \Rightarrow v = u \tag{148}$$

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

$$v^2 = u^2 + w^2 \tag{149}$$

At cutoff, $w = 0$. Remembering equation (119),

$$v^2 = u^2 + w^2 \Rightarrow a^2 k_0^2 (n_1^2 - n_2^2) = u^2 \tag{150}$$

Using (122), remembering that $\omega = 2\pi f$ and applying the square root to both the sides:

$$\begin{aligned} &a^2 \omega^2 \mu_0 \varepsilon_0 (n_1^2 - n_2^2) = u^2 \\ &a^2 4\pi^2 f^2 \mu_0 \varepsilon_0 (n_1^2 - n_2^2) = u^2 \\ &2a\pi f \sqrt{\mu_0 \varepsilon_0} \sqrt{n_1^2 - n_2^2} = u \end{aligned} \tag{151}$$

Remembering that $\sqrt{\mu_0 \varepsilon_0} = 1/c$ and equation (147),

$$\begin{aligned} &\frac{2a\pi f}{c} \sqrt{n_1^2 - n_2^2} = u \\ &\frac{2a\pi f}{c} \sqrt{n_1^2 - n_2^2} = n\pi \\ &\boxed{f = f_c = \frac{nc}{2a\sqrt{n_1^2 - n_2^2}}} \end{aligned} \tag{152}$$

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 (152), remembering that $\lambda = c/f$,

$$\boxed{\lambda_c = \frac{2a\sqrt{n_1^2 - n_2^2}}{n}} \tag{153}$$

Equation (138) shows the relation between the propagation constant k_z and the angle of incidence θ_1. Moreover:

$$
\begin{aligned}
k_z &= n_1 k_0 \sin\theta_1 \\
n_2 k_0 &= n_1 k_0 \sin\theta_1 \\
\sin\theta_1 &= \frac{n_2}{n_1}
\end{aligned} \tag{154}
$$

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 (131)): 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[18], k_z reaches the limit value of $k_1 = n_1 k_0$. Then,

$$
\begin{aligned}
k_z &= n_1 k_0 \sin\theta_1 \\
n_1 k_0 &= n_1 k_0 \sin\theta_1 \\
\sin\theta_1 &= 1 \\
\theta_1 &= \frac{\pi}{2}
\end{aligned} \tag{155}
$$

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:

[18]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.

$$\boxed{\theta_1^\ell \le \theta_1 \le \frac{\pi}{2}} \tag{156}$$

From another perspective, recalling Figure 7, a frequency value f of the sinusoidal signal will cause a *finite* number M of intersections 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, the frequency f allows for the same, finite number M of *angles of incidence*, **one for each active mode**, and no more. The above considerations apply to each of these modes.

The sinusoidal signal generated by the source is then allowed to assume only M field configurations inside the core[19], each one with its angle of incidence θ_1 and with its propagation constant k_z. For a given frequency f, only these allowed values of θ_1 will satisfy the self-consistence condition listed in the Introduction, letting the mode propagating along z.

[19] Its energy will split between those available modes and will propagate in M different ways, with M different speeds, along z.

Appendices

A subset of Second order differential equations

Introduction

The general form of a linear, second order, homogeneous, ordinary differential equation with constant coefficients is:

$$my''(t) + by'(t) + ky(t) = 0 \tag{157}$$

with $t \in \mathbb{R}$ and $m, b, k \in \mathbb{R}$ constants, not depending on t.

The solution of (157) is not obtained by developing the equation, nor by direct integration. It is simply *guessed* that a solution of the form

$$y(t) = e^{rt} \tag{158}$$

can exist, with $r \in \mathbb{C}$ in the most general case. This solution is, at the present time, not necessarily the only one. To verify if (158) is really a solution, it must be substituted in (157):

$$mr^2e^{rt} + bre^{rt} + ke^{rt} = 0 \tag{159}$$

Observing that, even with $r \in \mathbb{C}$,

$$e^{rt} \neq 0,\ \forall t \in \mathbb{R} \tag{160}$$

equation (159) can be rewritten as

$$mr^2 + br + k = 0 \tag{161}$$

which is known as the *characteristic polynomial*[20] of equation (157).

[20] Alternatively named *characteristic equation*.

Its roots

$$r_{1,2} = \frac{-b \pm \sqrt{b^2 - 4km}}{2m} \tag{162}$$

determine the form of the guessed solution (158). The functions

$$\begin{aligned} f_1(t) &= e^{r_1 t} \\ f_2(t) &= e^{r_2 t} \end{aligned} \tag{163}$$

are solutions of (157) if and only if $r_{1,2}$ are solutions of the characteristic polynomial. Being equation (157) *linear*, being the therein contained derivatives *linear operators*, being its coefficients m, b, k constant, if $f_1(t)$ and $f_2(t)$ are solutions of (157), then any *linear combination* of these functions

$$f(t) = C_1 f_1(t) + C_2 f_2(t) \tag{164}$$

with $C_1, C_2 \in \mathbb{C}$ is still a solution of (157). The fact that coefficients C_1 and C_2 can be complex numbers does not change the specified linearity features of (157).

Two linearly independent functions are required to determine the general solution of a second order, linear, differential, homogeneous, ordinary differential equation like (157). If $f_1(t)$ and $f_2(t)$ generate a non-zero Wronskian[21], they are linearly independent. If $f_1(t)$ and $f_2(t)$ are linearly independent, they constitute a fundamental set of solutions of equation (157): this implies that *any* solution of (157) can be expressed in the form (164).

Only after these verifications it can be stated that the initial guess made with (158) is correct: not only, it provides a full characterization of the solutions of (157).

The roots (162) can be:

1. Both real, when $b^2 - 4km > 0$: $r_1, r_2 \in \mathbb{R}, r_1 \neq r_2$. It can be proved that the general solution to (157) is:

$$f(t) = C_1 e^{r_1 t} + C_2 e^{r_2 t} \tag{165}$$

[21] It will not be further discussed here, but it is described in most of textbooks about differential equations.

2. Real and repeated, when $b^2 - 4km = 0$: $r_1, r_2 \in \mathbb{R}, r_1 = r_2$. The general solution is:

$$f(t) = C_1 e^{rt} + C_2 t e^{rt},\ r = -\frac{b}{2m} \tag{166}$$

3. Complex conjugate, when $b^2 - 4km < 0$: $r_1, r_2 \in \mathbb{C}, r_1 = r_2^*$. The general solution can initially be written in the same form of the first case:

$$f(t) = C_1 e^{r_1 t} + C_2 e^{r_2 t} \tag{167}$$

Note that, according to the values of C_1 and C_2, this can lead to a *real* $f(t)$, as well as to a *complex* $f(t)$. So, even a *real* initial equation like (157) can lead to a *complex* solution.

A specific case

Helmholtz wave equation can be considered as a specific case of (157):

$$y''(t) + ky(t) = 0 \tag{168}$$

with $m = 1, b = 0$. Roots of the characteristic polynomial are

$$r_{1,2} = \pm j\sqrt{\frac{k}{m}} \tag{169}$$

and the couple of solutions is

$$f_1(t) = e^{j\sqrt{\frac{k}{m}}t}, f_2(t) = e^{-j\sqrt{\frac{k}{m}}t} \tag{170}$$

It can be proved through the Wronskian that they are linearly independent and they are a fundamental set of solutions. So, any linear combination of (170) is also a solution to (168).

Using $C_1 = C_2 = 1$ will lead to a complex solution:

$$f(t) = e^{j\sqrt{\frac{k}{m}}t} + e^{-j\sqrt{\frac{k}{m}}t} \tag{171}$$

Different choices of the values of these coefficient can instead lead to real solutions. Not only:

$$
\begin{aligned}
f_1(t) &= \frac{1}{2}e^{j\sqrt{\frac{k}{m}}t} + \frac{1}{2}e^{-j\sqrt{\frac{k}{m}}t} = \cos\left(\sqrt{\frac{k}{m}}t\right) \\
f_2(t) &= \frac{1}{2j}e^{j\sqrt{\frac{k}{m}}t} - \frac{1}{2j}e^{-j\sqrt{\frac{k}{m}}t} = \sin\left(\sqrt{\frac{k}{m}}t\right)
\end{aligned} \tag{172}
$$

$$
f(t) = C_1\cos\left(\sqrt{\frac{k}{m}}t\right) + C_2\sin\left(\sqrt{\frac{k}{m}}t\right) \tag{173}
$$

A real, alternative fundamental set of solutions has been determined, by an alternative choice for the coefficients in (170). Of course, the fundamental set (170) itself can be obtained back from (173), using Euler's identity. Set (173) is sometimes preferred, because it is composed by *real* functions.

It can also be observed that the coefficients C_1, C_2 (their number is the same as the order of the differential equation) have a double use:

- If the fundamental set of solutions is not the desired one, appropriate values of these coefficients can generate an *alternative* fundamental set of solutions. An example of this procedure has already been made in (172) to switch from set (170) to set (173).
- If instead the fundamental set of solutions is accepted, the differential equation (168) can be fully solved, using the initial conditions $y(t_0)$ and $y'(t_0)$, which will give a specific value to both the coefficients C_1 and C_2: then, for that problem, the specific solution of the equation is fully determined.

The fundamental set (170) does not only represent complex exponentials. In fact, for all the cases when the ratio k/m is a negative, real number:

$$
\begin{aligned}
&f_1(t) = e^{j\sqrt{\frac{k}{m}}t}, f_2(t) = e^{-j\sqrt{\frac{k}{m}}t},\ \frac{k}{m} < 0 \\
&f_1(t) = e^{j\sqrt{-\left|\frac{k}{m}\right|}t} = e^{j\cdot j\sqrt{\left|\frac{k}{m}\right|}t}, f_2(t) = e^{-j\sqrt{-\left|\frac{k}{m}\right|}t} = e^{-j\cdot j\sqrt{\left|\frac{k}{m}\right|}t} \\
&f_1(t) = e^{-\sqrt{\left|\frac{k}{m}\right|}t}, f_2(t) = e^{\sqrt{\left|\frac{k}{m}\right|}t}
\end{aligned} \tag{174}
$$

$$f(t) = C_1 e^{-\sqrt{|\frac{k}{m}|}t} + C_2 e^{\sqrt{|\frac{k}{m}|}t} \tag{175}$$

These functions produce a non-zero Wronskian: therefore, they constitute a fundamental set of solutions. According to the physical reality that equation (168) represents, one of them can be unacceptable, so one of the constants C can be forced to be 0.

Summary

An equation of the form (168) can have three different sets of fundamental solutions:

- a couple of complex conjugate exponentials: set (170);
- for the case $k/m < 0$, set (170) becomes a couple of respectively decreasing and increasing exponentials: set (175);
- a couple of sinusoidal functions: set (173).

According to the form of the initial equation and to the physical problem it describes, it could be useful to choose one of these sets rather than the others, to better fit the desired solution.

Bibliography

David M. Pozar. *Microwave engineering.* Wiley, Hoboken (NJ), 3rd edition, 2005.