

10

Surface Waveguides

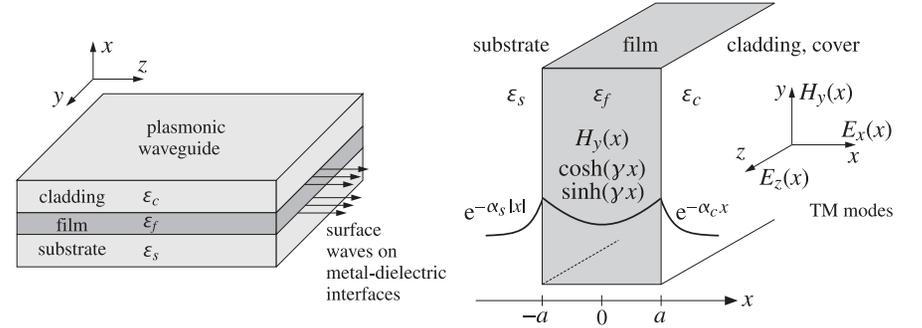


Fig. 10.1.1 Plasmonic waveguide depicting TM modes in either a DMD or MDM configuration.

In this chapter we discuss a variety of surface waveguides, such as plasmonic waveguides, which have the ability to confine light at sub-wavelength scales, and the Sommerfeld wire and Goubau lines, in which there is renewed interest for THz applications. We also use the Sommerfeld wire as the ideal example to explain the skin effect in conductors.

10.1 Plasmonic Waveguides

Surface plasmons have a large number of applications in the field of nanophotonic devices, waveguides, and nanocircuits, and in the area of biological and chemical sensors, and other applications [593–631]. Their distinguishing feature is their ability to confine light at sub-wavelength scales and guide it at long (relative to nanoscale) distances.

In this section, we consider plasmonic waveguides [934–981] from the point of view the longitudinal-transverse waveguide decompositions that we developed in this chapter. The asymmetric dielectric waveguide problem of Sec. 9.12 is very similar—a very significant difference being that in the plasmonic case at least one the layers is metallic with a dielectric constant having negative real-part in the operating frequency range (typically, infrared to optical).

Fig. 10.1.1 depicts a typical plasmonic waveguide consisting of a thin film ϵ_f , sandwiched between a cladding cover ϵ_c and a substrate ϵ_s . We discuss three cases: (a) single interface between a dielectric ϵ_c and a metal ϵ_f , (b) metal-dielectric-metal (MDM) configuration in which ϵ_f is a lossless dielectric and ϵ_c, ϵ_s are metals, (c) dielectric-metal-dielectric (DMD) configuration in which ϵ_f is a metal and ϵ_c, ϵ_s are lossless dielectrics.[†]

Here, the quantities $\epsilon_c, \epsilon_f, \epsilon_s$ denote that *relative* permittivities of the media, that is, $\epsilon_i = \epsilon_i/\epsilon_0$, $i = c, f, s$, where ϵ_i is the permittivities of the i th medium and ϵ_0 , the permittivity of vacuum.

In the geometry of Fig. 10.1.1, propagation is in the z -direction, the transverse confinement is along the x -direction, the layers have infinite extent along the y -direction, and the film thickness is $2a$. We look for field solutions that have t, z propagation dependence of $e^{j\omega t - j\beta z}$, and decay exponentially away from the interfaces so that the

transverse attenuation coefficients α_s, α_c in the substrate and cladding have *positive* real parts, thus, the x -dependence of the fields is as follows for $x < -a$ and $x > a$:

$$e^{-\alpha_s|x|} \text{ for } x < -a \text{ (substrate) and } e^{-\alpha_c|x|} \text{ for } x > a \text{ (cladding)}$$

As in the asymmetric dielectric guide of the previous section,

$$\begin{aligned} \alpha_s^2 &= \beta^2 - k_0^2 \epsilon_s \\ \alpha_c^2 &= \beta^2 - k_0^2 \epsilon_c \end{aligned} \quad (10.1.1)$$

where $k_0 = 2\pi/\lambda_0$ is the vacuum wavenumber and λ_0 the vacuum wavelength, with operating frequency $f = c_0/\lambda_0$ in Hz. Acceptable solutions must have z -propagation wavenumber β with *negative* imaginary part, $\beta = \beta_R - j\beta_I$, $\beta_I \geq 0$, so that the wave attenuates exponentially as it propagates in the positive z -direction:

$$e^{-j\beta z} = e^{-j(\beta_R - j\beta_I)z} = e^{-j\beta_R z} e^{-\beta_I z}$$

A measure of the effective propagation distance is $L = \beta_I^{-1}$, or $L = (2\beta_I)^{-1}$ if it is referred to power, and the attenuation is $20 \log_{10}(e) \beta_I = 8.686\beta_I$ in units of dB/m.

Within the film ϵ_f , the transverse cutoff wavenumber satisfies $k_f^2 = k_0^2 \epsilon_f - \beta^2$. If the film is a metal with a dielectric constant ϵ_f with negative real part, then k_f^2 will be essentially negative or k_f imaginary, and therefore, it makes sense to work with the “attenuation” coefficient defined by $\gamma = jk_f$, so that $\gamma^2 = -k_f^2$,

$$\gamma^2 = \beta^2 - k_0^2 \epsilon_f \quad (10.1.2)$$

Thus, within the film, instead of assuming an x -dependence that is a linear combination of oscillatory $\cos(k_f x)$ and $\sin(k_f x)$ terms, we may work with a linear combination of hyperbolic terms $\sinh(\gamma x)$ and $\cosh(\gamma x)$ —referred to as plasmonic solutions.

Another way to justify the change in notation to $\gamma = jk_f$, is to note that because the metals are lossy, the fields cannot penetrate too deeply into the metal and will be essentially surface waves that are concentrated primarily at the metal-dielectric interfaces (i.e., at $x = \pm a$) and attenuating away from them. By contrast, in the dielectric

[†] Also known as metal-insulator-metal (MIM) and insulator-metal-insulator (IMI) configurations.

waveguide of the previous sections, the fields typically peak in the center of the film. In other words, in plasmonic waveguides most of the field energy is carried at or near the metal-dielectric interfaces instead of at the center of the film.

We will consider only TM plasmonic modes, although in more complicated media, such as magnetic materials and metamaterials, TE modes are also possible. The recent book [934] includes examples of all possible types of media and modes.

The TM modes are obtained by solving Eqs. (9.3.10) in each region and applying the boundary conditions. Thus, we must solve in each region:

$$(\partial_x^2 - \gamma^2)E_z = 0, \quad E_x = -\frac{j\beta}{-\gamma^2} \partial_x E_z, \quad H_y = \frac{1}{\eta_{TM}} E_x, \quad \eta_{TM} = \frac{\beta}{\omega \epsilon_0 \epsilon}, \quad \gamma^2 = \beta^2 - k_0^2 \epsilon$$

or, more specifically,

$$\text{for } |x| \leq a, \quad (\partial_x^2 - \gamma^2)E_z = 0, \quad \gamma^2 = \beta^2 - k_0^2 \epsilon_f, \quad \eta_{TM} = \beta / (\omega \epsilon_0 \epsilon_f)$$

$$\text{for } x \geq a, \quad (\partial_x^2 - \alpha_c^2)E_z = 0, \quad \alpha_c^2 = \beta^2 - k_0^2 \epsilon_c, \quad \eta_{TM} = \beta / (\omega \epsilon_0 \epsilon_c)$$

$$\text{for } x \leq -a, \quad (\partial_x^2 - \alpha_s^2)E_z = 0, \quad \alpha_s^2 = \beta^2 - k_0^2 \epsilon_s, \quad \eta_{TM} = \beta / (\omega \epsilon_0 \epsilon_s)$$

The solutions for $E_z(x)$ that automatically satisfy the tangential E -field boundary conditions at $x = \pm a$ can be expressed as follows, where E_0 is a constant and ψ , a parameter to be determined:

$$E_z(x) = \begin{cases} E_0 \sinh(\gamma x + \psi), & |x| \leq a \\ E_0 \sinh(\gamma a + \psi) e^{-\alpha_c(x-a)}, & x \geq a \\ -E_0 \sinh(\gamma a - \psi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (10.1.3)$$

The complete space-time dependence is $E_z(x) e^{j\omega t - j\beta z}$. Eq. (10.1.3) results in the following transverse E -field:

$$E_x(x) = \begin{cases} -\frac{j\beta}{-\gamma^2} \partial_x E_z = E_0 \frac{j\beta}{\gamma} \cosh(\gamma x + \psi), & |x| \leq a \\ -\frac{j\beta}{-\alpha_c^2} \partial_x E_z = -E_0 \frac{j\beta}{\alpha_c} \sinh(\gamma a + \psi) e^{-\alpha_c(x-a)}, & x \geq a \\ -\frac{j\beta}{-\alpha_s^2} \partial_x E_z = -E_0 \frac{j\beta}{\alpha_s} \sinh(\gamma a - \psi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (10.1.4)$$

The corresponding transverse magnetic fields are then obtained from $H_y = E_x / \eta_{TM}$.

$$H_y(x) = \begin{cases} j\omega \epsilon_0 E_0 \frac{\epsilon_f}{\gamma} \cosh(\gamma x + \psi), & |x| \leq a \\ -j\omega \epsilon_0 E_0 \frac{\epsilon_c}{\alpha_c} \sinh(\gamma a + \psi) e^{-\alpha_c(x-a)}, & x \geq a \\ -j\omega \epsilon_0 E_0 \frac{\epsilon_s}{\alpha_s} \sinh(\gamma a - \psi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (10.1.5)$$

which, using the continuity conditions (10.1.7), can also be written as,

$$H_y(x) = \begin{cases} H_0 \cosh(\gamma x + \psi), & |x| \leq a \\ H_0 \cosh(\gamma a + \psi) e^{-\alpha_c(x-a)}, & x \geq a \\ H_0 \cosh(\gamma a - \psi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (10.1.6)$$

where $H_0 = j\omega \epsilon_0 E_0 \epsilon_f / \gamma$. The continuity of the tangential H -field of Eq. (10.1.5) at the interfaces is equivalent to the continuity of the normal D -field, that is, ϵE_x , resulting into the two conditions:

$$\begin{aligned} \frac{\epsilon_f}{\gamma} \cosh(\gamma a + \psi) &= -\frac{\epsilon_c}{\alpha_c} \sinh(\gamma a + \psi) \\ \frac{\epsilon_f}{\gamma} \cosh(\gamma a - \psi) &= -\frac{\epsilon_s}{\alpha_s} \sinh(\gamma a - \psi) \end{aligned} \Rightarrow \begin{cases} \tanh(\gamma a + \psi) = -\frac{p_c \alpha_c}{\gamma} \\ \tanh(\gamma a - \psi) = -\frac{p_s \alpha_s}{\gamma} \end{cases} \quad (10.1.7)$$

where we defined as in Eq. (9.12.23):

$$p_c = \frac{\epsilon_f}{\epsilon_c}, \quad p_s = \frac{\epsilon_f}{\epsilon_s} \quad (10.1.8)$$

The two conditions (10.1.7), together with,

$$\begin{cases} \gamma^2 = \beta^2 - k_0^2 \epsilon_f \\ \alpha_c^2 = \beta^2 - k_0^2 \epsilon_c \\ \alpha_s^2 = \beta^2 - k_0^2 \epsilon_s \end{cases} \quad (10.1.9)$$

allow the determination of the parameters $\beta, \gamma, \alpha_c, \alpha_s$, and ψ . Eqs. (10.1.7) can also be written in the forms:

$$e^{2(\gamma a + \psi)} = \frac{\gamma - p_c \alpha_c}{\gamma + p_c \alpha_c}, \quad e^{2(\gamma a - \psi)} = \frac{\gamma - p_s \alpha_s}{\gamma + p_s \alpha_s} \quad (10.1.10)$$

which immediately decouple into,

$$e^{4\gamma a} = \frac{(\gamma - p_c \alpha_c)(\gamma - p_s \alpha_s)}{(\gamma + p_c \alpha_c)(\gamma + p_s \alpha_s)} \quad (10.1.11)$$

$$e^{4\psi} = \frac{(\gamma - p_c \alpha_c)(\gamma + p_s \alpha_s)}{(\gamma + p_c \alpha_c)(\gamma - p_s \alpha_s)} \quad (10.1.12)$$

and can also be written in the equivalent forms:

$$\tanh(2\gamma a) = -\frac{\gamma(p_c \alpha_c + p_s \alpha_s)}{\gamma^2 + p_c \alpha_c p_s \alpha_s} \quad (10.1.13)$$

$$\tanh(2\psi) = -\frac{\gamma(p_c \alpha_c - p_s \alpha_s)}{\gamma^2 - p_c \alpha_c p_s \alpha_s} \quad (10.1.14)$$

The quantities $\beta, \gamma, \alpha_c, \alpha_s$ may be determined from Eqs. (10.1.9) and (10.1.11), or (10.1.13). Then, the parameter ψ can be calculated from Eq. (10.1.14). The quantity ψ is defined up to an integer multiple of $j\pi/2$ because of the identity $\tanh(2\psi \pm jm\pi) = \tanh(2\psi)$, and the integer m serves to label particular modes. For example, a possible way of introducing the integer m in the inverse of Eqs. (10.1.7) is:

$$\begin{aligned} \gamma a + \psi &= \operatorname{atanh}\left(-\frac{p_c \alpha_c}{\gamma}\right) \\ \gamma a - \psi &= \operatorname{atanh}\left(-\frac{p_s \alpha_s}{\gamma}\right) - jm\pi \end{aligned}$$

which result into:

$$\begin{aligned} \gamma a &= \frac{1}{2} \operatorname{atanh}\left(-\frac{p_c \alpha_c}{\gamma}\right) + \frac{1}{2} \operatorname{atanh}\left(-\frac{p_s \alpha_s}{\gamma}\right) - \frac{1}{2} jm\pi \\ \psi &= \frac{1}{2} \operatorname{atanh}\left(-\frac{p_c \alpha_c}{\gamma}\right) - \frac{1}{2} \operatorname{atanh}\left(-\frac{p_s \alpha_s}{\gamma}\right) + \frac{1}{2} jm\pi \end{aligned} \quad (10.1.15)$$

Eq. (10.1.13) has a very rich solution structure because, in general, the quantities $\beta, \gamma, \alpha_c, \alpha_s$ are complex-valued when the media are lossy. See, for example, Refs. [941-944,954,955] for a discussion of the variety of possible solutions.

The above choice of the m -terms was made so that the cases $m = 0$ and $m = 1$ will correspond to the TM_0 and TM_1 modes, respectively. Additional $j\pi/2$ terms may arise depending on the values of the parameters. For example, if ϵ_c, ϵ_s are real and negative and ϵ_f is positive as in an MDM case, and if $|p_c|\alpha_c/\gamma > 1$ and $|p_s|\alpha_s/\gamma > 1$, then using the identity,

$$\operatorname{atanh}\left(\frac{1}{x}\right) = \operatorname{atanh}(x) + \operatorname{sign}(x) \frac{j\pi}{2}, \quad x \text{ real, } |x| < 1$$

we may rewrite Eq. (10.1.15) in the form:

$$\begin{aligned} \gamma a &= \frac{1}{2} \operatorname{atanh}\left(-\frac{\gamma}{p_c \alpha_c}\right) + \frac{1}{2} \operatorname{atanh}\left(-\frac{\gamma}{p_s \alpha_s}\right) - \frac{1}{2} j(m-1)\pi \\ \psi &= \frac{1}{2} \operatorname{atanh}\left(-\frac{\gamma}{p_c \alpha_c}\right) - \frac{1}{2} \operatorname{atanh}\left(-\frac{\gamma}{p_s \alpha_s}\right) + \frac{1}{2} jm\pi \end{aligned} \quad (10.1.16)$$

Oscillatory Modes

In addition to the *plasmonic* waveguide modes, there may also be *oscillatory* modes. The plasmonic solutions have a value for γ that is predominantly real-valued, with the fields peaking at the metal-dielectric interfaces, whereas the oscillatory modes have a γ that is dominantly imaginary. In fact, replacing, $\gamma = jk_f$, in Eq. (10.1.13) and using the identity, $\tanh(jx) = j \tan(x)$, we obtain the characteristic equation for the oscillatory modes:

$$\tan(2k_f a) = \frac{k_f(p_c \alpha_c + p_s \alpha_s)}{k_f^2 - p_c \alpha_c p_s \alpha_s} \quad (\text{oscillatory modes}) \quad (10.1.17)$$

where $k_f = \sqrt{k_0^2 \epsilon_f - \beta^2}$. In an MDM configuration, k_f is predominantly real, whereas in an DMD configuration that has $\operatorname{Re}(\epsilon_f) < 0$, k_f will be predominantly imaginary, i.e., of the plasmonic type. Thus, Eq. (10.1.17) applies to the MDM case and admits both TE and TM type of solutions, with the TE ones obtained by replacing, $p_c = p_s = 1$, as in Sec. 9.12. Setting $\psi = j\phi$, the two matching conditions (10.1.7) read now

$$\begin{aligned} \tan(k_f a + \phi) &= \frac{p_c \alpha_c}{k_f} \\ \tan(k_f a - \phi) &= \frac{p_s \alpha_s}{k_f} \end{aligned} \quad (10.1.18)$$

which can be inverted and solved in terms of the mode number m :

$$\begin{aligned} k_f a &= \frac{1}{2} \operatorname{atan}\left(\frac{p_c \alpha_c}{k_f}\right) + \frac{1}{2} \operatorname{atan}\left(\frac{p_s \alpha_s}{k_f}\right) + \frac{1}{2} m\pi \\ \phi &= \frac{1}{2} \operatorname{atan}\left(\frac{p_c \alpha_c}{k_f}\right) - \frac{1}{2} \operatorname{atan}\left(\frac{p_s \alpha_s}{k_f}\right) + \frac{1}{2} m\pi \end{aligned} \quad (10.1.19)$$

These apply to both the TE and TM cases. For the TM case of the MDM configuration, because p_c, p_s are dominantly negative-real, we may use the identity,

$$\operatorname{atan}(x) = \frac{\pi}{2} + \operatorname{atan}\left(-\frac{1}{x}\right), \quad x > 0$$

applied, for example, with $x = -k_f/p_c \alpha_c$, to rewrite Eq. (10.1.19) in the form:

$$\begin{aligned} k_f a &= \frac{1}{2} \operatorname{atan}\left(-\frac{k_f}{p_c \alpha_c}\right) + \frac{1}{2} \operatorname{atan}\left(-\frac{k_f}{p_s \alpha_s}\right) + \frac{1}{2} (m-1)\pi \\ \phi &= \frac{1}{2} \operatorname{atan}\left(-\frac{k_f}{p_c \alpha_c}\right) - \frac{1}{2} \operatorname{atan}\left(-\frac{k_f}{p_s \alpha_s}\right) + \frac{1}{2} m\pi \end{aligned} \quad (10.1.20)$$

We will use this form to determine the cutoff waveguide thicknesses of certain modes. Replacing $\gamma = jk_f$ and $\psi = j\phi$ in Eq. (10.1.6), we also obtain the transverse magnetic field of the oscillatory TM modes:

$$H_y(x) = \begin{cases} H_0 \cos(k_f x + \phi), & |x| \leq a \\ H_0 \cos(k_f a + \phi) e^{-\alpha_c(x-a)}, & x \geq a \\ H_0 \cos(k_f a - \phi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (10.1.21)$$

where $H_0 = \omega \epsilon_0 E_0 \epsilon_f / k_f$. In the examples below, we look at the special roles played by the TM_0 plasmonic mode ($m = 0$), and the TM_1 plasmonic and oscillatory modes ($m = 1$).

Relationship to Surface Plasmon Resonance

We note also that Eq. (10.1.11) is equivalent to the surface plasmon resonance condition discussed in Sec. 8.5. The reflection coefficient Γ of the Kretschmann-Raether configuration of Fig. 8.5.2 is given by Eq. (8.5.7),

$$\Gamma = \frac{\rho_a + \rho_b e^{-2jk_z d}}{1 + \rho_a \rho_b e^{-2jk_z d}}, \quad \rho_a = \frac{k_z \varepsilon_a - k_{za} \varepsilon}{k_z \varepsilon_a + k_{za} \varepsilon}, \quad \rho_b = \frac{k_{zb} \varepsilon - k_z \varepsilon_b}{k_{zb} \varepsilon + k_z \varepsilon_b}$$

The surface plasmon resonance condition [608] corresponds to a pole of Γ , that is, the vanishing of the denominator, $1 + \rho_a \rho_b e^{-2jk_z d} = 0$, which is indeed the same as Eq. (10.1.11). To see this, we map the notation of Sec. 8.5 to that of the present section:

$$x \rightarrow z, \quad z \rightarrow x, \quad k_z \rightarrow -j\gamma, \quad k_{za} \rightarrow -j\alpha_s, \quad k_{zb} \rightarrow -j\alpha_c, \quad \varepsilon \rightarrow \varepsilon_f, \quad \varepsilon_a \rightarrow \varepsilon_s, \quad \varepsilon_b \rightarrow \varepsilon_c$$

which imply,

$$\rho_a \rightarrow \frac{\gamma - p_s \alpha_s}{\gamma + p_s \alpha_s}, \quad \rho_b \rightarrow -\frac{\gamma - p_c \alpha_c}{\gamma + p_c \alpha_c}$$

$$1 + \rho_a \rho_b e^{-2jk_z d} = 0 \rightarrow 1 - \left(\frac{\gamma - p_s \alpha_s}{\gamma + p_s \alpha_s} \right) \left(\frac{\gamma - p_c \alpha_c}{\gamma + p_c \alpha_c} \right) e^{-4\gamma a} = 0$$

Symmetric Configuration

When the cladding and substrate media are the same, i.e., $\varepsilon_c = \varepsilon_s$, then $\alpha_c = \alpha_s$ and $p_c = p_s$, and the characteristic equations (10.1.7) simplify further. Indeed, we note from Eq. (10.1.14) that $e^{4\psi} = 1$, which implies the two basic cases $\psi = 0$ and $\psi = j\pi/2$. Noting the identity $\tanh(x \pm j\pi/2) = \coth(x)$, the two characteristic equations (10.1.7) become a single one, given as follows for the two cases of ψ :

$$\begin{aligned} \text{(even)} \quad \psi = 0, \quad e^{2\gamma a} &= +\frac{\gamma - p_c \alpha_c}{\gamma + p_c \alpha_c}, \quad \tanh(\gamma a) = -\frac{p_c \alpha_c}{\gamma} \\ \text{(odd)} \quad \psi = \frac{j\pi}{2}, \quad e^{2\gamma a} &= -\frac{\gamma - p_c \alpha_c}{\gamma + p_c \alpha_c}, \quad \coth(\gamma a) = -\frac{p_c \alpha_c}{\gamma} \end{aligned} \quad (10.1.22)$$

The labeling as even or odd (symmetric or antisymmetric) refers to the symmetry or antisymmetry of the *transverse* electric and magnetic fields $E_x(x)$ and $H_y(x)$ as functions of x within the film region.[†] Indeed, it is evident from Eq. (10.1.4) that if $\psi = 0$, then $E_x(x)$ is proportional to $\cosh(\gamma x)$, an even function of x , and if $\psi = j\pi/2$, then because of the identity $\cosh(x + j\pi/2) = j \sinh(x)$, the field $E_x(x)$ will be proportional to $\sinh(\gamma x)$, an odd function of x .

The nomenclature carries over loosely to the asymmetric waveguide that has $\varepsilon_c \neq \varepsilon_s$, with the ‘‘symmetric/even’’ case corresponding to a value of ψ that is predominantly real, and the ‘‘antisymmetric/odd’’ one corresponding to a ψ whose imaginary part is near $j\pi/2$. Examples of these are given below.

[†]In the literature, sometimes the labeling is reversed, referring instead to the symmetry or antisymmetry of the *longitudinal* field $E_z(x)$ which is the opposite of that of $E_x(x)$.

Power Transfer

The z -component of the Poynting vector gives the power flow in the z -direction per unit xy area. Since $E_x = \eta_{TM} H_y$, $\eta_{TM} = \beta / (\omega \varepsilon)$, we obtain from Eq. (10.1.6),

$$\mathcal{P}_z(x) = \frac{1}{2} \text{Re}[E_x(x) H_y^*(x)] = \frac{1}{2} \text{Re}(\eta_{TM}) |H_y(x)|^2$$

and since $k_0 = \omega/c_0$, we may write η_{TM} in the form,

$$\eta_{TM} = \frac{\beta}{\omega \varepsilon} = \frac{\beta}{k_0 c_0 \varepsilon_0 \varepsilon} = \eta_0 \frac{\beta}{k_0 \varepsilon}, \quad \eta_0 = \sqrt{\frac{\mu_0}{\varepsilon_0}}$$

The Poynting vector is then in the three regions,

$$\mathcal{P}_z(x) = \frac{1}{2} \eta_0 |H_0|^2 \begin{cases} \text{Re} \left(\frac{\beta}{k_0 \varepsilon_f} \right) |\cosh(\gamma x + \psi)|^2, & |x| \leq a \\ \text{Re} \left(\frac{\beta}{k_0 \varepsilon_c} \right) |\cosh(\gamma a + \psi)|^2 e^{-2\alpha_{cR}(x-a)}, & x \geq a \\ \text{Re} \left(\frac{\beta}{k_0 \varepsilon_s} \right) |\cosh(\gamma a - \psi)|^2 e^{2\alpha_{sR}(x+a)}, & x \leq -a \end{cases} \quad (10.1.23)$$

where α_{cR}, α_{sR} denote the real parts of α_c, α_s . Integrating (10.3.2) over an xy area, $dS = dx \cdot (1 \text{ m along } y)$, we obtain the powers transmitted (per unit y -length) along the z -direction within each region. Dropping the overall unimportant factor $\eta_0 |H_0|^2/2$, and denoting the real and imaginary parts of γ, ψ by γ_R, ψ_R and γ_I, ψ_I , we have,

$$\begin{aligned} P_f &= \text{Re} \left(\frac{\beta}{k_0 \varepsilon_f} \right) \left[\frac{\sinh(2\gamma_R a) \cosh(2\psi_R)}{2\gamma_R} + \frac{\sin(2\gamma_I a) \cos(2\psi_I)}{2\gamma_I} \right] \\ P_c &= \text{Re} \left(\frac{\beta}{k_0 \varepsilon_c} \right) \frac{|\cosh(\gamma a + \psi)|^2}{2\alpha_{cR}} \\ P_s &= \text{Re} \left(\frac{\beta}{k_0 \varepsilon_s} \right) \frac{|\cosh(\gamma a - \psi)|^2}{2\alpha_{sR}} \end{aligned} \quad (10.1.24)$$

and the net power is,

$$P = P_f + P_c + P_s \quad (10.1.25)$$

These expressions must be multiplied by $e^{-2\beta_I z}$, where $\beta = \beta_R - j\beta_I$, for the attenuation of power with propagation distance z . These also apply to the lossless case, where we must replace $\sin(2\gamma_I a) / (2\gamma_I) |_{\gamma_I=0} = a$. Note that we may also write,

$$|\cosh(\gamma a \pm \psi)|^2 = \frac{1}{2} [\cosh(2\gamma_R a \pm 2\psi_R) + \cos(2\gamma_I a \pm 2\psi_I)]$$

The MATLAB function, `pwgpower`, implements Eqs. (10.1.24) and (10.1.25):

`[P, Pf, Pc, Ps] = pwgpower(a, ef, ec, es, be, m);`

% transmitted power in plasmonic waveguide

where a and β are in units of k_0^{-1} and k_0 , respectively, and $m = 0, 1$, for TM_0 or TM_1 modes, where ψ is calculated from Eq. (10.1.15).

10.2 Single Metal-Dielectric Interface

The case of a single metal-dielectric interface can be thought of as the limit of a DMD configuration when the film thickness tends to infinity, $a \rightarrow \infty$. It is depicted in Fig. 10.2.1.

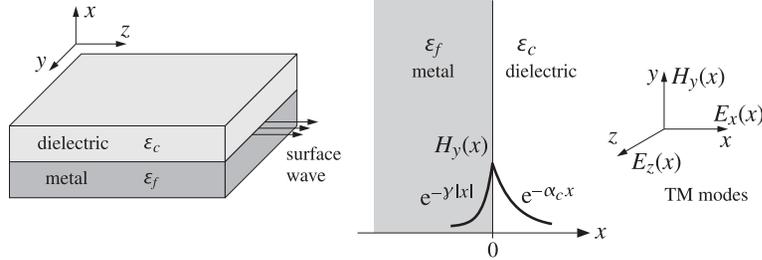


Fig. 10.2.1 Surface plasmon wave propagating along metal-dielectric interface.

Because y has positive real part, the left-hand side of Eq. (10.1.11) tends to infinity as $a \rightarrow \infty$, and this requires the vanishing of the denominator of the right-hand side, that is, one of the conditions:

$$y = -p_c \alpha_c, \quad \text{or,} \quad y = -p_s \alpha_s \quad (10.2.1)$$

Let us consider the first one. Because ϵ_f is a metal with a negative real part, the condition $y = -p_c \alpha_c = -\epsilon_f \alpha_c / \epsilon_c$ is consistent with the requirement that both y and α_c have positive real parts. Assuming that the interface is positioned at $x = 0$, the longitudinal electric field will be for TM modes:

$$E_z(x) = \begin{cases} E_0 e^{-\alpha_c x}, & x \geq 0 \\ E_0 e^{y x} = E_0 e^{-y|x|}, & x \leq 0 \end{cases} \quad (10.2.2)$$

with complete space-time dependence $E_z(x) e^{j\omega t - j\beta z}$. The E_x component will be:

$$E_x(x) = \begin{cases} -\frac{j\beta}{-\alpha_c^2} \partial_x E_z = -E_0 \frac{j\beta}{\alpha_c} e^{-\alpha_c x}, & x \geq 0 \\ -\frac{j\beta}{-y^2} \partial_x E_z = E_0 \frac{j\beta}{y} e^{y x}, & x \leq 0 \end{cases} \quad (10.2.3)$$

The corresponding magnetic field is obtained from $H_y = E_x / \eta_{TM}$, where $\eta_{TM} = \beta / (\omega \epsilon)$,

$$H_y(x) = \begin{cases} -j\omega \epsilon_0 E_0 \frac{\epsilon_c}{\alpha_c} e^{-\alpha_c x}, & x \geq 0 \\ j\omega \epsilon_0 E_0 \frac{\epsilon_f}{y} e^{y x}, & x \leq 0 \end{cases} \quad (10.2.4)$$

The continuity of either $D_x = \epsilon E_x$, or H_y at $x = 0$ gives the first condition in Eq. (10.2.1). These results are equivalent to those of Sec. 7.11. Indeed, the propagation wavenumber β , as well as y, α_c , are obtained from:

$$y = -\frac{\epsilon_f \alpha_c}{\epsilon_c} \Rightarrow \beta^2 - k_0^2 \epsilon_f = \frac{\epsilon_f^2}{\epsilon_c^2} (\beta^2 - k_0^2 \epsilon_c) \quad (10.2.5)$$

or, solving for β, y, α_c ,

$$\beta = k_0 \sqrt{\frac{\epsilon_f \epsilon_c}{\epsilon_f + \epsilon_c}}, \quad y = \frac{-k_0 \epsilon_f}{\sqrt{-\epsilon_f - \epsilon_c}}, \quad \alpha_c = \frac{k_0 \epsilon_c}{\sqrt{-\epsilon_f - \epsilon_c}} \quad (10.2.6)$$

where the square root signs have been selected to satisfy the requirements that β have negative imaginary part and β, y, α_c , positive real parts. To see this in more detail, set $\epsilon_f = -\epsilon_R - j\epsilon_I$, with $\epsilon_R > 0$, and assume $\epsilon_R > \epsilon_c$, as in Sec. 7.11. Then, a first-order calculation in ϵ_I yields the expressions:

$$\begin{aligned} \beta &= k_0 \sqrt{\frac{\epsilon_R \epsilon_c}{\epsilon_R - \epsilon_c}} \left[1 - j \frac{\epsilon_c \epsilon_I}{2\epsilon_R (\epsilon_R - \epsilon_c)} \right] \\ y &= \frac{k_0 \epsilon_R}{\sqrt{\epsilon_R - \epsilon_c}} \left[1 + j \frac{(\epsilon_R - 2\epsilon_c) \epsilon_I}{2\epsilon_R (\epsilon_R - \epsilon_c)} \right] \\ \alpha_c &= \frac{k_0 \epsilon_c}{\sqrt{\epsilon_R - \epsilon_c}} \left[1 - j \frac{\epsilon_I}{2(\epsilon_R - \epsilon_c)} \right] \end{aligned} \quad (10.2.7)$$

Eq. (10.2.7) shows explicitly how the condition $\epsilon_R > \epsilon_c$ guarantees the existence of plasmonic waves with $\text{Re}(\beta) > 0$ and $\text{Im}(\beta) < 0$. We note also that $\text{Re}(y) / \text{Re}(\alpha_c) = \epsilon_R / \epsilon_c$, which is typically much greater than unity, $\epsilon_R / \epsilon_c \gg 1$. Therefore, the attenuation length within the metal is typically much shorter than that in the dielectric, i.e., $1 / \text{Re}(y) \ll 1 / \text{Re}(\alpha_c)$. This is depicted in Figs. 10.2.1 and 10.2.2.

Example 10.2.1: Following Example 7.11.1, we use the value $\epsilon_f = -16 - 0.5j$ for silver at $\lambda_0 = 0.632 \mu\text{m}$, and air $\epsilon_c = 1$. Then, $k_0 = 2\pi / \lambda_0 = 9.94 \text{ rad}/\mu\text{m}$ and Eqs. (10.2.6) give the following values for the parameters β, y, α_c and corresponding propagation and penetration distances:

$$\begin{aligned} \beta &= 10.2674 - 0.0107j \text{ rad}/\mu\text{m}, & L_z &= -\frac{1}{\text{Im}(\beta)} = 93.5969 \mu\text{m} \\ y &= 41.0755 + 0.5989j \text{ rad}/\mu\text{m}, & L_f &= \frac{1}{\text{Re}(y)} = 0.0243 \mu\text{m} \\ \alpha_c &= 2.5659 - 0.0428j \text{ rad}/\mu\text{m}, & L_c &= \frac{1}{\text{Re}(\alpha_c)} = 0.3897 \mu\text{m} \end{aligned}$$

Thus, the fields extend more into the dielectric than the metal, but at either side they are confined to distances that are less than their free-space wavelength. The transverse magnetic field $H_y(x)$, which is continuous at the interface, is shown in Fig. 10.2.2. It was plotted with the MATLAB code:

```
ec = 1; ef = -16-0.5*j; la0 = 0.632; k0 = 2*pi/la0;
b = k0*sqrt(ef*ec/(ef+ec));
```

```

ga = -k0*ef/sqrt(-ef-ec);      % or, ga = sqrt(b^2 - k0^2*ef);
ac = k0*ec/sqrt(-ef-ec);      % or, ac = sqrt(b^2 - k0^2*ec);
x = linspace(-0.4, 1, 141);
Hy = abs(exp(ga*x).*(x<0) + exp(-ac*x).*(x>=0));
plot(x,Hy);

```

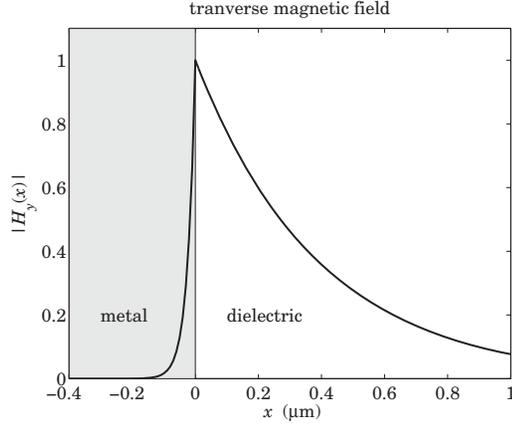


Fig. 10.2.2 Surface plasmon magnetic field profile.

10.3 Power Transfer, Energy & Group Velocities

Replacing $\beta = \beta_R - j\beta_I$, $\epsilon_f = -\epsilon_R - j\epsilon_I$, and $\alpha_c = \alpha_R - j\alpha_I$, $\gamma = \gamma_R + j\gamma_I$ into Eq. (10.2.7), we note also the following relationships, with the second following by equating real parts in Eq. (10.2.5):

$$2\beta_R\beta_I = 2\alpha_R\alpha_I = \frac{k_0^2 \epsilon_c^2 \epsilon_I}{(\epsilon_R - \epsilon_c)^2 + \epsilon_I^2}, \quad \epsilon_c \gamma_R = \epsilon_R \alpha_R + \epsilon_I \alpha_I \quad (10.3.1)$$

and since $\epsilon_I \geq 0$, Eq. (10.3.1) implies that $\beta_R \geq 0$ and $\alpha_I \geq 0$ since $\beta_I \geq 0$ and $\alpha_R \geq 0$.

The power flow along the propagation direction is described by the z -component of the Poynting vector. From Eqs. (10.2.3) and (10.2.4), we obtain,

$$\mathcal{P}_z(x) = \frac{1}{2} \operatorname{Re}[E_x(x)H_y^*(x)] = \frac{1}{2} \omega \epsilon_0 |E_0|^2 \begin{cases} \frac{\operatorname{Re}[\beta \epsilon_c^*]}{|\alpha_c|^2} e^{-2\alpha_R x}, & x \geq 0 \\ \frac{\operatorname{Re}[\beta \epsilon_f^*]}{|\gamma|^2} e^{2\gamma_R x}, & x \leq 0 \end{cases} \quad (10.3.2)$$

Integrating (10.3.2) over an xy area, $dS = dx \cdot (1 \text{ m along } y)$, we obtain the net power transmitted (per unit y -length) along the z -direction:

$$P_T = \int_{-\infty}^{\infty} \mathcal{P}_z dx = \frac{1}{4} \omega \epsilon_0 |E_0|^2 \left[\frac{\beta_R \epsilon_c}{\alpha_R |\alpha_c|^2} - \frac{\beta_R \epsilon_R - \beta_I \epsilon_I}{\gamma_R |\gamma|^2} \right] \quad (10.3.3)$$

where the first term represents the power flow within the dielectric, and the second, the power flow within the metal. Since typically, $\beta_R \epsilon_R \gg \beta_I \epsilon_I$, the second term is negative, representing power flow along the negative z direction within the metal. However, the sum of the two terms is positive, so that the net power flow is along the positive z -direction. In fact, since also, $|\gamma| \gg |\alpha_c|$, the second term will be much smaller than the first one, which is to be expected since the fields tend to be attenuate more quickly within the metal than in the dielectric. Indeed, for the values given in Example 10.2.1, we find the terms in the brackets to be:

$$\left[\frac{\beta_R \epsilon_c}{\alpha_R |\alpha_c|^2} - \frac{\beta_R \epsilon_R - \beta_I \epsilon_I}{\gamma_R |\gamma|^2} \right] = 0.6076 - 0.0024 = 0.6053$$

Using the second of Eqs. (10.3.1), one can show that the term in brackets in (10.3.3) becomes

$$\left[\frac{\beta_R \epsilon_c}{\alpha_R |\alpha_c|^2} - \frac{\beta_R \epsilon_R - \beta_I \epsilon_I}{\gamma_R |\gamma|^2} \right] = \frac{\beta_R \alpha_R \epsilon_R (\epsilon_R^2 - \epsilon_c^2 + \epsilon_I^2) + \beta_R \alpha_I \epsilon_I (\epsilon_R^2 + \epsilon_I^2) + \beta_I \alpha_R \epsilon_I \epsilon_c^2}{\alpha_R \gamma_R \epsilon_c^2 |\gamma|^2}$$

which is positive since we assumed $\epsilon_R > \epsilon_c$, and all other terms are non-negative. The power, of course, attenuates exponentially with distance z , and the expression (10.3.3) for P_T must be multiplied by the factor $e^{-2\beta_I z}$.

The positivity of P_T implies also the positivity of the energy velocity, $v_{\text{en}} = P_T/W'$, where W' is the energy density per unit z -length (see Sec. 9.7), which is always positive. In the *lossless* case, we show below that the group velocity, $v_g = d\omega/d\beta$, is equal to the energy velocity, and hence v_g will also be positive. This can also be seen from Fig. 7.11.2 depicting the ω - β dispersion relationship, which has a positive slope $d\omega/d\beta$.

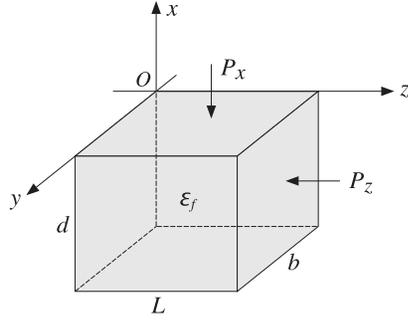
However, when losses are taken into account, it can be shown [973], using for example the Drude model for the metal permittivity, that there are frequency regions in which the group velocity is negative. Similar situations arise in MDM and DMD configurations in which, depending on the media permittivities and layer thicknesses, there can exist modes that have negative group velocities—see, for example, Refs. [943] and [965-981].

Next, we determine the x -component of the Poynting vector, which describes power flow from the dielectric into the metal, and we show that the power entering the metal and the power flowing in the z -direction in the metal are completely dissipated into heat as ohmic and dielectric losses. We have,

$$\mathcal{P}_x(x) = -\frac{1}{2} \operatorname{Re}[E_z(x)H_y^*(x)] = \frac{1}{2} \omega \epsilon_0 |E_0|^2 \begin{cases} -\operatorname{Im}\left(\frac{\epsilon_c}{\alpha_c}\right) e^{-2\alpha_R x}, & x \geq 0 \\ \operatorname{Im}\left(\frac{\epsilon_f}{\gamma}\right) e^{2\gamma_R x}, & x \leq 0 \end{cases} \quad (10.3.4)$$

The power flow \mathcal{P}_x is continuous across the interface at $x = 0$ because of the condition $\epsilon_c/\alpha_c = -\epsilon_f/\gamma$. The negative sign on the dielectric side simply means that power flows towards the negative x direction into the metal. The z -dependence of \mathcal{P}_x is obtained by multiplying Eq. (10.3.4) by $e^{-2\beta_I z}$.

To account for the power dissipation within the metal, we consider a rectangular volume of sides L, d, b along the z, x, y directions lying within the metal below the yz interface plane, as shown below.



The net power entering and leaving the top and bottom sides is obtained by restoring the factor $e^{-2\beta_I z}$ and integrating (10.3.4) over the $L \times b$ area and subtracting the values at $x = 0$ and $x = -d$,

$$P_{x,\text{net}} = \frac{1}{2} \omega \epsilon_0 |E_0|^2 \text{Im} \left(\frac{\epsilon_c}{\alpha_c} \right) \frac{(1 - e^{-2\beta_I L})}{2\beta_I} (1 - e^{-2\gamma_R d}) b \quad (10.3.5)$$

Similarly, the net power entering and leaving the left and right sides is obtained by integrating (10.3.2) over the area $d \times b$ and subtracting the values at $z = 0$ and $z = L$,

$$P_{z,\text{net}} = -\frac{1}{2} \omega \epsilon_0 |E_0|^2 \frac{\text{Re}[\beta \epsilon_f^*]}{|y|^2} \frac{(1 - e^{-2\gamma_R d})}{2\gamma_R} (1 - e^{-2\beta_I L}) b \quad (10.3.6)$$

The sum of these two powers must equal the power loss within the volume $L \times d \times b$,

$$P_{\text{loss}} = \frac{1}{2} \text{Re}(j\omega \epsilon_f) \int_V \mathbf{E} \cdot \mathbf{E}^* dV = \frac{1}{2} \epsilon_0 \omega \epsilon_I \int_V (|E_x|^2 + |E_z|^2) dV$$

Restoring the factor $e^{-j\beta z}$ in E_x, E_z , we obtain from Eqs. (10.2.2) and (10.2.3):

$$P_{\text{loss}} = \frac{1}{2} \omega \epsilon_0 |E_0|^2 \epsilon_I \left(1 + \frac{|\beta|^2}{|y|^2} \right) \frac{(1 - e^{-2\gamma_R d}) (1 - e^{-2\beta_I L})}{2\gamma_R \cdot 2\beta_I} b \quad (10.3.7)$$

Energy conservation requires that, $P_{x,\text{net}} + P_{z,\text{net}} = P_{\text{loss}}$. Canceling some common factors, this condition is equivalent to,

$$2\gamma_R \text{Im} \left(\frac{\epsilon_c}{\alpha_c} \right) - 2\beta_I \frac{\text{Re}[\beta \epsilon_f^*]}{|y|^2} = \epsilon_I \left(1 + \frac{|\beta|^2}{|y|^2} \right) \quad (10.3.8)$$

The proof of this result is left for Problem 10.1, however, we note that this calculation is an example of the more general result stated in Problem 1.5. We note also that (10.3.8) is trivially satisfied in the lossless case that has $\beta_I = \alpha_I = \epsilon_I = 0$.

We conclude this section by showing the equality $v_g = v_{\text{en}}$ in the lossless case and assuming an arbitrary frequency dependence of the (real-valued) metal permittivity.

We assume that $\epsilon_f = -\epsilon_R$ is a real negative function of frequency ω , ϵ_c is positive constant such that $\epsilon_R > \epsilon_c$, and both media are non-magnetic, $\mu = \mu_0$. Then, the

quantities β, α_c, γ are given by the real parts of Eqs. (10.2.7). Inserting these expressions into P_T , we find:

$$P_T = \frac{1}{4} \epsilon_0 |E_0|^2 \frac{c_0^2 (\epsilon_R + \epsilon_c) (\epsilon_R - \epsilon_c)^2}{\omega \epsilon_R^{3/2} \epsilon_c^{3/2}} \quad (10.3.9)$$

For lossless non-magnetic media, the time-averaged energy density is given by

$$w_{\text{en}} = \frac{1}{4} (\omega \epsilon)' |E|^2 + \frac{1}{4} \mu_0 |\mathbf{H}|^2, \quad (\omega \epsilon)' \equiv \frac{d(\omega \epsilon)}{d\omega}$$

or,

$$w_{\text{en}} = \frac{1}{4} (\omega \epsilon)' (|E_x|^2 + |E_z|^2) + \frac{1}{4} \mu_0 |H_y|^2$$

Using Eqs. (10.2.2)-(10.2.4), we obtain,

$$w_{\text{en}} = \frac{1}{4} \epsilon_0 |E_0|^2 \begin{cases} \left[\epsilon_c \left(1 + \frac{\beta^2}{\alpha_c^2} \right) + \frac{k_0^2 \epsilon_c^2}{\alpha_c^2} \right] e^{-\alpha_c x}, & x \geq 0 \\ \left[(\omega \epsilon_f)' \left(1 + \frac{\beta^2}{y^2} \right) + \frac{k_0^2 \epsilon_f^2}{y^2} \right] e^{\gamma x}, & x \leq 0 \end{cases}$$

Integrating over x , we find the energy density per unit z -length and unit y -length:

$$W' = \int_{-\infty}^{\infty} w_{\text{en}} dx = \frac{1}{4} \epsilon_0 |E_0|^2 \left\{ \left[\epsilon_c \left(1 + \frac{\beta^2}{\alpha_c^2} \right) + \frac{k_0^2 \epsilon_c^2}{\alpha_c^2} \right] \frac{1}{2\alpha_c} + \left[(\omega \epsilon_f)' \left(1 + \frac{\beta^2}{y^2} \right) + \frac{k_0^2 \epsilon_f^2}{y^2} \right] \frac{1}{2\gamma} \right\}$$

The derivative term $(\omega \epsilon_f)'$ can be replaced by

$$(\omega \epsilon_f)' = -(\omega \epsilon_R)' = -\epsilon_R - \omega \epsilon_R', \quad \epsilon_R' \equiv \frac{d\epsilon_R}{d\omega}$$

By differentiating with respect to ω both sides of the equation,

$$\beta^2 = \frac{\omega^2}{c_0^2} \frac{\epsilon_R \epsilon_c}{\epsilon_R - \epsilon_c}$$

we may relate the derivative ϵ_R' to the derivative $\beta' = d\beta/d\omega$, as follows:

$$\epsilon_R' = 2 \left(\frac{\beta}{\omega} - \beta' \right) \frac{\epsilon_R^2 \omega^2}{\beta^3 c_0^2} \quad (10.3.10)$$

Substituting this into W' , we find after some algebra,

$$W' = \frac{1}{4} \epsilon_0 |E_0|^2 \frac{\beta' c_0^2 (\epsilon_R + \epsilon_c) (\epsilon_R - \epsilon_c)^2}{\omega \epsilon_R^{3/2} \epsilon_c^{3/2}} \quad (10.3.11)$$

This implies that the energy velocity will be equal to the group velocity:

$$v_{\text{en}} = \frac{P_T}{W'} = \frac{1}{\beta'} = \frac{d\omega}{d\beta} = v_g$$

10.4 MDM Configuration - Lossless Case

An MDM waveguide is depicted in Fig. 10.4.1. To gain an understanding of the properties of the propagating modes, we will assume initially that all three media are *lossless*, with ϵ_f real positive, and ϵ_c, ϵ_s real negative, so that we can set $\epsilon_c = -|\epsilon_c|$ and $\epsilon_s = -|\epsilon_s|$. Without loss of generality, we will also assume that $|\epsilon_c| \geq |\epsilon_s|$. Following [943], we distinguish three cases:

1. $|\epsilon_s| \leq \epsilon_f \leq |\epsilon_c| \Rightarrow |p_c| \leq 1 \leq |p_s|$
2. $|\epsilon_s| \leq |\epsilon_c| \leq \epsilon_f \Rightarrow 1 \leq |p_c| \leq |p_s|$
3. $\epsilon_f \leq |\epsilon_s| \leq |\epsilon_c| \Rightarrow |p_c| \leq |p_s| \leq 1$

where $p_c = \epsilon_f/\epsilon_c$ and $p_s = \epsilon_f/\epsilon_s$. These cases define three regions labeled 1,2,3, on the p_c, p_s parameter plane, as shown in Fig. 10.4.2. Regions 1',2',3' are obtained by interchanging the roles of ϵ_c and ϵ_s . For typical metals like silver and gold at optical frequencies, the relevant region is 3. For example, the permittivity of silver at $\lambda = 650$ nm is $\epsilon_s = -15.48 - 1.15j$, while typically, ϵ_f is less than 3.

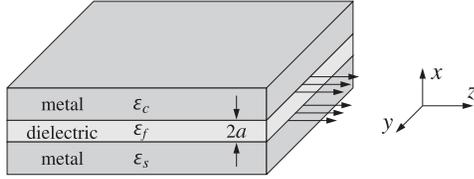


Fig. 10.4.1 MDM plasmonic waveguide.

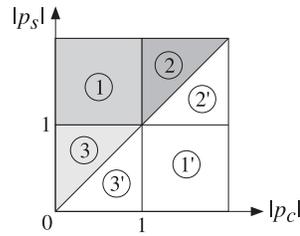


Fig. 10.4.2 Possible regions in MDM lossless case.

The characteristic equations (10.1.7), (10.1.11) and (10.1.13) can be written in the forms,

$$\tanh(\gamma a + \psi) = -\frac{p_c \alpha_c}{\gamma} = \frac{|p_c| \alpha_c}{\gamma} \quad (10.4.2)$$

$$\tanh(\gamma a - \psi) = -\frac{p_s \alpha_s}{\gamma} = \frac{|p_s| \alpha_s}{\gamma}$$

$$e^{4\gamma a} = \frac{(\gamma - p_c \alpha_c)(\gamma - p_s \alpha_s)}{(\gamma + p_c \alpha_c)(\gamma + p_s \alpha_s)} = \frac{|p_c| \alpha_c + \gamma}{|p_c| \alpha_c - \gamma} \cdot \frac{|p_s| \alpha_s + \gamma}{|p_s| \alpha_s - \gamma} \quad (10.4.3)$$

$$\tanh(2\gamma a) = -\frac{\gamma(p_c \alpha_c + p_s \alpha_s)}{\gamma^2 + p_c \alpha_c p_s \alpha_s} = \frac{\gamma(|p_c| \alpha_c + |p_s| \alpha_s)}{\gamma^2 + |p_c| \alpha_c |p_s| \alpha_s} \quad (10.4.4)$$

We note also that Eqs. (10.1.9) can be written as,

$$\begin{aligned} \gamma &= \sqrt{\beta^2 - k_0^2 \epsilon_f} \\ \alpha_c &= \sqrt{\beta^2 - k_0^2 \epsilon_c} = \sqrt{\gamma^2 + k_0^2 (\epsilon_f + |\epsilon_c|)} \\ \alpha_s &= \sqrt{\beta^2 - k_0^2 \epsilon_s} = \sqrt{\gamma^2 + k_0^2 (\epsilon_f + |\epsilon_s|)} \end{aligned} \quad (10.4.5)$$

Because we are looking for plasmonic solutions that have real and positive $\beta, \gamma, \alpha_c, \alpha_s$, it follows that β and γ must be restricted to the ranges $\beta \geq k_0 \sqrt{\epsilon_f}$ and $\gamma \geq 0$. Taking the limit of (10.4.4) as $\gamma \rightarrow 0$, and using the Taylor series approximation $\tanh(x) \simeq x$, valid for small x , we obtain the cutoff thickness of the dielectric layer,

$$\tanh(2\gamma a) \simeq 2\gamma a = \frac{\gamma(|p_c| \alpha_c + |p_s| \alpha_s)}{\gamma^2 + |p_c| \alpha_c |p_s| \alpha_s} \Rightarrow 2a = \frac{1}{|p_c| \alpha_c} + \frac{1}{|p_s| \alpha_s} \Big|_{\gamma=0}, \quad \text{or,}$$

$$2k_0 a_{\text{cutoff}} = \frac{|\epsilon_c|}{\epsilon_f \sqrt{\epsilon_f + |\epsilon_c|}} + \frac{|\epsilon_s|}{\epsilon_f \sqrt{\epsilon_f + |\epsilon_s|}} \quad (10.4.6)$$

But the upper limit of β , and whether a_{cutoff} is an upper or a lower cutoff, will depend on which region in Fig. 10.4.2 we are. A plot of β versus the film thickness a can be made by solving Eq. (10.4.4) and varying β over its allowed range,

$$a = \frac{1}{2\gamma} \operatorname{atanh} \left(-\frac{\gamma(p_c \alpha_c + p_s \alpha_s)}{\gamma^2 + p_c \alpha_c p_s \alpha_s} \right) \quad (10.4.7)$$

Depending on the mode and region, a may be an increasing or a decreasing function of β . This can be determined from the derivative of a with respect to β , given by,

$$\frac{da}{d\beta} = \frac{\beta}{\gamma} \frac{da}{d\gamma} = \frac{\beta |p_c| (\epsilon_f + |\epsilon_c|)}{2\gamma^2 \alpha_c (p_c^2 \alpha_c^2 - \gamma^2)} + \frac{\beta |p_s| (\epsilon_f + |\epsilon_s|)}{2\gamma^2 \alpha_s (p_s^2 \alpha_s^2 - \gamma^2)} - \frac{\beta a}{\gamma^2} \quad (10.4.8)$$

Region 1

Region-1 is defined by the conditions, $|\epsilon_s| \leq \epsilon_f \leq |\epsilon_c|$, or, $|p_c| \leq 1 \leq |p_s|$. Because $|p_s| \geq 1$ and $\alpha_s > \gamma$, it follows that $|p_s| \alpha_s / \gamma \geq 1$. But this means that the equation, $\tanh(\gamma a - \psi) = |p_s| \alpha_s / \gamma \geq 1$, can be satisfied only if ψ has a $j\pi/2$ imaginary part, that is, $\psi = \chi + j\pi/2$, with real χ , so that

$$\tanh(\gamma a - \psi) = \tanh \left(\gamma a - \chi - \frac{j\pi}{2} \right) = \coth(\gamma a - \chi) = \frac{|p_s| \alpha_s}{\gamma} \geq 1$$

But then $|p_c| \alpha_c / \gamma$ is also forced to be greater than one because,

$$\tanh(\gamma a + \psi) = \tanh \left(\gamma a + \chi + \frac{j\pi}{2} \right) = \coth(\gamma a + \chi) = \frac{|p_c| \alpha_c}{\gamma}$$

Thus, with y, χ real, we have $|\coth(ya + \chi)| \geq 1$ so that $|p_c|\alpha_c/y \geq 1$. The same conclusion is reached by inspecting Eq. (10.4.3), which requires both denominator factors to have the same sign since e^{4ya} is real and positive. The condition $|p_c|\alpha_c/y \geq 1$ imposes an upper limit on β , indeed, we have $y^2 \leq p_c^2\alpha_c^2$, or,

$$\beta^2 - k_0^2\epsilon_f \leq p_c^2(\beta^2 + k_0^2|\epsilon_c|) \Rightarrow \beta^2 \leq k_0^2 \frac{p_c^2|\epsilon_c| + \epsilon_f}{1 - p_c^2} = k_0^2 \frac{|\epsilon_c|\epsilon_f}{|\epsilon_c| - \epsilon_f}, \text{ or,}$$

$$\beta \leq \beta_{c,\infty} \equiv k_0 \sqrt{\frac{|\epsilon_c|\epsilon_f}{|\epsilon_c| - \epsilon_f}} = k_0 \sqrt{\frac{\epsilon_c\epsilon_f}{\epsilon_c + \epsilon_f}} \quad (10.4.9)$$

The limit $\beta_{c,\infty}$ is recognized as the wavenumber of a surface plasmon at the ϵ_f - ϵ_c dielectric-metal interface obtained in the limit of infinite thickness for ϵ_f , in fact, the interface condition $y = |p_c|\alpha_c = -p_c\alpha_c$ is realized exactly at $\beta = \beta_{c,\infty}$. Thus, the allowed range of β is,

$$k_0\sqrt{\epsilon_f} \leq \beta \leq \beta_{c,\infty} \quad (\text{region 1}) \quad (10.4.10)$$

Moreover, for this range of β , the film thickness a is an increasing function of β , so that its range will be $a_{\text{cutoff}} \leq a < \infty$, therefore, a_{cutoff} will be a lower cutoff. Next, we show that the corresponding field solution will be an antisymmetric-like TM₁ mode. The characteristic equations,

$$\coth(ya + \chi) = \frac{|p_c|\alpha_c}{y}, \quad \coth(ya - \chi) = \frac{|p_s|\alpha_s}{y}$$

can be inverted,

$$\begin{aligned} \tanh(ya + \chi) = \frac{y}{|p_c|\alpha_c} \leq 1 & \Rightarrow ya + \chi = \operatorname{atanh}\left(\frac{y}{|p_c|\alpha_c}\right) \\ \tanh(ya - \chi) = \frac{y}{|p_s|\alpha_s} \leq 1 & \Rightarrow ya - \chi = \operatorname{atanh}\left(\frac{y}{|p_s|\alpha_s}\right) \end{aligned}$$

and separated,

$$\begin{aligned} ya &= \frac{1}{2} \operatorname{atanh}\left(\frac{y}{|p_c|\alpha_c}\right) + \frac{1}{2} \operatorname{atanh}\left(\frac{y}{|p_s|\alpha_s}\right) \\ \psi &= \chi + \frac{j\pi}{2} = \frac{1}{2} \operatorname{atanh}\left(\frac{y}{|p_c|\alpha_c}\right) - \frac{1}{2} \operatorname{atanh}\left(\frac{y}{|p_s|\alpha_s}\right) + \frac{j\pi}{2} \end{aligned} \quad (10.4.11)$$

These have exactly the form of Eqs. (10.1.16) with $m = 1$, so that we have a TM₁ mode. The magnetic field is given by Eq. (10.1.6),

$$H_y(x) = \begin{cases} H_0 \cosh(yx + \psi), & |x| \leq a \\ H_0 \cosh(ya + \psi)e^{-\alpha_c(x-a)}, & x \geq a \\ H_0 \cosh(ya - \psi)e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (10.4.12)$$

where $H_0 = j\omega\epsilon_0 E_0 \epsilon_f / y$. Using the identity $\cosh(x + j\pi/2) = j \sinh(x)$, we see that within the dielectric film, H_y resembles an antisymmetric solution (provided χ is small),

$$H_y(z) = H_0 \cosh(yx + \psi) = H_0 \cosh(yx + \chi + j\pi/2) = jH_0 \sinh(yx + \chi)$$

Given a film thickness a , the characteristic equation (10.4.4) can be solved by writing it in the following form,

$$y = -\frac{y^2 + p_c\alpha_c p_s\alpha_s}{p_c\alpha_c + p_s\alpha_s} \tanh(2ya)$$

and replacing it by the iteration,

$$y_{n+1} = -\frac{y_n^2 + p_c\alpha_c p_s\alpha_s}{p_c\alpha_c + p_s\alpha_s} \tanh(2y_n a), \quad n = 0, 1, 2, \dots \quad (10.4.13)$$

and initialized at a value of β that lies somewhere in the interval $k_0\sqrt{\epsilon_f} < \beta \leq \beta_{c,\infty}$ (the value $\beta = k_0\sqrt{\epsilon_f}$ should be not be used to initialize because it corresponds to $y = 0$ and the iteration will remain stuck at $y = 0$.)

The iteration can be stopped when a desired level of accuracy is reached, that is, when $|y_{n+1} - y_n| \leq \text{tol}$, for some desired error tolerance, such as $\text{tol} = 10^{-12}$. Alternative iterative methods for this type of problem can be found in [963].

Example 10.4.1: We choose a region-1 example from Ref. [943] to verify our approach. Consider the permittivity values:

$$\epsilon_c = -2.2^2 = -4, \quad \epsilon_f = 1.5^2 = 2.25, \quad \epsilon_s = -1.3^2 = -1.69$$

and the two normalized film thickness $k_0a = 0.6$ and $k_0a = 0.8$. The following MATLAB code illustrates the iteration (10.4.13), randomly initialized within the interval $k_0\sqrt{\epsilon_f} < \beta \leq \beta_{c,\infty}$, and demonstrates the calculation and plotting of the magnetic field profile.

```
ec = -2^2; ef = 1.5^2; es = -1.3^2; % region-1 has |ec|>ef>|es|
k0 = 1; % use normalized units k0=1
a = 0.6; % half-thickness of dielectric film
% a = 0.8; % uncomment for the case k0*a = 0.8

pc = ef/ec; ps = ef/es; % |pc|=0.5625, |ps|=1.3314, |pc|<1<|ps|

b0 = sqrt(ef); % b0 = 1.5, lower limit of beta
bcinf = sqrt(ef*ec/(ef+ec)); % bcinf = 2.2678, upper limit of beta

rng(101); % initialize random number generator
binit = (bcinf-b0)*rand(1) + b0; % binit = 1.8965, random initial value
ac = sqrt(binit^2 - ec); % initialize ac,as,ga
as = sqrt(binit^2 - es);
ga = sqrt(binit^2 - ef);

tol = 1e-12; n = 0; % tolerance and starting iteration index

while 1 % forever loop
    ga_new = -(ga^2 + pc*ac*ps*as)/(pc*ac+ps*as)*tanh(2*ga*a*k0);
    if abs(ga_new-ga) < tol; break; end
```

```

ga = ga_new;
be = sqrt(ga^2 + ef);
as = sqrt(be^2 - es);
ac = sqrt(be^2 - ec);
n = n+1;
end

% upon exit from loop, print the number of iterations and the effective index beta
n      % n = 203 when k0*a = 0.6, and n = 113 when k0*a = 0.8
be     % be = 1.9394 when k0*a = 0.6, and be = 2.2127 when k0*a = 0.8

psi = atan(-ga/pc/ac)/2 - atanh(-ga/ps/as)/2 + j*pi/2; % psi = 0.3193 + j*pi/2

% computational error of characteristic equation
E = abs(tanh(2*k0*a*ga) + ga*(pc*ac+ps*as)/(ga^2+pc*ac*ps*as)) % 6.95e-13

% be = pwga(2*pi,ef,ec,es,a,binit,1,tol); % alternative calculation using PWGA

% magnetic field profile
x = linspace(-3,3,601)*a; % x in units of a

Hy = j*cosh(ga*k0*a - psi).*exp(k0*as*(x+a)).*(x<-a) + ...
     j*cosh(ga*k0*a + psi).*exp(-k0*ac*(x-a)).*(x>a) + ...
     j*cosh(k0*ga*x + psi).*(abs(x)<=a); % up to an overall constant

fill([-3, -1, -1, -3], [-4 -4, 1, 1], [0.9 0.9 0.9]); hold on
fill([1, 3, 3, 1], [-4 -4, 1, 1], [0.9 0.9 0.9]);

plot(x/a,real(Hy), 'linewidth',2);

axis([-3,3,-4,1]);
xlabel('x/a'); ylabel('H_y(x)');
title(['magnetic field profile, k_0 a = ', num2str(a)]);

line([-3,3],[0,0],'linestyle','--','linewidth',0.5);
line([0,0],[-4,1],'linestyle','--','linewidth',0.5);

text(-2.1,-2,'epsilon_s');
text(-0.6,-2,'epsilon_f');
text(1.9,-2,'epsilon_c');

```

where the quantity, E , measures the computational error of the characteristic equation (10.4.4), and the final n is the number of iterations to converge. The computed values are shown in the comments of the above code segment and printed below.

$k_0 a$	β/k_0	γ/k_0	α_c/k_0	α_s/k_0	ψ	n	E
0.6	1.9394	1.2294	2.7859	2.3348	$0.3193 + j\pi/2$	203	6.95×10^{-13}
0.8	2.2127	1.6266	2.9826	2.5663	$0.7834 + j\pi/2$	113	5.64×10^{-13}

The quantities $\beta, \gamma, \alpha_c, \alpha_s$ are in units of k_0 . The magnetic field is shown in Fig. 10.4.3 for the cases $k_0 a = 0.6$ and $k_0 a = 0.8$. Because $H_0 = j\omega\epsilon_0 E_0 \epsilon_f / \gamma$, we have chosen the (completely arbitrary) constant $H_0 = j$ for the magnetic field. Note that as the thickness increases, the magnetic field tends to be more concentrated on the ϵ_f - ϵ_c dielectric-metal interface, because as we saw, the solution tends to the single-interface solution as $a \rightarrow$

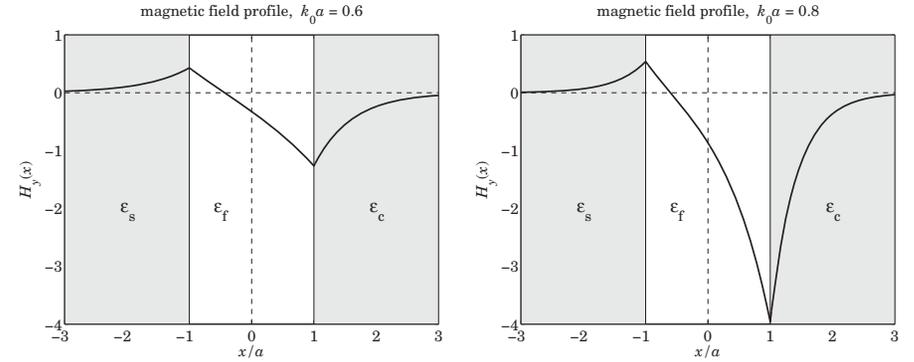


Fig. 10.4.3 Magnetic field profiles, $H_y(x)$ versus x for antisymmetric-like TM_1 mode.

∞ . The ϵ_f - ϵ_s interface cannot support a surface plasmon because it does not meet the necessary condition $|\epsilon_s| > \epsilon_f$, but the ϵ_f - ϵ_c interface does because $|\epsilon_c| > \epsilon_f$.

Fig. 10.4.4 shows a plot of the effective index β/k_0 versus the normalized thickness $k_0 a$. The computed values of β for the two cases $k_0 = 0.6$ and $k_0 a = 0.8$ are shown on the graph, as is the lower cutoff thickness $k_0 a_{\text{cutoff}} = 0.5448$ computed from Eq. (10.4.6). It is evident from this graph that the thickness a is an increasing function of β .

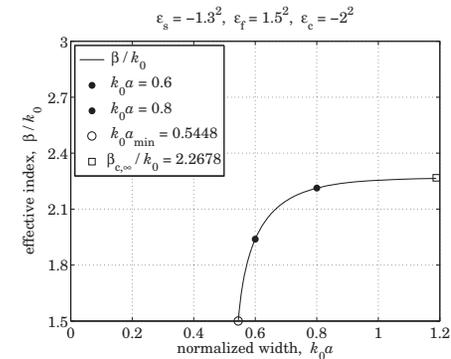


Fig. 10.4.4 Effective index, β/k_0 , versus normalized thickness, $k_0 a$. Region 1.

The following MATLAB code was used to produce the graph.

```

ac = sqrt(ef-ec);
as = sqrt(ef-es);
k0a_min = -1/2*(1/pc/ac + 1/ps/as) % 0.5448 - lower cutoff

be = linspace(b0, 0.9985*bcinf, 401); % plot almost up to bcinf

ga = sqrt(be.^2 - ef);

```

```

ac = sqrt(be.^2 - ec);
as = sqrt(be.^2 - es);

% solve tanh(2*k0*a*ga) = -ga.*(pc*ac+ps*as)./(ga.^2 + pc*ps*ac.*as) for k0*a
k0a = atanh(-ga.*(pc*ac+ps*as)./(ga.^2 + pc*ps*ac.*as))/2./ga;

plot(k0a,be); hold on           % be vs. k0a
plot(k0*a1,be1,'. ');         % k0*a1 = 0.6, be1 = 1.9394, found above
plot(k0*a2,be2,'. ');         % k0*a2 = 0.8, be2 = 2.2127, found above
plot(k0a_min,b0,'o');         % a_cutoff corresponds to b0
plot(k0a(end),bcinf,'s');     % add bcinf to the last value of k0a

axis([0,1.2,1.5,3]);
xlabel('normalized width, {\itk}_0{\itita}');
ylabel('effective index, \beta / {\itk}_0');

```

Also, shown is the asymptotic value $\beta_{c,\infty}$, which was added to the last value of the k_0a array as a reference, although it is actually realized at $k_0a = \infty$. \square

Region 2

Region-2 is defined by the conditions, $|\varepsilon_s| \leq |\varepsilon_c| \leq \varepsilon_f$, or, $1 \leq |p_c| \leq |p_s|$. Therefore, both $|p_c|\alpha_c/\gamma \geq 1$ and $|p_s|\alpha_s/\gamma \geq 1$ are automatically satisfied, and hence, there is no restriction on β other than $\beta \geq k_0\sqrt{\varepsilon_f}$. Thus its range is,

$$\boxed{k_0\sqrt{\varepsilon_f} \leq \beta < \infty} \quad (\text{region 2}) \quad (10.4.14)$$

Moreover, for this range of β , the thickness a is a decreasing function of β , so that it varies over, $0 \leq a \leq a_{\text{cutoff}}$, and therefore, a_{cutoff} is an upper cutoff. The characteristic equations and magnetic field, Eqs. (10.4.11) and (10.4.12), remain the same as in the region-1 case, so that again, we have an antisymmetric-like TM_1 mode.

As $a \rightarrow 0$, we may derive a simplified approximation of the characteristic equation. Since β is large in that limit, it follows that $\gamma, \alpha_c, \alpha_s$ are all approximately equal to β , and we obtain, after canceling some common factors of β in the right-hand side,

$$e^{4\gamma a} = \frac{|p_c|\alpha_c + \gamma}{|p_c|\alpha_c - \gamma} \cdot \frac{|p_s|\alpha_s + \gamma}{|p_s|\alpha_s - \gamma} \Rightarrow e^{4\beta a} = \frac{|p_c| + 1}{|p_c| - 1} \cdot \frac{|p_s| + 1}{|p_s| - 1}, \quad \text{or,}$$

$$\beta \approx \frac{1}{4a} \ln \left[\frac{|p_c| + 1}{|p_c| - 1} \cdot \frac{|p_s| + 1}{|p_s| - 1} \right], \quad a \rightarrow 0 \quad (10.4.15)$$

For region 2, the characteristic equation (10.4.4) can be solved for any given value of $a < a_{\text{cutoff}}$ by rearranging it in the form,

$$\gamma = \sqrt{-\gamma(p_c\alpha_c + p_s\alpha_s) \coth(2\gamma a) - p_c\alpha_c p_s\alpha_s}$$

and replacing it by the iteration, for $n = 0, 1, 2, \dots$, initialized just above $\beta = k_0\sqrt{\varepsilon_f}$,

$$\gamma_{n+1} = \sqrt{-\gamma_n(p_c\alpha_{cn} + p_s\alpha_{sn}) \coth(2\gamma_n a) - p_c\alpha_{cn} p_s\alpha_{sn}} \quad (10.4.16)$$

Example 10.4.2: Consider the following permittivity values from Ref. [943],

$$\varepsilon_c = -1.4^2 = -1.96, \quad \varepsilon_f = 1.5^2 = 2.25, \quad \varepsilon_s = -1.3^2 = -1.69$$

and the two normalized film thickness $k_0a = 0.3$ and $k_0a = 0.1$. The cutoff thickness is $k_0a_{\text{cutoff}} = 0.4015$, calculated from Eq. (10.4.6). Fig. 10.4.5 shows the magnetic field profiles. Fig. 10.4.6 plots the effective index β/k_0 versus film thickness a over the range $0 \leq a \leq a_{\text{cutoff}}$, and demonstrates that β is a decreasing function of a . Superimposed on that graph is the approximation of Eq. (10.4.15).

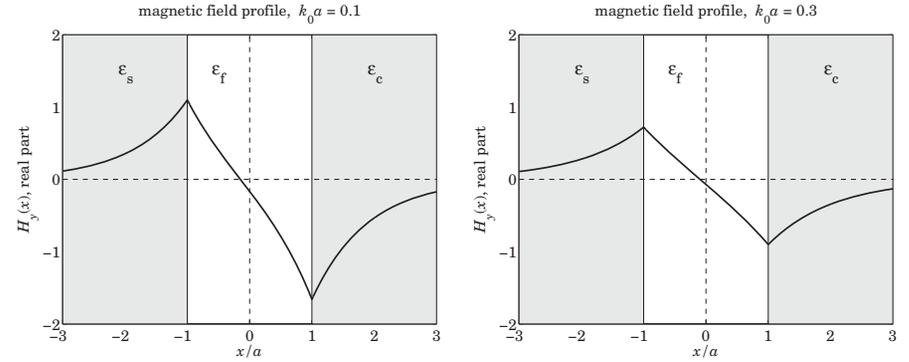


Fig. 10.4.5 Magnetic field profiles, $H_y(x)$ versus x for region 2.

The following MATLAB code illustrates the iteration (10.4.16). The code for generating Fig. 10.4.5 is not given—it is the same as that in Example 10.4.1.

```

ef = 1.5^2; ec = -1.4^2; es = -1.3^2;
k0 = 1;
a = 0.3;
% a = 0.1;

pc = ef/ec; ps = ef/es;

b0 = 1.01*sqrt(ef); % could choose any b0 above b_min, e.g., b0 = 10*sqrt(ef)
ga = sqrt(b0^2 - ef); % initialize ga, ac, as
as = sqrt(b0^2 - es);
ac = sqrt(b0^2 - ec);

tol = 1e-12; n = 1;

while 1
    ga_new = sqrt(-pc*ac.*as*ps - ga.*(pc*ac+ps*as).*coth(2*ga*a*k0));
    if abs(ga_new-ga) < tol, break; end
    ga = ga_new;
    be = sqrt(ga^2 + ef);
    as = sqrt(be^2 - es);
    ac = sqrt(be^2 - ec);
    n=n+1;
end

```

```

psi = atanh(-ga*(pc*ac-ps*as)/(ga^2 - pc*ac*ps*as))/2 + j*pi/2;

% computational error of characteristic equation
E = abs(tanh(2*k0*a*ga) + ga*(pc*ac+ps*as)/(ga^2+pc*ac*ps*as));

% be = pwga(2*pi,ef,ec,es,a,b0,1,tol); % alternative calculation using PWGA

```

where the quantity, E , measures the computational error of the characteristic equation, and the final n is the number of iterations to converge. The calculated values for the two thicknesses are as follows,

k_0a	β/k_0	γ/k_0	α_c/k_0	α_s/k_0	ψ	n	E
0.3	2.8886	2.4686	3.2100	3.1677	$0.0700 + j\pi/2$	99	1.78×10^{-13}
0.1	11.2518	11.1513	11.3385	11.3266	$0.1658 + j\pi/2$	205	6.01×10^{-14}

The following MATLAB code generates Fig. 10.4.6.

```

be0 = sqrt(ef);
as = sqrt(ef-es);
ac = sqrt(ef-ec);
k0a_c = -1/2*(1/pc/ac + 1/ps/as); % upper cutoff

be = linspace(be0, 30, 1001);

ga = sqrt(be.^2 - ef);
ac = sqrt(be.^2 - ec);
as = sqrt(be.^2 - es);

k0a = atanh(-ga.*(pc*ac+ps*as)./(ga.^2 + pc*ps*ac.*as))/2./ga;

bapp = log((1-pc)/(1+pc)*(1-ps)/(1+ps))/4./k0a; % approximate solution

figure; plot(k0a,be,'-', k0a,bapp,'--') % graph annotations are omitted

```

Even though both surface plasmon conditions $|\varepsilon_c| > \varepsilon_f$ and $|\varepsilon_s| > \varepsilon_f$ are violated in region 2, and therefore, separate surface plasmons cannot exist at those interfaces, yet, the solution for finite thickness still exhibits peaks at the two interfaces. \square

Region 3

In region-3, we have the conditions, $\varepsilon_f \leq |\varepsilon_s| \leq |\varepsilon_c|$, or, $|p_c| \leq |p_s| \leq 1$. We will see that in this region, there are two types of plasmonic modes, a symmetric-like TM_0 mode that has no cutoffs, and an antisymmetric-like TM_1 mode with a lower cutoff. In addition, there are TE and TM oscillatory modes with certain lower cutoffs. In particular, the TM_1 mode is special in the sense that its oscillatory version also has an upper cutoff beyond which the mode becomes plasmonic.

In the limit of infinite thickness, both the ε_f - ε_c and ε_f - ε_s interfaces can support surface plasmons with corresponding wavenumbers,

$$\beta_{c,\infty} = k_0 \sqrt{\frac{\varepsilon_c \varepsilon_f}{\varepsilon_c + \varepsilon_f}}, \quad \beta_{s,\infty} = k_0 \sqrt{\frac{\varepsilon_s \varepsilon_f}{\varepsilon_s + \varepsilon_f}} \quad (10.4.17)$$

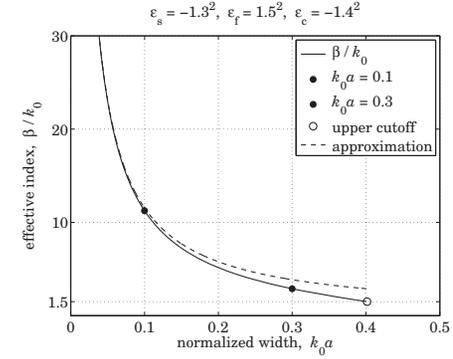


Fig. 10.4.6 Effective index, β/k_0 , versus normalized thickness, k_0a . Region 2.

Because of the assumption, $\varepsilon_f \leq |\varepsilon_s| \leq |\varepsilon_c|$, it follows that $k_0\sqrt{\varepsilon_f} < \beta_{c,\infty} \leq \beta_{s,\infty}$. We note also the relationships,

$$\begin{aligned} p_c^2 \alpha_c^2 - \gamma^2 &= (1 - p_c^2) (\beta_{c,\infty}^2 - \beta^2) \\ p_s^2 \alpha_s^2 - \gamma^2 &= (1 - p_s^2) (\beta_{s,\infty}^2 - \beta^2) \end{aligned} \quad (10.4.18)$$

Because $|p_c| \leq |p_s| \leq 1$, it is possible for the ratios $|p_c|\alpha_c/\gamma$ and $|p_s|\alpha_s/\gamma$ to be either both greater than one, or both less than one, with the former case leading to the TM_1 plasmonic mode, and the latter, to the TM_0 mode. That both ratios must be simultaneously greater or less than one follows from Eq. (10.4.3), which requires that the denominator be a positive quantity. The corresponding ranges of β are easily determined from Eq. (10.4.18). If $|p_c|\alpha_c \geq \gamma$, then, $\beta \leq \beta_{c,\infty} \leq \beta_{s,\infty}$, and if, $|p_s|\alpha_s \leq \gamma$, then, $\beta \geq \beta_{s,\infty} \geq \beta_{c,\infty}$. Thus, the possible β ranges for the two plasmonic modes are,

$$\begin{aligned} |p_{c,s}|\alpha_{c,s} \leq \gamma &\Rightarrow \begin{cases} \beta_{s,\infty} \leq \beta < \infty & (TM_0 \text{ plasmonic mode}) \\ k_0\sqrt{\varepsilon_f} \leq \beta \leq \beta_{c,\infty} & (TM_1 \text{ plasmonic mode}) \end{cases} \\ |p_{c,s}|\alpha_{c,s} \geq \gamma &\Rightarrow \end{aligned} \quad (10.4.19)$$

For the TM_0 mode, the film width a is a decreasing function of β , and varies over $0 \leq a \leq \infty$, where $a = \infty$ at $\beta = \beta_{s,\infty}$. For the TM_1 mode, the width a is an increasing function of β varying over the range, $a_{\text{cutoff}} \leq a \leq \infty$, with a_{cutoff} realized at $\beta = k_0\sqrt{\varepsilon_f}$ and given by Eq. (10.4.6), with $a = \infty$ realized at $\beta = \beta_{c,\infty}$.

For the TM_1 case, the characteristic equations are the same as in Eq. (10.4.11), while those for the TM_0 mode can be obtained by inverting,

$$\tanh(\gamma a + \psi) = \frac{|p_c|\alpha_c}{\gamma} \leq 1, \quad \tanh(\gamma a - \psi) = \frac{|p_s|\alpha_s}{\gamma} \leq 1$$

These imply that ψ must be real, thus, inverting and separating, we have,

$$\begin{aligned} \gamma a &= \frac{1}{2} \operatorname{atanh} \left(-\frac{p_c \alpha_c}{\gamma} \right) + \frac{1}{2} \operatorname{atanh} \left(-\frac{p_s \alpha_s}{\gamma} \right) \\ \psi &= \frac{1}{2} \operatorname{atanh} \left(-\frac{p_c \alpha_c}{\gamma} \right) - \frac{1}{2} \operatorname{atanh} \left(-\frac{p_s \alpha_s}{\gamma} \right) \end{aligned} \quad (10.4.20)$$

which are recognized to have the form of Eqs. (10.1.15) with $m = 0$, that is, a TM_0 mode. The characteristic equation can be solved using the iteration (10.4.13) for the TM_1 mode, and (10.4.16), for the TM_0 mode.

Example 10.4.3: Fig. 10.4.7 shows the dependence of β on the width a for the two modes, for the following permittivity and width parameters from [943],

$$\varepsilon_f = 1.5^2 = 2.25, \quad \varepsilon_s = -1.8^2 = -3.24, \quad \varepsilon_c = -2^2 = -4, \quad k_0 a = 0.8$$

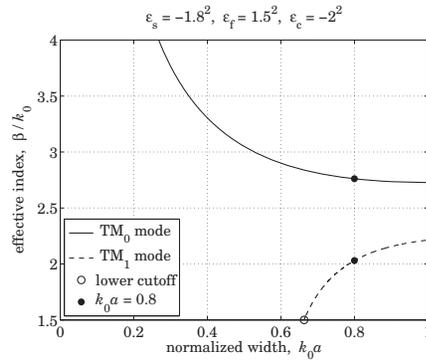


Fig. 10.4.7 Effective index, β/k_0 , versus normalized thickness, $k_0 a$.

The lower cutoff, $k_0 a_{\text{cutoff}} = 0.6628$, was computed from Eq. (10.4.6). Fig. 10.4.8 shows the magnetic field profiles of the two modes, with the TM_0 one being symmetric-like, and the TM_1 , antisymmetric-like. The computed values of β , γ , α_c , α_s , ψ are in the two cases,

mode	β/k_0	γ/k_0	α_c/k_0	α_s/k_0	ψ	n	E
TM_0	2.7612	2.3182	3.4094	3.2961	-0.6751	427	4.68×10^{-13}
TM_1	2.0301	1.3680	2.8498	2.7132	$0.1741 + j\pi/2$	339	6.78×10^{-13}

Also shown are, n, E , the number of iterations to converge with a tolerance of $tol = 10^{-12}$, and the computational error of the characteristic equation defined as

$$E = \left| \tanh(2\gamma a) + \frac{\gamma(p_c \alpha_c + p_s \alpha_s)}{\gamma^2 + p_c \alpha_c p_s \alpha_s} \right|$$

The MATLAB code for this example is given below, where one should comment/uncomment the appropriate lines to generate the TM_0 and TM_1 solutions.

```

ef = 1.5^2; ec = -2^2; es = -1.8^2;
k0 = 1; a = 0.80;
pc = ef/ec; ps = ef/es;

ac = sqrt(ef-ec); as = sqrt(ef-es);
acutoff = -(1/ac/pc + 1/as/ps)/2; % units of k0=1

binit = 1.1*sqrt(ef);
ga = sqrt(binit^2-ef); as = sqrt(binit^2-es); ac = sqrt(binit^2-ec);

tol = 1e-12; n = 1;

while 1
    ga_new = sqrt(-pc*ac*as*ps - ga*(pc*ac+ps*as)*coth(2*ga*a*k0)); % TM0
    % ga_new = -(ga^2 + pc*ac*as*ps)/(pc*ac+ps*as) * tanh(2*ga*a*k0); % TM1
    if abs(ga_new-ga) < tol, break; end
    ga = ga_new;
    be = sqrt(ga^2 + ef);
    as = sqrt(be^2 - es);
    ac = sqrt(be^2 - ec);
    n=n+1;
end

E = abs(tanh(2*k0*a*ga) + ga*(pc*ac+ps*as)/(ga^2+pc*ac*ps*as));
m=0; % set m=0 for TM0, m=1 for TM1
psi = atanh(-pc*ac/ga)/2 - atanh(-ps*as/ga)/2 + j*m*pi/2;

% be = pwga(2*pi,ef,ec,es,a,binit,0,tol); % alternative calculation using PWGA
% be = pwga(2*pi,ef,ec,es,a,binit,1,tol);

```

And, for the magnetic field profiles, using the calculated values of γ , α_c , α_s , ψ from above,

```

x = linspace(-3,3,601)*a;

Hy = j*cosh(ga*k0*a - psi) .* exp(k0*as*(x+a)) .* (x<-a) + ...
    j*cosh(ga*k0*a + psi) .* exp(-k0*ac*(x-a)) .* (x>a) + ...
    j*cosh(k0*ga*x + psi) .* (abs(x)<=a);

plot(x/a, imag(Hy)); % TM0
% plot(x/a, real(Hy)); % TM1

```

The H -field constant was arbitrarily set to $H_0 = j$. But then, in the TM_0 case because ψ is real, H_y is effectively imaginary, whereas in the TM_1 case, because ψ has a $j\pi/2$ imaginary part, H_y becomes real. This explains the above plotting choices for H_y .

The calculation and plotting of the β - a dispersion curves was done by the MATLAB code,

```

be0 = sqrt(ef);
bcinf = sqrt(ef*ec/(ef+ec));
bsinf = sqrt(ef*es/(ef+es));

be = linspace(1.001*bsinf, 40, 2001); % TM0 mode
ga = sqrt(be.^2 - ef); ac = sqrt(be.^2 - ec); as = sqrt(be.^2 - es);

k0a = atanh(-ga.*(pc*ac+ps*as)./(ga.^2 + pc*ps*ac.*as))/2./ga;

```

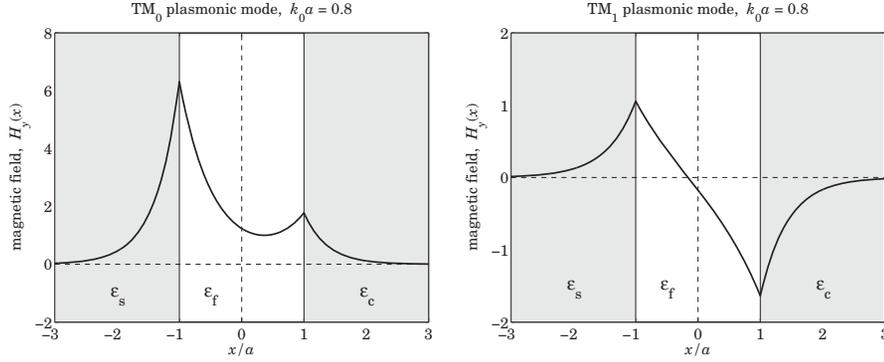


Fig. 10.4.8 Magnetic field profiles of TM_0 and TM_1 modes for $k_0a = 0.80$. Region 3.

figure; plot(k0a,be); hold on

```
be = linspace(1.00001*be0, 0.995*bcinf, 2001); % TM1 mode
ga = sqrt(be.^2 - ef); ac = sqrt(be.^2 - ec); as = sqrt(be.^2 - es);
```

```
k0a = atanh(-ga.*(pc*ac+ps*as)./(ga.^2 + pc*ps*ac.*as))./2./ga;
```

```
plot(k0a,be, '--'); axis([0,1,0,4.5]);
```

For the TM_0 mode, a range of values over $\beta_{s,\infty} \leq \beta < \infty$ was used to evaluate the corresponding thickness a , and a range over $k_0\sqrt{\epsilon_f} \leq \beta \leq \beta_{c,\infty}$ was used for the TM_1 mode. \square

10.5 Oscillatory Modes

In addition to the plasmonic modes, there are also oscillatory TE and TM modes characterized by an imaginary $y = jk_f$, with $k_f = \sqrt{k_0^2\epsilon_f - \beta^2}$. Because k_f is real, the wave-number β is restricted to the values $\beta \leq k_0\sqrt{\epsilon_f}$.

In the lossless asymmetric waveguide case of Sec. 9.12, the positivity of the decay parameters α_c, α_s , imposed also a lower limit on β , see Eq. (9.12.3). However here, $\alpha_{c,s} = \sqrt{\beta^2 - k_0^2\epsilon_{c,s}} = \sqrt{\beta^2 + k_0^2|\epsilon_{c,s}|}$, and α_c, α_s remain positive even at $\beta = 0$. Thus, the range of β s for all oscillatory modes is

$$0 \leq \beta \leq k_0\sqrt{\epsilon_f} \quad (\text{oscillatory modes}) \quad (10.5.1)$$

The TE_m modes are obtained by setting $p_c = p_s = 1$ in the characteristic equations (10.1.19), for $m = 0, 1, 2, \dots$,

$$k_f a = \frac{1}{2} \operatorname{atan} \left(\frac{\alpha_c}{k_f} \right) + \frac{1}{2} \operatorname{atan} \left(\frac{\alpha_s}{k_f} \right) + \frac{1}{2} m \pi \quad (TE_m) \quad (10.5.2)$$

$$\phi = \frac{1}{2} \operatorname{atan} \left(\frac{\alpha_c}{k_f} \right) - \frac{1}{2} \operatorname{atan} \left(\frac{\alpha_s}{k_f} \right) + \frac{1}{2} m \pi$$

The TM_m modes are indexed by $m = 1, 2, 3, \dots$, and Eqs. (10.1.20) are more appropriate because p_c, p_s are negative,

$$k_f a = \frac{1}{2} \operatorname{atan} \left(-\frac{k_f}{p_c \alpha_c} \right) + \frac{1}{2} \operatorname{atan} \left(-\frac{k_f}{p_s \alpha_s} \right) + \frac{1}{2} (m-1) \pi \quad (TM_m) \quad (10.5.3)$$

$$\phi = \frac{1}{2} \operatorname{atan} \left(-\frac{k_f}{p_c \alpha_c} \right) - \frac{1}{2} \operatorname{atan} \left(-\frac{k_f}{p_s \alpha_s} \right) + \frac{1}{2} m \pi$$

All modes have a lower cutoff found by setting $\beta = 0$, $k_f = k_0\sqrt{\epsilon_f}$, $\alpha_{c,s} = k_0\sqrt{-\epsilon_{c,s}}$, in the characteristic equations (10.5.2) and (10.5.3), resulting basically in the same expression,

$$2k_0 a_{\min} = \frac{1}{\sqrt{\epsilon_f}} \left[\operatorname{atan} \left(\sqrt{\frac{-\epsilon_c}{\epsilon_f}} \right) + \operatorname{atan} \left(\sqrt{\frac{-\epsilon_s}{\epsilon_f}} \right) + m \pi \right] \quad (TE_m) \quad (10.5.4)$$

$$2k_0 a_{\min} = \frac{1}{\sqrt{\epsilon_f}} \left[\operatorname{atan} \left(\sqrt{\frac{-\epsilon_c}{\epsilon_f}} \right) + \operatorname{atan} \left(\sqrt{\frac{-\epsilon_s}{\epsilon_f}} \right) + (m-1) \pi \right] \quad (TM_m)$$

For all, but the TM_1 oscillatory mode, there is no upper cutoff thickness, that is, the upper limit $\beta = k_0\sqrt{\epsilon_f}$ is reached at infinite width, $a \rightarrow \infty$. Indeed, using the property that $\operatorname{atan}(x) \rightarrow \pm\pi/2 - 1/x$ as $x \rightarrow \pm\infty$, we find from (10.1.19) that as $\beta \rightarrow k_0\sqrt{\epsilon_f}$ and $a \rightarrow \infty$, then $k_f \rightarrow 0$ and $\alpha_{c,s} \rightarrow k_0\sqrt{\epsilon_f - \epsilon_{c,s}}$, and

$$k_f a \rightarrow \pm \frac{\pi}{4} \pm \frac{\pi}{4} - \frac{1}{2} k_f \left[\frac{1}{k_0 p_c \sqrt{\epsilon_f - \epsilon_c}} + \frac{1}{k_0 p_s \sqrt{\epsilon_f - \epsilon_s}} \right] + \frac{m \pi}{2}, \quad \text{or,}$$

$$k_f \approx \frac{(m \pm 1) \pi}{2(a - a_c)}, \quad a_c = -\frac{1}{2k_0} \left[\frac{1}{p_c \sqrt{\epsilon_f - \epsilon_c}} + \frac{1}{p_s \sqrt{\epsilon_f - \epsilon_s}} \right] \quad (10.5.5)$$

where '+' corresponds to TE_m and '-' to TM_m . This leads to the approximation for β ,

$$\beta = \sqrt{k_0^2 \epsilon_f - k_f^2} \approx k_0 \sqrt{\epsilon_f} - \frac{k_f^2}{2k_0 \sqrt{\epsilon_f}} = k_0 \sqrt{\epsilon_f} - \frac{(m \pm 1)^2 \pi^2}{8(a - a_c)^2 k_0 \sqrt{\epsilon_f}}, \quad a \rightarrow \infty \quad (10.5.6)$$

The quantity a_c is recognized as a_{cutoff} of Eq. (10.4.6) in the TM case, but in the TE case, a_c is negative and has no special meaning. For large a , one could simply replace $(a - a_c)^2$ by a^2 in (10.5.6).

The TM_1 oscillatory case is special because it also has an upper cutoff, which is the same as the lower cutoff of the TM_1 plasmonic mode, that is, a_{cutoff} of Eq. (10.4.6). Indeed, when $m = 1$, a common factor of k_f can be canceled from both sides of Eq. (10.5.3) in the limit $k_f \rightarrow 0$, leading to the expression (10.4.6). Thus, the TM_1 oscillatory mode exists only for thicknesses $a_{\min} \leq a \leq a_{\text{cutoff}}$ and for $0 \leq \beta \leq k_0\sqrt{\epsilon_f}$, and then it becomes plasmonic for $a_{\text{cutoff}} \leq a < \infty$ and $k_0\sqrt{\epsilon_f} \leq \beta \leq \beta_{c,\infty}$.

Example 10.5.1: Fig. 10.5.1 shows the dependence of β on the width a for the oscillatory modes $TM_1, TM_2, TM_3, TE_0, TE_1, TE_2$ for the same permittivity values as in Example 10.5.1.

The graphs were generated by calculating the values of a from Eqs. (10.5.2) and (10.5.3) over the range of values $0 \leq \beta \leq k_0\sqrt{\epsilon_f}$. For example, the MATLAB code for calculating the TM_1, TE_0, TM_2 , and TE_1 curves was,

```

ef = 1.5^2; ec = -2^2; es = -1.8^2;

pc = ef/ec; ps = ef/es;

be = linspace(0, 0.995*sqrt(ef), 201);
kf = sqrt(ef - be.^2); ac = sqrt(be.^2 - ec); as = sqrt(be.^2 - es);

a_tm1 = (atan(-kf./ac/pc) + atan(-kf./as/ps))./kf/2;
a_tm2 = (atan(-kf./ac/pc) + atan(-kf./as/ps) + pi)./kf/2;
a_te0 = (atan(ac./kf) + atan(as./kf))./kf/2;
a_te1 = (atan(ac./kf) + atan(as./kf) + pi)./kf/2;

plot(a_tm1, be, '-', a_te0, be, '--', a_tm2, be, '-', a_te1, be, '--'); xlim([0,5]);

```

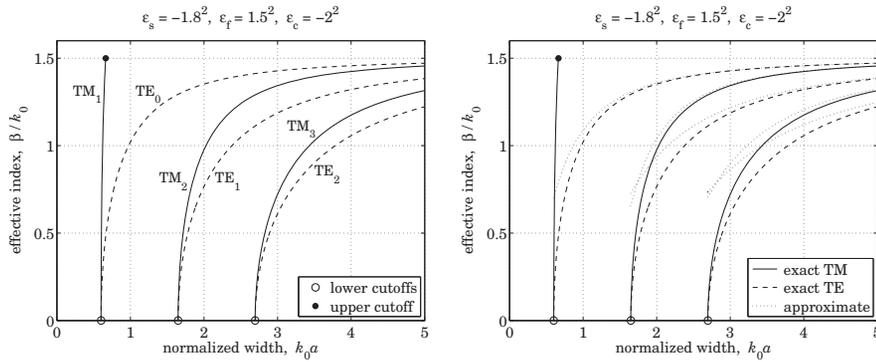


Fig. 10.5.1 Oscillatory TE and TM modes, and the approximation of Eq. (10.5.6).

The lower cutoffs and the upper cutoff of the TM_1 mode were,

$$\begin{aligned}
 k_0 a_{\text{cutoff}} &= 0.6628, & \text{upper cutoff, } TM_1 \text{ mode} \\
 k_0 a_{\text{min}} &= 0.6011, & \text{lower cutoff, } TM_1, TE_0 \text{ modes} \\
 k_0 a_{\text{min}} &= 1.6483, & \text{lower cutoff, } TM_2, TE_1 \text{ modes} \\
 k_0 a_{\text{min}} &= 2.6955, & \text{lower cutoff, } TM_3, TE_2 \text{ modes}
 \end{aligned}$$

The graph on the right compares the large- a approximation of Eq. (10.5.6) with the exact solutions. The approximation is applicable to all but the TM_1 mode. \square

Example 10.5.2: In this example we look at the TM_0 plasmonic, and TM_1 oscillatory and plasmonic modes. The permittivity parameters are as in Example 10.4.3. In that example, we determined the field profiles for $k_0 a = 0.8$, which lies in the plasmonic range for TM_1 since $k_0 a > k_0 a_{\text{cutoff}}$. The allowed thickness range of the TM_1 oscillatory mode was found in the previous example to be rather narrow, $k_0 [a_{\text{min}}, a_{\text{cutoff}}] = [0.6011, 0.6628]$. Here, we determine β and the field profiles for an intermediate value of a , such as, $k_0 a = 0.63$, which lies in the oscillatory range for TM_1 , while TM_0 remains plasmonic. Fig. 10.5.2 displays the β - a relationships.

We observe how the oscillatory TM_1 mode switches to its plasmonic version at $a = a_{\text{cutoff}}$. The TM_0 effective index and ψ parameter were $\beta/k_0 = 2.8685$ and $\psi = -0.4292$, and were

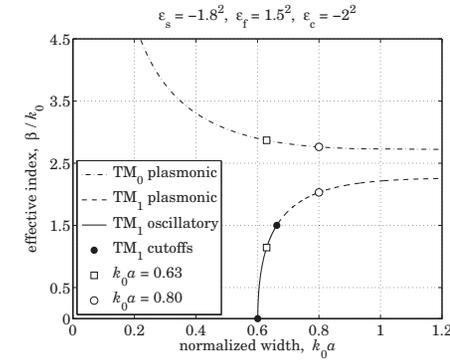


Fig. 10.5.2 Effective index of oscillatory and plasmonic modes.

computed by the same code as that in Example 10.4.3 with $k_0 a = 0.63$. The oscillatory TM_1 index and phase factor were $\beta/k_0 = 1.1434$ and $\phi = 1.6021$, and were computed by the following MATLAB code, including its magnetic field profile,

```

ef = 1.5^2; ec = -2^2; es = -1.8^2; % a = 0.63
k0 = 1; a = 0.63;

pc = ef/ec; ps = ef/es;

binit = 0.5*sqrt(ef);
kf = sqrt(ef-binit^2); as = sqrt(binit^2-es); ac = sqrt(binit^2-ec);

tol = 1e-12; n = 1;

while 1
    kf_new = (atan(-kf/ac/pc)+atan(-kf/as/ps))/2/a;
    if abs(kf_new-kf) < tol, break; end
    kf = kf_new;
    be = sqrt(ef - kf^2);
    as = sqrt(be^2 - es);
    ac = sqrt(be^2 - ec);
    n=n+1;
end

E = abs(tan(2*k0*a*kf) - kf*(pc*ac+ps*as)/(kf^2 - pc*ac*ps*as));
phi = (atan(pc*ac/kf) - atan(ps*as/kf) + pi)/2

% be = pwga(2*pi,ef,ec,es,a,binit,0,tol); % alternative calculation using PWGA
% be = pwga(2*pi,ef,ec,es,a,binit,1,tol);

x = linspace(-3,3,601)*a;

Hy = cos(kf*k0*a - phi).*exp(k0*as*(x+a)).*(x<-a) + ...
    cos(kf*k0*a + phi).*exp(-k0*ac*(x-a)).*(x>a) + ...
    cos(kf*k0*x + phi).*(abs(x)<=a);

plot(x/a, real(Hy));

```

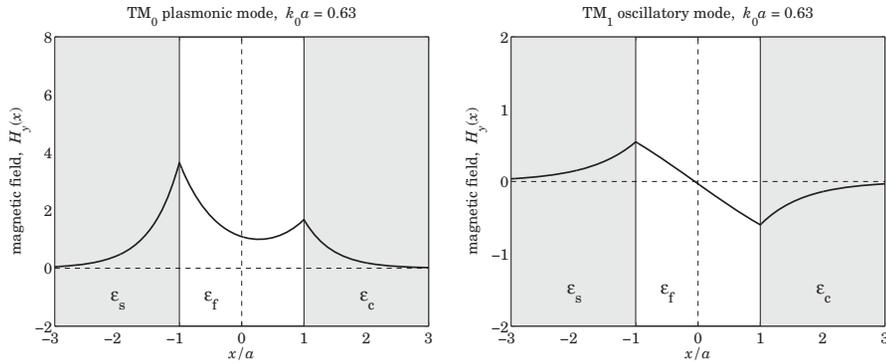


Fig. 10.5.3 Magnetic field profiles of TM_0 and oscillatory TM_1 modes for $k_0a = 0.63$.

The iterative method was the same as that discussed in Sec. 9.12 for the asymmetric dielectric waveguides. The number of iterations to converge and the computational error of the characteristic equation (10.1.17) were $n = 267$ and $E = 1.01 \times 10^{-11}$.

The MATLAB code for calculating the β - a dispersion curves for the TM_1 oscillatory and plasmonic portions is the same as in Examples 10.4.3 and 10.5.1.

Figures 10.5.4, 10.5.5, and 10.5.6 show the dispersion curves and magnetic field profiles for the cases $k_0a = 0.8$ and $k_0a = 0.63$ for a symmetric MDM configuration with permittivities $\epsilon_f = 1.5^2$ and $\epsilon_s = \epsilon_c = -2.2^2$. The magnetic field profiles are now either completely symmetric or antisymmetric. The dispersion curves have the same asymptotic limit for large a , that is, $\beta_{c,\infty} = \beta_{s,\infty} = 2.2678k_0$. The calculated (normalized by k_0) propagation wavenumbers were $\beta = 2.4586$ and $\beta = 1.8197$ for the TM_0 and TM_1 modes at $k_0a = 0.80$, and $\beta = 2.5951$ and $\beta = 0.6448$ for the case $k_0a = 0.63$. □

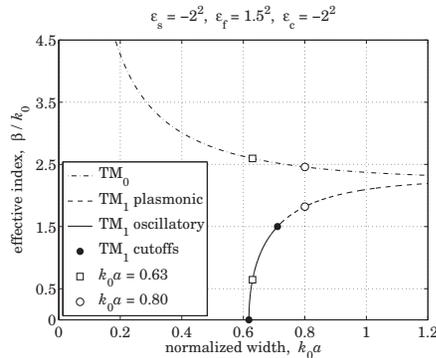


Fig. 10.5.4 Effective index of oscillatory and plasmonic modes. Symmetric guide.

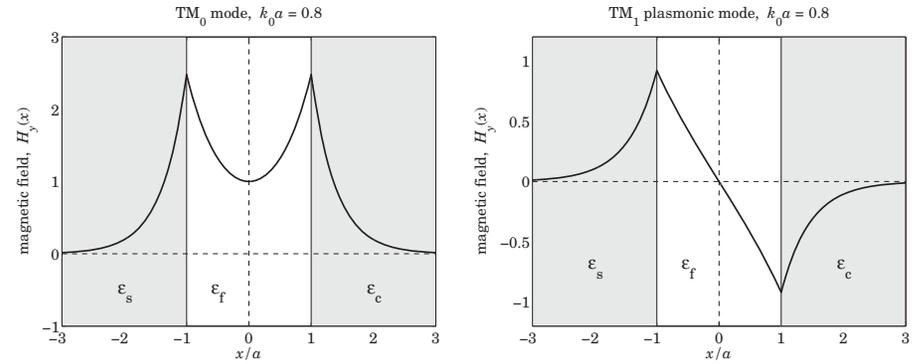


Fig. 10.5.5 Magnetic field profiles of TM_0 and TM_1 modes for $k_0a = 0.80$. Symmetric guide.

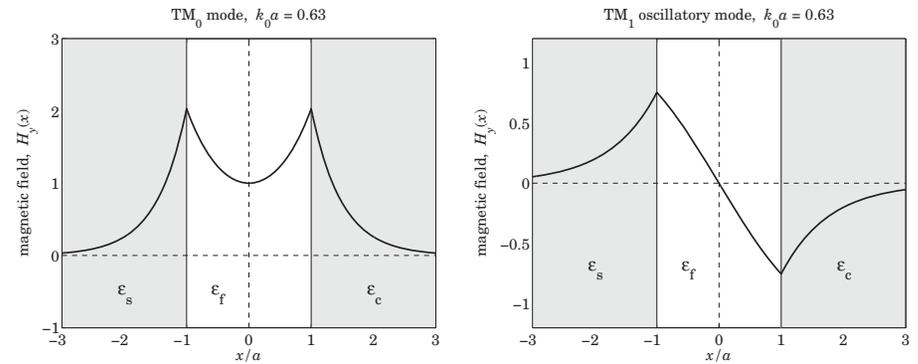


Fig. 10.5.6 Magnetic field profiles of TM_0 and TM_1 modes for $k_0a = 0.63$. Symmetric guide.

Complex Modes

It should be noted that complex-valued solutions for β (with negative imaginary part) also exist (and there is an infinity of them), even though we have assumed lossless media [952]. For example, for the symmetric configuration of Example 10.5.2 at $k_0a = 0.8$, the following are all solutions of the TM_0 characteristic equation, $\gamma \tanh(\gamma a) + p_c \alpha_c = 0$,

q	β/k_0	γ/k_0	α_c/k_0	E
0	2.4586	1.9480	3.1694	0.022×10^{-14}
1	$0.6391 - 3.5638j$	$0.5902 - 3.8589j$	$0.7645 - 2.9794j$	0.076×10^{-14}
2	$0.7585 - 7.7008j$	$0.7446 - 7.8442j$	$0.7851 - 7.4393j$	0.562×10^{-14}
3	$0.7790 - 11.6825j$	$0.7727 - 11.7780j$	$0.7906 - 11.5108j$	0.289×10^{-14}
4	$0.7862 - 15.6351j$	$0.7827 - 15.7067j$	$0.7927 - 15.5069j$	6.671×10^{-14}
5	$0.7896 - 19.5770j$	$0.7873 - 19.6343j$	$0.7938 - 19.4747j$	1.212×10^{-14}

where $E = |\gamma \tanh(\gamma a) + p_c \alpha_c|$ is the computational error. They were computed with the help of the MATLAB function, `pwg`, discussed in the next section. The MATLAB code was as follows,

```

1a0 = 2*pi; k0 = 1; ef = 2.25; ec = -4; es = ec; a = 0.8; % units of k0=1
pc = ef/ec; ps = ef/es;
bcinf = sqrt(ec*ef/(ec+ef));

M = 5; q = (0:M)';
ga0 = 1/k0/a*atanh(-pc) - j*pi*q/a/k0; % justified in next section
be0 = sqrt(ga0.^2 + ef); % vector of initial search points

be = pwg(1a0,ef,ec,a,be0); % uses built-in FSOLVE to find the solutions

ga = sqrt(be.^2-ef);
ac = sqrt(be.^2 -ec);
E = abs(ga.*tanh(ga*a) + pc*ac);

```

The first one is the real-valued solution of Example 10.5.2. All of the complex ones do have positive real parts for γ, α_c . However, the imaginary parts of β , albeit negative, are very large and therefore, these modes are highly damped and cannot be considered as propagating. In fact, for the lossless case, such modes carry no net power in the propagation direction [952]—see Problem 10.4. For such modes, the power flowing forward in the dielectric film cancels the power flowing backward in the metal sides. Indeed, using the MATLAB function, `pwgpower`, or the results of Problem 10.4, the net power and the powers in each medium are as follows for the above example,

q	P	P_f	P_c	P_s
0	2.8401	4.0337	-0.59682	-0.59682
1	0	0.2586	-0.12928	-0.12928
2	0	0.3378	-0.16891	-0.16891
3	0	0.3531	-0.17654	-0.17654
4	0	0.3585	-0.17927	-0.17927
5	0	0.3611	-0.18055	-0.18055

where $P = P_f + P_c + P_s$. The units of P are arbitrary. The MATLAB code was,

```
[P,Pf,Pc,Ps] = pwgpower(k0*a,ef,ec,es,be,0);
```

where the input `be` is the vector of β/k_0 calculated above. We discuss such anomalous modes further in Sec. 10.9.

10.6 MDM Configuration – Lossy Case

In the previous section, we ignored losses in the metal cladding and substrate of an MDM guide in order to simplify the problem and gain some insight into the possible types of propagating plasmonic modes—their essential feature being the subwavelength confinement of the fields in the transverse direction.

Here, we assume that the metals are lossy with permittivities that have both negative real and imaginary parts, $\epsilon_c = \epsilon_{cR} - j\epsilon_{cI}$, where $\epsilon_{cR} < 0$ and $\epsilon_{cI} \geq 0$, and similarly,

$\epsilon_s = \epsilon_{sR} - j\epsilon_{sI}$, where $\epsilon_{sR} < 0$ and $\epsilon_{sI} \geq 0$. The dielectric film will still be assumed to be lossless, $\epsilon_f > 0$. Moreover, as is typically the case in practice, we will assume that the real parts satisfy the region-3 conditions, $\epsilon_f \leq |\epsilon_{sR}| \leq |\epsilon_{cR}|$, which imply the existence of single-interface surface plasmons at both interfaces in the case of infinite thickness.

The propagation parameters, $\beta, \gamma, \alpha_c, \alpha_s$, become complex-valued, and in particular, we require that α_c, α_s have positive real parts, and β , negative imaginary part, $\beta = \beta_R - j\beta_I$, with $\beta_I \geq 0$, so that the wave attenuates as it propagates along the positive z -direction, that is, $e^{-j\beta z} = e^{-j\beta_R z} e^{-\beta_I z}$, while it remains confined in the transverse direction, e.g., $e^{-\alpha_c x} = e^{-\alpha_{cR} x} e^{-j\alpha_{cI} x}$, for $x \geq a$. The field expressions and characteristic equations (10.1.3)–(10.1.16) remain valid.

Let us consider the symmetric case, $\epsilon_c = \epsilon_s$, whose characteristic equations for the symmetric TM_0 mode and antisymmetric TM_1 mode are ,

$$\tanh(\gamma a + \psi) = -\frac{p_c \alpha_c}{\gamma}, \quad \psi = \frac{j m \pi}{2}, \quad m = 0, 1 \quad (10.6.1)$$

where $p_c = \epsilon_f / \epsilon_c$, and $\gamma = \sqrt{\beta^2 - k_0^2 \epsilon_f}$, $\alpha_c = \sqrt{\beta^2 - k_0^2 \epsilon_c}$. Equivalently,

$$e^{2\gamma a + 2\psi} = \frac{\gamma - p_c \alpha_c}{\gamma + p_c \alpha_c}, \quad \psi = \frac{j m \pi}{2}, \quad m = 0, 1 \quad (10.6.2)$$

In the limit of large separation, $a \rightarrow \infty$, both the TM_0 and TM_1 modes, tend to the single-interface surface plasmon solution. Indeed, because γ has a positive real part, we have $\tanh(\gamma a + \psi) \rightarrow 1$, and thus, $-p_c \alpha_c = \gamma$, which corresponds to,

$$\beta_{c,\infty} = k_0 \sqrt{\frac{\epsilon_f \epsilon_c}{\epsilon_f + \epsilon_c}} \quad (10.6.3)$$

A variety of approaches can be taken to solving the characteristic equation (10.6.1). We discuss three.

First, if the losses are small, $\epsilon_{cI} \ll |\epsilon_{cR}|$, the solutions for the TM_0 and TM_1 modes evolve smoothly from the corresponding ones of the lossless case, as we verify in the example below. The results of Problem 10.5 can be used to construct the lossy solutions from the lossless ones by a first-order approximation in the parameter ϵ_{cI} .

Second, Eq. (10.6.1) can be solved iteratively by turning it into the following iteration [963], initialized at $\beta = \beta_{c,\infty}$,

$$\gamma_{n+1} = -p_c \alpha_{c,n} \coth(\gamma_n a + \psi), \quad n = 0, 1, 2, \dots \quad (10.6.4)$$

and stopping when two successive iterates get closer to each other than some specified error tolerance, $|\gamma_{n+1} - \gamma_n| \leq tol$, such as $tol = 10^{-12}$.

Third, we have written a MATLAB function, `pwg`, which uses the built-in function, `fsolve`, to solve the system of two equations consisting of the real and imaginary parts of the characteristic equation,

$$E = \gamma \tanh(\gamma a + \psi) + p_c \alpha_c = 0 \quad (10.6.5)$$

in the two unknowns, the real and imaginary parts of β . It has usage:

```
[be,E] = pwg(la0,ef,ec,a,be0,mode,tol) % plasmonic modes for symmetric guides
```

```
la0 = operating wavelength, k0 = 2*pi/la0 = free-space wavenumber
ef,ec = permittivities of film and cladding/substrate
a = half-width of film in same units as la0
be0 = starting search point in units of k0 - can be a vector of choices
mode = 0,1 for TM0 or TM1 mode, default mode=0
tol = computational error tolerance, default tol=1e-12
```

```
be = propagation wavenumber in units of k0 - same size as be0
E = value of the characteristic equation (10.6.5) - same size as be0
```

There is also a version, **pwga**, for asymmetric guides, with similar usage:

```
[be,E] = pwga(la0,ef,ec,es,a,be0,mode,tol) % plasmonic modes for asymmetric guides
```

The iterative method generally works well for the primary TM_0 and TM_1 modes. However, it is not capable of finding the anomalous complex-mode solutions that are highly damped. Those can be found using **pwg** with properly choosing the initial search points, as we discuss later.

Example 10.6.1: This example, with parameters taken from [953], illustrates the three numerical approaches. Consider a symmetric silver-air-silver waveguide at an operating free-space wavelength of $\lambda_0 = 650$ nm and permittivity of silver[†] $\epsilon_c = \epsilon_s = -19.6224 - 0.443j$, and $\epsilon_f = 1$ for an air gap of width $2a = 100$ nm.

The following MATLAB code computes the propagation parameters β, γ, α_c for the symmetric TM_0 mode ($\psi = 0$) using the **pwg** function and the 1st-order approximation using Eq. (10.21.3) of Problem 10.5.

```
la0 = 650; k0 = 2*pi/la0; a = 100/2; % la0,a in nm
ef = 1; ec = -19.6224-0.443*j; es = ec;

pc = ef/ec; ps = ef/es;
bcinf = sqrt(ec*ef/(ec+ef)); % bcinf = 1.0265 - 0.0006i
tol = 1e-12;
mode=0;
be = pwg(la0,ef,ec,a,bcinf,mode,tol); % be = 1.2261 - 0.0026i

ac = sqrt(be^2 - ec); as = ac; % ac = as = 4.5965 + 0.0475i
ga = sqrt(be^2 - ef); % ga = 0.7095 - 0.0045i

E = abs(pc*ac + ga*tanh(ga*k0*a)); % E = 5.5518e-17

[P,Pf,Pc,Ps] = pwgpower(k0*a,ef,ec,es,be,0); % power flow in the three media
Powers = [P,Pf,Pc,Ps]; % Powers = [1.2175, 1.2328, -0.0076, -0.0076]

% ---- 1st order calculation -----

ecr = real(ec); % lossless case, ecr = -19.6224
bcinr = sqrt(ecr*ef/(ecr+ef)); % bcinr = 1.0265
```

```
pcr = ef/ecr;
ber = pwg(la0,ef,ecr,a,bcinr,0,tol); % lossless solution
gr = sqrt(ber^2-ef); acr = sqrt(ber^2-ecr);
Dec = imag(ec)*j; % correction to ec

Dg = pcr*gr*(acr^2/ecr+1/2)*dec/(k0*a*(gr^2*acr)*(1-pcr^2*acr^2/gr^2)-pcr*(ef-ecr));
Db = gr/ber*Dg;

be1 = ber + Db; % approximate lossy solution

num2str([be;be1], '%2.7f') % be = 1.2261233 - 0.0026097i % exact from PWG
% be1 = 1.2261699 - 0.0026106i % 1st order
```

The numerical values are given in the comments. As expected, energy flow is negative in the metals and positive in the dielectric. We note that the 1st-order approximation is very good because $\epsilon_{cl} \ll |\epsilon_{cR}|$. The iterative version is implemented by the following code segment and results into the same solution to within the error tolerance *tol*.

```
binit = bcinf; % initialize iteration
ga = sqrt(binit^2 - ef);
ac = sqrt(binit^2 - ec);

N = 1; % number of iterations to converge
while 1 % forever while loop
    ga_new = -pc*ac*coth(ga*a*k0);
    if abs(ga_new-ga) < tol, break; end
    ga = ga_new;
    be = sqrt(ga^2 + ef);
    ac = sqrt(be^2 - ec);
    N=N+1;
end

N, be % N = 270, be = 1.2261233 - 0.0026097i
```

The iteration was initialized at $\beta = \beta_{c,\infty}$, but the algorithm is very insensitive to the initial choice, for example, it converges equally fast with $\beta = 0$ or $\beta = 10\beta_{c,\infty}$.

The corresponding magnetic field profile $H_y(x)$ is plotted in Fig. 10.6.1. The following MATLAB code generates the graph (annotations are omitted.)

```
x = linspace(-200,200,401); % units of nm

Hy = cosh(ga*k0*a)*exp(-k0*ac*(abs(x)-a)).*(abs(x)>a) + cosh(k0*ga*x).*(abs(x)<=a);

fill([-200, -50, -50, -200],[0, 0, 1.5, 1.5], [0.9 0.9 0.9]); hold on
fill([50, 200, 200, 50],[0, 0, 1.5, 1.5], [0.9 0.9 0.9]);
plot(x,real(Hy), 'linewidth',2);
```

Next, we calculate and plot the β - a dispersion relationship by solving for β over the range of thicknesses, $5 \leq 2a \leq 3500$ nm. Fig. 10.6.2 plots the effective phase index $\text{Re}(\beta)/k_0$ and propagation length $L = -[2 \text{Im}(\beta)]^{-1}$ versus a , where L is plotted in units of μm . The previously computed solution for $2a = 100$ nm is also added to the graph. It is interesting

[†]Our MATLAB function DRUDE, from Sec. 1.12, based on [163], gives $\epsilon_c = -15.4758 - 1.1513j$.

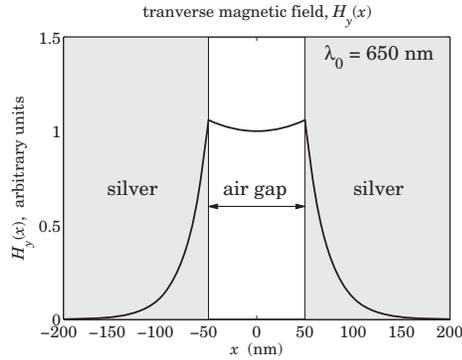


Fig. 10.6.1 Magnetic field profile of silver-air-silver guide ($\lambda_0 = 650$ nm, $2a = 100$ nm).

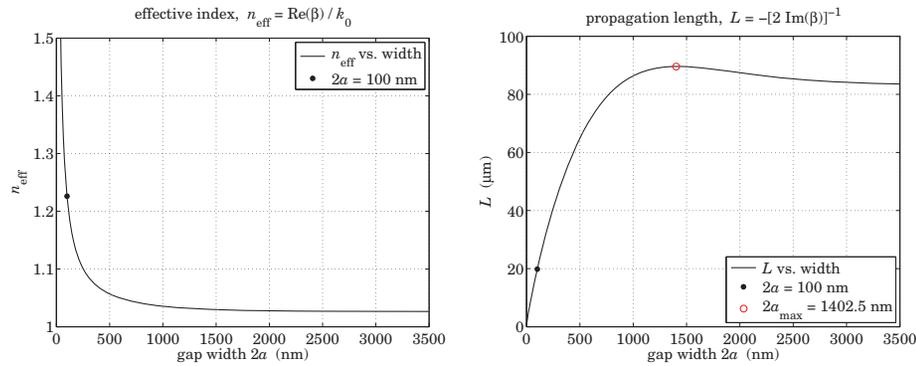


Fig. 10.6.2 Effective index and propagation distance ($\lambda_0 = 650$ nm).

to note that L exhibits a maximum before settling towards its asymptotic limit, with the maximum occurring approximately at [957],

$$2a_{\max}k_0 = \pi\sqrt{|1 + \text{Re}(\epsilon_c)|}$$

The maximum was added in the graph. The following MATLAB code illustrates the computation using the `pwg` function as well as the iterative method (again, graph annotation details are omitted.)

```
w1 = 2*a; be1 = be; % save previous results for 2a=100
L1 = -1/2./imag(be1)/k0/1000; % propagation length in microns

wmax = pi/k0*sqrt(abs(real(ec)+1)) % max L at wmax = 1402.49 nm
bemax = pwg(la0,ef,ec,wmax/2,bcinf,0,tol); % bemax = 1.030361 - 0.000577i
Lmax = -1/2./imag(bemax)/k0/1000; % Lmax = 89.5901 microns
```

```
w = (5:10:3500); a = w/2; % vector of gap thicknesses w in nm

for i=1:length(w) % PWG method
    be(i) = pwg(la0,ef,ec,a(i),bcinf,0,tol); % la0,ef,ec,bcinf,tol, as above
end

L = -1/2./imag(be)/k0/1000; % propagation length in micrometers

ac = sqrt(be.^2 - ec); ga = sqrt(be.^2 - ef);
E = abs(pc*ac + ga.*tanh(ga.*a*k0)); % vector of computational errors

figure; plot(w,real(be),'-', w1,real(be1),''); % plot effective index
figure; plot(w,L,'-', w1,L1,'.', wmax,Lmax,'ro'); % plot propagation length

% ----- iterative version -----

for i=1:length(w)
    binit = bcinf; % initialize iteration
    ga = sqrt(binit^2 - ef);
    ac = sqrt(binit^2 - ec);
    N(i) = 1; % number of iterations for i-th thickness
    while 1
        ga_new = -pc*ac*coth(ga*a(i)*k0);
        if abs(ga_new-ga) < tol, break; end
        ga = ga_new;
        be(i) = sqrt(ga^2 + ef); % beta for i-th thickness
        ac = sqrt(be(i)^2 - ec);
        N(i) = N(i)+1;
    end % end while-loop
end % end for-loop

ac = sqrt(beit.^2 - ec); ga = sqrt(beit.^2 - ef);

Eit = abs(pc*ac + ga.*tanh(ga.*a*k0)); % vector of errors

norm(be-beit) % compare PWG and iterative methods, ans = 2.5818e-12
```

Fig. 10.6.3 plots the computational error $E = |\gamma \tanh(\gamma a) + p_c \alpha_c|$ versus thickness $w = 2a$, for both the PWG and the iterative methods, showing that it remains below the specified error tolerance of 10^{-12} . The figure also shows a plot versus w of the number of iterations N required for convergence of the iterative method, where N is saved during the iteration for each thickness. \square

10.7 Gap Surface Plasmons

The TM_0 symmetric mode illustrated by the above example is the only mode that exists for the symmetric MDM guide for all values of the gap thickness $2a$, and is known as a *gap surface plasmon polariton* (G-SPP) mode [981]. For small values of a , the propagation wavenumber β becomes very large. Indeed, assuming that $|\gamma a| \ll 1$ and using the approximation $\tanh(x) \approx x$, the characteristic equation (10.6.1) (with $\psi = 0$) becomes,

$$\gamma a = -\frac{p_c \alpha_c}{\gamma} \Rightarrow \gamma^2 a = -p_c \sqrt{\gamma^2 + (\epsilon_f - \epsilon_c) k_0^2} \quad (10.7.1)$$

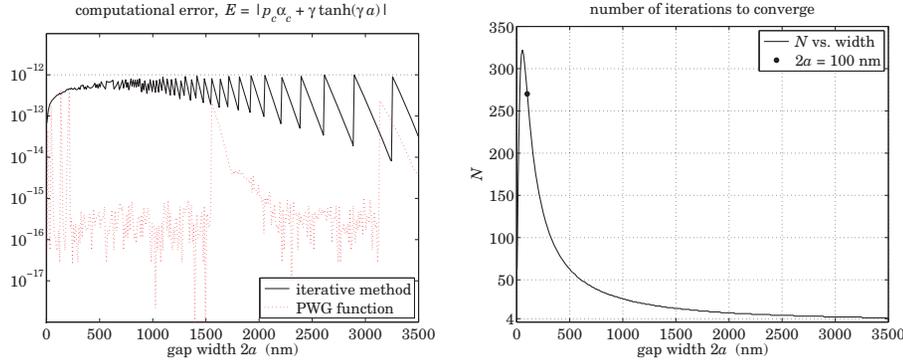


Fig. 10.6.3 Computational error and number of iterations ($\lambda_0 = 0.650$ nm, $tol = 10^{-12}$).

This is a second-order algebraic equation in the variable y^2 with solution for y^2 and β ,

$$y^2 = \frac{p_c^2}{2a^2} + \sqrt{\frac{p_c^4}{4a^4} + \frac{p_c^2 k_0^2}{a^2} (\epsilon_f - \epsilon_c)} \quad (10.7.2)$$

$$\beta = \sqrt{k_0^2 \epsilon_f + y^2} = \sqrt{k_0^2 \epsilon_f + \frac{p_c^2}{2a^2} + \sqrt{\frac{p_c^4}{4a^4} + \frac{p_c^2 k_0^2}{a^2} (\epsilon_f - \epsilon_c)}}$$

For very small a , the quartic term p_c^4/a^4 wins inside the above square roots, so that $y^2 \approx p_c^2/a^2$. Its square root is taken with negative sign, $y = -p_c/a$, in order to guarantee a positive real part for y . This results in the simpler expression for $\beta \approx y$,

$$\beta = -\frac{p_c}{a}, \quad \text{for } k_0 a \ll 1 \quad (10.7.3)$$

The same can also be derived from the characteristic equation (10.7.1) when y is large,

$$y^2 a = -p_c \sqrt{y^2 + (\epsilon_f - \epsilon_c) k_0^2} \approx -p_c y \Rightarrow y = -\frac{p_c}{a}$$

On the other hand, if a is small but not too small so that the quartic term p_c^4/a^4 can be ignored inside the square roots of Eq. (10.7.2), then we obtain another simple expression which works well [981] over a substantial range of a ,

$$y^2 \approx \sqrt{\frac{p_c^2 k_0^2}{a^2} (\epsilon_f - \epsilon_c)} = -\frac{p_c k_0}{a} \sqrt{\epsilon_f - \epsilon_c} \quad (10.7.4)$$

$$\beta = k_0 \sqrt{\epsilon_f - \frac{p_c}{k_0 a} \sqrt{\epsilon_f - \epsilon_c}}$$

TM₁ modes

We saw in Sec. 10.4 that, in the lossless case, region-3 supported antisymmetric TM₁ plasmonic modes for $k_0 \sqrt{\epsilon_f} \leq \beta \leq \beta_{c,\infty}$, corresponding to the film's half-width range

$a_{\text{cutoff}} \leq a < \infty$, where for symmetric guides ($\epsilon_c = \epsilon_s$), the cutoff thickness is given by Eq. (10.4.6) and is obtained by setting, $\beta = k_0 \sqrt{\epsilon_f}$, or $y = 0$, in the TM₁ characteristic equation,

$$\tanh(\gamma a) = -\frac{y}{p_c \alpha_c} \Big|_{y=0} \Rightarrow k_0 a_{\text{cutoff}} = -\frac{1}{p_c \sqrt{\epsilon_f - \epsilon_c}} - \frac{\epsilon_c}{\epsilon_f \sqrt{\epsilon_f - \epsilon_c}}$$

In the lossy case, because ϵ_c and β are complex-valued, an approximate value for the cutoff with may be obtained by taking the real part of the above expression,

$$k_0 a_{\text{cutoff}} = -\text{Re} \left[\frac{1}{p_c \sqrt{\epsilon_f - \epsilon_c}} \right] \quad (10.7.5)$$

Example 10.7.1: Here, we look at the TM₁ modes of Example 10.6.1 that had operating free-space wavelength of $\lambda_0 = 650$ nm and permittivity of silver $\epsilon_c = \epsilon_s = -19.6224 - 0.443j$, and $\epsilon_f = 1$ for air. The cutoff width calculated from Eq. (10.7.5) is found to be

$$w_{\text{cutoff}} = 2a_{\text{cutoff}} = 894.0795 \text{ nm}$$

The TM₁ modes were calculated over the thickness range, $w_{\text{cutoff}} \leq 2a \leq 3500$ nm, and the TM₀ modes over, $5 \leq 2a \leq 3500$ nm. Fig. 10.7.1 shows the effective index $\text{Re}(\beta)/k_0$ and propagation distance $L = -[2 \text{Im}(\beta)]^{-1}$ in units of nm. The solutions for the specific value, $2a = 1000$ nm, that lies above the cutoff, were also placed on the graphs. The following MATLAB code segment illustrates the computations.

```

lambda0 = 650; k0 = 2*pi/lambda0; ef = 1; ec = -19.6224-0.443*j; es=ec;
pc = ef/ec; ps = ef/es;
bcinf = sqrt(ec*ef/(ec+ef));
tol = 1e-12;
acut = -real(1/pc/sqrt(ef-ec))/k0;
wc = 2*acut;
nc = sqrt(ef);

w = 5:10:3500; % thickness range for TMO
w1 = 1.001*wc:10:3500; % thickness range for TM1

for i=1:length(w)
    be0(i) = pwg(lambda0,ef,ec,w(i)/2,bcinf,0,tol); % TMO
end

for i=1:length(w1)
    be1(i) = pwg(lambda0,ef,ec,w1(i)/2,bcinf,1,tol); % TM1
end

neff0 = real(be0); L0 = -1/2./imag(be0)/k0/1000; % TMO index & distance
neff1 = real(be1); L1 = -1/2./imag(be1)/k0/1000; % TM1 index & distance

w2=1000; a = w2/2; % specific solutions
be02 = pwg(lambda0,ef,ec,a2,bcinf,0,tol); % TMO
be12 = pwg(lambda0,ef,ec,a2,bcinf,1,tol); % TM1

neff02 = real(be02); L02 = -1/2/imag(be02)/k0/1000;
neff12 = real(be12); L12 = -1/2/imag(be12)/k0/1000;

```

```
figure; plot(w,neff0,'-', w1,neff1,'r--', ...
            [w2,w2],[neff02,neff12],'o', wc,nc,'s', 'markersize',9);
```

```
yaxis(1, 1.15, 1:0.05:1.15);
xaxis(0,3500, 0:500:3500);
```

```
figure; plot(w,L0,'-', w1,L1,'r--', [w2,w2],[L02,L12],'o', 'markersize',9);
yaxis(0,100,0:20:100);
xaxis(0,3500, 0:500:3500);
```

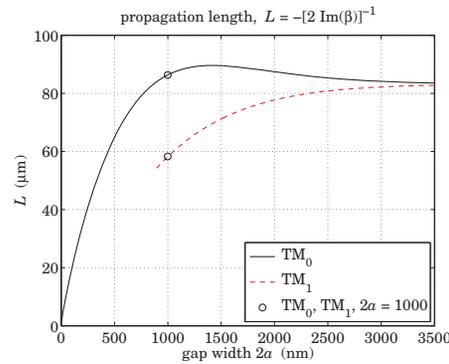
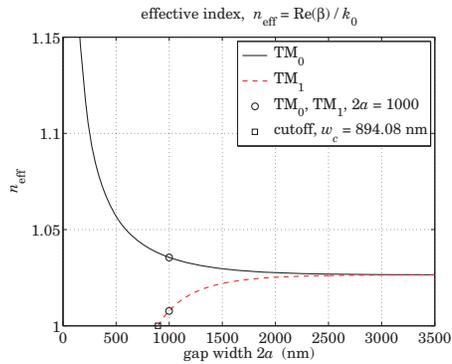


Fig. 10.7.1 Effective index and propagation distance for TM_0 and TM_1 plasmonic modes.

The symmetric TM_0 and antisymmetric TM_1 magnetic field profiles $H_y(x)$ for the case $2a = 1000$ nm are shown in Fig. 10.7.2. They were computed and plotted by the following MATLAB code segment (annotations are omitted.)

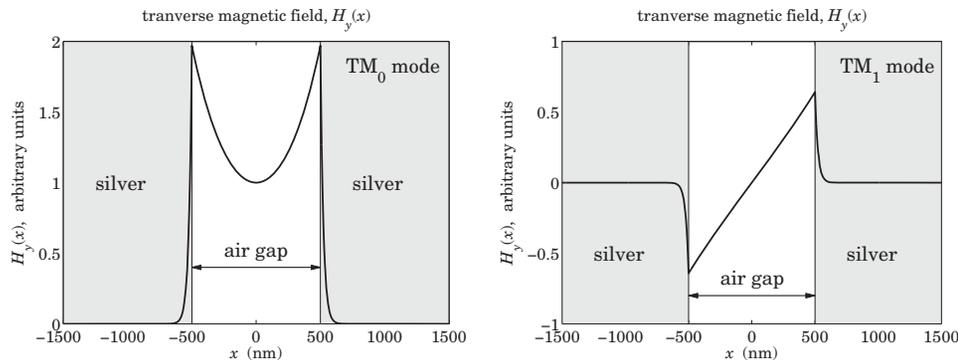


Fig. 10.7.2 Transverse magnetic field profiles for TM_0 and TM_1 plasmonic modes.

```
x = linspace(-1500,1500,601);
```

```
ga = sqrt(be02^2 - ef); % transverse wavenumbers for TM0
ac = sqrt(be02^2 - ec); % ga = 0.2690-0.0023i; ac = 4.5494+0.0486i
psi = 0;
Hy = cosh(ga*k0*a+psi)*exp(-k0*ac*(x-a)).*(x > a) + ... % TM0 field
     cosh(ga*k0*a-psi)*exp(k0*ac*(x+a)).*(x <-a) + ... % here, a=1000/2
     cosh(k0*ga*x+psi).*(abs(x)<=a);
```

```
figure; fill([-1500, -500, -500, -1500],[0, 0, 2, 2], [0.9 0.9 0.9]); hold on
fill([500, 1500, 1500, 500],[0, 0, 2, 2], [0.9 0.9 0.9]);
plot(x,real(Hy), 'linewidth',2); % plot TM0
```

```
xaxis(-1500, 1500, -1500:500:1500); yaxis(0,2, 0:0.5:2);
```

```
ga = sqrt(be12^2 - ef); % transverse wavenumbers for TM1
ac = sqrt(be12^2 - ec); % ga = 0.1251-0.0072i; ac = 4.5432+0.0486i
psi = j*pi/2;
Hy = cosh(ga*k0*a+psi)*exp(-k0*ac*(x-a)).*(x > a) + ... % TM1 field
     cosh(ga*k0*a-psi)*exp(k0*ac*(x+a)).*(x <-a) + ...
     cosh(k0*ga*x+psi).*(abs(x)<=a);
```

```
figure; fill([-1500, -500, -500, -1500],[-1, -1, 1, 1], [0.9 0.9 0.9]); hold on
fill([500, 1500, 1500, 500],[-1, -1, 1, 1], [0.9 0.9 0.9]);
plot(x,imag(Hy), 'linewidth',2); % plot TM1
```

```
xaxis(-1500, 1500, -1500:500:1500); yaxis(-1,1, -1:0.5:1);
```

We note that because of the rather large thickness of 1000 nm, the TM_0 field resembles two separate single-interface surface plasmons at the two silver-air interfaces. □

10.8 PEC Limit

Plasmonic waveguides operate near the visible and infrared spectrum and excel at guiding light at decent propagation distances, while being laterally confined at subwavelength distances.

At the microwave regime, MDM waveguides behave very much like ordinary parallel-plate TEM transmission lines, or like parallel-plate waveguides supporting TM and TE modes. The connection to the plasmonic case can be seen in the perfect-electric conductor (PEC) limit [961]. We recall from Chapter 1 that the effective dielectric constant of a conductor is related to its conductivity $\sigma(\omega)$ by

$$\epsilon_c = 1 - j \frac{\sigma(\omega)}{\epsilon_0 \omega} \quad (10.8.1)$$

The PEC limit assumes that the metals are perfect conductors ($\sigma \rightarrow \infty$), so that $\epsilon_c \rightarrow \infty$, and therefore, $p_c = \epsilon_f / \epsilon_c \rightarrow 0$. In this limit, the characteristic equation (10.6.1) of a symmetric MDM guide becomes for the even and odd TM modes,

$$\tanh(\gamma a + \psi) = -\frac{p_c \alpha_c}{\gamma} \rightarrow \frac{\epsilon_f k_0 \sqrt{-\epsilon_c}}{\epsilon_c \gamma} \rightarrow 0, \quad \psi = \frac{j m \pi}{2}, \quad m = 0, 1$$

where we replaced $\alpha_c = \sqrt{\beta^2 - k_0^2 \epsilon_c} \rightarrow k_0 \sqrt{-\epsilon_c}$. This condition then requires that $\gamma a + \psi = j\pi q$, with integer q , resulting in the quantized values of γ ,

$$\gamma = jk_c, \quad k_c = \frac{\pi(2q + m)}{2a}, \quad q = 0, 1, 2, \dots, \quad m = 0, 1 \quad (10.8.2)$$

The quantities k_c are recognized as the cutoff wavenumbers of the TM modes of a parallel-plate waveguide with perfectly conducting walls. The corresponding dispersion relationship is obtained from $\beta = \sqrt{\gamma^2 + k_0^2 \epsilon_f} = \sqrt{k_0^2 \epsilon_f - k_c^2}$, and requires that the guide be operated above cutoff, $k_0 \sqrt{\epsilon_f} \geq k_c$.

The case $q = m = 0$, or, $\gamma = 0$, is special and corresponds to the parallel-plate TEM transmission line discussed in the next chapter. The TEM property follows from Eq. (10.1.3), which shows that $E_z = 0$ in this case. Moreover, E_x, H_y become constants, independent of x , and $E_x/H_y = \eta_0$.

TE modes, which generally do not exist in the plasmonic case, correspond to setting $p_c = 1$ in the characteristic equation (10.6.1), which would have the PEC limit,

$$\tanh(\gamma a \pm \psi) = -\frac{\alpha_c}{\gamma} \rightarrow \frac{k_0 \sqrt{-\epsilon_c}}{\gamma} \rightarrow \infty, \quad \psi = \frac{j m \pi}{2}, \quad m = 0, 1$$

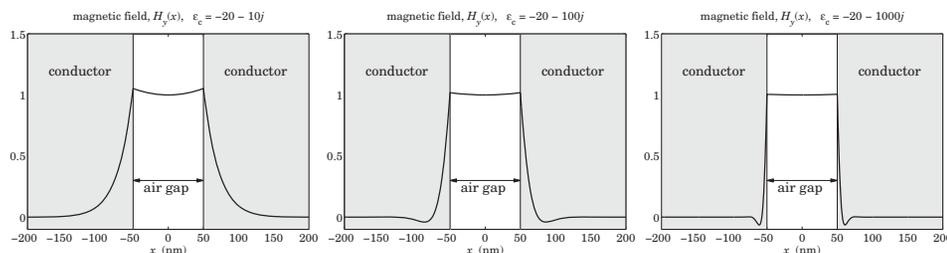
This condition gives the quantized values for γ and TE cutoff wavenumbers:

$$\gamma = jk_c, \quad k_c = \frac{\pi(2q + m + 1)}{2a}, \quad q = 0, 1, 2, \dots, \quad m = 0, 1 \quad (10.8.3)$$

We note that $\gamma = 0$ is not possible for TE modes.

Example 10.8.1: The following graphs show the magnetic field profiles for a fictitious symmetric MDM guide with $\lambda_0 = 650$ nm, $\epsilon_f = 1$, with film width of 100 nm, and the successive values of ϵ_c having increasingly larger imaginary parts:

$$\begin{aligned} \epsilon_c &= -20 - 10j \\ \epsilon_c &= -20 - 100j \\ \epsilon_c &= -20 - 1000j \end{aligned}$$



The MATLAB code is the same as in Example 10.6.1. We observe how the magnetic field tends to the TEM limit, becoming progressively constant, and more quickly decaying in the metal. The computed values for β, γ, α_c were,

β/k_0	γ/k_0	α_c/k_0
1.20520 - 0.04899j	0.67656 - 0.08727j	4.74695 + 1.04087j
1.07807 - 0.06220j	0.42772 - 0.15677j	7.84987 + 6.36099j
1.02333 - 0.02255j	0.23702 - 0.09736j	22.59670 + 22.12610j

10.9 Anomalous Complex Modes

As we mentioned at the end of Sec. 10.5, there exist complex-modes even in the lossless case, that are anomalous in the sense that they are stable and laterally confined, having propagation wavenumber, $\beta = \beta_R - j\beta_I$, with $\beta_I > 0$, but with β_I so large that they cannot be considered as propagating, nor are they necessarily evanescent, which would have $\beta_R = 0$ and $\beta_I > 0$, although such evanescent modes do exist.

The mode spectrum of MDM guides for the lossless and lossy cases has been studied in various references such as [952-955,965].

Consider a symmetric configuration ($\epsilon_c = \epsilon_s$), whose even and odd TM modes must satisfy the characteristic equation (10.6.1),

$$\tanh(\gamma a \pm \psi) = -\frac{p_c \alpha_c}{\gamma}, \quad \psi = \frac{j m \pi}{2}, \quad m = 0, 1 \quad (10.9.1)$$

Because we are looking for anomalous modes that have large values of β , or, $\gamma = \sqrt{\beta^2 + k_0^2 \epsilon_f}$, we have, $\alpha_c = \sqrt{\gamma^2 + k_0^2 (\epsilon_f - \epsilon_c)} \approx \gamma$. It follows that the right-hand-side of Eq. (10.9.1) tends to a constant,

$$\tanh(\gamma a \pm \psi) \approx -p_c, \quad \text{for } |\beta| \gg k_0$$

But this equation does not uniquely define γ , since,

$$\tanh(\gamma a \pm \psi) = \tanh(\gamma a \pm \psi \pm j\pi q) = -p_c$$

for any integer q . Replacing $\psi = j m \pi / 2$, for $m = 0, 1$, and inverting, we obtain an infinite family of possible solutions, which will have large $|\gamma|$ for large values of q ,

$$\gamma = \frac{1}{a} \operatorname{atanh}(-p_c) \pm \frac{j\pi(2q + m)}{2a}, \quad m = 0, 1, \quad q = 0, 1, 2, \dots \quad (10.9.2)$$

This is similar to Eq. (10.8.2). By using these approximate γ s as the initial search points of the function, **pwg**, we may derive the exact values of the anomalous modes.

To clarify the role of the \pm sign choices, we note that in the lossless case (real ϵ_c, ϵ_f), because the characteristic equation has real parameters then if β is a solution, so will be the complex conjugate, β^* . Moreover, since the dependence on β comes through β^2 , it follows that given a solution β , all four choices, $\pm\beta, \pm\beta^*$, would also be solutions. However, if $\beta = \beta_R - j\beta_I$ with $\beta_I \geq 0$, then the only other physically acceptable choice would be, $-\beta^* = -\beta_R - j\beta_I$, that also has negative imaginary part. Thus, in the lossless case the anomalous complex modes come in pairs, $\{\beta, -\beta^*\}$, or, $\pm\beta_R - j\beta_I$ [952]. The \pm signs in Eq. (10.9.2) typically correspond to such pair $\beta, -\beta^*$. In the lossy case, however, this conjugate symmetry is broken, making the two values in the pair $\beta, -\beta^*$ slightly different.

Example 10.9.1: Consider again the silver-air-silver guide of Example 10.6.1 with free-space wavelength of $\lambda_0 = 650$ nm, silver permittivity, $\epsilon_c = \epsilon_s = -19.6224 - 0.443j$, and $\epsilon_f = 1$ for an air gap of width $2a = 100$ nm.

The following table shows the solutions for the even modes ($m = 0$), for $q = 0, 1, 2, 3$ and the two \pm sign choices. The computational error vector, $E = \gamma \tanh(\gamma a + \psi) + p_c \alpha_c$, has

norm $\|E\| = 4.3687 \times 10^{-14}$. The second table shows the odd solutions ($m = 1$). Its error norm was $\|E\| = 9.1312 \times 10^{-14}$.

The $q = 0$ even solution is the fundamental G-SPP mode that we considered in Example 10.6.1. For the rest of the even cases and all the odd ones, we observe that all solutions are acceptable, in the sense that they have $\beta_I = -\text{Im}(\beta) > 0$, and $\text{Re}(\alpha_c) > 0$, however, they cannot be considered as propagating because of their large value of β_I . We note also that they have a γ that is predominantly imaginary, so that those modes may be viewed as oscillatory as opposed to plasmonic.

q		β/k_0 (even modes)	γ/k_0	α_c/k_0
0		1.226123 - 0.002610j	0.709501 - 0.004510j	4.596523 + 0.047492j
1	-	0.076360 - 6.424289j	0.075451 - 6.501643j	0.057829 - 4.652592j
1	+	-0.076321 - 6.419272j	0.075411 + 6.496685j	0.153066 + 4.647824j
2	-	0.099118 - 12.963768j	0.098824 - 13.002278j	0.087286 - 12.183377j
2	+	-0.099113 - 12.958989j	0.098819 + 12.997513j	0.123651 + 12.178607j
3	-	0.102709 - 19.476698j	0.102573 - 19.502352j	0.093794 - 18.966223j
3	+	-0.102707 - 19.471926j	0.102572 + 19.497586j	0.117154 + 18.961452j

q		β/k_0 (odd modes)	γ/k_0	α_c/k_0
0		0.000249 - 2.976418j	0.000236 - 3.139915j	3.281442 + 0.067275j
1	-	0.093819 - 9.700739j	0.093325 - 9.752140j	0.079792 - 8.630150j
1	+	-0.093808 - 9.695937j	0.093313 + 9.747364j	0.131131 + 8.625381j
2	-	0.101461 - 16.221536j	0.101269 - 16.252329j	0.091276 - 15.604931j
2	+	-0.101459 - 16.216762j	0.101267 + 16.247564j	0.119668 + 15.600160j
3	-	0.103452 - 22.730378j	0.103352 - 22.752364j	0.095539 - 22.294531j
3	+	-0.103451 - 22.725607j	0.103351 + 22.747597j	0.115411 + 22.289760j

The tables were produced by the following MATLAB code:

```

la0 = 650; k0 = 2*pi/la0; a = 100/2; % la0,a in nm
ef = 1; ec = -19.6224 - 0.443*j; % set ec = -16.6224 for lossless case
pc = ef/ec; tol = 1e-12;
q = (0:3)'; m=0; % set m=0,1 for even or odd modes

ga1 = 1/k0/a*atanh(-pc) - j*pi/2/a/k0*(m+2*q);
ga2 = 1/k0/a*atanh(-pc) + j*pi/2/a/k0*(m+2*q);

gin = [ga1.'; ga2.']; gin = gin(:); % list them as -/+ pairs
q = [q';q']; q = q(:);
gin(1)=[]; q(1)=[]; % eliminate duplicate q=0 case

bin = sqrt(gin.^2 + ef); % initial search vector

be = pwg(la0,ef,ec,a,bin,m,tol); % solution vector for beta

ga = sqrt(be.^2 - ef); ac = sqrt(be.^2 - ec); % gamma and alpha_c vectors
E = norm(pc*ac + ga.*tanh(ga*k0*a + j*m*pi/2)) % computational error norm

num2str([be,ga,ac], '%12.6f') % table

```

We note also that the \pm pairs for the same value of q are very close to each other. If we execute the same code for the lossless case ($\epsilon_c = -19.6224$), we will find that the pairs are numerically equal, that is, $\pm\beta_R - j\beta_I$. \square

Example 10.9.2: This example is from [955], and we have also added the odd modes. A symmetric silver-air-silver MDM waveguide at free-space wavelength $\lambda_0 = 1550$ nm, has film width $2a = \lambda_0/4$, film permittivity $\epsilon_f = 1$, and silver permittivity $\epsilon_c = -143.497 - 9.517j$, or, $\epsilon_c = -143.417$ in the lossless limit.[†]

```

la0 = 1550; k0=2*pi/la0; a = la0/8;
ef=1; ec=-143.497 - j*9.517*0; % ec=-143.497 - j*9.517 in lossy case
pc = ef/ec;
tol = 1e-12;

q = (0:4)'; m=0; % even modes TM0, TM2, TM4, TM6, TM8

gin = 1/k0/a*atanh(-pc) - j*pi/2/a/k0*(m+2*q);
bin = sqrt(gin.^2 + ef);

be_even = pwg(la0,ef,ec,a,bin,m,tol); % even mode solutions

m=1; % odd modes TM1, TM3, TM5, TM7, TM9

gin = 1/k0/a*atanh(-pc) - j*pi/2/a/k0*(m+2*q);
bin = sqrt(gin.^2 + ef);

be_odd = pwg(la0,ef,ec,a,bin,m,tol); % odd mode solutions

be = [be_even, be_odd].'; be = be(:); % list modes sequentially

ga = sqrt(be.^2 - ef); ac = sqrt(be.^2 - ec);

num2str([be,ga,ac], '%11.5f') % table

M = repmat([0;1], length(q),1);

E = norm(pc*ac + ga.*tanh(ga*k0*a + j*M*pi/2)) % error norm of all modes

```

The calculated error norms for the lossless and lossy cases were $\|E\| = 5.3579 \times 10^{-14}$ and 2.6947×10^{-14} , respectively. The following two tables list the solutions for the two cases.

mode	β/k_0 (lossless case)	γ/k_0	α_c/k_0
0	1.05313	0.33027	12.02523
1	-1.66934j	-1.94594j	11.86214
2	-3.84683j	-3.97468j	11.34456
3	-5.90040j	5.98454j	10.42508
4	-7.92720j	7.99003j	8.98089
5	-9.94391j	9.99407j	6.67949
6	-11.95773j	-11.99947j	0.71391
7	0.00456 - 13.96424j	0.00455 - 14.00000j	0.00887 - 7.17656j
8	0.00587 - 15.96872j	0.00586 - 16.00000j	0.00887 - 10.55950j
9	0.00661 - 17.97220j	0.00660 - 18.00000j	0.00887 - 13.39787j

[†]Note that our DRUDE function gives here $\epsilon_c = -103.332 - 8.130j$.

mode	β/k_0 (lossy case)	γ/k_0	α_c/k_0
0	1.05304 – 0.00176j	0.33004 – 0.00563j	12.03172 + 0.39534j
1	0.00210 – 1.66945j	0.00180 – 1.94603j	11.86889 + 0.40063j
2	0.00077 – 3.84686j	0.00075 – 3.97471j	11.35228 + 0.41891j
3	0.00035 – 5.90041j	0.00035 – 5.98455j	10.43503 + 0.45581j
4	0.00008 – 7.92719j	0.00008 – 7.99001j	8.99646 + 0.52886j
5	–0.00023 – 9.94387j	0.00023 + 9.99402j	6.71706 + 0.70877j
6	–0.00146 – 11.95650j	0.00145 + 11.99824j	2.24792 + 2.12461j
7	–0.00453 – 13.96352j	0.00452 + 13.99928j	0.66911 + 7.20628j
8	–0.00583 – 15.96808j	0.00582 + 15.99936j	0.45906 + 10.56850j
9	–0.00658 – 17.97159j	0.00657 + 17.99939j	0.36388 + 13.40199j

The listed mode numbers were the quantities, $2q + m$, for $q = 0, 1, 2, 3, 4$, $m = 0, 1$. The $q = 0$ mode is the fundamental G-SPP mode. For the lossless case, modes 1-6 are evanescent in the sense that they have $\beta_R = 0$ and $\beta_I = -\text{Im}(\beta) > 0$, but they are laterally confined because α_c is positive-real. These modes carry no power in any of the three media, because the media were assumed to be lossless, so that the factors that appear in the expressions Eq. (10.1.24) for the transmitted powers are, $\text{Re}(\beta/\varepsilon) = \text{Re}(\beta)/\varepsilon = 0$. The numerical values for the even modes given in [955] differ slightly from ours, but that is probably due to using different calculation procedures. \square

In this section, we have discussed only the modes of the discrete spectrum. Modes belonging to the *continuous spectrum* are not laterally confined. If $\beta = \beta_R - j\beta_I$, then, the lateral decay parameter is,

$$\alpha_c = \sqrt{\beta^2 - k_0^2 \varepsilon_c} = \sqrt{\beta_R^2 - \beta_I^2 - 2j\beta_R\beta_I - k_0^2 \varepsilon_c}$$

Continuous modes would have an imaginary α_c , or, $\alpha_c^2 < 0$. For the lossless case ($\varepsilon_c = -|\varepsilon_c|$), this would require that

$$\alpha_c^2 = \beta_R^2 - \beta_I^2 + k_0^2 |\varepsilon_c| - 2j\beta_R\beta_I < 0 \quad \Rightarrow \quad \begin{cases} \beta_R\beta_I = 0 \\ \beta_R^2 - \beta_I^2 + k_0^2 |\varepsilon_c| < 0 \end{cases}$$

The two conditions can only be satisfied with $\beta_R = 0$ and $\beta_I^2 > k_0^2 |\varepsilon_c|$, or, $\beta_I > k_0 \sqrt{|\varepsilon_c|}$. Such modes have been discussed in [952,955]. and are unbounded laterally, but decay along the propagation direction as $e^{-\beta_I z}$.

10.10 DMD Configuration – Lossless Case

A DMD plasmonic waveguide consists of a metal film ε_f , surrounded by two possibly different dielectrics $\varepsilon_c, \varepsilon_s$, as shown in Fig. 10.10.1. It is operated at optical or infrared frequencies where the metal has permittivity with negative real part, $\varepsilon_f = -\varepsilon_R - j\varepsilon_I$, $\varepsilon_R > 0$ and $\varepsilon_I \geq 0$. The dielectrics may be assumed to be lossless. Moreover, without loss of generality, we will assume that $\varepsilon_c \geq \varepsilon_s > 0$.

The field and characteristic equations were given by Eqs. (10.1.4)–(10.1.16). The various solution modes have been discussed extensively in the literature, see for example, [939–945] and the reviews [962,981].

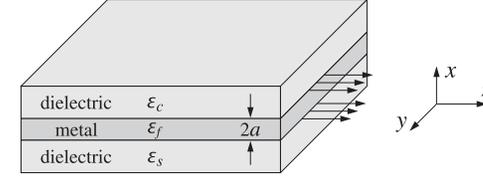


Fig. 10.10.1 DMD plasmonic waveguide.

As in the MDM case, we begin our discussion by first considering the lossless case, $\varepsilon_I = 0$, so that we may write, $\varepsilon_f = -|\varepsilon_f|$. Since $\varepsilon_c \geq \varepsilon_s$, the quantities, $p_c = \varepsilon_f/\varepsilon_c$, $p_s = \varepsilon_f/\varepsilon_s$, satisfy the inequality $|p_c| \leq |p_s|$. We note the following relationships,

$$\begin{aligned} \gamma &= \sqrt{\beta^2 - k_0^2 \varepsilon_f} = \sqrt{\beta^2 + k_0^2 |\varepsilon_f|} \\ \alpha_c &= \sqrt{\beta^2 - k_0^2 \varepsilon_c} & \Rightarrow & \begin{aligned} \gamma &= \sqrt{\alpha_c^2 + k_0^2 (\varepsilon_c + |\varepsilon_f|)} \\ \alpha_s &= \sqrt{\beta^2 - k_0^2 \varepsilon_s} & \gamma &= \sqrt{\alpha_s^2 + k_0^2 (\varepsilon_s + |\varepsilon_f|)} \end{aligned} \end{aligned} \quad (10.10.1)$$

which imply that, $\alpha_c \leq \alpha_s < \gamma$.[†] Note also,

$$\begin{aligned} p_c^2 \alpha_c^2 - \gamma^2 &= (p_c^2 - 1) (\beta^2 - \beta_{c,\infty}^2) \\ p_s^2 \alpha_s^2 - \gamma^2 &= (p_s^2 - 1) (\beta^2 - \beta_{s,\infty}^2) \end{aligned} \quad (10.10.2)$$

$$\beta_{s,\infty}^2 - k_0^2 \varepsilon_c = k_0^2 |\varepsilon_f| \left[\frac{1}{|p_s| - 1} - \frac{1}{|p_c|} \right] \quad (10.10.3)$$

where $\beta_{c,\infty}, \beta_{s,\infty}$ are the would-be plasmonic wavenumbers of the single f - c and f - s metal-dielectric interfaces,

$$\beta_{c,\infty}^2 = \frac{k_0^2 \varepsilon_c \varepsilon_f}{\varepsilon_c + \varepsilon_f} = \frac{k_0^2 |\varepsilon_f|}{|p_c| - 1}, \quad \beta_{s,\infty}^2 = \frac{k_0^2 \varepsilon_s \varepsilon_f}{\varepsilon_s + \varepsilon_f} = \frac{k_0^2 |\varepsilon_f|}{|p_s| - 1} \quad (10.10.4)$$

Clearly, for the single interfaces to support surface plasmons, we must have $|p_c| > 1$ and $|p_s| > 1$. The characteristic equations (10.1.10) can be written in the forms:

$$e^{2\gamma a + 2\psi} = \frac{\gamma + |p_c| \alpha_c}{\gamma - |p_c| \alpha_c}, \quad e^{2\gamma a - 2\psi} = \frac{\gamma + |p_s| \alpha_s}{\gamma - |p_s| \alpha_s} \quad (10.10.5)$$

$$e^{4\gamma a} = \frac{\gamma + |p_c| \alpha_c}{\gamma - |p_c| \alpha_c} \cdot \frac{\gamma + |p_s| \alpha_s}{\gamma - |p_s| \alpha_s} \quad \Rightarrow \quad \tanh(2\gamma a) = \frac{\gamma (|p_c| \alpha_c + |p_s| \alpha_s)}{\gamma^2 + |p_c| \alpha_c |p_s| \alpha_s} \quad (10.10.6)$$

We are looking for plasmonic solutions that have *real and positive* $\beta, \alpha_c, \alpha_s, \gamma$ parameters. There exist, of course, complex-mode solutions as in the MDM case, which can be derived by taking the large- γ limit of Eq. (10.10.6), that is, up to any integer q ,

$$e^{4\gamma a} = e^{4\gamma a - 2\pi i q} = \frac{1 + |p_c|}{1 - |p_c|} \cdot \frac{1 + |p_s|}{1 - |p_s|} \quad \Rightarrow \quad \gamma = \frac{1}{4a} \ln \left[\frac{1 + |p_c|}{1 - |p_c|} \cdot \frac{1 + |p_s|}{1 - |p_s|} \right] + \frac{j\pi q}{2a}$$

[†]In the MDM lossless case, we had the complementary inequality, $\alpha_c \geq \alpha_s > \gamma$.

For all plasmonic solutions with positive real β , it is evident from (10.10.1) that the lateral confinement conditions $\alpha_c \geq 0$ and $\alpha_s \geq 0$ require that $\beta \geq k_0\sqrt{\epsilon_c}$ and $\beta \geq k_0\sqrt{\epsilon_s}$, which, because $\epsilon_c \geq \epsilon_s$, reduce into one,

$$\beta \geq k_0\sqrt{\epsilon_c} \quad (10.10.7)$$

For such solutions, because γ is real, it follows from (10.10.6) that in order for the quantity $e^{4\gamma a}$ to be positive, the quantities, $(\gamma - |p_c|\alpha_c)$ and $(\gamma - |p_s|\alpha_c)$, must either be both positive or both negative. If they are both positive, then, it follows from Eq. (10.10.5) that ψ must be real, so the solution is a TM₀ even-like mode, and when they are both negative, then, ψ must have a $j\pi/2$ imaginary part, so that the solution is a TM₁ odd-like mode. We summarize,

$$\begin{aligned} \gamma - |p_{c,s}|\alpha_{c,s} > 0, \quad \text{Im}[\psi] = 0, \quad (\text{TM}_0 \text{ mode}) \\ \gamma - |p_{c,s}|\alpha_{c,s} < 0, \quad \text{Im}[\psi] = \frac{j\pi}{2}, \quad (\text{TM}_1 \text{ mode}) \end{aligned} \quad (10.10.8)$$

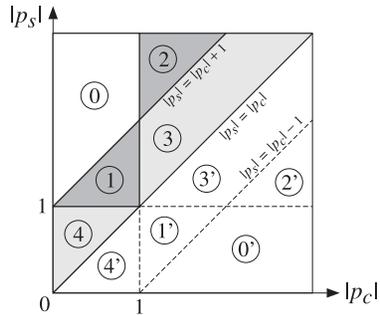
Thus, ψ is calculated from Eq. (10.1.7), with $m = 0, 1$, for TM₀ or TM₁,

$$\psi = \frac{1}{2} \text{atanh} \left(-\frac{p_c \alpha_c}{\gamma} \right) - \frac{1}{2} \text{atanh} \left(-\frac{p_s \alpha_s}{\gamma} \right) + \frac{1}{2} j m \pi \quad (10.10.9)$$

Finally, we recall that the magnetic field is given by,

$$H_y(x) = \begin{cases} H_0 \cosh(\gamma x + \psi), & |x| \leq a \\ H_0 \cosh(\gamma a + \psi) e^{-\alpha_c(x-a)}, & x \geq a \\ H_0 \cosh(\gamma a - \psi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (10.10.10)$$

The possible plasmonic solutions can be understood with the aid of the following diagram that divides the p_c - p_s plane into five regions [943], as defined below,[†]



0. $|p_c| \leq 1 \leq |p_s|$, $|p_s| > |p_c| + 1$
1. $|p_c| \leq 1 \leq |p_s|$, $|p_s| \leq |p_c| + 1$
2. $1 \leq |p_c| \leq |p_s|$, $|p_s| > |p_c| + 1$
3. $1 \leq |p_c| \leq |p_s|$, $|p_s| \leq |p_c| + 1$
4. $|p_c| \leq |p_s| \leq 1$

The regions below the diagonal are obtained by switching the roles of ϵ_c, ϵ_s . The one of most practical interest is region-3, because most metals (like gold and silver) have

[†]Because we are working with the variables p_c, p_s , our geometry of the regions differs from that of [943], but the results are equivalent.

$|\epsilon_f| > \epsilon_c \geq \epsilon_s$, at optical frequencies and using typical dielectrics with a small difference, $\epsilon_c - \epsilon_s$. Next, we determine the nature of solutions in each region, but summarize the results here,

0. no solution exists
1. TM₀ mode with lower cutoff, $k_0\sqrt{\epsilon_c} \leq \beta < \beta_{s,\infty}$, $a_{\text{cutoff}} \leq a < \infty$
2. TM₁ mode with no cutoff, $\beta > \beta_{c,\infty}$, $0 < a < \infty$
3. TM₀ mode with lower cutoff, $k_0\sqrt{\epsilon_c} \leq \beta < \beta_{s,\infty}$, $a_{\text{cutoff}} \leq a < \infty$
TM₁ mode with no cutoff, $\beta > \beta_{c,\infty}$, $0 < a < \infty$
4. TM₀ mode with upper cutoff, $\beta_{\text{max}} \leq \beta < \infty$, $a_{\text{max}} \geq a > 0$
TM₀ mode with lower cutoff, $k_0\sqrt{\epsilon_c} \leq \beta \leq \beta_{\text{max}}$, $a_{\text{cutoff}} \leq a \leq a_{\text{max}}$

Region - 0

Because $|p_c| \leq 1$ and $\alpha_c < \gamma$, we have $(|p_c|\alpha_c - \gamma) < 0$, so necessarily we must also have, $(|p_s|\alpha_s - \gamma) < 0$. It follows that the first of Eqs. (10.10.2) is automatically satisfied because $\beta_{c,\infty}^2 < 0$ and $p_c^2 < 1$. But, because $\beta_{s,\infty}^2 > 0$, the second constraint in (10.10.2) requires that $\beta < \beta_{s,\infty}$. On the other hand, in region-0, Eq. (10.10.3) implies that $\beta_{s,\infty} < k_0\sqrt{\epsilon_f}$, and combining the two inequalities, $\beta < \beta_{s,\infty} < k_0\sqrt{\epsilon_f}$. Thus, it is not possible to fulfill condition (10.10.7) and, therefore, no solutions exist in this region.

Region - 1

The only difference with region-0 is that now Eq. (10.10.3) implies that $\beta_{s,\infty} > k_0\sqrt{\epsilon_c}$, and therefore, there can exist a solution with β spanning the range,

$$k_0\sqrt{\epsilon_c} \leq \beta \leq \beta_{s,\infty} \quad (10.10.11)$$

The upper limit, $\beta = \beta_{s,\infty}$, is reached in the limit of infinite thickness $a \rightarrow \infty$. The lower limit implies a lower cutoff thickness obtained from Eq. (10.10.6) by setting $\beta = k_0\sqrt{\epsilon_c}$, and hence, $\alpha_c = 0$, $\gamma = k_0\sqrt{\epsilon_c + |\epsilon_f|}$, and $\alpha_s = k_0\sqrt{\epsilon_c - \epsilon_s}$, resulting in,

$$\begin{aligned} \tanh(2k_0 a_{\text{cutoff}} \sqrt{\epsilon_c + |\epsilon_f|}) &= \frac{|p_s| \sqrt{\epsilon_c - \epsilon_s}}{\sqrt{\epsilon_c + |\epsilon_f|}} = \frac{|\epsilon_f| \sqrt{\epsilon_c - \epsilon_s}}{\epsilon_s \sqrt{\epsilon_c + |\epsilon_f|}} \\ 2k_0 a_{\text{cutoff}} &= \frac{1}{\sqrt{\epsilon_c + |\epsilon_f|}} \text{atanh} \left[\frac{|\epsilon_f| \sqrt{\epsilon_c - \epsilon_s}}{\epsilon_s \sqrt{\epsilon_c + |\epsilon_f|}} \right], \quad \beta_{\text{cutoff}} = k_0\sqrt{\epsilon_c} \end{aligned} \quad (10.10.12)$$

The argument of the arc-tanh function must be less than unity, but this follows from the condition $\beta_{s,\infty} > k_0\sqrt{\epsilon_c}$, which implies the allowed range for ϵ_c , or for ϵ_s , [941,942],

$$\epsilon_s \leq \epsilon_c \leq \frac{|\epsilon_f| \epsilon_s}{|\epsilon_f| - \epsilon_s} \Leftrightarrow \frac{|\epsilon_f| \epsilon_c}{|\epsilon_f| + \epsilon_c} \leq \epsilon_s \leq \epsilon_c \quad (10.10.13)$$

To decide whether this solution is even-like or odd-like, we look at Eqs. (10.10.5). Because both denominators are positive, then $e^{2\gamma a + 2\psi}$ will be positive and greater than

unity, so that ψ must be real. Therefore, this mode is TM_0 even-like. As the film thickness increases, it resembles more and more a single surface plasmon at the f - s interface, while the f - c interface does not support a surface plasmon because $\beta_{c,\infty}^2 < 0$.

Example 10.10.1: We consider an example from [943] in order to compare our results. It has parameters, $\epsilon_c = 2.2^2$, $\epsilon_s = 1.7^2$, $\epsilon_f = -2^2$, and half-width, $k_0 a = 0.3$. It belongs to region-1 because $|p_c| = 0.8264$, $|p_s| = 0.3841$, and $|p_s| - |p_c| = 0.5576 < 1$. The left graph in Fig. 10.10.2 shows the transverse magnetic field $H_y(x)$. The calculated mode parameters were,

$k_0 a$	β/k_0	γ/k_0	α_c/k_0	α_s/k_0	ψ	E
0.3	2.9165	3.5364	1.9147	2.3699	-0.5794	1.1102×10^{-16}

where E is the computational error for the characteristic equation (10.10.6),[†]

$$E = \left| \tanh(2\gamma a) - \frac{\gamma(|p_c|\alpha_c + |p_s|\alpha_s)}{\gamma^2 + |p_c|\alpha_c|p_s|\alpha_s} \right|$$

The solution was obtained with the help of the MATLAB function, **pwga**,

```
k0 = 1; la0 = 2*pi/k0; a = 0.3;
ec = 2.2^2; es = 1.7^2; ef = -2;
pc = ef/ec; ps = ef/es;
tol = 1e-12;

bsinf = sqrt(ef*es/(ef+es));

[be,E] = pwga(la0,ef,ec,es, a, bsinf, 0, tol); % TM0 mode

ga = sqrt(be^2-ef); as = sqrt(be^2-es); ac = sqrt(be^2-ec);
```

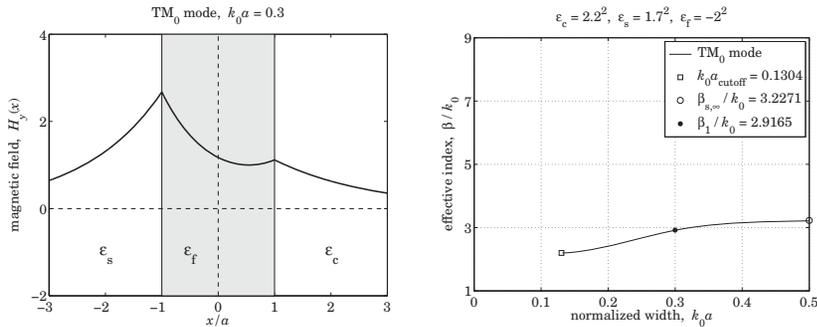


Fig. 10.10.2 DMD region-1, TM_0 mode.

It can also be computed recursively by transforming the characteristic equation into the iteration:

$$\gamma = -\frac{\gamma^2 + p_c \alpha_c p_s \alpha_s}{p_c \alpha_c + p_s \alpha_s} \tanh(2\gamma a) \Rightarrow \gamma_{n+1} = -\frac{\gamma_n^2 + p_c \alpha_c p_s \alpha_s}{p_c \alpha_c + p_s \alpha_s} \tanh(2\gamma_n a)$$

[†]The achieved E is actually half of MATLAB's machine epsilon, $\text{eps} = 2.2204 \times 10^{-16}$.

The magnetic field was computed and plotted by the code,

```
psi = atanh(-pc*ac/ga)/2 - atanh(-ps*as/ga)/2

x = linspace(-3,3,601)*a;

Hy = cosh(ga*a - psi).*exp(as*(x+a)).*(x<-a) + ...
      cosh(ga*a + psi).*exp(-ac*(x-a)).*(x>a) + ...
      cosh(ga*x + psi).*(abs(x)<=a);

fill([-1, 1, 1, -1], [-2 -2, 4, 4], [0.9, 0.9, 0.9]);
hold on
plot(x/a, real(Hy), 'linewidth',2);
```

The right graph in Fig. 10.10.1 shows the β - a dispersion relationship, with the half-width varying over $a_{\text{cutoff}} \leq a < \infty$, as β varies over $k_0 \sqrt{\epsilon_c} \leq \beta < \beta_{s,\infty}$. The calculated values were,

$$k_0 a_{\text{cutoff}} = 0.1304, \quad \sqrt{\epsilon_c} = 2.2, \quad \beta_{s,\infty} = 3.2271$$

and the graph was generated by the following code, where for plotting convenience, we evaluated the solution only up to $k_0 a_{\text{max}} = 0.5$ where β has effectively converged to $\beta_{s,\infty}$,

```
acut = 1/2/sqrt(ec-ef) * atanh(-ef*sqrt(ec-es)/es/sqrt(ec-ef));
amax = 0.5;

a = linspace(1.01*acut,amax, 101); % span interval a_cut < a < amax

for i=1:length(a)
    be(i) = pwga(la0,ef,ec,es, a(i), bsinf, 0, tol);
end

be = abs(real(be)); % make sure beta is positive real

plot(a,be, acut,sqrt(ec),'s', amax,bsinf,'o');
hold on
plot(a1,be1,'.'); % here, a1 = 0.3, be1 = 2.9165
xlim([0,0.5]); ylim([1,9])
```

Region - 2

Here, it follows from Eqs. (10.10.3) and (10.10.4) that, $0 < \beta_{s,\infty} < k_0 \sqrt{\epsilon_c} < \beta_{c,\infty}$. This excludes a solution that would have, $|p_s|\alpha_s - \gamma < 0$, because it would require from Eqs. (10.10.2) that $\beta < \beta_{s,\infty}$ and therefore, it could not satisfy (10.10.7). On the other hand, a solution exists that satisfies, $|p_{c,s}|\alpha_{c,s} - \gamma > 0$, which requires from (10.10.2) that $\beta \geq \beta_{c,\infty}$ and $\beta \geq \beta_{s,\infty}$, the latter being satisfied a fortiori from the former.

Thus, such solution will have range $\beta \geq \beta_{c,\infty}$, with $\beta = \beta_{c,\infty}$ achieved at infinite thickness $a \rightarrow \infty$, and at the other end, $\beta \rightarrow \infty$ as the thickness tends to zero, $a \rightarrow 0$. Interestingly, although the s - f interface can support a surface plasmon, it is not the limit of one of the modes.

Because $|p_{c,s}|\alpha_{c,s} - \gamma > 0$, it follows from Eqs. (10.10.5) that the left-hand sides $e^{2\gamma a \pm 2\psi}$ will be negative, and that can only happen if ψ has a $j\pi/2$ imaginary part.

Therefore, this solution is a TM_1 odd-like mode, with ψ given by (10.1.16),

$$\psi = \frac{1}{2} \operatorname{atanh} \left(-\frac{p_c \alpha_c}{\gamma} \right) - \frac{1}{2} \operatorname{atanh} \left(-\frac{p_s \alpha_s}{\gamma} \right) + \frac{1}{2} j\pi \quad (10.10.14)$$

Example 10.10.2: This example is from [943]. It has parameters, $\epsilon_c = 1.9^2$, $\epsilon_s = 1.3^2$, $\epsilon_f = -2^2$, and half-width, $k_0 a = 0.15$. It belongs to region-2 because $|p_c| = 1.1080$, $|p_s| = 2.3669$, and $|p_s| - |p_c| = 1.2588 > 1$. The left graph in Fig. 10.10.3 shows the transverse magnetic field $H_y(x)$. The calculated mode parameters were,

$k_0 a$	β/k_0	γ/k_0	α_c/k_0	α_s/k_0	ψ	E
0.15	7.8273	8.0788	7.5932	7.7186	$0.7368 + j\pi/2$	1.1102×10^{-16}

where E is the computational error for the characteristic equation (10.10.6),

$$E = \left| \tanh(2\gamma a) - \frac{\gamma(|p_c|\alpha_c + |p_s|\alpha_s)}{\gamma^2 + |p_c|\alpha_c|p_s|\alpha_s} \right|$$

The solution was obtained with the help of the MATLAB function, **pwga**,

```
k0 = 1; la0 = 2*pi/k0; a = 0.15;
ec = 1.9^2; es = 1.3^2; ef = -2^2;
pc = ef/ec; ps = ef/es;
tol = 1e-12;

bcinf = sqrt(ef*ec/(ef+ec));

[be,E] = pwga(la0,ef,ec,es, a, bcinf, 1, tol); % run with mode=1

ga = sqrt(be^2-ef); as = sqrt(be^2-es); ac = sqrt(be^2-ec);
```

It can also be computed recursively by transforming the characteristic equation into the iteration:

$$\gamma = \sqrt{-p_c \alpha_c p_s \alpha_s - \gamma(p_c \alpha_c + p_s \alpha_s) \coth(2\gamma a)}$$

$$\gamma_{n+1} = \sqrt{-p_c \alpha_{cn} p_s \alpha_{sn} - \gamma(p_c \alpha_{cn} + p_s \alpha_{sn}) \coth(2\gamma_n a)}$$

The magnetic field was computed and plotted by the code,

```
psi = atanh(-pc*ac/ga)/2 - atanh(-ps*as/ga)/2 + j*pi/2

x = linspace(-3,3,601)*a;

Hy = cosh(ga*a - psi).*exp(as*(x+a)).*(x<-a) + ...
     cosh(ga*a + psi).*exp(-ac*(x-a)).*(x>a) + ...
     cosh(ga*x + psi).*(abs(x)<=a);

fill([-1, 1, 1, -1], [-2 -2, 4, 4], [0.9, 0.9, 0.9]);
hold on
plot(x/a, imag(Hy), 'linewidth',2); % because cosh(ga*x+psi) = imaginary
```

The right graph in Fig. 10.10.2 shows the β - a dispersion relationship, with the half-width varying over $0 < a < \infty$, as β varies over $\infty > \beta > \beta_{c,\infty}$, with $\beta_{c,\infty}/k_0 = 6.0849$. The graph was generated by the following code, where for plotting convenience, the solution was evaluated only over the interval $a_{\min} \leq a \leq a_{\max}$, with $k_0 a_{\min} = 0.1$ and $k_0 a_{\max} = 0.5$.

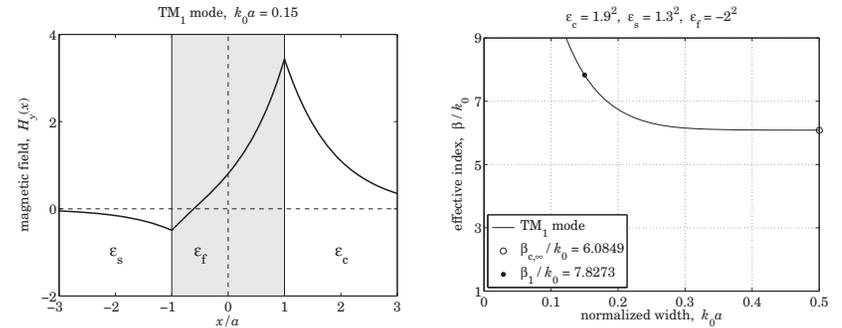


Fig. 10.10.3 DMD region-2, TM_1 mode.

```
amin = 0.1; amax = 0.5;

a = linspace(amin, amax, 101);

for i=1:length(a),
    [be(i),Err(i)] = pwga(la0,ef,ec,es, a(i), bcinf, 1, tol);
end

be = abs(real(be)); % make sure beta is positive real

plot(a,be, amax,bcinf, 'o', 'markersize',8);
hold on
plot(a1,be1, '.', 'markersize',20); % here a1 = 0.15, be1 = 7.8273
xlim([0,0.5]); ylim([1,9])
```

Region - 3

In this case, it follows from Eqs. (10.10.3) and (10.10.4) that, $k_0 \sqrt{\epsilon_c} < \beta_{s,\infty} < \beta_{c,\infty}$. Therefore, the solution that had been excluded in region-2, that would have, $|p_s|\alpha_s - \gamma < 0$, is now allowed and will have a range, $k_0 \sqrt{\epsilon_c} \leq \beta < \beta_{s,\infty}$. Moreover, because $|p_{c,s}|\alpha_{c,s} - \gamma < 0$, then Eq. (10.10.5) implies that ψ must be real-valued, and hence, this is a TM_0 even-like mode. Setting $\beta = k_0 \sqrt{\epsilon_c}$, defines a lower cutoff thickness, given exactly by Eq. (10.10.12).

Similarly, the solution that satisfies, $|p_{c,s}|\alpha_{c,s} - \gamma > 0$, also exists, and as in region-2, it will have range, $\beta_{c,\infty} < \beta < \infty$, and will be a TM_1 odd-like mode. To summarize, there are two types of solutions:

$$TM_1 \text{ with no cutoff, } 0 < a < \infty, \quad \beta_{c,\infty} < \beta < \infty$$

$$TM_0 \text{ with lower cutoff, } a_{\text{cutoff}} < a < \infty, \quad k_0 \sqrt{\epsilon_c} < \beta < \beta_{s,\infty}$$

Example 10.10.3: This example is also from [943] and has parameters, $\epsilon_c = 1.9^2$, $\epsilon_s = 1.8^2$, $\epsilon_f = -2^2$, $k_0 a = 0.3$. It belongs to region-3 because $|p_c| = 1.1080$, $|p_s| = 1.2346$, $|p_s| - |p_c| = 0.1265 < 1$. The calculated asymptotic values for the TM_0 and TM_1 modes

are, $\beta_{s,\infty}/k_0 = 4.1295$, $\beta_{c,\infty}/k_0 = 6.0849$, and the lower cutoff, $k_0 a_{\text{cutoff}} = 0.0506$. The calculated wavenumbers are,

mode	β/k_0	γ/k_0	α_c/k_0	α_s/k_0	ψ	E
TM ₀	3.1487	3.7302	2.5109	2.5835	-0.1556	1.1102×10^{-16}
TM ₁	6.4132	6.7178	6.1253	6.1554	$0.6215 + j\pi/2$	1.1102×10^{-16}

The MATLAB code is identical to that of Example 10.10.1 for the TM₀ mode, and to that of Example 10.10.2 for the TM₁ mode. The following recursions can also be used for the calculation,

$$\text{TM}_0 : \quad y_{n+1} = -\frac{\gamma_n^2 + p_c \alpha_{cn} p_s \alpha_{sn}}{p_c \alpha_{cn} + p_s \alpha_{sn}} \tanh(2\gamma_n a)$$

$$\text{TM}_1 : \quad y_{n+1} = \sqrt{-p_c \alpha_{cn} p_s \alpha_{sn} - \gamma(p_c \alpha_{cn} + p_s \alpha_{sn})} \coth(2\gamma_n a)$$

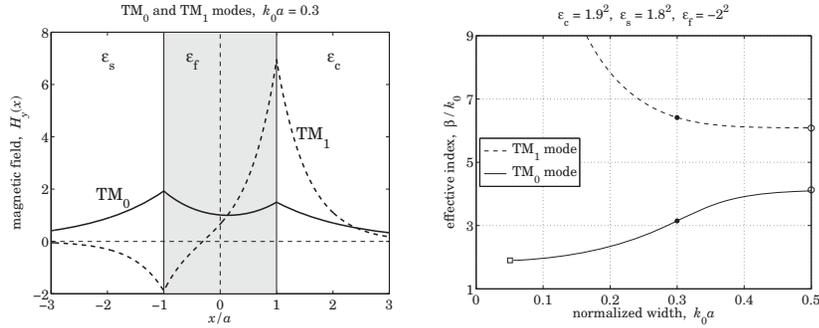


Fig. 10.10.4 DMD region-3, TM₀ and TM₁ modes.

Fig. 10.10.4 shows the magnetic field profiles and the β - a relationships, displaying also the cutoff point, the specific solutions for $k_0 a = 0.3$, and the asymptotic values. \square

Region-3 is the most relevant case in practice, and in fact, as we shall see in Sec. 10.11, the TM₀ mode becomes in the lossy case the so-called *long-range surface plasmon polariton* (LR-SPP) mode that has received a lot of attention in the literature [962,981].

Region – 4

In region-4, because $|p_c| \leq |p_s| \leq 1$, neither metal-dielectric interface can support a separate surface plasmon because both $\beta_{s,\infty}^2, \beta_{c,\infty}^2$ are negative. Nevertheless, for sufficiently narrow widths, there is a solution. Because $\alpha_c \leq \alpha_s \leq \gamma$, it follows that $|p_{c,s}| \alpha_{c,s} - \gamma < 0$, and therefore from Eq. (10.10.8), the solution will be a TM₀ mode.

Eqs. (10.10.2) and (10.10.3) are automatically satisfied and impose no further restrictions on the range of β . Thus, β ranges over, $k_0 \sqrt{\epsilon_c} \leq \beta < \infty$. Over this range, we may think of Eq. (10.10.6) as defining the width a as a function of β , indeed,

$$a = \frac{1}{4\gamma} \ln \left[\frac{(\gamma + |p_c| \alpha_c)(\gamma + |p_s| \alpha_s)}{(\gamma - |p_c| \alpha_c)(\gamma - |p_s| \alpha_s)} \right] \equiv F(\beta) \quad (10.10.15)$$

This function has the property that, starting at $\beta = k_0 \sqrt{\epsilon_c}$ with $a = a_{\text{cutoff}}$ of Eq. (10.10.12), it first increases up to a maximum value, say, a_{max} at some $\beta = \beta_{\text{max}}$, and then decreases to zero for $\beta > \beta_{\text{max}}$. Therefore, we may distinguish two solution branches:

$$\begin{aligned} \text{(lower branch),} \quad & k_0 \sqrt{\epsilon_c} \leq \beta \leq \beta_{\text{max}}, \quad a_{\text{cutoff}} \leq a \leq a_{\text{max}} \\ \text{(upper branch),} \quad & \beta_{\text{max}} \leq \beta < \infty, \quad a_{\text{max}} \geq a > 0 \end{aligned} \quad (10.10.16)$$

The maximum point, $\beta_{\text{max}}, a_{\text{max}}$, can easily be determined using MATLAB's built-in function, **fminbnd**. Once β_{max} is known, the two solutions can be found using the function, **pwga**, by initializing it above and below β_{max} .

Example 10.10.4: This example, from [943], is defined by the parameters, $\epsilon_c = 2.2^2$, $\epsilon_s = 2.1^2$, $\epsilon_f = -2^2$, $k_0 a = 0.15$, and has $|p_c| = 0.8264$, $|p_s| = 0.9070$, and $k_0 a_{\text{cutoff}} = 0.0341$, and imaginary $\beta_{s,\infty}/k_0 = 6.5593j$, $\beta_{c,\infty}/k_0 = 4.8008j$.

The maximum point is, $\beta_{\text{max}}/k_0 = 4.2090$ and $k_0 a_{\text{max}} = 0.1759$, and was calculated by the following MATLAB code, which defines the function $F(\beta)$ and passes its negative into the **fminbnd** function, with a search interval $[\sqrt{\epsilon_c}, \infty)$, where ∞ is numerically replaced by the inverse of the machine epsilon,

```
k0 = 1; la0 = 2*pi/k0; a = 0.15;
ec = 2.2^2; es = 2.1^2; ef = -2^2;
pc = ef/ec; ps = ef/es;

ga = @(b) sqrt(b.^2 - ef); % auxiliary functions
ac = @(b) sqrt(b.^2 - ec);
as = @(b) sqrt(b.^2 - es);
F = @(b) 1/4./ga(b) .* log((ga(b)-pc*ac(b)).*(ga(b)-ps*as(b))./...
(ga(b)+pc*ac(b))./(ga(b)+ps*as(b)));

bmax = fminbnd(@(b) -F(b), sqrt(ec), 1/eps);
amax = F(bmax);
```

With β_{max} at hand, we search above and below it for the TM₀ solutions corresponding to the value $k_0 a = 0.15$. This is implemented by the code segment,

```
tol = 1e-12;

[be1,E1] = pwga(la0,ef,ec,es, a, bmax/2, 0, tol); % lower
be1 = abs(real(be1));
ga1 = sqrt(be1^2-ef); as1 = sqrt(be1^2-es); ac1 = sqrt(be1^2-ec);

[be2,E2] = pwga(la0,ef,ec,es, a, bmax*2, 0, tol); % upper
be2 = abs(real(be2));
ga2 = sqrt(be2^2-ef); as2 = sqrt(be2^2-es); ac2 = sqrt(be2^2-ec);

[Ptot,Pf,Pc,Ps] = pwgpower(a,ef,ec,es,[be1; be2], 0); % transmitted powers
```

The computed values were,

mode	β/k_0	γ/k_0	α_c/k_0	α_s/k_0	ψ	E
lower	2.9271	3.5452	1.9309	2.5835	-0.0469	1.1102×10^{-16}
upper	6.9626	7.2441	6.6059	6.6383	-0.1053	7.6494×10^{-14}

Using the function, `pwgpower`, the above code segment also calculates the transmitted powers within each medium, as well as the total power for the two solutions,

mode	P_{tot}	P_f	P_c	P_s
lower	0.1780	-0.2420	0.1964	0.2236
upper	-0.1570	-0.7937	0.2520	0.3847

As noted in [943], the total power is negative for the upper mode, implying that it has negative group velocity. See Refs. [965–980] for more on the issue of backward waves and negative group velocity in plasmonic waveguides.

Fig. 10.10.5 shows the corresponding magnetic field profiles computed from Eq. (10.10.10), as well as the β – a dispersion relationship, with the above computed points included.

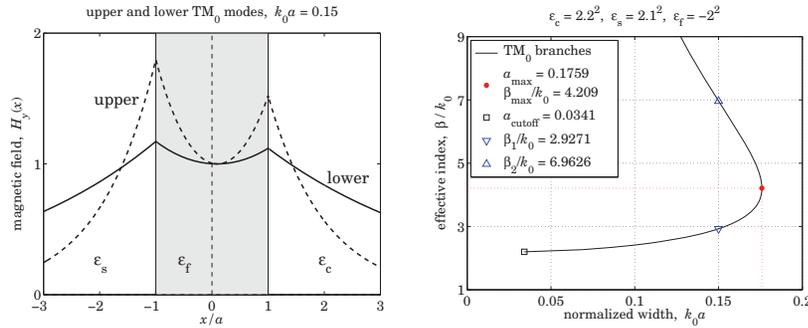


Fig. 10.10.5 DMD region-4, upper and lower TM₀ modes.

10.11 DMD Configuration – Lossy Case

We assume now that the metal film is lossy, $\epsilon_f = -\epsilon_R - j\epsilon_I$, with $\epsilon_R > 0$ and $\epsilon_I > 0$. The dielectric claddings will be assumed to be lossless with $\epsilon_c \geq \epsilon_s > 0$. We will also assume that $\epsilon_R > \epsilon_c \geq \epsilon_s$, which is satisfied by typical metals and typical dielectrics in the optical and infrared regimes.

The presence of losses causes β to develop a negative imaginary part, $\beta = \beta_R - j\beta_I$, with the wave attenuating like, $e^{-j\beta z} = e^{-j\beta_R z} e^{-\beta_I z}$, as it propagates in the positive z -direction. The effective index, propagation length, and propagation loss in dB/m are defined by (β_I is in units of nepers/m),

$$n_{\text{eff}} = \frac{\text{Re}(\beta)}{k_0} = \frac{\beta_R}{k_0}, \quad L = -\frac{1}{2\text{Im}(\beta)} = \frac{1}{2\beta_I}, \quad \text{dB} = 20 \log_{10}(e) \beta_I \quad (10.11.1)$$

We will discuss the impact of losses, as well as the impact of asymmetry ($\epsilon_c \neq \epsilon_s$), on the two basic plasmonic modes, the even-like TM₀ and the odd-like TM₁ modes.

These modes have drastically different behavior as the film thickness is varied, with the TM₀ mode having an *increasing* propagation length as the film width decreases, while

the TM₁ mode has a *decreasing* propagation length. Because of this property, the TM₀ and TM₁ modes are called the *long-range* (LRSP) and *short-range* (SRSP) surface plasmon modes. Equivalently, the propagation losses, quantified by the value of β_I , tend to zero for the LRSP mode, and to infinity for SRSP, as the film width tends to zero for symmetric guides (for asymmetric ones, there is a lower cutoff width).

A tradeoff to the LRSP property, however, is that as the film width becomes smaller and the propagation length longer, the guided wave becomes less confined laterally, penetrating more deeply into the dielectric claddings. We discuss this below.

The properties of long-range surface plasmons in DMD waveguides have been reviewed by Berini [962] with an extensive bibliography, including the impact of asymmetry. Another review is [981] and earlier [957]. Some of the earliest references on long-range surface plasmons are [937–942], on which we have based some of our computational examples.

10.12 Symmetric DMD Waveguides

We begin with the symmetric case ($\epsilon_c = \epsilon_s$) for which the TM₀ and TM₁ modes are exactly even or odd. The characteristic equations (10.1.22) can be written in the equivalent forms:

$$\frac{y - p_c \alpha_c}{y + p_c \alpha_c} = e^{2ya + 2\psi} = \begin{cases} e^{2ya}, & \psi = 0, & \text{TM}_0 \text{ mode, even, (LRSP)} \\ -e^{2ya}, & \psi = \frac{j\pi}{2}, & \text{TM}_1 \text{ mode, odd, (SRSP)} \end{cases} \quad (10.12.1)$$

$$\frac{p_c \alpha_c}{y} = -\tanh(\gamma a + \psi) = \begin{cases} -\tanh(\gamma a), & \psi = 0, & \text{(LRSP)} \\ -\coth(\gamma a), & \psi = \frac{j\pi}{2}, & \text{(SRSP)} \end{cases} \quad (10.12.2)$$

The even or odd labeling refers to the transverse magnetic field,

$$H_y(x) = \begin{cases} H_0 \cosh(\gamma x + \psi), & |x| \leq a \\ H_0 \cosh(\gamma a + \psi) e^{-\alpha_c(x-a)}, & x \geq a \\ H_0 \cosh(\gamma a - \psi) e^{\alpha_c(x+a)}, & x \leq -a \end{cases} \quad (10.12.3)$$

The reduced losses for the LRSP case can be intuitively understood by noting that the ohmic losses are due to the longitudinal electric field $E_z(x)$, which has the opposite symmetry than $H_y(x)$, and it is odd for the LRSP mode, thus, having a zero crossing within the metal film reducing its strength.

A very efficient way of solving the characteristic equations (10.12.2) is by turning them into the following iterative algorithm, which converges extremely fast, requiring very few iterations, like 5–6,

$$\alpha_c = -\frac{1}{p_c} y \tanh(\gamma a + \psi) \Rightarrow \begin{cases} y_n = \sqrt{\alpha_{c,n}^2 + k_0^2(\epsilon_c - \epsilon_f)} \\ \alpha_{c,n+1} = -\frac{1}{p_c} y_n \tanh(\gamma_n a + \psi) \end{cases} \quad (10.12.4)$$

The iteration can be initialized at $\alpha_{c,0} = 0$. In fact, just one step of the iteration provides an excellent approximation [981] and results in a simple closed-form expression. Indeed, from $\alpha_{c,0} = 0$, we have, $\gamma_0 = \sqrt{\alpha_{c,0}^2 + k_0^2(\epsilon_c - \epsilon_f)} = k_0\sqrt{\epsilon_c - \epsilon_f}$, and the next α_c and corresponding β are,

$$\alpha_{c1} = -\frac{1}{p_c} \gamma_0 \tanh(\gamma_0 a + \psi)$$

$$\beta = \sqrt{k_0^2 \epsilon_c + \alpha_{c1}^2} = k_0 \sqrt{\epsilon_c + \frac{\epsilon_c^2}{\epsilon_f^2} (\epsilon_c - \epsilon_f) \tanh^2(k_0 a \sqrt{\epsilon_c - \epsilon_f} + \psi)}$$

or, specifically, in the LRSP and SRSP cases, ($\psi = 0$, $\psi = j\pi/2$),

$$\beta = k_0 \sqrt{\epsilon_c + \frac{\epsilon_c^2}{\epsilon_f^2} (\epsilon_c - \epsilon_f) \tanh^2(k_0 a \sqrt{\epsilon_c - \epsilon_f})} \quad (\text{LRSP})$$

$$\beta = k_0 \sqrt{\epsilon_c + \frac{\epsilon_c^2}{\epsilon_f^2} (\epsilon_c - \epsilon_f) \coth^2(k_0 a \sqrt{\epsilon_c - \epsilon_f})} \quad (\text{SRSP})$$

(10.12.5)

The following MATLAB function, **dmds**, implements the iteration (10.12.4), with the iteration stopping when two successive values of α_c become closer to each other than a specified error tolerance, such as, $tol = 10^{-12}$. It has usage,

```
[be,E,N] = dmds(la0,ef,ec,a,mode,tol,be0); % symmetric DMD guide - iterative solution

la0 = operating wavelength, k0 = 2*pi/la0 = free-space wavenumber
ef,ec = permittivities of film and cladding/substrate
a = vector of half-widths of film, in same units as la0
mode = 0,1 for TMO or TMI mode
tol = computational error tolerance, default tol = 1e-12
be0 = starting search point in units of k0 - default be0 = sqrt(ec)

be = vector of propagation wavenumbers in units of k0 - size(a)
E = vector of computational errors of characteristic equation - size(a)
N = number of iterations to converge to within tol
```

In the limit of small thickness a , Eqs. (10.12.5) simplify further by making the small- x approximation, $\tanh(x) \approx x$,

$$\beta = k_0 \sqrt{\epsilon_c + \frac{\epsilon_c^2}{\epsilon_f^2} (\epsilon_c - \epsilon_f)^2 (k_0 a)^2} \quad (\text{LRSP})$$

$$\beta = k_0 \sqrt{\epsilon_c + \frac{\epsilon_c^2}{\epsilon_f^2} (k_0 a)^2} \quad (\text{SRSP})$$

(10.12.6)

These can also be obtained by applying the same approximation, directly to the characteristic equations (10.12.2), that is,

$$\gamma a \approx -\frac{p_c \alpha_c}{\gamma} \quad (\text{LRSP}) \quad \text{and} \quad \frac{1}{\gamma a} \approx -\frac{p_c \alpha_c}{\gamma} \quad (\text{SRSP}), \quad \text{or,}$$

$$\gamma^2 a = -p_c \alpha_c \quad (\text{LRSP}) \quad \text{and} \quad \alpha_c = -\frac{1}{p_c a} \quad (\text{SRSP})$$

For the SRSP case, substituting α_c into $\beta = \sqrt{k_0^2 \epsilon_c + \alpha_c^2}$ leads to Eq. (10.12.6). For the LRSP case, an additional step is required by writing,

$$-p_c \alpha_c = \gamma^2 a = [\alpha_c^2 + k_0^2 (\epsilon_c - \epsilon_f)] a$$

and noting that this implies that α_c is already small and order- a , and therefore, the second-order term α_c^2 can be ignored on the right side, so that, $-p_c \alpha_c = k_0^2 (\epsilon_c - \epsilon_f) a$,

$$\alpha_c = -\frac{1}{p_c} k_0^2 (\epsilon_c - \epsilon_f) a = k_0 \frac{\epsilon_c (\epsilon_f - \epsilon_c)}{\epsilon_f} (k_0 a) \quad (10.12.7)$$

resulting in Eq. (10.12.6). Making the further approximation, $\sqrt{1+x} \approx 1+x/2$, we obtain the more explicit relationships for the real and imaginary parts of β in the LRSP case,

$$\beta = k_0 \sqrt{\epsilon_c + \frac{\epsilon_c^2}{\epsilon_f^2} (\epsilon_c - \epsilon_f)^2 (k_0 a)^2} \approx k_0 \sqrt{\epsilon_c} \left[1 + \frac{\epsilon_c (\epsilon_c - \epsilon_f)^2}{2\epsilon_f^2} (k_0 a)^2 \right] \quad (10.12.8)$$

Setting $\beta = \beta_R - j\beta_I$ and $\epsilon_f = -\epsilon_R - j\epsilon_I$, and noting that,

$$\frac{(\epsilon_c - \epsilon_f)^2}{\epsilon_f^2} = \frac{\epsilon_f^{*2} (\epsilon_c - \epsilon_f)^2}{|\epsilon_f|^4} = \frac{(|\epsilon_f|^2 - \epsilon_c \epsilon_f^*)^2}{|\epsilon_f|^4} = \frac{(|\epsilon_f|^2 + \epsilon_c \epsilon_R - j\epsilon_c \epsilon_I)^2}{|\epsilon_f|^4}$$

$$= \frac{(|\epsilon_f|^2 + \epsilon_c \epsilon_R)^2 - \epsilon_c^2 \epsilon_I^2 - 2j\epsilon_c \epsilon_I (|\epsilon_f|^2 + \epsilon_c \epsilon_R)}{|\epsilon_f|^4}$$

where $|\epsilon_f|^2 = \epsilon_R^2 + \epsilon_I^2$, we obtain the real and imaginary parts of β [942],

$$\beta_R = k_0 \sqrt{\epsilon_c} \left[1 + \frac{(|\epsilon_f|^2 + \epsilon_c \epsilon_R)^2 - \epsilon_c^2 \epsilon_I^2}{2|\epsilon_f|^4} \epsilon_c (k_0 a)^2 \right] \quad (\text{LRSP})$$

$$\beta_I = k_0 \sqrt{\epsilon_c} \frac{\epsilon_c^2 \epsilon_I (|\epsilon_f|^2 + \epsilon_c \epsilon_R)}{|\epsilon_f|^4} (k_0 a)^2 \quad (10.12.9)$$

The long-range property follows from the fact that β_I tends to zero for small a . This remarkable property comes with the *tradeoff* that as a decreases and the range increases, the wave becomes less confined laterally. This follows from Eq. (10.12.7) which shows that the lateral decay constant α_c also goes to zero with a . Its real and imaginary parts can be given explicitly by,

$$\alpha_c = k_0 \frac{\epsilon_c (|\epsilon_f|^2 - \epsilon_c \epsilon_f^*)}{|\epsilon_f|^2} (k_0 a) = k_0 \frac{\epsilon_c (|\epsilon_f|^2 + \epsilon_c \epsilon_R - j\epsilon_c \epsilon_I)}{|\epsilon_f|^2} (k_0 a) \quad (10.12.10)$$

The factor, $|\epsilon_f|^2 + \epsilon_c \epsilon_R$, is positive since we assumed $\epsilon_R > 0$, therefore, both β_I and the real-part of α_c are non-negative, as required for stability and lateral confinement.

For the short-range mode, both β_I and the real part of α_c increase like $1/a$, resulting in shorter propagation distances, but more confinement. These follow from the approximation (10.12.6), as $a \rightarrow 0$,

$$\beta \approx \alpha_c = -\frac{\epsilon_c}{\epsilon_f a} = k_0 \frac{\epsilon_c (\epsilon_R - j\epsilon_I)}{|\epsilon_f|^2 (k_0 a)} \quad (\text{SRSP}) \quad (10.12.11)$$

Example 10.12.1: This example is from [981] and demonstrates the primary features of the LRSP and SRSP modes. The operating wavelength is 775 nm, the metal film is gold with $\epsilon_f = -23.5 - 1.69j$,[†] surrounded by air, $\epsilon_c = \epsilon_s = 1$, and the film thickness $2a$ is varied over the range $1 \leq 2a \leq 210$ nm. Fig. 10.12.1 shows the effective index $n_{\text{eff}} = \beta_R/k_0$ as a function of a for both the LR and SP modes, as well as the propagation length $L = 1/(2\beta_I)$ in units of μm .

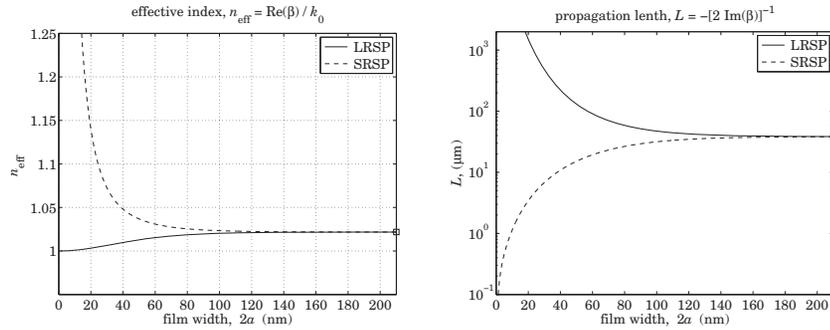


Fig. 10.12.1 Effective index and propagation length of long and short-range modes.

The graphs were generated with the help of the function, `dmads`, by the MATLAB code,

```

la0 = 775; k0 = 2*pi/la0; ec = 1; ef = -23.6 - 1.69*j;
pc = ef/ec;
bcinf = sqrt(ef*ec/(ef+ec)); % asymptotic value
ninf = real(bcinf);
Linf = -1/2/k0./imag(bcinf)/1000; % units of microns

w = 1:210; a = w/2; % film thicknesses

tol = 1e-12; % error tolerance

for i = 1:length(w)
    [be0(i), E0(i), N0(i)] = dmads(la0,ef,ec,a(i),0,tol); % LRSP
    [be1(i), E1(i), N1(i)] = dmads(la0,ef,ec,a(i),1,tol); % SRSP
end

% [be0, E0, N0] = dmads(la0,ef,ec,a,0,tol); % vectorized computation

neff0 = real(be0); L0 = -1/2/k0./imag(be0)/1000; % LRPP
neff1 = real(be1); L1 = -1/2/k0./imag(be1)/1000; % SRSP

```

[†]at 775 nm, the DRUDE function produces, $\epsilon_f = -18.57 - 2.008j$.

```

figure; plot(w,neff0,'-', w,neff1,'--', w(end),ninf,'s'); % eff. index
figure; semilogy(w,L0,'-', w,L1,'--', w(end),Linf,'s'); % propag. length

g0 = sqrt(ec-ef);

be0a = sqrt(ec + g0^2/pc^2 * tanh(g0*k0*a).^2); % approximation
be1a = sqrt(ec + g0^2/pc^2 * coth(g0*k0*a).^2);

diff0 = norm(be0 - be0a); % diff0 = 4.7377e-04
diff1 = norm(be1 - be1a); % diff1 = 0.0076

```

Because of the symmetric geometry, LRSP and SRSP both converge to the same asymptotic value for large a ,

$$\frac{\beta_{c,\infty}}{k_0} = \sqrt{\frac{\epsilon_c \epsilon_f}{\epsilon_c + \epsilon_f}} = 1.0218 - 0.0016j, \quad n_{\text{eff}} = \frac{\beta_R}{k_0} = 1.0218, \quad L = \frac{1}{2\beta_I} = 38.3023 \mu\text{m}$$

The code also computes the approximate solutions using Eq. (10.12.5), but does not plot them, because they are visually indistinguishable from the exact ones. As a measure of the approximation accuracy, the norms of the difference between the vector of the exact β s and the approximate ones, $\|\beta - \beta_{\text{approx}}\|$, was computed for the two cases - the values are shown in the comments. Fig. 10.12.2 shows the computational errors as defined in the function `dmads`,

$$E = |p_c \alpha_c + \gamma \tanh(\gamma a)| \quad (\text{LRSP})$$

$$E = |p_c \alpha_c + \gamma \coth(\gamma a)| \quad (\text{SRSP})$$

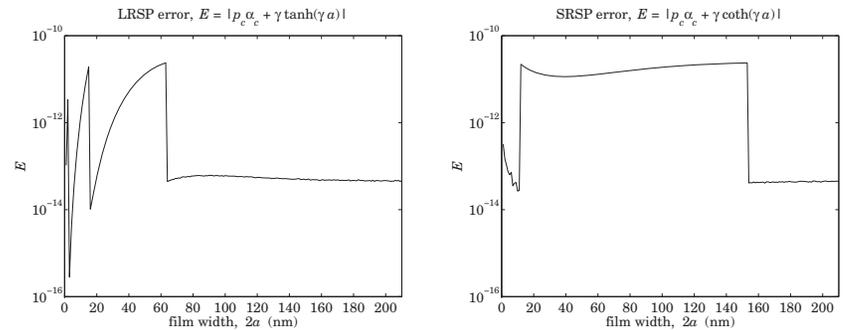


Fig. 10.12.2 Computational error.

The error values are small but not quite as small as the specified tolerance, $tol = 10^{-12}$, because tol measures the closeness of two successive iterates of α_c , not the value of the error E . The norms of the error vectors were $\|E\| = 8.0235 \times 10^{-11}$ and $\|E\| = 2.0993 \times 10^{-10}$.

The function `dmads` also calculates the number of iterations required to converge to within the specified error tolerance. The number of iterations for each thickness are plotted in Fig. 10.12.3. One observes how quick the iterative method is. \square

Example 10.12.2: This example, also from [981], illustrates the more realistic case of a gold film at the operating wavelength of 1550 nm, with permittivity, $\epsilon_f = -131.9475 - 12.65j$,[†]

[†] ϵ_f is from Palik [162], but the DRUDE function gives, $\epsilon_f = -103.3325 - 8.1301j$.

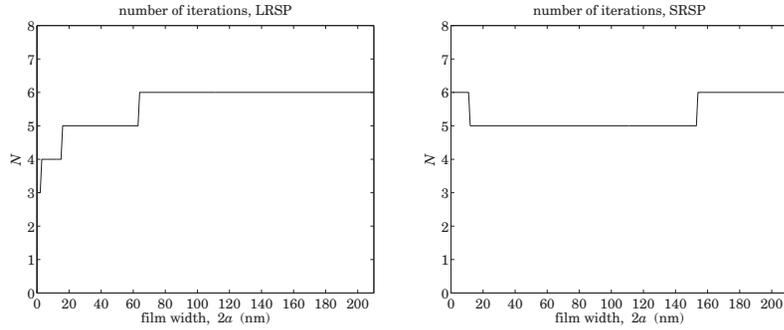


Fig. 10.12.3 Number of iterations to converge.

surrounded by the often-used benzocyclobutene (BCB) polymer dielectric, with $\epsilon_c = \epsilon_s = 1.535^2$. The film thickness is varied over the range $10 \leq 2a \leq 110$ nm. The MATLAB code is identical to that of the previous example, with the change in first two lines,

```
1a0 = 1550; k0 = 2*pi/1a0;
ec = 1.535^2; ef = -131.9475 - 12.65j;
```

Fig. 10.12.4 shows the effective index and the propagation loss measured in dB/mm, and computed by the formula, $\text{dB} = 20 \log_{10}(e) \beta_I = 8.68589 \beta_I$. The approximations (10.12.5) are still visually indistinguishable from the exact values, and are not plotted, nor are the number of iterations which are of the order of 4-5.

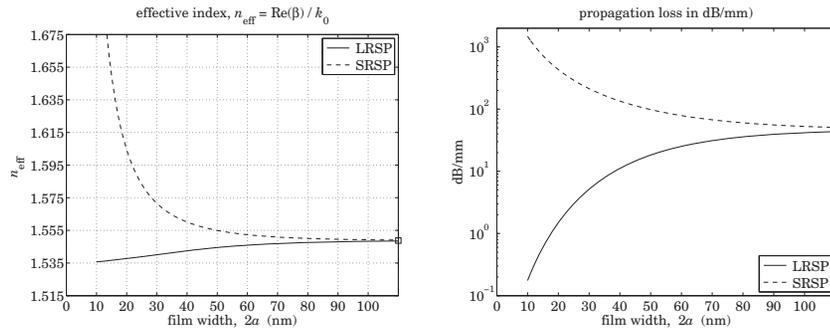


Fig. 10.12.4 Effective index and propagation length of long and short-range modes.

Fig. 10.12.5 shows the computational error for the LRSP and SRSP modes, and on the right, the lateral penetration depth into the two dielectric sides defined by $L_c = 1/\text{Re}(\alpha_c)$. We observe the basic tradeoff of DMD guides that even though the gold film has thickness of nanometers, and the propagation loss is fairly small, the fields are not very well confined laterally, penetrating at hundreds of micro-meters into the dielectrics. In this regard, MDM guides provide perhaps a better solution that results in much better lateral confinement, at the expense of shorter propagation distances. □

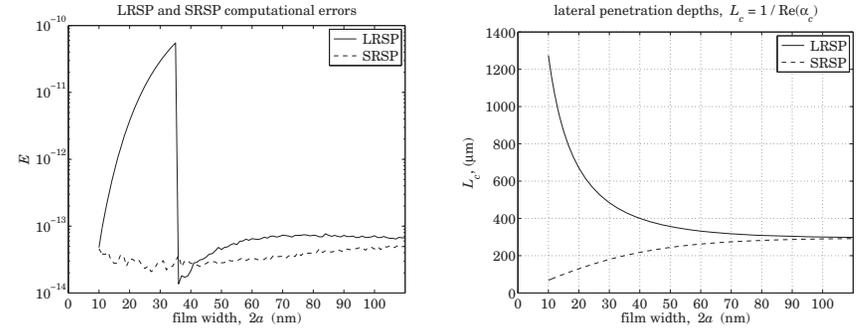


Fig. 10.12.5 Computational error.

Example 10.12.3: To illustrate the basic tradeoff between propagation distance and lateral confinement, we compare three waveguides, all operating at 1550 nm wavelength: (a) the LRSP mode of the DMD guide of the previous example with a 20 nm gold film embedded in BCB polymer, (b) the complementary MDM waveguide with a 50 nm BCB film and gold claddings, and (c) a surface plasmon on a single BCB-gold interface. For each case, we calculate β , α_c , and the propagation attenuation in dB/mm, $\text{dB} = 8.686 \beta_I$, the propagation distance, $L = (2\beta_I)^{-1}$, and the lateral penetration depth into the dielectric side, $L_c = 1/\text{Re}(\alpha_c)$. The following table shows the results.

guide	β/k_0	α_c/k_0	dB/mm	L (μm)	L_c (μm)
DMD	$1.537673 - 0.000042j$	$0.090632 - 0.000721j$	1.50	2904.47	2.72
single	$1.548762 - 0.001337j$	$0.206246 - 0.010042j$	47.09	92.23	1.20
MDM	$2.102705 - 0.024595j$	$11.690006 + 0.536637j$	865.97	5.02	0.02

They were computed by the following MATLAB code,

```
1a0 = 1550; k0 = 2*pi/1a0;
a1 = 20/2; a3 = 50/2; % DMD and MDM half-widths
ec = 1.535^2; ef = -131.9475-12.65j; % BCB and gold

Ec = [ec; ec; ef]; % interchange ec,ef for MDM
deb = 20*log10(exp(1)); % conversion factor to dB
tol = 1e-12; % error tolerance

be1 = dmdu(1a0,ef,ec,a1,0,tol); % DMD, 20 nm
be2 = sqrt(ef*ec/(ef+ec)); % single interface
be3 = pwg(1a0,ec,ef,a3,be2,0,tol); % MDM, 50 nm, interchange ec,ef

be = [be1; be2; be3]; % propagation wavenumbers
ac = sqrt(be.^2 - Ec); % lateral wavenumbers

dBmm = - imag(be) * k0 * deb * 1e6; % propagation loss in dB/mm
L = -1/2/k0./imag(be)/1000; % propagation length in microns
Lc = 1/k0./real(ac)/1000; % lateral depth in microns
```

```
num2str([be,ac], '%10.6f')
num2str([dBmm,L,Lc], ' %10.2f') % table
```

They are listed in order increasing propagation loss or decreasing propagation distance, and decreasing lateral confinement or increasing lateral penetration. □

Example 10.12.4: One of the earliest investigations of LRSP and SRSP modes in symmetric DMD guides was Kovacs [937]. The metal film was indium (In) at $\lambda_0 = 450$ nm wavelength, with permittivity $\epsilon_f = -20.358 - 6.019j$, symmetrically surrounded by magnesium fluoride (MgF_2) dielectric with $\epsilon_c = \epsilon_s = 1.382^2 = 1.9099$.

Kovacs computed the propagation parameters for both the LRSP and SRSP modes at three film thicknesses, $w = [10, 30, 50]$ nm, and noted their basic trends, namely, that the propagation length increases (decreases) for the LRSP (SRSP) mode as w gets smaller, and that both LRSP and SRSP tend to the same value (i.e., $\beta_{c,\infty}$) as w increases. The following MATLAB code reproduces the results of [937],

```
la0 = 450; k0=2*pi/la0; ef=-20.358-6.019j; es=1.382^2;
tol = 1e-12;

w = [10; 30; 50];

[be0,E0,N0] = dmds(la0,ef,ec,w/2,0,tol); % LRSP
ga0 = sqrt(be0.^2 - ef); % E0 = computational error
ac0 = sqrt(be0.^2 - ec); % N0 = number of iterations

L0 = -1/2/k0./imag(be0)/1000; % units of microns
Lc0 = 1./real(ac0)/k0; % units of nm

num2str([w, real(be0), imag(be0), L0, Lc0], '%11.5f') % make table

[be1,E1,N1] = dmds(la0,ef,ec,w/2,1,tol); % SRSP
ga1 = sqrt(be1.^2 - ef);
ac1 = sqrt(be1.^2 - ec);

L1 = -1/2/k0./imag(be1)/1000;
Lc1 = 1./real(ac1)/k0;

num2str([w, real(be1), imag(be1), L1, Lc1], '%11.5f')

bcinf = sqrt(ec*ef/(ec+ef)); % w = inf
gcinf = sqrt(bcinf^2 - ef);
acinf = sqrt(bcinf^2 - ec);

Linf = -1/2/k0./imag(bcinf)/1000;
Lcinf = 1./real(acinf)/k0;

num2str([real(bcinf), imag(bcinf), Linf, Lcinf], '%11.5f')
```

The computed values of $\beta = \beta_R - j\beta_I$, propagation length, $L = 1/(2\beta_I)$, and lateral penetration depth into the dielectric, $L_c = 1/\text{Re}(\alpha_c)$, are given below.

		LRSP	
w (nm)	β/k_0	L (μm)	L_c (nm)
10	$1.38905 - 0.00047j$	76.88	512.00
20	$1.42100 - 0.00658j$	5.44	215.87
50	$1.43873 - 0.01443j$	2.48	177.69
∞	$1.44535 - 0.02017j$	1.76	167.26

		SRSP	
w (nm)	β/k_0	L (μm)	L_c (nm)
10	$1.86878 - 0.25421j$	0.14	55.68
20	$1.48238 - 0.04757j$	0.75	130.38
50	$1.45247 - 0.02707j$	1.32	157.60
∞	$1.44535 - 0.02017j$	1.76	167.26

where $\beta = \beta_{c,\infty} = \sqrt{\epsilon_c \epsilon_f / (\epsilon_c + \epsilon_f)}$ at $w = \infty$. The norm of the computational error (for the three thicknesses) and the number of iterations to converge were $\|E_0\| = 1.5035 \times 10^{-14}$, $N_0 = 7$ for the LRSP case, and $\|E_1\| = 4.9127 \times 10^{-15}$, $N_1 = 7$ for SRSP. □

10.13 Asymmetric DMD Waveguides

The main result for asymmetric guides is that, just like the lossless case, the asymmetry introduces a lower cutoff thickness for the TM_0 , LRSP, mode, and that the propagation length increases dramatically as the thickness approaches the cutoff from above, but at the expense of becoming less confined laterally. For a given film thickness, there is a critical value of the permittivity ϵ_c that achieves much larger propagation lengths as compared to the symmetric case ($\epsilon_c = \epsilon_s$) of the same thickness. Of course, one could also lower the thickness of the symmetric case to increase the propagation length, as we saw in the previous section.

To simplify the discussion, we will assume that the metal, $\epsilon_f = -\epsilon_R - j\epsilon_I$, where $\epsilon_R > 0$, $\epsilon_I \geq 0$, and lossless dielectrics, $\epsilon_c, \epsilon_s > 0$, satisfy the condition $\epsilon_R > \epsilon_c \geq \epsilon_s$, as well as the region-3 condition of the lossless case, that is,

$$\frac{\epsilon_R}{\epsilon_s} \leq \frac{\epsilon_R}{\epsilon_c} + 1 \quad (10.13.1)$$

which implies the following range restrictions on ϵ_s, ϵ_c , also given in Eq. (10.10.13),

$$\epsilon_s \leq \epsilon_c \leq \frac{\epsilon_R \epsilon_s}{\epsilon_R - \epsilon_s} \Leftrightarrow \frac{\epsilon_R \epsilon_c}{\epsilon_R + \epsilon_c} \leq \epsilon_s \leq \epsilon_c \quad (10.13.2)$$

We recall from the region-3 discussion that in the lossless case ($\epsilon_f = -\epsilon_R$), the quantity $\beta_{s,\infty}$ is equal to $\epsilon_R \epsilon_s / (\epsilon_R - \epsilon_s)$, and that Eq. (10.13.2) implies the existence of two modes: a TM_0 mode with range, $k_0 \sqrt{\epsilon_c} \leq \beta < \beta_{s,\infty}$, with a lower thickness cutoff, $a_{\text{cutoff}} \leq a < \infty$, and a TM_1 mode with range, $\beta \geq \beta_{c,\infty}$, and no cutoff. The cutoff width

is realized at $\beta = k_0\sqrt{\varepsilon_c}$ and is given by Eq. (10.10.12),

$$2k_0a_{\text{cutoff}} = \frac{1}{\sqrt{\varepsilon_c + |\varepsilon_f|}} \operatorname{atanh} \left[\frac{|\varepsilon_f|\sqrt{\varepsilon_c - \varepsilon_s}}{\varepsilon_s\sqrt{\varepsilon_c + |\varepsilon_f|}} \right], \quad \beta_{\text{cutoff}} = k_0\sqrt{\varepsilon_c} \quad (10.13.3)$$

For a small difference, $0 \leq (\varepsilon_c - \varepsilon_s) \ll \varepsilon_c, \varepsilon_s$, we obtain a simpler expression, using the approximation $\operatorname{atanh}(x) \approx x$,

$$2k_0a_{\text{cutoff}} = \frac{|\varepsilon_f|\sqrt{\varepsilon_c - \varepsilon_s}}{\varepsilon_s|\varepsilon_f - \varepsilon_c|} \quad (10.13.4)$$

We will see below that this approximation is also valid in the lossy case with ε_f being replaced by its complex-valued version, $\varepsilon_f = -\varepsilon_R - j\varepsilon_I$.

In the lossy case, because β is complex, $\beta = \beta_R - j\beta_I$, one can no longer set $\beta = k_0\sqrt{\varepsilon_c}$ in the characteristic equation to find the cutoff width. Instead, because the cutoff corresponds to infinite propagation length, one sets $\beta_I = 0$, or, $\beta = \beta_R$, to obtain a system of two equations (the real and imaginary parts of the characteristic equation) in the two unknowns, $a_{\text{cutoff}}, \beta_R$,

$$\begin{aligned} \tanh(2\gamma a_{\text{cutoff}}) &= -\frac{\gamma(p_c\alpha_c + p_s\alpha_s)}{\gamma^2 + p_c\alpha_c p_s\alpha_s}, & \gamma &= \sqrt{\beta_R^2 + k_0^2(\varepsilon_R + j\varepsilon_I)} \\ \alpha_c &= \sqrt{\beta_R^2 - k_0^2\varepsilon_c} & \alpha_c &= \sqrt{\beta_R^2 - k_0^2\varepsilon_c} \\ \alpha_s &= \sqrt{\beta_R^2 - k_0^2\varepsilon_s} & \alpha_s &= \sqrt{\beta_R^2 - k_0^2\varepsilon_s} \end{aligned} \quad (10.13.5)$$

These can be solved numerically. However, a simple estimate can be obtained by simply setting $\beta_R = k_0\sqrt{\varepsilon_c}$ and taking the complex absolute value of the solution,[†]

$$w_{\text{cut}} = 2a_{\text{cutoff}} \approx \frac{1}{k_0} \left| \frac{1}{\sqrt{\varepsilon_c - \varepsilon_f}} \operatorname{atanh} \left(\frac{-\varepsilon_f\sqrt{\varepsilon_c - \varepsilon_s}}{\varepsilon_s\sqrt{\varepsilon_c - \varepsilon_f}} \right) \right| \quad (10.13.6)$$

This provides a better estimate than Eq. (10.13.4), which is valid for thin films. The MATLAB function, **dmdcut**, implements Eq. (10.13.6),

$$\text{wcut} = \text{dmdcut}(1a_0, \text{ef}, \text{ec}, \text{es}); \quad \% \text{ cutoff film width (vectorized in ec) - same units as } \lambda_0$$

Next, we discuss the numerical solution of the characteristic equation (10.1.13) for an asymmetric DMD guide of a given film width $2a$,

$$\tanh(2\gamma a) = -\frac{\gamma(p_c\alpha_c + p_s\alpha_s)}{\gamma^2 + p_c\alpha_c p_s\alpha_s} \quad (10.13.7)$$

Once this is solved, one can calculate the ψ parameter from Eq. (10.1.15), by setting $m = 0$, or $m = 1$, for the LRSP or SRSP modes,

$$\psi = \frac{1}{2} \operatorname{atanh} \left(-\frac{p_c\alpha_c}{\gamma} \right) - \frac{1}{2} \operatorname{atanh} \left(-\frac{p_s\alpha_s}{\gamma} \right) + \frac{1}{2} jm\pi \quad (10.13.8)$$

[†]Alternatively, one can simply set $\varepsilon_f = -\varepsilon_R$ in Eq. (10.13.6), which would lead back to (10.13.3), see for example [940].

and determine the transverse magnetic field,

$$H_y(x) = \begin{cases} H_0 \cosh(\gamma x + \psi), & |x| \leq a \\ H_0 \cosh(\gamma a + \psi) e^{-\alpha_c(x-a)}, & x \geq a \\ H_0 \cosh(\gamma a - \psi) e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (10.13.9)$$

Eq. (10.13.7) can be solved efficiently using the following iteration proposed in [963],

for $n = 0, 1, 2, \dots$, do:

$$\begin{aligned} \hat{A} &= -\gamma_n \coth(2\gamma_n a) \pm \sqrt{B_n^2 + \gamma_n^2 (\coth^2(2\gamma_n a) - 1)} \\ \hat{B} &= \sqrt{\gamma_n^2 + 2\gamma_n \hat{A} \coth(2\gamma_n a) + \hat{A}^2} \\ \alpha_c &= \frac{1}{p_c} (\hat{A} + \hat{B}) \\ \gamma_{n+1} &= \sqrt{\alpha_c^2 + k_0^2 (\varepsilon_c - \varepsilon_f)} \\ \alpha_s &= \sqrt{\gamma_{n+1}^2 + k_0^2 (\varepsilon_f - \varepsilon_s)} \\ B_{n+1} &= \frac{1}{2} (p_c \alpha_c - p_s \alpha_s) \end{aligned} \quad (10.13.10)$$

where \pm correspond to the TM₀ and TM₁ modes, respectively. The iteration may be initialized at $\beta = k_0\sqrt{\varepsilon_c}$. For a symmetric guide, we have $B = 0$ and the iteration reduces to that of Eq. (10.12.4). This follows from the hyperbolic trigonometric identity,

$$-\cosh(2x) \pm \sqrt{\coth^2(2x) - 1} = -\coth(2x) \pm \frac{1}{\sinh(2x)} = \begin{cases} -\tanh(x), & + \text{ sign} \\ -\coth(x), & - \text{ sign} \end{cases}$$

The iteration (10.13.10) can be justified by defining the following quantities A, B ,

$$\begin{aligned} A &= \frac{1}{2} (p_c \alpha_c + p_s \alpha_s) & \Rightarrow & \alpha_c = \frac{1}{p_c} (A + B) \\ B &= \frac{1}{2} (p_c \alpha_c - p_s \alpha_s) & & p_c \alpha_c p_s \alpha_s = A^2 - B^2 \end{aligned}$$

and rewriting the characteristic equation in the form,

$$\tanh(2\gamma a) = -\frac{\gamma(p_c\alpha_c + p_s\alpha_s)}{\gamma^2 + p_c\alpha_c p_s\alpha_s} = -\frac{2\gamma A}{\gamma^2 + A^2 - B^2}$$

and then, solving it for A in terms of B , or, for B in terms of A ,

$$\begin{aligned} A &= -\gamma \coth(2\gamma a) \pm \sqrt{B^2 + \gamma^2 (\coth^2(2\gamma a) - 1)} \\ B &= \sqrt{\gamma^2 + 2\gamma A \coth(2\gamma a) + A^2} \end{aligned} \quad (10.13.11)$$

The following MATLAB function, **dmda**, implements the iteration (10.13.10), with the iteration stopping when two successive values of γ_n become closer to each other than a specified error tolerance, such as, $tol = 10^{-12}$, that is, $|\gamma_{n+1} - \gamma_n| < tol$. The function accepts a vector of widths a , and calculates the corresponding vector of wavenumbers β . It has usage,

```
[be,E,N] = dmda(la0,ef,ec,es,a,mode,tol,be0); % asymmetric DMD guide - iterative solution
```

la0 = operating wavelength, $k_0 = 2\pi/la0$ = free-space wavenumber
 ef,ec,es = permittivities of metal film, dielectric cladding and substrate
 a = vector of half-widths of film in the same units as la0
 mode = 0,1 for TMO or TM1 mode
 tol = computational error tolerance, default tol=1e-12
 be0 = starting search point in units of k_0 - default be0 = sqrt(ec)

be = vector of propagation constants in units of k_0 - size(a)
 E = vector of computational errors of the characteristic equation - size(a)
 N = number of iterations to convergence, until $\text{norm}(ga_new - ga) < tol$

with the computational error defined as,

$$E = \left| \tanh(2\gamma a) + \frac{\gamma(p_c \alpha_c + p_s \alpha_s)}{\gamma^2 + p_c \alpha_c p_s \alpha_s} \right| \quad (10.13.12)$$

As discussed in [611], the existence of the cutoff for the LRSP mode, and the properties of the solutions near it, can be demonstrated by deriving an approximate analytical solution of the characteristic equation in the case of thin films and small permittivity difference ($\epsilon_c - \epsilon_s$). To simplify the algebra, we will work with units of $k_0 = 1$, and restore k_0 at the end. Denoting the normalized width by $w = 2k_0 a$, the characteristic equation reads

$$\tanh(\gamma w) = -\frac{\gamma(p_c \alpha_c + p_s \alpha_s)}{\gamma^2 + p_c \alpha_c p_s \alpha_s}, \quad \begin{aligned} \gamma &= \sqrt{\beta^2 - \epsilon_f} = \sqrt{\beta^2 + \epsilon_R + j\epsilon_I} \\ \alpha_c &= \sqrt{\beta^2 - \epsilon_c} \\ \alpha_s &= \sqrt{\beta^2 - \epsilon_s} \end{aligned} \quad (10.13.13)$$

Because we are interested in the LRSP mode near cutoff, we will study the solution near $\beta \approx \sqrt{\epsilon_c}$, or, equivalently, when α_c is small. Because, $\alpha_s = \sqrt{\alpha_c^2 + \epsilon_c - \epsilon_s}$, if we assume that $(\epsilon_c - \epsilon_s) \ll 1$, then, α_s will also be small. On the other hand, γ is not small because $\gamma = \sqrt{\alpha_c^2 + \epsilon_c - \epsilon_f}$, and $|\epsilon_f|$ is typically much larger than ϵ_c . If we now assume that the normalized width $w = 2k_0 a$ is small, $w \ll 1$, or, more accurately, $|\gamma w| \ll 1$, we may use the approximation, $\tanh(x) \approx x$, to simplify the characteristic equation,

$$\tanh(\gamma w) \approx \gamma w = -\frac{\gamma(p_c \alpha_c + p_s \alpha_s)}{\gamma^2 + p_c \alpha_c p_s \alpha_s} \Rightarrow w = -\frac{p_c \alpha_c + p_s \alpha_s}{\gamma^2 + p_c \alpha_c p_s \alpha_s}$$

Under our assumptions, γ^2 is much larger than the product $\alpha_c \alpha_s$, and we obtain the further simplification,

$$\begin{aligned} w &= -\frac{p_c \alpha_c + p_s \alpha_s}{\gamma^2} \\ p_c \alpha_c + p_s \alpha_s &= -\gamma^2 w \\ \frac{\alpha_s}{\epsilon_s} + \frac{\alpha_c}{\epsilon_c} &= -\frac{\gamma^2 w}{\epsilon_f} \end{aligned}$$

or,

$$\begin{aligned} \frac{\alpha_s}{\epsilon_s} &= -\frac{\alpha_c}{\epsilon_c} - \frac{\gamma^2 w}{\epsilon_f} = -\frac{\alpha_c}{\epsilon_c} - \frac{(\alpha_c^2 + \epsilon_c - \epsilon_f) w}{\epsilon_f} \\ \frac{\alpha_s^2}{\epsilon_s^2} &= \left(\frac{\alpha_c}{\epsilon_c} + \frac{\gamma^2 w}{\epsilon_f} \right)^2 = \frac{\alpha_c^2}{\epsilon_c^2} + \frac{2\alpha_c \gamma^2 w}{\epsilon_c \epsilon_f} + \frac{\gamma^4 w^2}{\epsilon_f^2} \\ \frac{\alpha_c^2 + \epsilon_c - \epsilon_s}{\epsilon_s^2} &= \frac{\alpha_c^2}{\epsilon_c^2} + \frac{2\alpha_c (\alpha_c^2 + \epsilon_c - \epsilon_f) w}{\epsilon_c \epsilon_f} + \frac{(\alpha_c^2 + \epsilon_c - \epsilon_f)^2 w^2}{\epsilon_f^2} \end{aligned}$$

Ignoring all terms that involve α_c^2 , these reduce further to,

$$\begin{aligned} \frac{\alpha_s}{\epsilon_s} &= -\frac{\alpha_c}{\epsilon_c} + \frac{(\epsilon_f - \epsilon_c) w}{\epsilon_f} \\ \frac{\epsilon_c - \epsilon_s}{\epsilon_s^2} &= -\frac{2\alpha_c (\epsilon_f - \epsilon_c) w}{\epsilon_c \epsilon_f} + \frac{(\epsilon_f - \epsilon_c)^2 w^2}{\epsilon_f^2} \end{aligned} \quad (10.13.14)$$

Let us define the quantities,

$$A = \frac{\epsilon_f - \epsilon_c}{\epsilon_f} = \frac{(\epsilon_f - \epsilon_c) \epsilon_f^*}{|\epsilon_f|^2} = \frac{|\epsilon_f|^2 - \epsilon_c \epsilon_f^*}{|\epsilon_f|^2} = \frac{|\epsilon_f|^2 + \epsilon_c \epsilon_R - j\epsilon_c \epsilon_I}{|\epsilon_f|^2} \quad (10.13.15)$$

$$w_c^2 = \frac{\epsilon_c - \epsilon_s}{\epsilon_s^2 |A|^2} = \frac{(\epsilon_c - \epsilon_s) |\epsilon_f|^2}{\epsilon_s^2 |\epsilon_f - \epsilon_c|^2} \Rightarrow w_c = \frac{\sqrt{\epsilon_c - \epsilon_s} |\epsilon_f|}{\epsilon_s |\epsilon_f - \epsilon_c|} \quad (10.13.16)$$

Setting $w_c = 2k_0 a_{\text{cutoff}}$, we see that Eq. (10.13.16) is the same as (10.13.4). Noting that, $(\epsilon_c - \epsilon_s)/\epsilon_s^2 = |A|^2 w_c^2 = AA^* w_c^2$, and canceling a common factor of A from the second of Eqs. (10.13.14), we may rewrite them in the form,

$$\begin{aligned} \frac{\alpha_s}{\epsilon_s} &= -\frac{\alpha_c}{\epsilon_c} + Aw \\ A^* w_c^2 &= -\frac{2\alpha_c w}{\epsilon_c} + Aw^2 \end{aligned} \quad (10.13.17)$$

and solve them for α_c, α_s ,

$$\begin{aligned} \frac{\alpha_c}{\epsilon_c} &= \frac{1}{2w} (Aw^2 - A^* w_c^2) \\ \frac{\alpha_s}{\epsilon_s} &= \frac{1}{2w} (Aw^2 + A^* w_c^2) \end{aligned} \quad (10.13.18)$$

and after separating their real and imaginary parts, $\alpha_c = \alpha_{cR} - j\alpha_{cI}$, $\alpha_s = \alpha_{sR} - j\alpha_{sI}$,

$$\begin{aligned} \frac{\alpha_{cR}}{\epsilon_c} &= \frac{A_R}{2W} (w^2 - w_c^2), & \frac{\alpha_{cI}}{\epsilon_c} &= \frac{A_I}{2W} (w^2 + w_c^2) \\ \frac{\alpha_{sR}}{\epsilon_s} &= \frac{A_R}{2W} (w^2 + w_c^2), & \frac{\alpha_{sI}}{\epsilon_c} &= \frac{A_I}{2W} (w^2 - w_c^2) \end{aligned} \quad (10.13.19)$$

where A_R and A_I are the real and imaginary parts of A , defined from Eq. (10.13.15) by,

$$A = A_R - jA_I, \quad A_R = \frac{|\epsilon_f|^2 + \epsilon_c \epsilon_R}{|\epsilon_f|^2} = 1 + \frac{\epsilon_c \epsilon_R}{|\epsilon_f|^2}, \quad A_I = \frac{\epsilon_c \epsilon_I}{|\epsilon_f|^2} \quad (10.13.20)$$

Using these results, we can now obtain an approximation to the propagation wave number $\beta = \beta_R - j\beta_I$,

$$\beta = \sqrt{\epsilon_c + \alpha_c^2} = \sqrt{\epsilon_c} \sqrt{1 + \frac{\alpha_c^2}{\epsilon_c}} \approx \sqrt{\epsilon_c} \left(1 + \frac{\alpha_c^2}{2\epsilon_c} \right) = \sqrt{\epsilon_c} \left(1 + \frac{\epsilon_c}{2} \frac{\alpha_{cR}^2 - \alpha_{cI}^2 - 2j\alpha_{cR}\alpha_{cI}}{\epsilon_c^2} \right)$$

Using Eqs. (10.13.19), the real and imaginary parts are, after restoring the factor k_0 ,

$$\begin{aligned} \beta_R &= k_0 \sqrt{\epsilon_c} \left[1 + \frac{\epsilon_c}{2} \frac{A_R^2 (w^2 - w_c^2)^2 - A_I^2 (w^2 + w_c^2)^2}{4w^2} \right] \\ \beta_I &= k_0 \sqrt{\epsilon_c} \epsilon_c \frac{A_R A_I (w^4 - w_c^4)}{4w^2} \end{aligned} \quad (10.13.21)$$

where, $w = 2k_0a$, and, $w_c = 2k_0a_{\text{cutoff}}$. For a symmetric guide ($\epsilon_c = \epsilon_s$), we have $w_c = 0$, and Eqs. (10.13.21) reduce to (10.12.9).

Although the approximations (10.13.19) and (10.13.21) are valid for small w and small w_c , they capture the essential properties near cutoff. Above cutoff, both β_I and α_{cR} are positive as they should be for a proper solution (α_{sR} is positive, too). As cutoff is approached from above, β_I tends to zero resulting in infinite propagation length, but at the same time the lateral confinement decreases since α_{cR} also tends to zero. Below cutoff, both α_{cR} and β_I become negative, resulting in unbounded waves. We note also that at cutoff, $w = w_c$, the real part β_R is somewhat less than $k_0\sqrt{\epsilon_c}$, as has been observed in [941].

Example 10.13.1: The nomenclature “long-range mode” originates with Sarid [939]. Here, we reproduce the results of that reference. The DMD guide, operated at $\lambda_0 = 632.8$ nm, consists of a 20-nm silver film of permittivity, $\epsilon_f = (0.0657 - 4j)^2 = -15.9957 - 0.5256j$, surrounded by dielectrics $\epsilon_s = 1.5^2 = 2.25$ and $\epsilon_c = 1.55^2 = 2.4025$.

The cutoff thickness calculated from Eq. (10.13.6) is $w_{\text{cutoff}} = 2a_{\text{cutoff}} = 18.0929$ nm. Therefore, the chosen thickness of 20 nm, lies close to the cutoff, and we expect long propagation lengths for the LRSP mode, and poor lateral confinement. Indeed, the calcu-

lated propagation parameters are given below and agree closely with those of [939],

	LRSP	SRSP
β/k_0	1.550707 - 0.000164j	2.184165 - 0.035423j
γ/k_0	4.290002 + 0.061199j	4.557045 + 0.040691j
α_c/k_0	0.047132 - 0.005394j	1.539269 - 0.050264j
α_s/k_0	0.393310 - 0.000646j	1.587985 - 0.048722j
ψ	-0.3525 - 0.0131j	0.0249 + 1.5716j
error, E	1.0745×10^{-10}	1.2194×10^{-14}
iterations, N	34	5
$L = 1/(2\beta_I)$	307.15 μm	1.42 μm
$L_c = 1/\alpha_{cR}$	2136.82 nm	65.43 nm
$L_s = 1/\alpha_{sR}$	256.07 nm	63.42 nm

We note the large value of L for the LRSP case, but also its large lateral penetration depth L_c (relative to the film width). By contrast, the propagation lengths of the single-interface plasmons at the s - f and c - f interfaces are much shorter:

$$\begin{aligned} \frac{\beta_{c,\infty}}{k_0} &= \sqrt{\frac{\epsilon_c \epsilon_f}{\epsilon_c + \epsilon_f}} = 1.681224 - 0.004876j, & L &= -\frac{1}{\text{Im}(\beta_{c,\infty})} = 10.33 \mu\text{m} \\ \frac{\beta_{s,\infty}}{k_0} &= \sqrt{\frac{\epsilon_s \epsilon_f}{\epsilon_s + \epsilon_f}} = 1.617955 - 0.004346j, & L &= -\frac{1}{\text{Im}(\beta_{s,\infty})} = 11.59 \mu\text{m} \end{aligned}$$

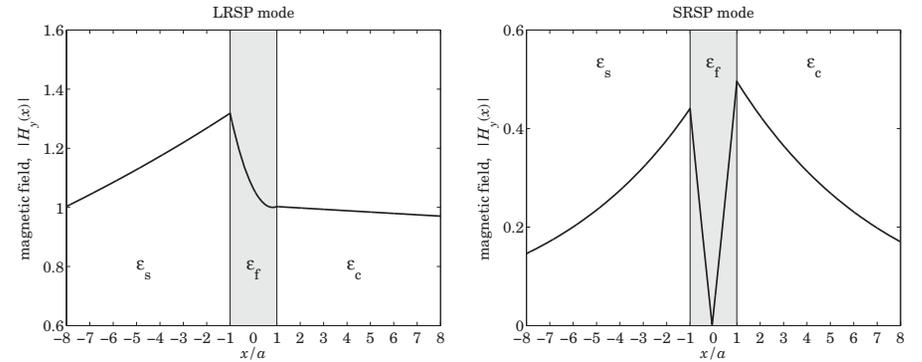


Fig. 10.13.1 Magnetic field profiles, $2a = 20$ nm film.

The corresponding magnetic field profiles, shown in Fig. 10.13.1, demonstrate the asymmetry as well as the poor lateral binding of the LRSP mode. The above were generated by the following MATLAB code.

```

la0 = 632.8; k0 = 2*pi/la0; a = 20/2; % la0,a in units of nm
es = 1.5^2; ec = 1.55^2; ef = (0.0657-4*j)^2; % error tolerance
tol = 1e-12;
[be0,E0,N0] = dmda(la0,ef,ec,es,a,0,tol); % LRSP
    
```

```

ga0 = sqrt(be0^2 - ef); % E0 = computational error
ac0 = sqrt(be0^2 - ec); % N0 = number of iterations
as0 = sqrt(be0^2 - es);

[be1,E1,N1] = dmda(la0,ef,ec,es,a,1,tol); % SRSP
ga1 = sqrt(be1^2 - ef);
ac1 = sqrt(be1^2 - ec);
as1 = sqrt(be1^2 - es);

num2str([be0;ga0;ac0;as0], [be1;ga1;ac1;as1]), '%12.6f') % make table

L0 = -1/2/k0/imag(be0)/1000; % units of microns
L1 = -1/2/k0/imag(be1)/1000;
Lc0 = 1/real(ac0)/k0; Ls0 = 1/real(as0)/k0; % units of nm
Lc1 = 1/real(ac1)/k0; Ls1 = 1/real(as1)/k0

pc = ef/ec; ps = ef/es;

psi0 = atanh(-pc*ac0/ga0)/2 - atanh(-ps*as0/ga0)/2;
psi1 = atanh(-pc*ac1/ga1)/2 - atanh(-ps*as1/ga1)/2 + j*pi/2;

x = linspace(-8,8,1601)*a; % units of half-width a

Hy0 = cosh(ga0*a*k0 - psi0).*exp(as0*(x+a)*k0).*(x<-a) + ...
      cosh(ga0*a*k0 + psi0).*exp(-ac0*(x-a)*k0).*(x>a) + ...
      cosh(ga0*x*k0 + psi0).*(abs(x)<=a);

Hy1 = cosh(ga1*a*k0 - psi1).*exp(as1*(x+a)*k0).*(x<-a) + ...
      cosh(ga1*a*k0 + psi1).*exp(-ac1*(x-a)*k0).*(x>a) + ...
      cosh(ga1*x*k0 + psi1).*(abs(x)<=a);

figure;
fill([-1 1 1 -1],[0.6 0.6, 1.6, 1.6], [0.9 0.9 0.9]); hold on
plot(x/a, abs(Hy0)); % LRSP

figure;
fill([-1 1 1 -1],[0 0, 0.6, 0.6], [0.9 0.9 0.9]); hold on
plot(x/a, abs(Hy1)); % SRSP

```

The effect of asymmetry can also be seen by the distribution of power within the media,

	P	P_f	P_c	P_s
LRSP	360.81	-1	296.26	65.55
SRSP	66.87	-1	37.34	30.53

They were computed with the **pwgpower** MATLAB function and normalized to unity within the metal,

```

[P,Pf,Pc,Ps] = pwgpower(a*k0,ef,ec,es,be0,0); P0 = [P,Pf,Pc,Ps]/-Pf;
[P,Pf,Pc,Ps] = pwgpower(a*k0,ef,ec,es,be1,1); P1 = [P,Pf,Pc,Ps]/-Pf;

```

For the LRSP case, most of the power is pushed into the ϵ_c dielectric and out of the metal, thus, having lower power losses, and longer propagation distances.

The estimated cutoff widths calculated from Eqs. (10.13.4) and (10.13.6) are $w_c = 15.1993$ and $w_c = 18.0929$ nm, respectively. If we use the latter value, $w_c = 18.0929$, we can compare the approximation (10.13.21) to the exact solution,

	exact LRSP	approximate
β/k_0	1.550707 - 0.000164j	1.550799 - 0.000069j
y/k_0	4.290002 + 0.061199j	4.290036 + 0.061233j
α_c/k_0	0.047132 - 0.005394j	0.049831 - 0.002139j
α_s/k_0	0.393310 - 0.000646j	0.393673 - 0.000271j

The approximation is fairly good even though $\epsilon_c - \epsilon_s$ is not that small. The following MATLAB code was used for this calculation,

```

A = (ef-ec)/ef; AR = real(A); AI = -imag(A);

wc_app = sqrt(ec-es) / es / abs(A);
wc = abs(1/sqrt(ec-ef)*atanh(-ef*sqrt(ec-es)/es/sqrt(ec-ef)));
w = 2*a*k0;

wc_app_nm = wc_app/k0 % cutoff in units of nm
wc_nm = wc/k0

beR = sqrt(ec)*(1 + ec/2 * (AR^2*(w^2 - wc^2)^2 - ...
                          AI^2*(w^2 + wc^2)^2)/4/w^2);
beI = sqrt(ec)*ec * AR*AI * (w^4 - wc^4)/4/w^2;

be = beR-j*beI;
ga = sqrt(be^2 - ef); ac = sqrt(be^2 - ec); as = sqrt(be^2 - es);

num2str([be0;ga0;ac0;as0], [be;ga;ac;as]), '%12.6f') % make table

```

Finally, Fig. 10.13.2 shows the LRSP and SRSP propagation wavenumbers as functions of the film width $2a$, starting at cutoff, $2a_{\text{cutoff}} = 18.0929$ nm. For large a , the LRSP β tends to the single-interface $\beta_{s,\infty}$ of the s - f interface, while the SRSP tends to $\beta_{c,\infty}$. Superimposed on the graph of the imaginary part, is the approximation (10.13.21) evaluated from just below cutoff for plotting purposes.

The graphs were generated by the following MATLAB code,

```

w = linspace(1.01*wcutoff,100, 1000); % start just above wcutoff

[be0,E0,N0] = dmda(la0,ef,ec,es, w/2, 0,tol); % LRSP
[be1,E1,N1] = dmda(la0,ef,ec,es, w/2, 1,tol); % SRSP

wa = linspace(0.8*wcutoff,100,1001); % start a bit below wcutoff

beR = sqrt(ec) * (1 + ec/2 * (AR^2*(wa.^2 - wcutoff.^2).^2 - ...
                          AI^2*(wa.^2 + wcutoff.^2).^2)/4./wa.^2 * k0^2);
beI = sqrt(ec) * ec * AR*AI * (wa.^4 - wcutoff.^4)/4./wa.^2 * k0^2;

figure; plot(w,real(be0),'-', w,real(be1),'--');
figure; semilogy(w,-imag(be0),'-', w,-imag(be1),'-.', wa,abs(beI),'r--')

```

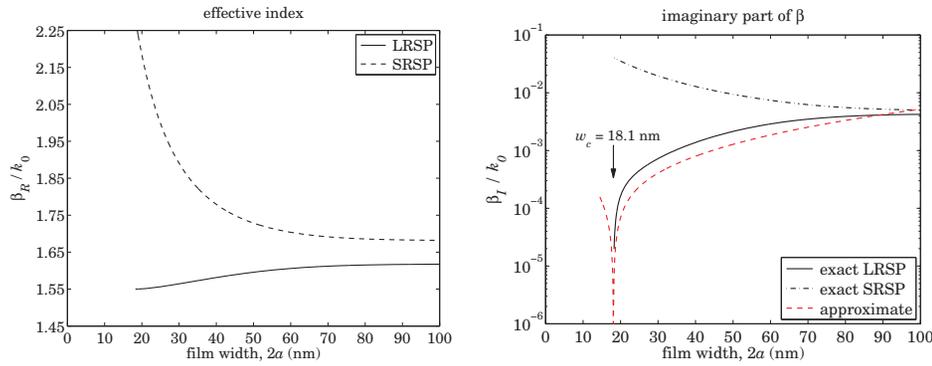


Fig. 10.13.2 Real and imaginary parts of wavenumber β .

The computational errors, as measured by the norms of the error vectors E_0 and E_1 , were $\|E_0\| = 3.1757 \times 10^{-10}$ and $\|E_1\| = 2.0696 \times 10^{-14}$, while the number of iterations to converge (for the entire vector of widths) were $N_0 = 38$ and $N_1 = 10$. \square

Example 10.13.2: To see how the approximation (10.13.21) improves for smaller permittivity difference ($\epsilon_c - \epsilon_s$), we consider the same example, but change the permittivity ϵ_c to the hypothetical value, $\epsilon_c = 1.51^2 = 2.2801$, with all other parameters kept the same. Figs. 10.13.3 and 10.13.4 show the magnetic field profiles of the LRSP and SRSP modes, and the dependence of β on the film width.

The estimated cutoff width calculated from Eqs. (10.13.4) and (10.13.6) is $w_c = 6.7978$ nm and $w_c = 6.9964$ nm, respectively. Using the latter value, we can compare the approximation (10.13.21) to the exact solution,

	exact LRSP	approximate
β/k_0	$1.524661 - 0.000226j$	$1.527059 - 0.000179j$
γ/k_0	$4.280658 + 0.061312j$	$4.281513 + 0.061316j$
α_c/k_0	$0.210935 - 0.001635j$	$0.227622 - 0.001199j$
α_s/k_0	$0.273116 - 0.001263j$	$0.286201 - 0.000954j$

Example 10.13.3: Another early reference on LRSP modes in asymmetric DMD guides is Wendler & Haupt [940], who considered the dependence of the propagation length as a function of the permittivity difference ($\epsilon_c - \epsilon_s$). Here, we reproduce their results. The DMD guide consists of a silver film with permittivity $\epsilon_f = -18 - 0.47j$ at $\lambda_0 = 632.8$ nm.[†] The substrate is fused silica with $\epsilon_s = 2.1211$, and the cover's permittivity will be varied over the range $1.9 \leq \epsilon_c \leq 2.35$.

The function **dmdcut** determines the cutoff width given the values of ϵ_c, ϵ_s . It can also be used to find the value of ϵ_c that would make a given width $w = 2a$ equal to the cutoff width. This can be done by solving for ϵ_c the equation,

$$\text{dmdcut}(\lambda_0, \epsilon_f, \epsilon_c, \epsilon_s) = w \quad (10.13.22)$$

[†]In the previous example, we used, $\epsilon_f = -15.9957 - 0.5256j$, our DRUDE function gives, $\epsilon_f = -14.4824 - 1.0946j$, and the value interpolated from Palik [162] at $\lambda_0 = 632.8$ nm is, $\epsilon_f = -15.8742 - 1.0728j$.

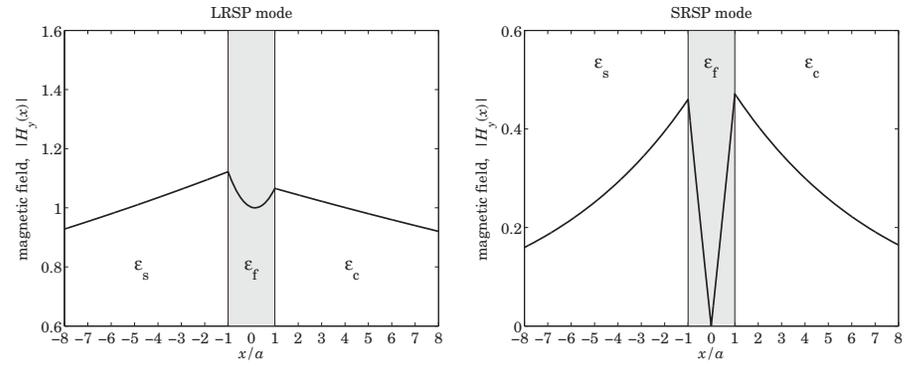


Fig. 10.13.3 Magnetic field profiles, $2a = 20$ nm film.

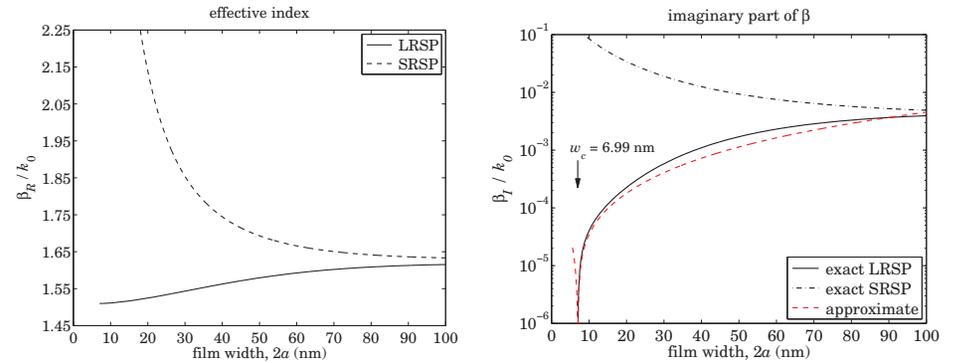


Fig. 10.13.4 Real and imaginary parts of wavenumber β .

It can be implemented easily in MATLAB, for example, using the **fzero** function, with ϵ_s as the initial search point,

```
ec = fzero(@(e) dmdcut(1a0,ef,e,es)-w, es)
```

Fig. 10.13.5 shows the propagation lengths (in cm) for the LRSP mode, for three different film thicknesses, $w = 2a = [10, 17, 30]$ nm, as a function of the variable ϵ_c . For each case, as ϵ_c comes close to its critical value calculated from Eq. (10.13.22) that renders w equal to the cutoff width, the propagation length tends to infinity. To prevent clutter, the graph has been plotted twice, with the one on the right showing the vertical lines at the calculated critical values of ϵ_c ,

$$\epsilon_{c,\text{cut}} = [1.9432, 2.0177, 2.0744, 2.1702, 2.2364, 2.3362]$$

where the left three values correspond to the part of the graph to the left of the dividing line $\epsilon_c = \epsilon_s$. The functions **dmda** and **dmdcut** assume that $\epsilon_c \geq \epsilon_s$. The cases $\epsilon_c < \epsilon_s$

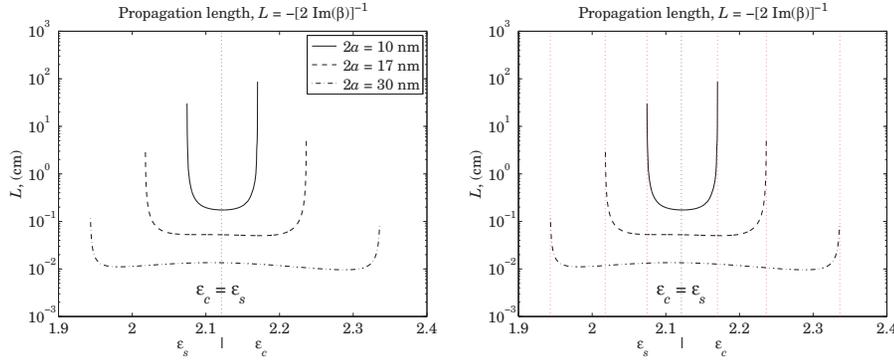


Fig. 10.13.5 Propagation length L vs. permittivity ϵ_c .

can be handled by interchanging the roles of ϵ_c, ϵ_s . The following MATLAB code segment generates the right graph,

```

la0 = 632.8; k0 = 2*pi/la0;           % la0 in nm
ef = -18-0.47j; es = 2.1211;
tol = 1e-12;

ec = linspace(1.9, 2.35, 2000);     % 1.9 <= ec <= 2.35

w = [10, 17, 30];

for i=1:length(ec),
    eci = max(ec(i),es); esi = min(ec(i),es); % interchange es,ec when ec<es
    be = dmada(la0,ef,eci,esi, w/2, 0, tol);
    L(i,:) = -1/2/k0./imag(be)/1e7; % units of cm
end

L(L<0) = NaN; % ignore all negative L's that lie beyond the ec-cutoffs

semilogy(ec, L(:,1), '-', ec, L(:,2), '--', ec,L(:,3),'-');

line([es,es],[10^-3, 10^3],'linestyle',':'); % es=ec dividing line

for i=1:length(w)
    % ec cutoff lines
    eccut(i) = fzero(@(e) dmada(la0,ef,e,es) - w(i), es);
    escut(i) = fzero(@(e) dmada(la0,ef,es,e) - w(i), es);
    line([eccut(i),eccut(i)],[10^-3,10^3],'linestyle',':', 'color','r')
    line([escut(i),escut(i)],[10^-3,10^3],'linestyle',':', 'color','r')
end
    
```

Fig. 10.13.6 shows the propagation lengths as functions of film thickness, with $\epsilon_s = 2.1121$ and five values of ϵ_c ,

$$\epsilon_c = [2.1211, 2.1256, 2.1700, 2.2200, 2.3000]$$

The graphs were generated by the MATLAB code,

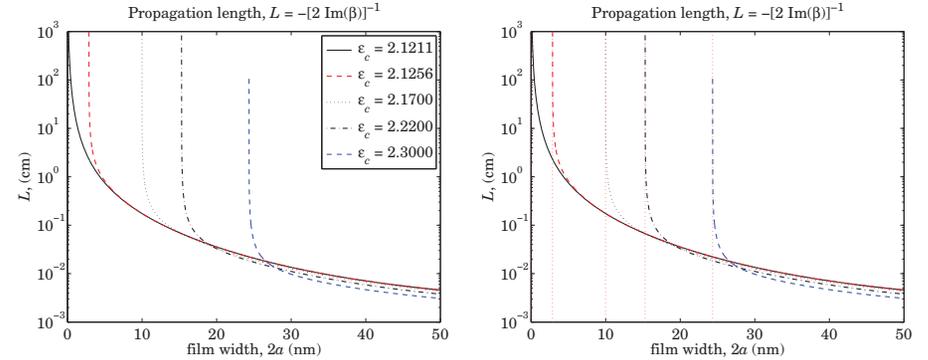


Fig. 10.13.6 Propagation length L vs. film width.

```

es = 2.1211;
ec = [2.1211, 2.1256, 2.1700, 2.2200, 2.3000];
tol = 1e-12;

wcut = dmada(la0,ef,ec,es); % dmada is vectorized in ec

style = {'-', 'r--', ':', '-.', 'b--'};

for i=1:length(ec),
    w = linspace(wcut(i), 50, 2001); % compute for w >= wcut only

    if i==1
        be = dmada(la0,ef,es, w/2, 0, tol); % use dmada when ec=es
    else
        be = dmada(la0,ef,ec(i),es, w/2, 0, tol);
    end

    L = 1/2/k0./abs(imag(be)) * 1e-7; % units of cm

    semilogy(w,L,style{i});

    line([wcut(i),wcut(i)],[10^-3,10^3],'linestyle',':', 'color','r');
end
    
```

The cutoff widths corresponding to the five ϵ_c values, calculated from Eq. (10.13.6), are,

$$w_{cut} = 2a_{cut} = [0, 2.8644, 9.9815, 15.2990, 24.3452] \text{ (nm)}$$

and have been placed on the right graph, which is the same as the left one. As the width comes close to one of the cutoff widths, the propagation length becomes very large. □

10.14 Note on Computations

The characteristic equations of plasmonic waveguides are generally difficult to solve accurately because of the complex-valued nature of the permittivity parameters. We

have been unable to come up with a single algorithm that works robustly for both MDM and DMD waveguides over a wide range of material parameters. Instead, we have presented a variety of solution approaches, which could be viewed as tools to be tried. Our PWG and PWGA functions may work, but one must choose a starting point that is near the true solution. The iterative methods that we discussed work under many circumstances, but not always. We tried to give sufficient code examples to clarify these methods. The plasmonic waveguide literature is somewhat lacking of published numerical examples that could be used as benchmarks, with some notable exceptions [937,939,940,953,955,956,963].

10.15 Sommerfeld Wire

The problem of a TM surface wave propagating along a cylindrical conductor was solved by Sommerfeld in 1899. He showed that the finite conductivity of the conductor was essential in localizing the wave near the surface of the conductor, while exhibiting very low attenuation along its length [982,983,1293]. The planar version of this problem is the Zenneck surface wave that we discussed in Sec. 7.10.

The lateral localization is not particularly good but can be improved by adding a dielectric coating on the surface of the conductor as considered first by Harms and studied later in detail by Goubau [984–999]—a configuration generally referred to as a Goubau line, or a Harms-Goubau line. We discuss this further in Sec. 10.19.

Even though the attenuation along the conductor is very low, these early waveguiding systems, envisioning propagation at long distances of tens of meters or even kilometers, did not catch on for use at RF, primarily because of the poor lateral localization, which could be disturbed by nearby objects, imperfections or bends along the line.

With the rapid development of terahertz applications in the past two decades, interest in the Sommerfeld and Goubau lines has been revived for use at THz frequencies, providing a viable means of waveguiding at relatively short distances (e.g., centimeters), with low attenuation and high degree of confinement. We discuss the interplay of frequency, conductor radius, and dielectric coating in Sec. 10.19.

The THz band typically spans the range of [0.1, 10] THz, or, [0.03, 3] millimeter, and lies between the microwave and infrared/optical bands. There are a large number of THz applications in astrophysics, remote sensing, plasma diagnostics, spectroscopy in chemistry and biology, gas identification, complex molecular dynamics, DNA signature detection, communications, medical imaging, and imaging for homeland security, drug enforcement, pharmaceuticals, biosensing, and non-destructive testing in manufacturing for quality and process control.

For example, Wien's radiation law relating black body temperature and the frequency of the maximum of the Planck spectrum is, $f = 58.7891 T$, where f is in GHz and T in degrees Kelvin. Thus, the Big Bang cosmic microwave background at 2.725 K peaks at $f = 160.2$ GHz, which lies at the low end of the THz band; similarly, interstellar gas clouds at temperatures of 30 K correspond to $f = 1.76$ THz.[†]

Interest in THz imaging for homeland security arises because several everyday materials, such as clothing, paper, plastics, wood, and ceramics are transparent to THz

[†]Some historical references on the CMB, and its data fitting, may be found in [129–137].

waves. Medical imaging at THz is also a very promising area because THz radiation does not have the same harmful effects as X-rays.

Several reviews and applications of THz technology may be found in the following (very incomplete list of) references [1017–1053].

The Sommerfeld wire is an infinitely long cylindrical conductor of radius a and finite conductivity σ , immersed in air, as shown in Fig. 10.15.1. The permittivities of air and conductor are taken to be $\epsilon_c = \epsilon_0$ and

$$j\omega\epsilon_c = j\omega\epsilon_0 + \sigma \Rightarrow \epsilon_c = \epsilon_0 - j\frac{\sigma}{\omega} \quad (10.15.1)$$

with relative permittivities,

$$\epsilon_c = \frac{\epsilon_c}{\epsilon_0} = 1 - j\frac{\sigma}{\omega\epsilon_0}, \quad \epsilon_a = \frac{\epsilon_a}{\epsilon_0} = 1 \quad (10.15.2)$$

The conductivity σ is assumed to be a constant in frequency, but it can also be taken to be frequency-dependent, as given for example by Drude's law (see Sec. 1.12),

$$\sigma = \frac{\sigma_{dc}}{1 + j\omega\tau}, \quad \tau = \text{collisional time}, \quad \sigma_{dc} = \text{conductivity at DC} \quad (10.15.3)$$

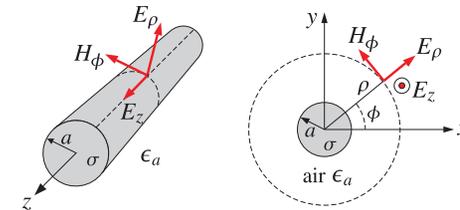


Fig. 10.15.1 Sommerfeld wire, shown in cylindrical coordinates.

Using cylindrical coordinates and assuming, with Sommerfeld, a TM wave with cylindrical symmetry, it follows from Eqs. (9.1.23) and (9.1.24) by setting $\partial_\phi = 0$ and $H_z = 0$ that only the field components E_z, E_ρ, H_ϕ will be non-zero and will be functions only of the radial distance ρ , satisfying the equations:

$$\begin{aligned} \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial E_z}{\partial \rho} \right) + \gamma^2 E_z &= 0, \quad \gamma^2 = \omega^2 \mu \epsilon - \beta^2 \\ E_\rho &= -\frac{j\beta}{\gamma^2} \frac{\partial E_z}{\partial \rho}, \quad H_\phi = \frac{1}{\eta_{TM}} E_\rho = -\frac{j\omega\epsilon}{\gamma^2} \frac{\partial E_z}{\partial \rho} \end{aligned} \quad (10.15.4)$$

where we assumed the usual t, z dependence, $e^{j\omega t - j\beta z}$, and replaced $\eta_{TM} = \beta/\omega\epsilon$. Eqs. (10.15.4) must be solved in each region (inside and outside the conductor) using the appropriate values of ϵ , and hence γ . The E and H tangential fields must be matched at the surface of the conductor (i.e., at $\rho = a$).

The most general solution of the above cylindrical Helmholtz equation for E_z is a linear combination of the 0th order Bessel functions of first and second kinds, $J_0(\gamma\rho)$

and $Y_0(\gamma\rho)$, or of the Hankel functions, $H_0^{(1,2)}(\gamma\rho) = J_0(\gamma\rho) \pm jY_0(\gamma\rho)$.[†] Let us denote any of these functions, or linear combinations thereof, by $C_0(\gamma\rho)$, and note that they satisfy the property $\partial C_0(\gamma\rho)/\partial\rho = -\gamma C_1(\gamma\rho)$, that is, the corresponding 1st order Bessel function. Then, the solutions of Eq. (10.15.4) can be written in general as,

$$E_z = C_0(\gamma\rho), \quad E_\rho = \frac{j\beta}{\gamma} C_1(\gamma\rho), \quad H_\phi = \frac{jk_0\varepsilon}{\eta_0\gamma} C_1(\gamma\rho) \quad (10.15.5)$$

where $\gamma^2 = \omega^2\mu_0\varepsilon - \beta^2 = k_0^2\varepsilon - \beta^2$, and $k_0 = \omega\sqrt{\mu_0\varepsilon_0}$, $\eta_0 = \sqrt{\mu_0/\varepsilon_0}$ are the vacuum wavenumber and impedance, and $\varepsilon = \varepsilon/\varepsilon_0$, and we assumed non-magnetic media. Note the equivalent expression for $k_0\varepsilon/\eta_0 = \omega\sqrt{\mu_0\varepsilon_0}\varepsilon/\sqrt{\mu_0/\varepsilon_0} = \omega\varepsilon_0\varepsilon = \omega\varepsilon$.

Within the conductor ($\rho \leq a$), only the function J_0 is acceptable because $Y_0(u)$ diverges for $u = 0$. Outside the conductor ($\rho \geq a$), because the fields must decay to zero for large radial distances ($\rho \rightarrow \infty$), the Hankel functions are the appropriate choice. In particular, we have the following asymptotic expansions for large $u = u_R + ju_I$,

$$\begin{aligned} H_0^{(1,2)}(u) &\simeq \sqrt{\frac{2}{\pi u}} e^{\pm j(u-\pi/4)} = \sqrt{\frac{2}{\pi(u_R + ju_I)}} e^{\pm j(u_R-\pi/4)} e^{\mp u_I} \\ H_1^{(1,2)}(u) &\simeq \sqrt{\frac{2}{\pi u}} e^{\pm j(u-3\pi/4)} = \sqrt{\frac{2}{\pi(u_R + ju_I)}} e^{\pm j(u_R-3\pi/4)} e^{\mp u_I} \end{aligned} \quad (10.15.6)$$

thus, $H_0^{(1)}(u)$, (resp. $H_0^{(2)}(u)$), is decaying somewhat faster than exponentially in u_I , provided $u_I > 0$, (resp. $u_I < 0$). With these choices, the fields inside and outside the conductor are as follows, with A, B to be determined:

$\begin{aligned} 0 \leq \rho \leq a \\ \hline E_z &= AJ_0(\gamma_c\rho) \\ E_\rho &= A\frac{j\beta}{\gamma_c}J_1(\gamma_c\rho) \\ H_\phi &= A\frac{jk_0\varepsilon_c}{\eta_0\gamma_c}J_1(\gamma_c\rho) \\ \gamma_c &= \sqrt{k_0^2\varepsilon_c - \beta^2} \end{aligned}$	$\begin{aligned} a \leq \rho < \infty \\ \hline E_z &= BH_0^{(1)}(\gamma\rho) \\ E_\rho &= B\frac{j\beta}{\gamma}H_1^{(1)}(\gamma\rho) \\ H_\phi &= B\frac{jk_0\varepsilon_a}{\eta_0\gamma}H_1^{(1)}(\gamma\rho) \\ \gamma &= \sqrt{k_0^2\varepsilon_a - \beta^2} \end{aligned}$
--	--

(10.15.7)

These expressions must be multiplied by the common factor $e^{j\omega t - j\beta z}$. It will turn out that γ has positive imaginary part, thus, justifying the choice of $H_1^{(1)}(u)$ instead of $H_1^{(2)}(u)$. The matching of the tangential fields, E_z and H_ϕ on the surface of the conductor, at $\rho = a$, yields the two conditions:

$$\begin{aligned} AJ_0(\gamma_c a) &= BH_0^{(1)}(\gamma a) \\ A\frac{jk_0\varepsilon_c}{\eta_0\gamma_c}J_1(\gamma_c a) &= B\frac{jk_0\varepsilon_a}{\eta_0\gamma}H_1^{(1)}(\gamma a) \end{aligned} \quad (10.15.8)$$

[†]For definitions and properties of Bessel functions, see [1790] or [1822], which are available online.

We note that the matching of the normal D -field, $D_\rho = \varepsilon E_\rho = \varepsilon\eta_{TM}H_\phi = H_\phi\beta/\omega$, is equivalent to the matching of H_ϕ . Dividing the two sides of Eqs. (10.15.8) results in the following characteristic equation from which the wavenumber β may be determined:

$$\begin{aligned} \gamma &= \sqrt{k_0^2\varepsilon_a - \beta^2} \\ \gamma_c &= \sqrt{k_0^2\varepsilon_c - \beta^2} \end{aligned} \quad \boxed{\frac{\gamma}{\varepsilon_a} \frac{H_0^{(1)}(\gamma a)}{H_1^{(1)}(\gamma a)} = \frac{\gamma_c}{\varepsilon_c} \frac{J_0(\gamma_c a)}{J_1(\gamma_c a)}} \quad (10.15.9)$$

For all good conductors, such as copper, aluminum, and gold, and for frequencies up to and including the THz band, the permittivity ε_c can be approximated accurately by its imaginary part, which is much larger than unity:

$$\varepsilon_c = 1 - j\frac{\sigma}{\omega\varepsilon_0} \approx -j\frac{\sigma}{\omega\varepsilon_0} = e^{-j\pi/2} \frac{\sigma}{\omega\varepsilon_0} \quad (10.15.10)$$

For example, we have for copper, $\sigma = 5.75 \times 10^7$ siemens/m, and with f in units of THz:

$$\varepsilon_c = 1 - j\frac{\sigma}{\omega\varepsilon_0} = 1 - j\frac{1.03 \times 10^6}{f}$$

so for all f up to 10 THz, the imaginary part is at least 5 orders of magnitude greater than the real part, e.g., at $f = 10$ THz, we have $\varepsilon_c = 1 - j1.03 \times 10^5$. The same approximation is also valid if one uses the Drude model for σ of Eq. (10.15.3) up to the THz range. Because the wavenumber β is of the order of k_0 , it follows that the lateral wavenumber γ_c within the conductor will also be very large and can be approximated by

$$\gamma_c = \sqrt{k_0^2\varepsilon_c - \beta^2} \approx k_0\sqrt{\varepsilon_c} = k_0e^{-j\pi/4} \sqrt{\frac{\sigma}{\omega\varepsilon_0}} \quad (10.15.11)$$

for example, for copper at 1 GHz, we find $\gamma_c = (1 - j)4.7645 \times 10^5$ per meter, and at 10 THz, $\gamma_c = (1 - j)4.7645 \times 10^7$. For typical conductor radii of millimeters, it follows that the quantity $\gamma_c a$ will be very large, e.g., for $a = 1$ mm, we have

$$\begin{aligned} \gamma_c a &= (1 - j)4.7645 \times 10^2, \quad \text{at 1 GHz} \\ \gamma_c a &= (1 - j)4.7645 \times 10^4, \quad \text{at 10 THz} \end{aligned}$$

For such large complex arguments, the ratio $J_0(\gamma_c a)/J_1(\gamma_c a)$ may be replaced by $J_0(\gamma_c a)/J_1(\gamma_c a) \approx j$. Indeed, using the asymptotic expansions [1790], valid for large complex argument u ,

$$J_0(u) \approx \sqrt{\frac{2}{\pi u}} \cos\left(u - \frac{\pi}{4}\right), \quad J_1(u) \approx \sqrt{\frac{2}{\pi u}} \cos\left(u - \frac{3\pi}{4}\right)$$

we find after setting $u = u_R + ju_I$ and using some trigonometric identities:

$$\frac{J_0(u)}{J_1(u)} = \frac{\cos(u - \frac{\pi}{4})}{\cos(u - \frac{3\pi}{4})} = -j \frac{\cos(u_R - \frac{\pi}{4}) \cosh(u_I) - j \sin(u_R - \frac{\pi}{4}) \sinh(u_I)}{\cos(u_R - \frac{\pi}{4}) \sinh(u_I) - j \sin(u_R - \frac{\pi}{4}) \cosh(u_I)} \rightarrow -js$$

in the limit $|u_I| \rightarrow \infty$, where $s = \text{sign}(u_I)$. Thus, the ratio is $+j$ for large u with negative imaginary part, like γ_c . A more accurate approximation, which we employ in our

numerical solution of the characteristic equation (10.15.9), follows from the asymptotic expansions in [1790]:

$$\frac{J_0(u)}{J_1(u)} \approx -js \frac{P_0(u) - jsQ_0(u)}{P_1(u) - jsQ_1(u)} \quad (10.15.12)$$

where again, $s = \text{sign}(\text{Im}(u))$, and we defined the polynomials in $1/u$:

$$\begin{aligned} P_0(u) &= 1 - \frac{9}{2(8u)^2} + \frac{3675}{8(8u)^4}, & Q_0(u) &= -\frac{1}{8u} + \frac{75}{2(8u)^3} \\ P_1(u) &= 1 + \frac{15}{2(8u)^2} - \frac{4725}{8(8u)^4}, & Q_1(u) &= \frac{3}{8u} - \frac{105}{2(8u)^3} \end{aligned} \quad (10.15.13)$$

The approximation (10.15.12) can be used whenever $|\text{Im}(u)| > 700$. For $|\text{Im}(u)| \leq 700$, the built-in MATLAB function `besselj` gives accurate results. The MATLAB function, `J01`, in the EWA-toolbox, implements the evaluation of the ratio $J_0(u)/J_1(u)$ for any value of u using this improved approximation. It has usage:

```
y = J01(u); % evaluates the ratio J0(u)/J1(u), vectorized in u
```

To solve the characteristic equation (10.15.9), Sommerfeld made two approximations: first, he replaced the Bessel function ratio by $J_0(\gamma_c a)/J_1(\gamma_c a) \approx j$, and γ_c by $\gamma_c = k_0 \sqrt{\epsilon_c}$, and second, he assumed that the quantity γa is small and used the following approximations [1790] for the Hankel functions, valid for small argument $u \rightarrow 0$:

$$\begin{aligned} J_0(u) &\approx 1, & Y_0(u) &\approx \frac{2}{\pi} \ln\left(\frac{e^C u}{2}\right), & H_0^{(1)}(u) &= J_0(u) + jY_0(u) \approx \frac{2j}{\pi} \ln\left(\frac{e^C u}{2j}\right) \\ J_1(u) &\approx \frac{u}{2}, & Y_1(u) &\approx -\frac{2}{\pi u}, & H_1^{(1)}(u) &= J_1(u) + jY_1(u) \approx -\frac{2j}{\pi u} \end{aligned}$$

and for the ratio:

$$\frac{H_0^{(1)}(u)}{H_1^{(1)}(u)} \approx -u \ln\left(\frac{e^C u}{2j}\right) \quad (10.15.14)$$

where C is the Euler-Mascheroni constant, $C = 0.577215\dots$, so that $e^C = 1.78107$. With these simplifications, Eq. (10.15.9) reduces to:

$$-y^2 a \ln\left(\frac{e^C \gamma a}{2j}\right) = \frac{j\epsilon_a k_0}{\sqrt{\epsilon_c}} \Rightarrow y^2 = -\frac{j\epsilon_a k_0}{a \sqrt{\epsilon_c} \ln\left(\frac{e^C \gamma a}{2j}\right)} \quad (10.15.15)$$

which Sommerfeld then proceeded to solve iteratively, for $n = 0, 1, 2, \dots$,

$$y_{n+1}^2 = -\frac{j\epsilon_a k_0}{a \sqrt{\epsilon_c} \ln\left(\frac{e^C \gamma_n a}{2j}\right)} \quad \text{Sommerfeld iteration} \quad (10.15.16)$$

initialized at $y_0^2 = -\frac{j\epsilon_a k_0}{a \sqrt{\epsilon_c}}$

The approximation and iteration works well for frequencies in the GHz range. But, as discussed by King and Wiltse [994], in the THz range the quantity γa is no longer small enough to justify the approximation. Our approach to solving Eq. (10.15.9) is to rewrite it in the following form:

$$y = \frac{H_1^{(1)}(\gamma a)}{H_0^{(1)}(\gamma a)} \frac{\epsilon_a \gamma_c}{\epsilon_c} \frac{J_0(\gamma_c a)}{J_1(\gamma_c a)} \quad (10.15.17)$$

Noting that $\gamma_c^2 - \gamma^2 = k_0^2(\epsilon_c - \epsilon_a)$, or, $\gamma_c = \sqrt{k_0^2(\epsilon_c - \epsilon_a) + \gamma^2}$, Eq. (10.15.17) can then be turned into the following iteration, for $n = 0, 1, 2, \dots$,

$$\begin{aligned} \gamma_c &= \sqrt{k_0^2(\epsilon_c - \epsilon_a) + \gamma_n^2} \\ \gamma_{n+1} &= \frac{H_1^{(1)}(\gamma_n a)}{H_0^{(1)}(\gamma_n a)} \frac{\epsilon_a \gamma_c}{\epsilon_c} \frac{J_0(\gamma_c a)}{J_1(\gamma_c a)} \end{aligned} \quad (10.15.18)$$

It can be initialized at $\beta_0 = 0.9k_0$, $\gamma_0 = \sqrt{k_0^2 \epsilon_a - \beta_0^2}$, or something similar. The iteration does not require any approximations beyond the use of our function `J01` for the ratio of the Bessel functions. It uses the built-in MATLAB function `besselh` for the evaluation of the Hankel functions. The following MATLAB function, `sommer`, in the EWA toolbox, implements the iteration (10.15.18):

```
[be,ga,gc,E,N] = sommer(a,f,sigma,tol,be0); % Sommerfeld wire
[be,ga,gc,E,N] = sommer(a,f,sigma,tol); % (be0 = 0.9*k0)
[be,ga,gc,E,N] = sommer(a,f,sigma); % (tol = 1e-10)
```

```
% a = wire radius in meters
% f = vector of frequencies in Hz
% sigma = wire conductivity in siemens/m, scalar or same size as f
% tol = computational tolerance, default, tol = 1e-10
% be0 = initializing vector, size(f), default, be0 = 0.9*k0
%
% be = vector of propagation wavenumbers, (rads/m), same size as f
% ga = lateral wavenumber in air, (1/m), size(f)
% gc = lateral wavenumber in conductor, (1/m), size(f)
% E = computational error of characteristic equation, size(f)
% N = number of iterations to converge, scalar, represents all f's
```

To clarify the construction of the function, we list below the essential part of the code:

```
% =====
c0 = 299792458; % speed of light, m/sec
ep0 = 8.854187817e-12; % vacuum permittivity, farad/m
tol=1e-10; % default error tolerance

w = 2*pi*f; k0 = w/c0; % f can be a vector
ec = 1 - j * sigma./w/ep0; % sigma must be scalar or size(f)
```

```

be0 = 0.9*k0;           % initialize iteration
ga = sqrt(k0.^2 - be0.^2);
N=1;

while 1                % loop forever
    gc = sqrt(ga.^2 + k0.^2.*(ec-1));
    gnew = besselh(1,1,ga*a)./besselh(0,1,ga*a).*J01(gc*a).*gc./ec;
    if norm(ga-gnew)<tol, break; end
    N = N+1;
    ga = gnew;
end

ga = gnew;             % converged values
gc = sqrt(ga.^2 + k0.^2.*(ec-1));
be = sqrt(k0.^2 - ga.^2);
E = abs(ga.*besselh(0,1,ga*a)./besselh(1,1,ga*a) - gc.*J01(gc*a)./ec);

% =====

```

We note that the computational error is defined from the last computed y as follows. Setting $y_c = \sqrt{k_0^2(\epsilon_c - \epsilon_a) + y^2}$,

$$E = \left| \frac{y}{\epsilon_a} \frac{H_0^{(1)}(ya)}{H_1^{(1)}(ya)} - \frac{y_c}{\epsilon_c} \frac{J_0(y_c a)}{J_1(y_c a)} \right| \quad (10.15.19)$$

where E is a vector when f is a vector of frequencies.

Example 10.15.1: We consider the two examples discussed by Sommerfeld in Ref. [983]. The first one is a copper wire with conductivity $\sigma = 5.75 \times 10^7$ siemens/m, and radius of 1 mm, at a frequency of 1 GHz. We iterated (10.15.16) five times, with the MATLAB code:

```

ep0 = 8.854187817e-12; % farad/m,
c0 = 299792458;        % m/sec
f = 1e9;               % 1 GHz
w = 2*pi*f;           % rads/sec
k0 = w/c0;             % k0 = 20.958450, rads/m
sigma = 5.75e7;        % siemens/m
ec = 1 + sigma./j./w/ep0;
a = 1e-3;              % 1 mm
C = 0.577215664901533; % Euler constant

g0 = sqrt(-j*k0/a./sqrt(ec)); % initialize
ga = g0; G=[ga];          % save iterates in G
K=5;

for i=1:K
    ga = g0./sqrt(log(exp(C)*ga*a/2/j));
    G = [G;ga];
end

num2str(G,' %1.6f')      % print table of iterates

```

generating the iterates:

```

y0 = 0.745951 - 0.308983j
y1 = 0.147818 + 0.255109j
y2 = 0.115696 + 0.255973j
y3 = 0.113805 + 0.256030j
y4 = 0.113692 + 0.256039j
y5 = 0.113685 + 0.256040j

```

It is evident that after the third iteration, y has effectively converged. Following Sommerfeld, we use the third iterate, $y = y_3 = 0.113805 + 0.256030j$, to calculate β :

$$\beta = \sqrt{k_0^2 - y^2} = 20.959705 - 0.001390j \Rightarrow \frac{\beta - k_0}{k_0} = (5.9874 - 6.6330j) \cdot 10^{-5}$$

which agrees with Sommerfeld's rounded result of $(\beta - k_0)/k_0 = (6.0 - 6.6j) \cdot 10^{-5}$. We see also that the phase velocity is essentially that of c_0 , $v_{ph} = \omega / \text{Re}(\beta) = 0.999940c_0$.

The attenuation constant $\beta_I = -\text{Im}(\beta)$ can be expressed in dB/m by $8.686\beta_I = 0.0121$ dB/m, which is very low and corresponds to an effective propagation length of $1/\beta_I = 719.34$ meters along the wire. On the other hand, the lateral attenuation in air is effectively measured by $1/\text{Im}(y) = 3.91$ meters, which is unacceptably large. The calculation using our function `sommer` is implemented by the MATLAB command:

```
[be,ga,gc,E,N] = sommer(a,f,sigma);
```

and produces comparable results:

$$\begin{aligned} \beta &= \beta_R - j\beta_I = 20.959706 - 0.001390j \text{ rads/m} \Rightarrow \frac{\beta - k_0}{k_0} = (5.9907 - 6.6333j) \cdot 10^{-5} \\ y &= 0.113788 + 0.256080j \\ y_c &= (4.7645 - 4.7645j) \cdot 10^5 \Rightarrow y_c - k_0\sqrt{\epsilon_c} = -(2.3054 + 2.3048j) \cdot 10^{-4} \\ E &= 1.94 \times 10^{-13} \\ N &= 186 \end{aligned}$$

Sommerfeld's second example is a platinum wire of radius of $a = 2 \mu\text{m}$ and conductivity one eighth that of copper, $\sigma = (5.75/8) \cdot 10^7$ siemens/m, at a vacuum wavelength of $\lambda_0 = 1$ meter, or frequency $f = c_0/\lambda_0$. With these changes in parameters, the results from the function `sommer` are:

$$\begin{aligned} \beta &= \beta_R - j\beta_I = 8.4603 - 6.2561j \text{ rads/m} \\ y &= 7.5210 + 7.0374j \Rightarrow \text{Im}(y)^{-1} = 0.1421 \text{ meters} \\ y_c &= (1 - j) \cdot 9.2231 \cdot 10^4 \Rightarrow y_c a = (1 - j) \cdot 0.1845 \\ E &= 1.95 \times 10^{-14}, \quad N = 278 \\ 8.686\beta_I &= 54.34 \text{ dB/m}, \quad \beta_I^{-1} = 0.1598 \text{ meters} \\ v_{ph} &= \frac{\omega}{\beta_R} = 0.7427c_0 \end{aligned}$$

This solution agrees with Sommerfeld's, $\beta = 8.5 - 6.5j$, which he obtained using a modified form of his iteration that does not assume large values of $y_c a$ and does not replace the ratio of Bessel functions by $J_0(y_c a)/J_1(y_c a) \approx j$, but still using the small-argument expansions of the Hankel functions. \square

Example 10.15.2: Here, we repeat the calculations of the previous example (for the copper wire) using the same MATLAB code with $f = 10$ THz. The results from the function `sommer` are:

$$\beta = \beta_R - j\beta_I = 2.0959 \cdot 10^5 - 1.9187j \text{ rads/m} \Rightarrow \frac{\beta - k_0}{k_0} = (3.2780 - 9.1549j) \cdot 10^{-6}$$

$$\gamma = (5.3212 + 7.5573j) \cdot 10^2 \Rightarrow \text{Im}(\gamma)^{-1} = 0.0013 \text{ meters}$$

$$\gamma_c = (4.7645 - 4.7645j) \cdot 10^7 \Rightarrow \gamma_c - k_0\sqrt{\epsilon_c} = -(2.3049 + 2.3048j) \cdot 10^2$$

$$E = 1.73 \times 10^{-11}, \quad N = 26$$

$$8.686\beta_I = 16.67 \text{ dB/m}, \quad \beta_I^{-1} = 0.5212 \text{ meters}$$

where $k_0 = 2.0958 \times 10^5$ rads/m. The Sommerfeld iteration does converge, but to the wrong limit, $(\beta - k_0)/k_0 = (-2.9511 - 6.2879j) \cdot 10^{-6}$. We note that the lateral decay length in air is now only, $\text{Im}(\gamma)^{-1} = 1.3$ millimeters, and although the attenuation of 16.67 dB/m, or, 0.1667 dB/cm, is much larger than in the previous case, it is still acceptable for propagation lengths of the order of cm.

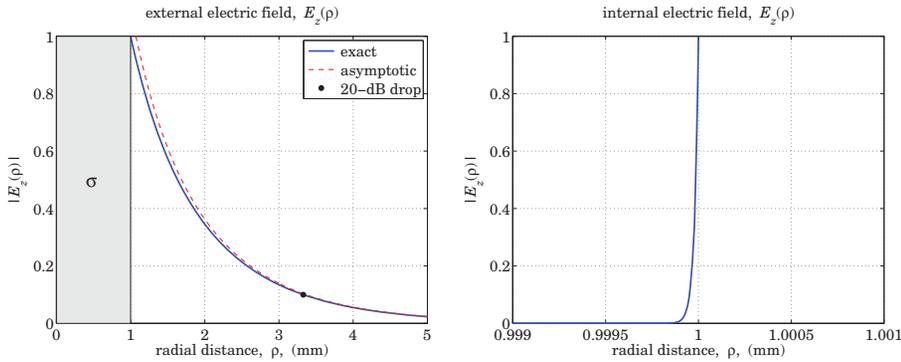


Fig. 10.15.2 Longitudinal electric field profiles, outside and inside the conductor.

Fig. 10.15.2 shows the profile of the electric field $|E_z(\rho)|$ outside and inside the conductor, normalized to unity at the surface. This normalization fixes the values of the coefficients A, B in Eq. (10.15.8) so that $AJ_0(\gamma_c a) = BH_0^{(1)}(\gamma a) = 1$. Thus, the field $E_z(\rho)$ is,

$$E_z(\rho) = \begin{cases} \frac{J_0(\gamma_c \rho)}{J_0(\gamma_c a)}, & \rho \leq a \\ \frac{H_0^{(1)}(\gamma \rho)}{H_0^{(1)}(\gamma a)}, & \rho \geq a \end{cases} \quad (10.15.20)$$

The field outside the conductor is easily computed and plotted, in the left graph, with the help of the MATLAB function:

```
Ez = @(r) abs(besselh(0,1,ga*r) / besselh(0,1,ga*a));
```

The graph also superimposes the asymptotic form of the Hankel function from Eq. (10.15.6), that is, for $\rho \gg a$, $E_z(\rho)$ can be approximated by,

$$E_z(\rho) \approx \frac{1}{H_0^{(1)}(\gamma a)} \sqrt{\frac{2}{\pi \gamma \rho}} e^{j(\gamma \rho - \pi/4)}$$

Moreover, the 20-dB point is also included on the graph. It can be determined by using the built-in function `fzero` to solve the equation $|E_z(\rho)|^2 = 10^{-2}$, which gives $\rho = 3.3$ mm. The MATLAB code is as follows and uses the previously constructed function `Ez(r)`:

```
r_dB = fzero(@(r) Ez(r)-1/10, a);
```

with initial search point near $\rho = a$. The calculation of the field inside the conductor is, on the other hand, very difficult numerically, because the quantity $\gamma_c a = (4.7645 - 4.7645j) \cdot 10^4$ is very large, and MATLAB returns the value $J_0(\gamma_c a) = \infty$. Within a narrow layer near the surface, $\rho \lesssim a$, we can use the large-argument asymptotic expansion of $J_0(u)$ from [1790] to approximate the ratio $J_0(\gamma_c \rho)/J_0(\gamma_c a)$ in terms of the polynomials defined in Eq. (10.15.13):

$$\frac{J_0(\gamma_c \rho)}{J_0(\gamma_c a)} \approx \frac{P_0(\gamma_c \rho) + jQ_0(\gamma_c \rho)}{P_0(\gamma_c a) + jQ_0(\gamma_c a)} \sqrt{\frac{a}{\rho}} e^{j\gamma_c(\rho-a)}, \quad \text{for } \rho \lesssim a \quad (10.15.21)$$

Within this narrow layer, the ratio of the polynomials is effectively equal to unity, as is the ratio $\sqrt{a/\rho}$, thus the entire expression is given by the exponential $e^{j\gamma_c(\rho-a)}$. We note the rapid exponential decay, $|e^{j\gamma_c(\rho-a)}| = e^{-\alpha_c(a-\rho)}$, where $\alpha_c = -\text{Im}(\gamma_c)$ is very large and positive. The graph on the right was obtained by applying Eq. (10.15.21) over the narrow range, $0.999a \leq \rho \leq a$.

The rapid attenuation is essentially a manifestation of the skin effect, indeed, α_c is related to the skin depth of a good conductor $\alpha_c = 1/\delta = \sqrt{\sigma\omega\mu_0/2}$, as discussed in Sec. 2.8. This follows by approximating γ_c by

$$\gamma_c \approx k_0\sqrt{\epsilon_c} = \omega\sqrt{\mu_0\epsilon_0} \sqrt{1 - j\frac{\sigma}{\omega\epsilon_0}} \approx \omega\sqrt{\mu_0\epsilon_0} \sqrt{-j\frac{\sigma}{\omega\epsilon_0}} = (1-j)\sqrt{\frac{\sigma\omega\mu_0}{2}} = \frac{1-j}{\delta}$$

To better visualize the fields inside and outside the conductor on the same plotting scale, we consider a different example that has, $f = 1$ THz and $a = 1 \mu\text{m}$, and the same $\sigma = 5.75 \times 10^7$. In this case, the solution obtained from the function `sommer` has parameters:

$$\begin{aligned} \gamma a &= 0.0008 + 0.0016j \\ \gamma_c a &= 15.0665 - 15.0665j \end{aligned}$$

Fig. 10.15.3 shows the field profile $|E_z(\rho)|$ evaluated using Eq. (10.15.20), plotted on a log-log scale and on an absolute scale. Since $J_0(0) = 1$, the value of the field at the center of the conductor ($\rho = 0$) is small but non-zero, $|E_z(0)| = 1/|J_0(\gamma_c a)| = 3.2979 \times 10^{-6}$. □

Example 10.15.3: Here, we compare three cases of a copper wire with conductivity, $\sigma = 5.75 \times 10^7$ siemens/m, and radii, $a = 0.01$ mm, $a = 0.1$ mm, and $a = 1$ mm, and study the dependence of the Sommerfeld wave on frequency over the range $10 \text{ GHz} \leq f \leq 10 \text{ THz}$. The solutions are obtained in the three cases by the MATLAB code:

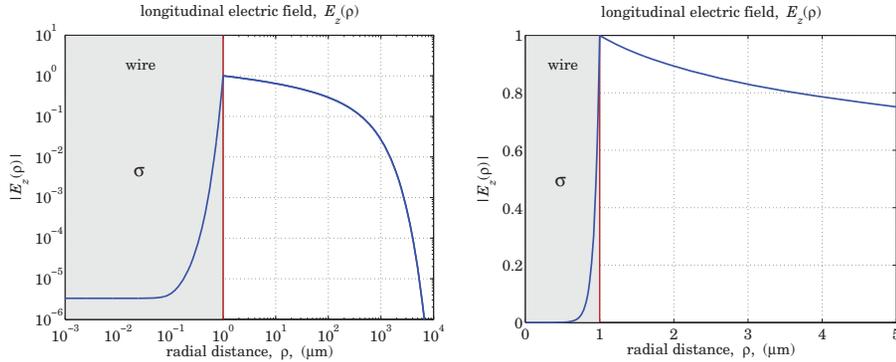


Fig. 10.15.3 Longitudinal electric field profiles.

```

c0 = 299792458; % m/sec
sigma = 5.75e7; % siemens/m

ff = logspace(-2,1,300); f = 1e12*ff; w = 2*pi*f; % THz

a1 = 1e-5; [be1,ga1,gc1,E1,N1] = sommer(a1,f,sigma); % case 1
a2 = 1e-4; [be2,ga2,gc2,E2,N2] = sommer(a2,f,sigma); % case 2
a3 = 1e-3; [be3,ga3,gc3,E3,N3] = sommer(a3,f,sigma); % case 3
    
```

The number of iterations to converge and the norms of the computational errors were:

$$N_1 = 239, \quad N_2 = 198, \quad N_3 = 159$$

$$\|E_1\| = 1.1130 \times 10^{-12}, \quad \|E_2\| = 0.9234 \times 10^{-12}, \quad \|E_3\| = 1.2927 \times 10^{-12}$$

Fig. 10.15.4 shows on the left the attenuation in dB per cm, given by, $\text{dB} = 0.08686 \cdot \beta_I$, defined in terms of the negative imaginary part of $\beta = \beta_R - j\beta_I$. The right graph shows the effective lateral radius of the field outside the conductor, defined as the distance where the field E_z has dropped to 1/10 its value at the surface, or, 20 dB down. It is obtained as the solution in ρ of the equation,

$$|E_z(\rho)| = \left| \frac{H_0^{(1)}(\gamma\rho)}{H_0^{(1)}(\gamma a)} \right| = \frac{1}{10} \quad (10.15.22)$$

which is solved using the built-in function `fzero`. The attenuations and 20-dB distances were computed and plotted by the following MATLAB code:

```

db = 20*log10(exp(1)); % 8.6859 dB/neper

dB1 = -db * imag(be1) / 100; % attenuations in dB/cm
dB2 = -db * imag(be2) / 100;
dB3 = -db * imag(be3) / 100;
    
```

```

figure; loglog(ff,dB1,'b-', ff,dB2,'r--', ff,dB3,'k:')
    
```

```

for i=1:length(f)
Ez1 = @(r) abs(besselh(0,1,ga1(i)*r) / besselh(0,1,ga1(i)*a1));
Ez2 = @(r) abs(besselh(0,1,ga2(i)*r) / besselh(0,1,ga2(i)*a2));
Ez3 = @(r) abs(besselh(0,1,ga3(i)*r) / besselh(0,1,ga3(i)*a3));

r1(i) = fzero(@(r) Ez1(r)-1/10, a1)*1000; % lateral radius in mm
r2(i) = fzero(@(r) Ez2(r)-1/10, a2)*1000;
r3(i) = fzero(@(r) Ez3(r)-1/10, a3)*1000;
end

figure; loglog(ff,r1,'b-', ff,r2,'r--', ff,r3,'k:')
    
```

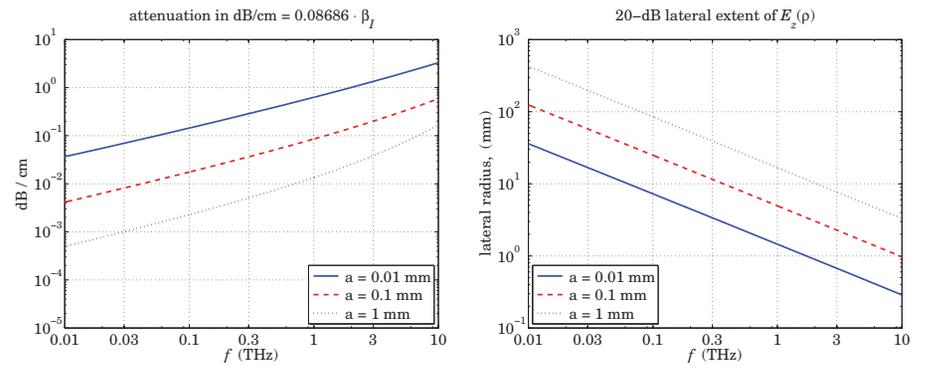


Fig. 10.15.4 Attenuations and 20-dB lateral distances vs. frequency.

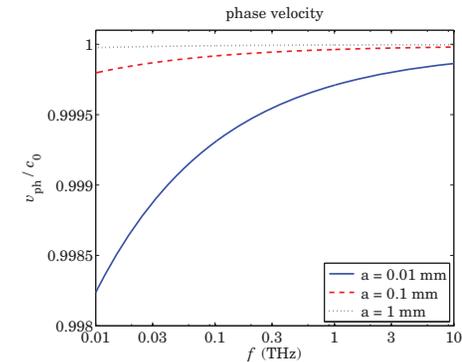


Fig. 10.15.5 Phase velocities vs. frequency.

Within the 1-10 THz range, the attenuations remain fairly small for propagating at distances of cm, while the lateral footprint of the wave is of the order of 1 cm or smaller. Finally, Fig. 10.15.5 plots the phase velocities $v_{ph} = \omega/\beta_R$ versus frequency, showing some dispersion, but not much, in the THz range. □

10.16 Power Transfer and Power Loss

The power transmitted along the wire in the positive z-direction is attenuating with the factor $|e^{-j\beta z}|^2 = e^{-2\beta_I z}$. The part that flows along the wire on the air side is the useful amount of power that gets transmitted. The part that flows inside the conductor, together with the amount of power that flows into the wire across its cylindrical surface, is entirely dissipated into Ohmic losses, as we verify below. The time-averaged Poynting vector has two components, one along z, and one along the radial direction inwards:

$$\mathcal{P} = \frac{1}{2} \text{Re}(\mathbf{E} \times \mathbf{H}^*) = \frac{1}{2} \text{Re}((\hat{z}E_z + \hat{\rho}E_\rho) \times (\hat{\phi}H_\phi^*)) = \hat{z} \frac{1}{2} \text{Re}(E_\rho H_\phi^*) - \hat{\rho} \frac{1}{2} \text{Re}(E_z H_\phi^*)$$

The total power transmitted is obtained by integrating \mathcal{P}_z over the cross-sectional area outside the wire, i.e, over $a \leq \rho < \infty$:

$$P_T = \int_a^\infty \mathcal{P}_z 2\pi\rho d\rho = \int_a^\infty \frac{1}{2} \text{Re}(E_\rho H_\phi^*) 2\pi\rho d\rho$$

Similarly, the amount of power flowing through an annular area, $a \leq \rho \leq r$, is given by,

$$P_T(r) = \int_a^r \frac{1}{2} \text{Re}(E_\rho H_\phi^*) 2\pi\rho d\rho = \frac{1}{2} |B|^2 \frac{\beta_R k_0 \epsilon_a}{\eta_0 |\gamma|^2} \int_a^r |H_1^{(1)}(\gamma\rho)|^2 2\pi\rho d\rho$$

where we used Eq. (10.15.7), and we must multiply this expression by the attenuation factor $e^{-2\beta_I z}$. Thus, the total transmitted power is $P_T = P_T(r) |_{r=\infty}$. Using the indefinite integral [1790],

$$\int |H_1^{(1)}(\gamma\rho) H_1^{(1)}(\gamma^*\rho)| 2\pi\rho d\rho = -2\pi\rho \frac{\text{Im}[\gamma H_0^{(1)}(\gamma\rho) H_1^{(1)}(\gamma^*\rho)]}{\text{Im}[\gamma^2]} \quad (10.16.1)$$

we obtain,

$$P_T(r) = \frac{1}{2} |B|^2 \frac{\beta_R k_0 \epsilon_a 2\pi}{\eta_0 |\gamma|^2} \frac{a \text{Im}[\gamma H_0^{(1)}(\gamma a) H_1^{(1)}(\gamma^* a)] - r \text{Im}[\gamma H_0^{(1)}(\gamma r) H_1^{(1)}(\gamma^* r)]}{\text{Im}[\gamma^2]}$$

Since the Hankel functions attenuate like $\rho^{-1/2} e^{-\rho \text{Im}(\gamma)}$ for large ρ , it follows that the total transmitted power will be given by,

$$P_T = \frac{1}{2} |B|^2 \frac{\beta_R k_0 \epsilon_a}{\eta_0 |\gamma|^2} \left[2\pi a \frac{\text{Im}[\gamma H_0^{(1)}(\gamma a) H_1^{(1)}(\gamma^* a)]}{\text{Im}[\gamma^2]} \right] \quad (10.16.2)$$

Dividing $P_T(r)$ by P_T , we obtain the fraction of the total power contained in the annular region, $a \leq \rho \leq r$, see also [985,994]:

$$\frac{P_T(r)}{P_T} = 1 - \frac{\text{Im}[\gamma r H_0^{(1)}(\gamma r) H_1^{(1)}(\gamma^* r)]}{\text{Im}[\gamma a H_0^{(1)}(\gamma a) H_1^{(1)}(\gamma^* a)]} \quad (10.16.3)$$

Example 10.16.1: For the copper wire of Example 10.15.2, Fig. 10.16.1 shows a plot of $P_T(r)/P_T$, which also includes the radius at which the power ratio is 95%, found to be, $r = 2.85$ mm, by solving the equation $P_T(r)/P_T = 0.95$ with respect to r . Assuming the same parameter values as in Example 10.15.2, the following MATLAB code generates the graph,

```
P = @(r) 1 - imag(ga*r.*besselh(0,1,ga*r).*conj(besselh(1,1,ga*r))) / ...
    imag(ga*a*besselh(0,1,ga*a).*conj(besselh(1,1,ga*a)));

r0 = fzero(@(r) P(r)-0.95, 2*a) % search near 2*a

plot(r/a,100*P(r),'b-', r0/a,100*P(r0),'r.');
```

The 95% power radius and the 20-dB radius discussed earlier are just alternative ways of defining a measure for the lateral extent of the wave. □

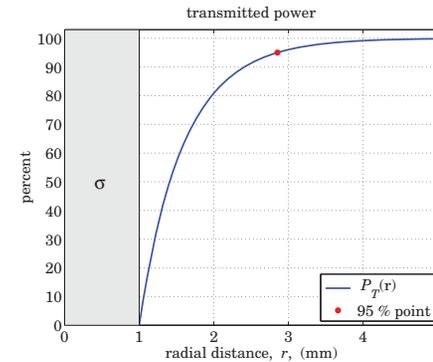


Fig. 10.16.1 Transmitted power vs. radial distance.

Next, we discuss power losses. In a lossy medium with complex-permittivity $\epsilon_c = \epsilon_R - j\epsilon_I$, the energy flux into a volume within the medium is dissipated completely into Ohmic losses. This is justified by the following result, first considered in Problem 1.5, that follows from Maxwell's equations:

$$P_{\text{in}} \equiv - \oint_S \frac{1}{2} \text{Re}[\mathbf{E} \times \mathbf{H}^*] \cdot d\mathbf{S} = \int_V \frac{1}{2} \text{Re}[\mathbf{E} \cdot \mathbf{J}_{\text{tot}}^*] dV \equiv P_{\text{loss}} \quad (10.16.4)$$

where $\mathbf{J}_{\text{tot}} = \mathbf{J} + j\omega\mathbf{D}$ is the effective current density in the medium that accounts for a conduction and a displacement current. For our wire, we may assume $\mathbf{J} = \sigma\mathbf{E}$ and $\mathbf{D} = \epsilon_0\mathbf{E}$ so that \mathbf{J}_{tot} is given by the effective permittivity defined through, $j\omega\epsilon_c = \sigma + j\omega\epsilon_0$,

$$\mathbf{J}_{\text{tot}} = \mathbf{J} + j\omega\mathbf{D} = \sigma\mathbf{E} + j\omega\epsilon_0\mathbf{E} = j\omega\epsilon_c\mathbf{E}$$

Noting that $\text{Re}(j\omega\epsilon_c) = \text{Re}(j\omega(\epsilon_R - j\epsilon_I)) = \omega\epsilon_I$, we may rewrite (10.16.4) as follows,

$$P_{\text{in}} \equiv - \frac{1}{2} \oint_S \text{Re}[\mathbf{E} \times \mathbf{H}^*] \cdot d\mathbf{S} = \frac{1}{2} \omega\epsilon_I \int_V \mathbf{E} \cdot \mathbf{E}^* dV \equiv P_{\text{loss}} \quad (10.16.5)$$

For our wire, $\omega\epsilon_I = \sigma$, but in general, if the dielectric displacement current term had a complex permittivity, e.g., $\mathbf{D} = \epsilon_d\mathbf{E} = (\epsilon_{dR} - j\epsilon_{dI})\mathbf{E}$, then it would also contribute to the effective ϵ_I , that is, $\omega\epsilon_I = \sigma + \omega\epsilon_{dI}$.

We will verify Eq. (10.16.5) explicitly for a length- L segment of the wire, as shown in Fig. 10.16.2. Using the solution (10.15.7), the power loss will arise from both components E_z, E_ρ . Because of the cylindrical symmetry, the volume element will be $dV = 2\pi\rho d\rho dz$, with the z -dependence given by $e^{-2\beta_I z}$. Then, we find for P_{loss} ,

$$\begin{aligned} P_{\text{loss}} &= \frac{1}{2} \omega \epsilon_I \int_V [|E_z|^2 + |E_\rho|^2] dV \\ &= \frac{1}{2} |A|^2 \omega \epsilon_I \int_0^L \int_0^a [|J_0(\gamma_c \rho)|^2 + \frac{|\beta|^2}{|\gamma_c|^2} |J_1(\gamma_c \rho)|^2] \cdot e^{-2\beta_I z} \cdot 2\pi\rho d\rho dz \\ &= \frac{1}{2} |A|^2 \omega \epsilon_I \left(\frac{1 - e^{-2\beta_I L}}{2\beta_I} \right) \int_0^a [|J_0(\gamma_c \rho)|^2 + \frac{|\beta|^2}{|\gamma_c|^2} |J_1(\gamma_c \rho)|^2] 2\pi\rho d\rho \end{aligned}$$

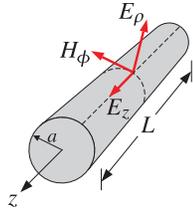


Fig. 10.16.2 Length- L section of Sommerfeld wire.

where we performed the z integration to get,

$$\int_0^L e^{-2\beta_I z} dz = \frac{1 - e^{-2\beta_I L}}{2\beta_I}$$

The Bessel function integrals can be done explicitly [1790]:

$$\begin{aligned} \int_0^a |J_0(\gamma_c \rho)|^2 2\pi\rho d\rho &= 2\pi a \frac{\text{Im}[\gamma_c^* J_0(\gamma_c^* a) J_1(\gamma_c a)]}{\text{Im}[\gamma_c^2]} \\ \int_0^a |J_1(\gamma_c \rho)|^2 2\pi\rho d\rho &= 2\pi a \frac{\text{Im}[\gamma_c J_1(\gamma_c a) J_0(\gamma_c^* a)]}{\text{Im}[\gamma_c^2]} \end{aligned} \quad (10.16.6)$$

It follows that P_{loss} will be,

$$\begin{aligned} P_{\text{loss}} &= \frac{1}{2} |A|^2 (1 - e^{-2\beta_I L}) \frac{\omega 2\pi a \epsilon_I}{2\beta_I} \cdot \\ &\cdot \frac{\text{Im}[\gamma_c^* J_0(\gamma_c^* a) J_1(\gamma_c a)] + \frac{|\beta|^2}{|\gamma_c|^2} \text{Im}[\gamma_c J_1(\gamma_c a) J_0(\gamma_c^* a)]}{\text{Im}[\gamma_c^2]} \end{aligned} \quad (10.16.7)$$

The power influx into the cylindrical volume of Fig. 10.16.2 consist of the following parts: the power entering the volume at the cross-sectional area at $z = 0$, minus

the power exiting the cross section at $z = L$, plus the power entering perpendicularly through the cylindrical area. Thus, using Eq. (10.15.7), with $H_\phi = Aj\omega\epsilon_c J_1(\gamma_c \rho) / \gamma_c$,

$$\begin{aligned} P_{\text{in}} &= - \oint_S \mathbf{P} \cdot d\mathbf{S} = \int_0^a \mathcal{P}_z 2\pi\rho d\rho \Big|_{z=0}^{z=L} - \int_0^L \mathcal{P}_\rho 2\pi a dz \\ &= \frac{1}{2} \int_0^a \text{Re}(E_\rho H_\phi^*) 2\pi\rho d\rho \Big|_{z=0}^{z=L} - \frac{1}{2} \int_0^L \text{Re}(E_z H_\phi^*) 2\pi a dz \\ &= \frac{1}{2} |A|^2 (1 - e^{-2\beta_I L}) \int_0^a \text{Re} \left[\frac{j\beta}{\gamma_c} J_1(\gamma_c \rho) \frac{-j\omega\epsilon_c^*}{\gamma_c^*} J_1(\gamma_c^* \rho) \right] 2\pi\rho d\rho + \\ &+ \frac{1}{2} |A|^2 \int_0^L \text{Re} \left[J_0(\gamma_c a) \frac{-j\omega\epsilon_c^*}{\gamma_c^*} J_1(\gamma_c^* a) \right] \cdot e^{-2\beta_I z} \cdot 2\pi a dz = \\ &= \frac{1}{2} |A|^2 (1 - e^{-2\beta_I L}) \frac{\omega \text{Re}[\beta\epsilon_c^*]}{|\gamma_c|^2} \int_0^a J_1(\gamma_c \rho) J_1(\gamma_c^* \rho) 2\pi\rho d\rho + \\ &+ \frac{1}{2} |A|^2 (1 - e^{-2\beta_I L}) \frac{\omega 2\pi a}{2\beta_I} \text{Im} \left[J_0(\gamma_c a) \frac{\epsilon_c^*}{\gamma_c^*} J_1(\gamma_c^* a) \right] \end{aligned}$$

where in the last term we used the identity $\text{Re}[-jC] = \text{Im}[C]$. Finally, using Eq. (10.16.6), we obtain the net power flowing into the cylindrical volume:

$$\begin{aligned} P_{\text{in}} &= \frac{1}{2} |A|^2 (1 - e^{-2\beta_I L}) \frac{\omega 2\pi a \text{Re}[\beta\epsilon_c^*]}{|\gamma_c|^2} \frac{\text{Im}[\gamma_c J_1(\gamma_c a) J_0(\gamma_c^* a)]}{\text{Im}[\gamma_c^2]} + \\ &+ \frac{1}{2} |A|^2 (1 - e^{-2\beta_I L}) \frac{\omega 2\pi a}{2\beta_I} \text{Im} \left[J_0(\gamma_c a) \frac{\epsilon_c^*}{\gamma_c^*} J_1(\gamma_c^* a) \right] \end{aligned} \quad (10.16.8)$$

It is left as an exercise in Problem 10.7 to show the equality of the two expressions in Eqs. (10.16.7) and (10.16.8). Canceling some common factors, this amounts to showing the equality of the following two sides:

$$\begin{aligned} \frac{2\beta_I \text{Re}[\beta\epsilon_c^*]}{|\gamma_c|^2} \frac{\text{Im}[\gamma_c J_1(\gamma_c a) J_0(\gamma_c^* a)]}{\text{Im}[\gamma_c^2]} + \text{Im} \left[J_0(\gamma_c a) \frac{\epsilon_c^*}{\gamma_c^*} J_1(\gamma_c^* a) \right] &= \\ = \epsilon_I \frac{\text{Im}[\gamma_c^* J_0(\gamma_c^* a) J_1(\gamma_c a)] + \frac{|\beta|^2}{|\gamma_c|^2} \text{Im}[\gamma_c J_1(\gamma_c a) J_0(\gamma_c^* a)]}{\text{Im}[\gamma_c^2]} \end{aligned} \quad (10.16.9)$$

10.17 Connection to Zenneck Surface Wave

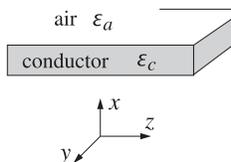
We mentioned earlier that the Zenneck surface wave is the planar limit of the Sommerfeld wave. Indeed, in the limit of large radius, $a \rightarrow \infty$, the ratios of both the Hankel functions and the Bessel functions in the characteristic equation (10.15.9) converge to $+j$, resulting in the equation:

$$\frac{\gamma}{\epsilon_a} = \frac{\gamma_c}{\epsilon_c} \quad (10.17.1)$$

which is precisely the characteristic equation of the Zenneck case that we considered in Sec. 7.10. Solving it for β gives rise to,

$$\beta = k_0 \sqrt{\frac{\epsilon_a \epsilon_c}{\epsilon_a + \epsilon_c}}, \quad \gamma = k_0 \sqrt{\frac{\epsilon_a^2}{\epsilon_a + \epsilon_c}}, \quad \gamma_c = k_0 \sqrt{\frac{\epsilon_c^2}{\epsilon_a + \epsilon_c}} \quad (10.17.2)$$

For completeness, we give the form of the fields above and below the interface at $x = 0$,



$x \leq 0$	$x \geq 0$
$E_z = E_0 e^{j\gamma_c x}$	$E_z = E_0 e^{j\gamma x}$
$E_x = E_0 \frac{\beta}{\gamma_c} e^{j\gamma_c x}$	$E_x = E_0 \frac{\beta}{\gamma} e^{j\gamma x}$
$H_y = E_0 \frac{k_0 \epsilon_c}{\eta_0 \gamma_c} e^{j\gamma_c x}$	$H_y = E_0 \frac{k_0 \epsilon_a}{\eta_0 \gamma} e^{j\gamma x}$

(10.17.3)

where E_0 is a constant. These agree with the results of Sec. 7.10 after remapping the notation by the replacements, $z \rightarrow -x$, $x \rightarrow z$, $k_x \rightarrow \beta$, $k_z \rightarrow \gamma$, $k'_z \rightarrow \gamma_c$.

They can be derived more directly by solving the TM propagation problem in cartesian coordinates using Eqs. (9.3.10), and applying the boundary conditions at the planar interface. They can also be derived as the large-radius limit of the wire solutions. To see this, let us renormalize Eqs. (10.15.7) with respect to their values at the surface of the conductor, i.e., by replacing A, B from the relationship, $A J_0(\gamma_c a) = B H_0^{(1)}(\gamma a) = E_0$,

$0 \leq \rho \leq a$	$a \leq \rho < \infty$
$E_z = E_0 \frac{J_0(\gamma_c \rho)}{J_0(\gamma_c a)}$	$E_z = E_0 \frac{H_0^{(1)}(\gamma \rho)}{H_0^{(1)}(\gamma a)}$
$E_\rho = E_0 \frac{j\beta}{\gamma_c} \frac{J_1(\gamma_c \rho)}{J_0(\gamma_c a)}$	$E_\rho = E_0 \frac{j\beta}{\gamma} \frac{H_1^{(1)}(\gamma \rho)}{H_0^{(1)}(\gamma a)}$
$H_\phi = E_0 \frac{jk_0 \epsilon_c}{\eta_0 \gamma_c} \frac{J_1(\gamma_c \rho)}{J_0(\gamma_c a)}$	$H_\phi = E_0 \frac{jk_0 \epsilon_a}{\eta_0 \gamma} \frac{H_1^{(1)}(\gamma \rho)}{H_0^{(1)}(\gamma a)}$
$\gamma_c = \sqrt{k_0^2 \epsilon_c - \beta^2}$	$\gamma = \sqrt{k_0^2 \epsilon_a - \beta^2}$

(10.17.4)

In this form, the boundary condition for E_z is automatically satisfied, whereas that for H_ϕ leads directly to the characteristic equation (10.15.9). Since x is measured with respect to the planar conductor surface, before we take the limit for large radius, let us make the replacement, $\rho = a + x$. Then, using the asymptotic expressions (10.15.6), we obtain the following limits, as $a \rightarrow \infty$ with fixed x ,

$$\frac{H_0^{(1)}(\gamma \rho)}{H_0^{(1)}(\gamma a)} \approx \sqrt{\frac{a}{\rho}} e^{j\gamma(\rho-a)} = \sqrt{\frac{a}{a+x}} e^{j\gamma x} \rightarrow e^{j\gamma x}$$

$$\frac{H_1^{(1)}(\gamma \rho)}{H_0^{(1)}(\gamma a)} \approx \sqrt{\frac{a}{\rho}} e^{j\gamma(\rho-a)} e^{-j\pi/2} = -j \sqrt{\frac{a}{a+x}} e^{j\gamma x} \rightarrow -j e^{j\gamma x}$$

from which the $x > 0$ expressions in Eq. (10.17.3) are obtained. The corresponding asymptotic expressions for $J_0(u), J_1(u)$ are given as follows [1790], assuming that u has a large and negative imaginary part,

$$J_0(u) = \sqrt{\frac{2}{\pi u}} \cos(u - \frac{\pi}{4}) \approx \sqrt{\frac{1}{2\pi u}} e^{j(u-\pi/4)}, \quad \text{Im}(u) < 0$$

$$J_1(u) = \sqrt{\frac{2}{\pi u}} \cos(u - \frac{3\pi}{4}) \approx \sqrt{\frac{1}{2\pi u}} e^{j(u-3\pi/4)}, \quad \text{Im}(u) < 0$$

(10.17.5)

and from these we obtain the limits, which give rise to the $x < 0$ expressions in (10.17.3),

$$\frac{J_0(\gamma_c \rho)}{J_0(\gamma_c a)} \approx \sqrt{\frac{a}{\rho}} e^{j\gamma_c(\rho-a)} = \sqrt{\frac{a}{a+x}} e^{j\gamma_c x} \rightarrow e^{j\gamma_c x}$$

$$\frac{J_1(\gamma_c \rho)}{J_0(\gamma_c a)} \approx \sqrt{\frac{a}{\rho}} e^{j\gamma_c(\rho-a)} e^{-j\pi/2} = -j \sqrt{\frac{a}{a+x}} e^{j\gamma_c x} \rightarrow -j e^{j\gamma_c x}$$

Like the Sommerfeld wave, the Zenneck wave exhibits a strong skin effect within the conductor, but weak bounding on the air side, and low attenuation in the direction of propagation. For a good conductor, we have $|\epsilon_c| \gg \epsilon_a$, and we can approximate ϵ_c and γ_c as in the wire case,

$$\epsilon_c \approx -j \frac{\sigma}{\omega \epsilon_0}, \quad \gamma_c = \frac{k_0 \sqrt{\epsilon_c}}{\sqrt{1 + \frac{\epsilon_a}{\epsilon_c}}} \approx k_0 \sqrt{\epsilon_c} \approx \frac{1-j}{\delta}, \quad \delta = \sqrt{\frac{2}{\sigma \mu_0 \omega}}$$

Thus, the fields are attenuating rapidly, within a skin depth δ , inside the conductor, $e^{j\gamma_c x} \approx e^{jx/\delta} e^{x/\delta} = e^{jx/\delta} e^{-|x|/\delta}$, for $x < 0$, but not so rapidly outside since $|\gamma| \ll k_0$. Indeed, we have in terms of the dimensionless parameter $\sigma / (\omega \epsilon_0) \gg 1$, setting $\epsilon_a = 1$,

$$\gamma = \frac{\gamma_c \epsilon_a}{\epsilon_c} \approx \frac{k_0 \sqrt{\epsilon_c}}{\epsilon_c} = \frac{k_0}{\sqrt{\epsilon_c}} = \gamma_R + j\gamma_I, \quad \gamma_R = \gamma_I = \frac{k_0}{\sqrt{\frac{2\sigma}{\omega \epsilon_0}}} \ll k_0$$

similarly, the propagation constant β has a very small imaginary part,

$$\beta = \beta_R - j\beta_I = k_0 \sqrt{\frac{\epsilon_c}{1 + \epsilon_c}} = k_0 \sqrt{\frac{1}{1 + \epsilon_c^{-1}}} \approx k_0 \left(1 - \frac{1}{2\epsilon_c}\right) = k_0 - jk_0 \frac{\omega \epsilon_0}{2\sigma}$$

10.18 Skin Effect for Round Wire

The Sommerfeld wave provides a nice illustration of the skin effect in a round wire. Within the wire, the total current density includes both the conduction and displacement currents,

$$\mathbf{J} = \sigma \mathbf{E} + j\omega \epsilon_0 \mathbf{E} = (\sigma + j\omega \epsilon_0) \mathbf{E} = j\omega \epsilon_c \mathbf{E}$$

For a good conductor, we typically ignore the displacement current, writing, $\mathbf{J} = \sigma \mathbf{E}$. For either \mathbf{J} , the current density in the z -direction is proportional to E_z , and therefore,

we have the same relationship as (10.15.20). Indeed, writing $E_z(\rho) = AJ_0(\gamma_c \rho)$ from Eq. (10.15.7), and setting, $J_z = j\omega\epsilon_c E_z$, we have,

$$\frac{J_z(\rho)}{J_z(a)} = \frac{E_z(\rho)}{E_z(a)} = \frac{J_0(\gamma_c \rho)}{J_0(\gamma_c a)}, \quad \rho \leq a \quad (10.18.1)$$

If we make the approximation $\gamma_c \approx k_0\sqrt{\epsilon_c} \approx (1-j)\sqrt{\sigma\mu_0\omega/2} = (1-j)/\delta$, then the results are equivalent to those obtained by the standard textbook treatments of the skin effect that (i) ignore the displacement current, (ii) assume no z-dependence, and (iii) consider the fields only inside the wire. Here, the z-dependence comes from the assumed factor $e^{-j\beta z}$ which could be set to unity for distances that are much less than a wavelength, $z \ll 1/k_0 \approx 1/\beta$.

A typical plot of Eq. (10.18.1) is depicted in Fig. 10.15.3. The numerical issues regarding the evaluation of the ratio (10.18.1) were discussed in Example 10.15.2. Within a very narrow layer of width of a few skin depths δ from the surface, $\rho \approx a$, the ratio decays exponentially with distance, $|e^{j\gamma_c(\rho-a)}| = e^{-(a-\rho)/\delta}$.

The total current I flowing in the z-direction is obtained by integrating $J_z(\rho) = j\omega\epsilon_c E_z(\rho) = j\omega\epsilon_c A J_0(\gamma_c \rho)$, over the cross-sectional area of the wire:

$$I = \int_0^a J_z(\rho) 2\pi\rho d\rho = 2\pi j\omega\epsilon_c A \int_0^a \rho J_0(\gamma_c \rho) d\rho = 2\pi a A \frac{j\omega\epsilon_c}{\gamma_c} J_1(\gamma_c a) \quad (10.18.2)$$

where we used the indefinite integral, $\int u J_0(u) du = u J_1(u)$, and the fact that $J_1(0) = 0$. From Eq. (10.15.7), we recognize this as, $I = 2\pi a H_\phi(a)$, which is a consequence of Ampère's law, obtained by integrating H_ϕ around the contour $\rho = a$. From the current I , we may derive the internal impedance of the wire (per unit wire length):

$$Z = \frac{E_z(a)}{I} = \frac{A J_0(\gamma_c a)}{2\pi a A \frac{j\omega\epsilon_c}{\gamma_c} J_1(\gamma_c a)} = \frac{\gamma_c J_0(\gamma_c a)}{2\pi a j\omega\epsilon_c J_1(\gamma_c a)}$$

and with the approximations, $\gamma_c \approx (1-j)/\delta = (1-j)\sqrt{\sigma\mu_0\omega/2}$, and, $j\omega\epsilon_c \approx \sigma$, we obtain the standard textbook formula,

$$Z = \frac{\gamma_c J_0(\gamma_c a)}{2\pi a \sigma J_1(\gamma_c a)} \quad (\text{impedance of round wire}) \quad (10.18.3)$$

In the low-frequency limit, $a/\delta = a\sqrt{\sigma\mu_0\omega/2} \rightarrow 0$, we may use the following second-order Taylor expansion, valid for small u ,

$$\frac{u J_0(u)}{2 J_1(u)} = 1 - \frac{u^2}{8}$$

to obtain,

$$Z = \frac{1}{\pi a^2 \sigma} \frac{\gamma_c a J_0(\gamma_c a)}{2 J_1(\gamma_c a)} \approx \frac{1}{\pi a^2 \sigma} \left[1 - \frac{(\gamma_c a)^2}{8} \right]$$

or, because, $\gamma_c^2 = (1-j)^2/\delta^2 = (1-j)^2 \omega\mu_0\sigma/2 = -j\omega\mu_0\sigma$,

$$Z_{\text{low}} = \frac{1}{\pi a^2 \sigma} + j\omega \frac{\mu_0}{8\pi} \quad (\text{low-frequency wire impedance}) \quad (10.18.4)$$

Comparing with $Z = R + j\omega L$, we obtain the standard low-frequency DC resistance and the internal *self-inductance* per unit length of the wire,

$$R = \frac{1}{\pi a^2 \sigma}, \quad L = \frac{\mu_0}{8\pi} \quad (10.18.5)$$

On the other hand, for high frequencies, $a/\delta = a\sqrt{\sigma\mu_0\omega/2} \rightarrow \infty$, we may replace the Bessel function ratio by $J_0/J_1 = j$, and $j\gamma_c = (1+j)/\delta$, to get,

$$Z_{\text{high}} = \frac{1+j}{2\pi a \delta \sigma} \quad (\text{high-frequency wire impedance}) \quad (10.18.6)$$

Eq. (10.18.6) is originally due to Rayleigh [1293]. Recalling from Sec. 2.8 that the surface resistance of a good conductor is $R_s = 1/(\sigma\delta)$, we may rewrite (10.18.6) as,

$$Z_{\text{high}} = \frac{1+j}{2\pi a \delta \sigma} = \frac{(1+j)R_s}{2\pi a} \quad (10.18.7)$$

The DC resistance per unit length of a wire of arbitrary cross-sectional area S is given by $R = 1/(S\sigma)$ and is derived by assuming that the current density is uniform over the area S . For the round wire, we have $S = \pi a^2$. At high frequencies, the same formula would imply from (10.18.6) that $S = 2\pi a\delta$, that is, the current is effectively confined to flow within a narrow ring of radius a and width δ , as shown in Fig. 10.18.1.

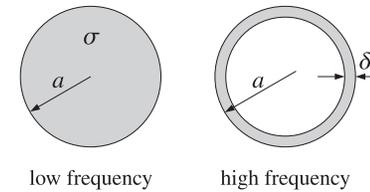


Fig. 10.18.1 Effective areas of current distribution at low and high frequencies.

We note in passing that the low frequency self-inductance L of Eq. (10.18.5) is usually derived using energy considerations, by considering the magnetic energy enclosed in a unit-length cylindrical volume of the conductor, and setting,

$$\frac{1}{2} L I^2 = \int_0^a \frac{1}{2} \mu_0 H_\phi^2 2\pi\rho d\rho$$

Since the current density is assumed to be uniform, the amount of current enclosed within a radius ρ is given by $J_z \pi \rho^2$, with total current $I = J_z \pi a^2$. Then, Ampère's law gives $J_z \pi \rho^2 = 2\pi\rho H_\phi(\rho)$, resulting in, $H_\phi(\rho) = J_z \rho/2 = I\rho/(2\pi a^2)$. Inserting this into the above energy relation we find,

$$\frac{1}{2} L I^2 = \frac{\mu_0 I^2}{4\pi a^4} \int_0^a \rho^3 d\rho = \frac{1}{2} \frac{\mu_0}{8\pi} I^2 \Rightarrow L = \frac{\mu_0}{8\pi}$$

The same expression for H_ϕ can also be obtained from (10.15.7) in the low-frequency limit, $\gamma_c \rightarrow 0$, by using the small argument expansion $J_1(\gamma_c a) \approx \gamma_c a/2$.

Finally, we comment on Ampère's law applied to the outside of the wire. In Fig. 10.15.3, the field E_z , outside the wire, is due to the displacement current only. If we apply Ampère's law to a contour of radius r , with $r > a$, the total enclosed current, $I_r = 2\pi r H_\phi(r)$, would consist of both the conduction and the displacement currents. The current density, $J_z = j\omega\epsilon E_z$, is given as follows inside and outside the conductor, from (10.15.7),

$$J_z(\rho) = \begin{cases} j\omega\epsilon_c A J_0(\gamma_c \rho), & \rho \leq a \\ j\omega\epsilon_0 B H_0^{(1)}(\gamma \rho), & \rho \geq a \end{cases}$$

The total current crossing the area of radius $r > a$ can be split into two parts:

$$\begin{aligned} I_r &= \int_0^r J_z 2\pi\rho d\rho = \int_0^a J_z 2\pi\rho d\rho + \int_a^r J_z 2\pi\rho d\rho \\ &= \int_0^a A J_0(\gamma_c \rho) 2\pi\rho d\rho + \int_a^r B H_0^{(1)}(\gamma \rho) 2\pi\rho d\rho \\ &= 2\pi a A \frac{j\omega\epsilon_c}{\gamma_c} J_1(\gamma_c a) + \left[2\pi r B \frac{j\omega\epsilon_0}{\gamma} H_1^{(1)}(\gamma r) - 2\pi a B \frac{j\omega\epsilon_0}{\gamma} H_1^{(1)}(\gamma a) \right] \\ &= 2\pi r B \frac{j\omega\epsilon_0}{\gamma} H_1^{(1)}(\gamma r) = 2\pi r H_\phi(r) \end{aligned}$$

where we used Eqs. (10.15.7), the boundary conditions (10.15.8), and the indefinite integral, $\int u H_0^{(1)}(u) du = u H_1^{(1)}(u)$.

The skin effect is discussed in most textbooks. A few references are [1054-1058], in which one can find additional ones, including historical references.

10.19 Goubau Line

The Goubau line, or Harms-Goubau line, refers to a TM surface wave propagating along a dielectric coated conductor, depicted in Fig. 10.19.1. It was considered first by Harms in 1907, and studied extensively in the 1950s by Goubau as an alternative to the Sommerfeld line that provides better lateral confinement [984-999]. Interest in the Goubau line has been renewed in the 2000s for use in THz applications [1000-1016].[†]

In the Sommerfeld wire, the finite conductivity of the conductor was essential in order to render the lateral wavenumber complex-valued for lateral confinement. In the Goubau line case, this is not necessary and we will initially assume that the conductor is perfect and that the dielectric coating is lossless. Ohmic and dielectric losses can then be taken into account approximately.

In Fig. 10.19.1, let a, b be the inner radius of the conductor and the outer radius of the dielectric coating, so that the coating thickness is, $d = b - a$, and let ϵ_d be the relative permittivity of the coating. Assuming cylindrical symmetry, the TM wave will

[†]Incidentally, our ECE Department at Rutgers awards yearly the "Georg Goubau Memorial Prize" to a graduating senior for excellence in the study of Electromagnetics.

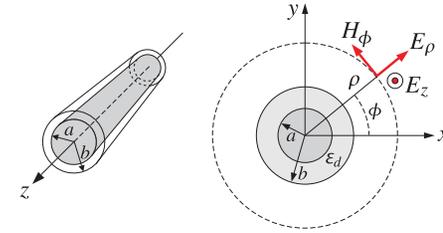


Fig. 10.19.1 Goubau line, shown in cylindrical coordinates.

have field components E_z, E_ρ, H_ϕ that satisfy Eq. (10.15.4),

$$\begin{aligned} \frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial E_z}{\partial \rho} \right) + \gamma^2 E_z &= 0, \quad \gamma^2 = k_0^2 \epsilon - \beta^2 \\ E_\rho &= -\frac{j\beta}{\gamma^2} \frac{\partial E_z}{\partial \rho}, \quad H_\phi = -\frac{jk_0 \epsilon}{\eta_0 \gamma^2} \frac{\partial E_z}{\partial \rho} \end{aligned} \tag{10.19.1}$$

with the assumed factor $e^{j\omega t - j\beta z}$, and solution given as in Eq. (10.15.5),

$$E_z = C_0(\gamma \rho), \quad E_\rho = \frac{j\beta}{\gamma} C_1(\gamma \rho), \quad H_\phi = \frac{jk_0 \epsilon}{\eta_0 \gamma} C_1(\gamma \rho) \tag{10.19.2}$$

where $C_n(\gamma \rho), n = 0, 1$, are linear combinations of the Bessel functions $J_n(\gamma \rho), Y_n(\gamma \rho)$, or, the Hankel functions, $H_n^{(1,2)}(\gamma \rho)$.

Because the conductor was assumed to be perfect, and the dielectric, lossless, we are looking for solutions that have a real-valued propagation wavenumber β . Let us denote the lateral wavenumber within the dielectric coating by $h = \sqrt{k_0^2 \epsilon_d - \beta^2}$. It may be assumed to be real-valued. But outside the coating, within the air, the wavenumber $\gamma_a = \sqrt{k_0^2 \epsilon_a - \beta^2}$ must necessarily be pure imaginary in order to guarantee that the fields decay rapidly at large radial distances. Setting for example, $\gamma_a = j\gamma$, with positive γ , it follows from Eq. (10.15.6) that the Hankel functions $H_n^{(1)}(\gamma_a \rho)$ will decay as follows with increasing $\rho \rightarrow \infty$,

$$H_n^{(1)}(\gamma_a \rho) = H_n^{(1)}(j\gamma \rho) \approx \sqrt{\frac{2}{\pi j \gamma \rho}} e^{-\gamma \rho} e^{-j(2n+1)\pi/4} = \frac{1}{j^{n+1}} \sqrt{\frac{2}{\pi \gamma \rho}} e^{-\gamma \rho} \tag{10.19.3}$$

Thus, $\gamma_a^2 = -\gamma^2 = k_0^2 \epsilon_a - \beta^2$, or, $\gamma = \sqrt{\beta^2 - k_0^2 \epsilon_a}$. Since both γ and h are real-valued, it follows that β must lie in the range,

$$k_0 \sqrt{\epsilon_a} \leq \beta \leq k_0 \sqrt{\epsilon_d} \tag{10.19.4}$$

The Hankel functions for imaginary argument are related to the modified Bessel functions of second kind, via the relationship [1790],

$$K_n(u) = \frac{\pi}{2} j^{n+1} H_n^{(1)}(ju), \quad u > 0 \tag{10.19.5}$$

which have the same asymptotic form as (10.19.3), for large positive u ,

$$K_n(u) \approx \sqrt{\frac{\pi}{2u}} e^{-u} \quad (10.19.6)$$

Thus, on the air side ($\rho > b$), instead of working with $H_n^{(1)}(j\gamma\rho)$, it proves simpler to express the field solutions in terms of $K_n(\gamma\rho)$. The Helmholtz equations (10.19.1) remain the same with the replacement, $y^2 \rightarrow -y^2$,

$$\frac{1}{\rho} \frac{\partial}{\partial \rho} \left(\rho \frac{\partial E_z}{\partial \rho} \right) - y^2 E_z = 0, \quad y^2 = \beta^2 - k_0^2 \epsilon_a \quad (10.19.7)$$

$$E_\rho = \frac{j\beta}{y^2} \frac{\partial E_z}{\partial \rho}, \quad H_\phi = \frac{jk_0 \epsilon_a}{\eta_0 y^2} \frac{\partial E_z}{\partial \rho}$$

The solutions are found after using the relationship, $\partial K_0(\gamma\rho) / \partial \rho = -\gamma K_1(\gamma\rho)$,

$$E_z = E_1 K_0(\gamma\rho), \quad E_\rho = -E_1 \frac{j\beta}{y} K_1(\gamma\rho), \quad H_\phi = -E_1 \frac{jk_0 \epsilon_a}{\eta_0 y} K_1(\gamma\rho) \quad (10.19.8)$$

where E_1 is a constant. Within the dielectric coating, $a \leq \rho \leq b$, the solution will be a linear combination of $J_0(h\rho)$, $Y_0(h\rho)$, say, $Z_0(h\rho) = J_0(h\rho) - AY_0(h\rho)$, and will be given by (10.19.2),

$$E_z = E_0 Z_0(h\rho), \quad E_\rho = E_0 \frac{j\beta}{h} Z_1(h\rho), \quad H_\phi = E_0 \frac{jk_0 \epsilon_d}{\eta_0 h} Z_1(h\rho) \quad (10.19.9)$$

where E_0 is a constant, and $Z_1(h\rho) = J_1(h\rho) - AY_1(h\rho)$. To determine A , we demand that E_z vanish on the surface of the perfect conductor, that is, at $\rho = a$,

$$E_z(a) = 0 \Rightarrow Z_0(ha) = J_0(ha) - AY_0(ha) = 0 \Rightarrow A = \frac{J_0(ha)}{Y_0(ha)} \quad (10.19.10)$$

In summary, the field solutions are given as follows, in the coating and in the air,

$a \leq \rho \leq b$	$b \leq \rho < \infty$
$E_z = E_0 Z_0(h\rho)$	$E_z = E_1 K_0(\gamma\rho)$
$E_\rho = E_0 \frac{j\beta}{h} Z_1(h\rho)$	$E_\rho = -E_1 \frac{j\beta}{y} K_1(\gamma\rho)$
$H_\phi = E_0 \frac{jk_0 \epsilon_d}{\eta_0 h} Z_1(h\rho)$	$H_\phi = -E_1 \frac{jk_0 \epsilon_a}{\eta_0 y} K_1(\gamma\rho)$
$h = \sqrt{k_0^2 \epsilon_d - \beta^2}$	$y = \sqrt{\beta^2 - k_0^2 \epsilon_a}$

(10.19.11)

and Z_0, Z_1 are defined as the functions,

$$\begin{aligned} Z_0(h\rho) &= J_0(h\rho) - \frac{J_0(ha)}{Y_0(ha)} Y_0(h\rho) \\ Z_1(h\rho) &= J_1(h\rho) - \frac{J_1(ha)}{Y_1(ha)} Y_1(h\rho) \end{aligned} \quad (10.19.12)$$

The boundary condition at the conductor surface ($\rho = a$) is built into the solution. The boundary conditions at the coating-air interface ($\rho = b$) are the continuity of the tangential electric and magnetic fields, E_z, H_ϕ , resulting in the conditions:

$$\begin{aligned} E_0 Z_0(hb) &= E_1 K_0(\gamma b) \\ E_0 \frac{jk_0 \epsilon_d}{\eta_0 h} Z_1(hb) &= -E_1 \frac{jk_0 \epsilon_a}{\eta_0 y} K_1(\gamma b) \end{aligned} \quad (10.19.13)$$

Dividing the two sides, we obtain the characteristic equation that determines the propagation wavenumber β ,

$$\begin{aligned} h &= \sqrt{k_0^2 \epsilon_d - \beta^2} \\ y &= \sqrt{\beta^2 - k_0^2 \epsilon_a} \end{aligned} \quad \boxed{\frac{h}{\epsilon_d} \frac{Z_0(hb)}{Z_1(hb)} = -\frac{y}{\epsilon_a} \frac{K_0(\gamma b)}{K_1(\gamma b)}} \quad (10.19.14)$$

where we may set $\epsilon_a = 1$ from now on. One way to solve this equation iteratively, is to first cast it in the form,

$$\frac{y}{h} = -\frac{1}{\epsilon_d} \frac{K_1(\gamma b)}{K_0(\gamma b)} \frac{Z_0(hb)}{Z_1(hb)} \equiv F(\beta) \quad (10.19.15)$$

and then solve the left-hand side for β ,

$$\begin{aligned} \frac{y}{h} = F(\beta) &\Rightarrow \frac{y^2}{h^2} = \frac{\beta^2 - k_0^2}{k_0^2 \epsilon_d - \beta^2} = F^2(\beta), \quad \text{or,} \\ \beta &= k_0 \sqrt{\frac{1 + \epsilon_d F^2(\beta)}{1 + F^2(\beta)}} \end{aligned} \quad (10.19.16)$$

This can be turned in the following recursion, for $n = 0, 1, 2, \dots$,

$$\beta_{n+1} = r k_0 \sqrt{\frac{1 + \epsilon_d F^2(\beta_n)}{1 + F^2(\beta_n)}} + (1 - r) \beta_n \quad (10.19.17)$$

where we introduced a relaxation parameter, $0 < r \leq 1$. The recursion can be initialized somewhere within the β range (10.19.4), for example, $\beta_0 = 0.999 k_0 \sqrt{\epsilon_d}$. The following MATLAB function, `goubau`, in the EWA toolbox, implements this iteration:

```
[be,ga,h,N,E,pd] = goubau(a,b,ed,f,r,tol,be0)
[be,ga,h,N,E,pd] = goubau(a,b,ed,f,r,tol)      (be0 = 0.999*k0*sqrt(ed))
[be,ga,h,N,E,pd] = goubau(a,b,ed,f,r)        (tol = 1e-10)
```

```
% a,b = inner and outer radii [meters]
% ed = relative dielectric constant of coating (ed>1)
% f = vector of frequencies [Hz]
% r = relaxation parameter (0 < r <= 1)
% tol = computational tolerance, default tol = 1e-10
% be0 = initializing vector, size(f), default, be0 = 0.999*k0*sqrt(ed)
%
% be = propagation wavenumber [rads/m], size(f)
% ga = lateral decay constant [1/m], size(f)
```

% h = lateral wavenumber in dielectric coating [rads/m], size(f)
 % N = number of iterations to converge, scalar, represents all f
 % E = computational error of characteristic equation, size(f)
 % pd = proportion of transmitted power in dielectric coating

The iteration is stopped when two successive iterates are to within a prescribed error tolerance, that is, $|\beta_{n+1} - \beta_n| \leq \text{tol}$, typically with, $\text{tol} = 10^{-10}$, or smaller. The output parameter p_d represents the proportion of transmitted power residing within the dielectric coating, and is a measure of confinement. Its calculation is explained below. The computational error output E is the difference between the two sides of Eq. (10.19.15) for the last computed iterates,

$$E = \left| \frac{Y}{h} - F(\beta) \right|$$

Goubau [1985] and King and Wiltse [1994] made certain approximations to the characteristic equation to cast it in a more manageable form. Such approximations are not necessary in the above iteration, which is implemented using MATLAB's built-in Bessel functions. The iteration is applicable over a wide range of frequencies, including THz.

Once the fields are determined for the ideal conductor and lossless dielectric case, they can be used to calculate the attenuation coefficient α (in nepers/m) along the line, following the procedures discussed in Sec. 9.2, that is,

$$\alpha = \frac{P'_{\text{loss}}}{2P_T} \quad (10.19.18)$$

where P'_{loss} is the power loss per unit conductor length, and P_T , the transmitted power.

The power loss P'_{loss} consists of a part due to the losses in the conductor, assuming a large but finite conductivity σ , and a part due to the losses in the dielectric coating, assuming a small (negative) imaginary part for the dielectric permittivity, ϵ_I . The two parts are given by,

$$P'_C = \oint_C \frac{1}{2} R_s |\mathbf{H}_{\text{tan}}|^2 dl, \quad R_s = \sqrt{\frac{\omega\mu_0}{2\sigma}} \quad (10.19.19)$$

$$P'_d = \frac{1}{2} \omega \epsilon_I \int_S |E|^2 dS$$

where C is the periphery of the conductor, i.e., the circle of radius $\rho = a$, and \mathbf{H}_{tan} is the tangential magnetic field on the surface, i.e., $H_\phi(\rho)$ for $\rho = a$, and R_s is the surface impedance of the conductor. For the dielectric, S is the cross-sectional annular area defined by $a \leq \rho \leq b$. Thus, the total loss is given by,

$$P'_{\text{loss}} = P'_C + P'_d = \frac{1}{2} R_s |H_\phi(a)|^2 (2\pi a) + \frac{1}{2} \omega \epsilon_I \int_a^b [|E_z|^2 + |E_\rho|^2] 2\pi\rho d\rho \quad (10.19.20)$$

The transmitted power is obtained by integrating the z-component of the Poynting vector on the cross-sectional areas of the coating and the air side,

$$P_T = \int_a^b \frac{1}{2} \text{Re}[E_\rho H_\phi^*] 2\pi\rho d\rho + \int_b^\infty \frac{1}{2} \text{Re}[E_\rho H_\phi^*] 2\pi\rho d\rho \quad (10.19.21)$$

From the solutions (10.19.11), we obtain,

$$\begin{aligned} P'_{\text{loss}} &= \frac{1}{2} R_s (2\pi a) \frac{k_0^2 \epsilon_d^2 |E_0|^2}{\eta_0^2 h^2} Z_1^2(ha) + \\ &+ \frac{1}{2} |E_0|^2 \omega \epsilon_I \left[\int_a^b Z_0^2(h\rho) 2\pi\rho d\rho + \frac{\beta^2}{h^2} \int_a^b Z_1^2(h\rho) 2\pi\rho d\rho \right] \\ P_T &= \frac{1}{2} \frac{k_0 \epsilon_d \beta |E_0|^2}{\eta_0 h^2} \int_a^b Z_1^2(h\rho) 2\pi\rho d\rho + \frac{1}{2} \frac{k_0 \epsilon_a \beta |E_1|^2}{\eta_0 y^2} \int_b^\infty K_1^2(y\rho) 2\pi\rho d\rho \end{aligned}$$

Let us define the following normalized integrals that can be derived with the help of related integrals in [1790], and after using the condition $Z_0(ha) = 0$,

$$\begin{aligned} U &\equiv \frac{1}{Z_1^2(ha)} \frac{2}{a^2} \int_a^b Z_1^2(h\rho) \rho d\rho = \frac{b^2}{a^2} \cdot \frac{Z_1^2(hb) - Z_0(hb)Z_2(hb)}{Z_1^2(ha)} - 1 \\ V &\equiv \frac{1}{Z_1^2(ha)} \frac{2}{a^2} \int_a^b Z_0^2(h\rho) \rho d\rho = \frac{b^2}{a^2} \cdot \frac{Z_1^2(hb) + Z_0^2(hb)}{Z_1^2(ha)} - 1 \\ W &\equiv \frac{Z_1^2(hb)}{Z_1^2(ha)} \frac{1}{K_1^2(yb)} \frac{2}{a^2} \int_b^\infty K_1^2(y\rho) \rho d\rho = \\ &= \frac{Z_1^2(hb)}{Z_1^2(ha)} \cdot \frac{b^2}{a^2} \cdot \frac{K_0(yb)K_2(yb) - K_1^2(yb)}{K_1^2(yb)} \end{aligned} \quad (10.19.22)$$

where the subscript 2 refers to the Bessel functions of order two. Using the relationships (10.19.13) and (10.19.14), and setting $\epsilon_a = 1$, we may express P'_{loss} and P_T in terms of the quantities U, V, W ,

$$\begin{aligned} P'_{\text{loss}} &= |E_0|^2 \pi a^2 \frac{k_0^2 \epsilon_d^2 Z_1^2(ha)}{\eta_0^2 h^2} \left[\frac{R_s}{a} + \frac{\eta_0 \tan \theta}{2k_0 \epsilon_d} (\beta^2 U + h^2 V) \right] \\ P_T &= \frac{1}{2} |E_0|^2 \pi a^2 \frac{\beta k_0 \epsilon_d^2 Z_1^2(ha)}{\eta_0 h^2} \left[\frac{1}{\epsilon_d} U + W \right] \end{aligned} \quad (10.19.23)$$

where we introduced the loss-tangent for the dielectric, $\epsilon_I = \epsilon_d \tan \theta$. Dividing, we obtain the total attenuation coefficient, with the R_s term representing the conductor losses, and the $\tan \theta$ term, the dielectric ones,

$$\alpha = \frac{P'_{\text{loss}}}{2P_T} = \frac{k_0}{\beta \eta_0} \cdot \frac{\frac{R_s}{a} + \frac{\eta_0 \tan \theta}{2k_0 \epsilon_d} (\beta^2 U + h^2 V)}{\frac{1}{\epsilon_d} U + W} \quad (10.19.24)$$

Eq. (10.19.24) is equivalent[†] to the results of [996]. By taking appropriate limits as $d/a = (b - a)/a \rightarrow 0$, we may obtain the approximation given by King-Wiltse [1994],

$$\alpha_{\text{app}} = \frac{k_0}{2\beta \eta_0} \cdot \frac{\frac{R_s}{a} + \frac{\eta_0 \tan \theta}{k_0 \epsilon_d} \frac{d}{a} \beta^2}{\frac{1}{\epsilon_d} \frac{d}{a} + \frac{1}{2\gamma b}} \quad (10.19.25)$$

[†] after using the Bessel function identities [1790], $C_0(u) + C_2(u) = \frac{2}{u} C_1(u)$, $K_0(u) - K_2(u) = \frac{2}{u} K_1(u)$.

where, to the same order of approximation, we can replace, $\frac{d}{a} = \frac{b-a}{a} \approx \ln\left(\frac{b}{a}\right)$.

It follows also from Eq. (10.19.23) that the proportion of the transmitted power within the dielectric (one of the outputs of the function goubau) is given by,

$$p_d = \frac{\frac{1}{\epsilon_d} U}{\frac{1}{\epsilon_d} U + W} = \text{proportion of power in dielectric coating} \quad (10.19.26)$$

The following MATLAB function, goubatt, implements Eqs. (10.19.24) and (10.19.25):

```
[atot,ac,ad,app,pd] = goubatt(a,b,ed,f,be,sigma,tand)

% a,b = inner and outer radii [meters]
% ed = relative dielectric constant of coating (ed>1)
% f = vector of frequencies [Hz]
% be = propagation wavenumber [rads/m], size(f), obtained from GOUBAU
% sigma = conductivity of inner conductor [siemens/m]
% tand = loss tangent of dielectric coating
%
% atot = total attenuation coefficient [nepers/m], size(f), atot = ac + ad
% ac = attenuation due to conductor [nepers/m], size(f)
% ad = attenuation due to dielectric [nepers/m], size(f)
% app = total attenuation coefficient [nepers/m], King-Wiltse-Goubau approx
% pd = proportion of transmitted power in dielectric coating
%
% conversions: dB/m = 8.68589 * atot = 20*log10(exp(1)) * atot
% dB/100ft = 8.68589 * 30.48 * atot
```

The Goubau TM mode discussed in this section is the lowest mode propagating along a dielectric-coated metal wire. Higher modes exist and their cutoff frequencies have been discussed in [995,998]. The cutoff condition occurs when the lateral wavenumber γ in air becomes zero and the wave is no longer bound to the vicinity of the wire, becoming a leaky mode. Since $\gamma^2 + h^2 = k_0^2(\epsilon_d - 1)$, it follows that when $\gamma = 0$, then h takes the value, $h_0 = k_0\sqrt{\epsilon_d - 1}$. The characteristic equation (10.19.14) then implies that $Z_0(h_0b) = 0$, or, equivalently [998],

$$Z_0(h_0b) = J_0(h_0b) - A Y_0(h_0b) = J_0(h_0b) - \frac{J_0(h_0a)}{Y_0(h_0a)} Y_0(h_0b) = 0, \quad \text{or,}$$

$$\boxed{J_0(h_0b) Y_0(h_0a) - J_0(h_0a) Y_0(h_0b) = 0}, \quad h_0 = k_0\sqrt{\epsilon_d - 1} \quad (10.19.27)$$

For given dimensions a, b , this determines the maximum operating frequency of the lowest TM mode, or, conversely, given a frequency f and inner radius a , it determines the largest radius b such that operation is restricted to the lowest mode. The MATLAB function, gcut, constructs this function, which may then be passed into fzero, to determine the cutoff frequency or cutoff radius b , or, into the function plot as an aid for choosing the initial search points for fzero,

```
G = gcut(a,b,ed,f) % Goubau line cutoff
```

```
% a,b = inner, outer radii [meters]
% ed = relative permittivity
% f = frequency [Hz]
%
% constructs the function:
% G = J0(h0*b).*Y0(h0*a) - J0(h0*a).*Y0(h0*b);
% vectorized either in b, or, in f
%
% to be used with FZERO to find cutoff frequency or cutoff radius b:
%
% fc = fzero(@(f) gcut(a,b,ed,f), f0); % search near f0
% bc = fzero(@(b) gcut(a,b,ed,f), b0); % search near b0
%
% a convenient initial search point is the lowest cutoff
% of the planar version: f0*d0 = c0/sqrt(ed-1)/2, b0 = a+d0
%
% the function can also be used to plot G vs. f or vs.b
```

Example 10.19.1: In order to get an idea of what the fields look like inside and outside the dielectric, consider an unrealistic example that has too large radii, $a = 0.9$ cm, $b = 1$ cm, to be of practical importance in THz applications. The dielectric permittivity is taken to be $\epsilon_d = 2.54$ (e.g., polystyrene rexolite), and the frequency, $f = 100$ GHz. Fig. 10.19.2 shows the longitudinal electric and magnetic fields, $E_z(\rho)$, $H_\phi(\rho)$ as functions of the radial distance ρ . Note that E_z vanishes at $\rho = a$, and both are normalized to unity at $\rho = b$. The following MATLAB code generates the graph of E_z :

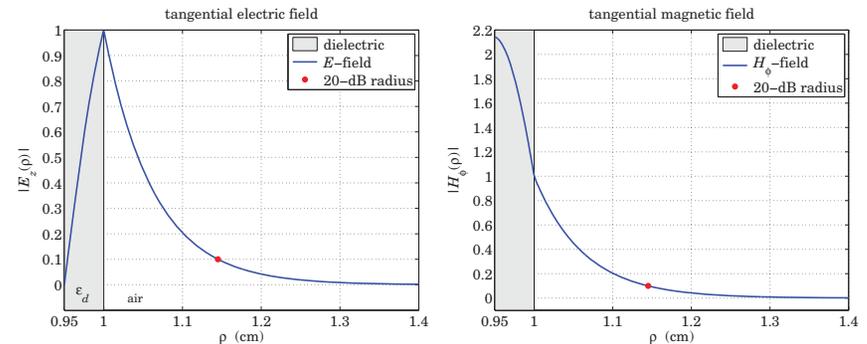


Fig. 10.19.2 Tangential electric and magnetic fields.

```
ed = 2.54;
a = 0.9e-2; b = 1.0e-2; % meters
f = 100e9; % 100 GHz
rel = 0.5; % relaxation parameter

[be,ga,h,N,E,pd] = goubau(a,b,ed,f,rel); % uses default tol=1e-10

A = besselj(0,h*a)/bessely(0,h*a);
Z0 = @(r) besselj(0,h*r) - A*bessely(0,h*r);
```

```

Ezd = @(r) abs(Z0(r)/Z0(b));           % Ez in dielectric
Eza = @(r) abs(besselk(0,ga*r)/besselk(0,ga*b)); % Ez in air

r20 = fzero(@(r) Eza(r)-1/10, b) * 100; % 20-dB radius in cm

r = linspace(a, 1.4*b, 400);          % plot range, a <= rho <= 1.4*b

Ez = Ezd(r).*(r<b) + Eza(r).*(r>=b);

figure; fill([0.9,0.9,1,1],[-0.1,1,1,-0.1], [0.9 0.9 0.9]); hold;
plot(r*100,Ez,'b-', 'linewidth',2);
plot(r20,1/10,'r.', 'markersize',22);
xlabel('\rho (cm)'); ylabel('\hat{E}_z(\rho)')

```

We observe how the electric field is concentrated in the vicinity of the dielectric surface at $\rho = b$. The percentage of the transmitted power residing in the dielectric coating is $100 p_d = 64.63\%$. A measure of the localization of the wave near the surface is the 20-dB radius defined as that distance ρ at which the E_z field in the air has dropped to one-tenth its value at the surface (i.e. by 20 dB). It is obtained by solving for ρ the equation,

$$\left| \frac{E_z(\rho)}{E_z(b)} \right| = \left| \frac{K_0(\gamma\rho)}{K_0(\gamma b)} \right| = \frac{1}{10} \quad (10.19.28)$$

We find, $\rho = 1.1451$ cm, that is, only 1.45 mm from the dielectric surface. The 20-dB radial distance for the magnetic field is found in a similar fashion by solving the following equation, whose solution is $\rho = 1.1449$ cm,

$$\left| \frac{H_\phi(\rho)}{H_\phi(b)} \right| = \left| \frac{K_1(\gamma\rho)}{K_1(\gamma b)} \right| = \frac{1}{10} \quad (10.19.29)$$

The computed wavenumbers were, $[\beta, \gamma, h] = [26.0121, 15.4070, 20.9542]$ cm^{-1} , the number of iterations of the iterative algorithm was $N = 12$, and the computational error, $E = 1.11 \times 10^{-15}$, with an assumed tolerance of, $\text{tol} = 10^{-10}$. \square

Example 10.19.2: This example is from King-Wiltse [994]. Consider a copper wire of radius, $a = 0.995$ cm, coated with a polystyrene dielectric of permittivity, $\epsilon_d = 2.54$, and thickness, $d = b - a = 0.005$ cm, so that the outer radius is, $b = a + d = 1$ cm. The conductivity of copper is $\sigma = 5.75 \times 10^7$ siemens/m. The frequency is varied over, $10 \leq f \leq 1000$ GHz.

The MATLAB code segment below, calculates the propagation wavenumber β , as well as the phase and group velocities, $v_{ph} = \omega/\beta$, $v_g = d\omega/d\beta$. The graphs are shown in Figs. 10.19.3 and 10.19.4.

```

c0 = 299792458;           % NIST value, m/sec, speed of light

ff = logspace(1,3,201); % log-spacing in GHz, 10^1 <= ff <= 10^3
f = 1e9*ff;              % Hz
w = 2*pi*f;

ed = 2.54;               % polystyrene
sig = 5.75e7;           % copper conductivity
a = 0.995e-2;          % meters
b = 1.000e-2;          % d = b-a = 0.005e-2

```

```

tol = 1e-10; rel = 0.3; % tolerance and relaxation parameters

[be,ga,h,N,E] = goubau(a,b,ed,f,rel,tol); % coated

figure; semilogx(ff,be/1000,'b-') % beta in rads/mm

vph = w./be; % phase velocity
figure; semilogx(ff, vph/c0, 'b-')

vg = diff(w)./diff(be); % group velocity
figure; semilogx(ff(1:end-1), vg/c0, 'b-')

```

For comparison, the uncoated Sommerfeld wire of the same radius, $b = 1$ cm, is calculated by the following code. In addition, the corresponding 20-dB radii of the coated and uncoated cases are computed and plotted in Fig. 10.19.5, in which the right graph shows an expanded view of the frequency sub-range, $500 \leq f \leq 1000$ GHz.

```

[bu,gau] = sommer(b,f,sig); % uncoated case, Sommerfeld wire

for i=1:length(f), % 20-dB radii
    Ezc = @(r) abs(besselk(0,ga(i)*r)/besselk(0,ga(i)*b));
    Ezu = @(r) abs(besselh(0,1,gau(i)*r)/besselh(0,1,gau(i)*a));

    rc(i) = fzero(@(r) Ezc(r)-1/10, b) * 1000; % coated
    ru(i) = fzero(@(r) Ezu(r)-1/10, b) * 1000; % uncoated
end

figure; loglog(ff,rc,'b-', ff,ru,'r--'); % coated and uncoated 20-dB radii

ffs = ff(ff>100); rcs = rc(ff>100); % 500 <= f <= 1000 GHz sub-range
figure; plot(ffs, rcs, 'b-'); % plot over sub-range

```

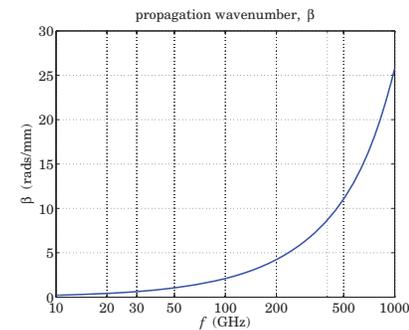


Fig. 10.19.3 Propagation wavenumber, β .

Fig. 10.19.6 shows the attenuation coefficients for the coated and uncoated cases, computed and plotted by the following MATLAB code. The King-Wiltse approximation was restricted to the range $f > 100$ GHz in order to more closely compare our results to those of [994]. The right graph shows separately the attenuations due to the conductor and the dielectric losses. The loss tangent for the dielectric was taken to be $\tan \theta = 0.001$.

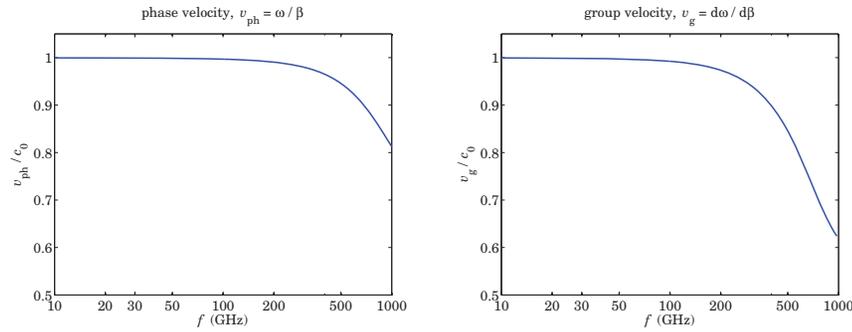


Fig. 10.19.4 Phase and group velocities in units of c_0 .

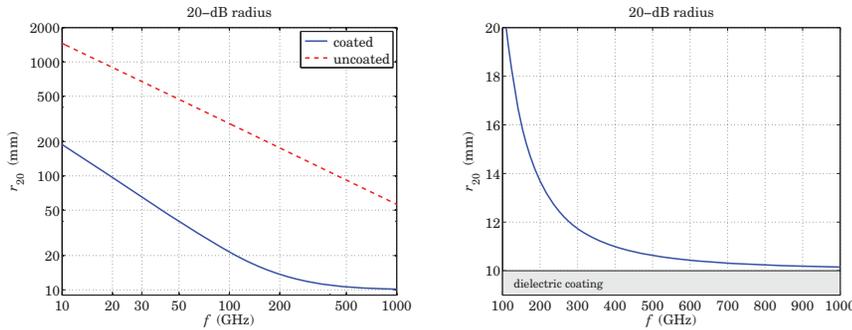


Fig. 10.19.5 Lateral 20-dB radii.

$\text{tanh} = 1e-3;$

`[atot,ac,ad,app] = goubatt(a,b,ed,f,be,sig,tanh);`

```
db100ft = 8.68589*30.48*atot; % convert to dB/100ft
dbc = 8.68589*30.48*ac;
dbd = 8.68589*30.48*ad;
dbapp = 8.68589*30.48*app;
f1 = ff(ff>=100);
db1 = dbapp(ff>=100); % restrict King-Wiltse to f>100 GHz

dBu = -8.68589*30.48*imag(bu); % uncoated case
```

```
figure; loglog(ff,db100ft,'b-', ff,dBu,'r--', f1,db1,'k:'); % coated
figure; loglog(ff,dbc,'b-', ff,dbd,'r--', ff,db100ft,'k:'); % uncoated
```

We observe that the coated case has substantially narrower lateral confinement, as quantified by the 20dB radii, than the uncoated case, but at the expense of much larger attenuation. However, in THz applications, lateral confinement is more important than attenuation, because the relevant propagation distances are short, i.e., centimeters. □

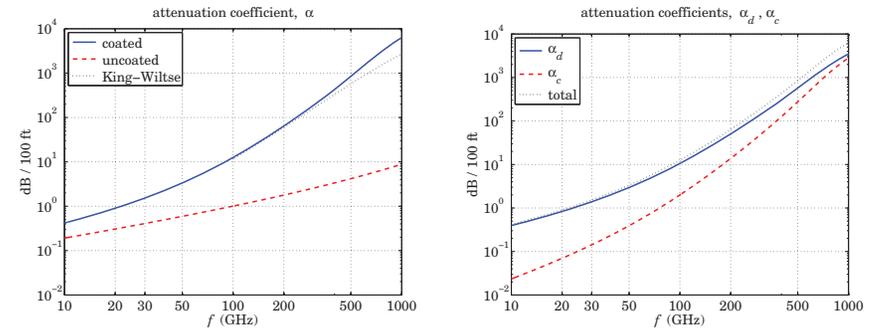


Fig. 10.19.6 Attenuation coefficient.

Example 10.19.3: Here, we explore the properties of the Goubau line over the THz frequency range, $0.1 \leq f \leq 10$ THz. Our choice of parameters a, b, ϵ_d , and results are similar to those of Refs. [1014,1027]. We choose $a = 100 \mu\text{m}$, and two values of the permittivity of the dielectric coating, $\epsilon_d = 2.54$, corresponding to a polymer, and a higher value, $\epsilon_d = 9$. For each ϵ_d , we compare two coating thicknesses d_1, d_2 , one thin and one larger near the cutoff thickness determined by (10.19.27),

$$d_1 = b_1 - a = 10 \mu\text{m}, \quad \text{for } \epsilon_d = 2.54 \text{ and } \epsilon_d = 9$$

$$d_2 = b_2 - a = 100 \mu\text{m}, \quad \text{for } \epsilon_d = 2.54$$

$$d_2 = b_2 - a = 50 \mu\text{m}, \quad \text{for } \epsilon_d = 9$$

For the computation of the attenuations, we will assume a copper inner conductor of conductivity $\sigma = 5.8 \times 10^7$ siemens/m, and a loss tangent of $\tan \theta = 0.001$ for both types of dielectric coatings.

Figs. 10.19.7 and (10.19.8) show the computed propagation wavenumber β , and the corresponding phase velocity $v_{\text{ph}} = \omega/\beta$.

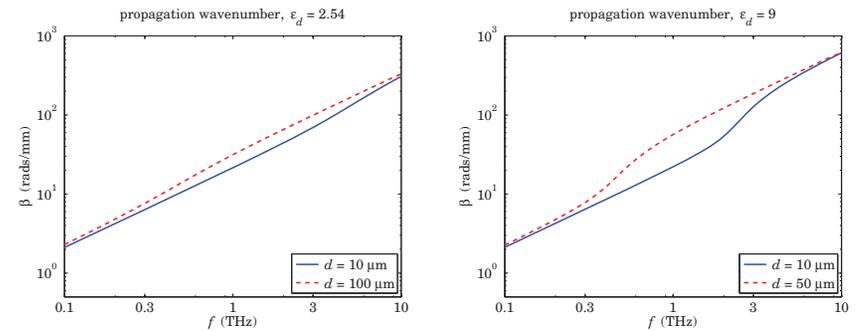


Fig. 10.19.7 Propagation wavenumber.

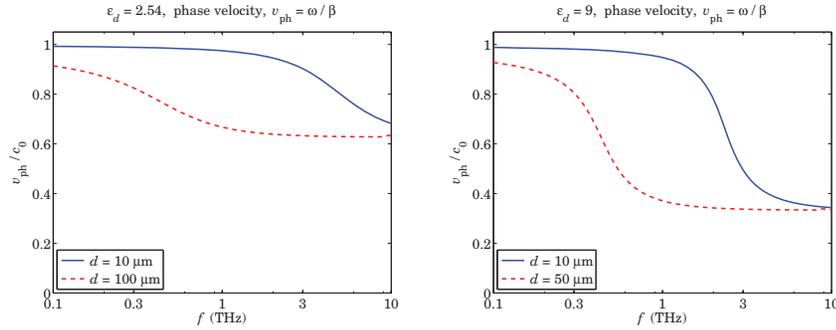


Fig. 10.19.8 Phase velocity.

We note that $\beta \approx k_0$ for low frequencies, switching gradually to $\beta \approx k_0\sqrt{\epsilon_d}$ at higher frequencies. This can be seen more clearly from the phase velocities, switching from near c_0 to near $c_0/\sqrt{\epsilon_d}$. Over the relatively flat parts of the phase velocity, the group velocity will also be flat and equal to the phase velocity, implying that there will be very little dispersion. The flat parts are wider for the thicker coatings and cover the 1-10 THz range. Indeed, from the definitions of the group and phase velocities, one can show the relationship,

$$v_{ph} = \frac{\omega}{\beta}, \quad v_g = \frac{d\omega}{d\beta} \Rightarrow v_g = \frac{v_{ph}}{1 - \beta v'_{ph}} \quad (10.19.30)$$

where $v'_{ph} = dv_{ph}/d\omega$, so that if v_{ph} is constant over a band of ω s, then so is v_g . The MATLAB code for computing β and generating Figs. 10.19.7 and (10.19.8) is as follows,

```

c0 = 299792458;           % NIST value, m/sec, speed of light

ff = logspace(-1,1,100); % log-spaced in THz, 0.1 <= ff <= 10
f = 1e12*ff; w = 2*pi*f; % f in Hz

a = 100e-6;              % 100 microns
ed = 2.54;               % uncomment this line for ed = 9
% ed = 9;

rel = 0.001;            % relaxation parameter

d1 = 10e-6;              % coating thickness
if ed==2.54, d2 = 100e-6; end
if ed==9.00, d2 = 50e-6; end

b1 = a+d1;
[be1,ga1,h1,N1,E1,Pd1] = goubau(a,b1,ed,f,rel); % uses tol=1e-10

b2 = a+d2;
[be2,ga2,h2,N2,E2,Pd2] = goubau(a,b2,ed,f,rel);

figure; loglog(ff,be1./1000,'b-', ff,be2./1000,'r--'); % rads/mm

v1 = w./be1/c0; v2 = w./be2/c0; % phase velocities
    
```

```
figure; semilogx(ff, v1, 'b-', ff, v2, 'r--');
```

The percentage of power in the dielectric coating is shown in Fig. 10.19.9. We note that for the thicker coating almost 100% of transmitted power is contained in the coating, over the 0.5-10 THz range. The graphs were generated by the code,

```
figure; semilogx(ff, 100*Pd1,'b-', ff, 100*Pd2,'r--');
```

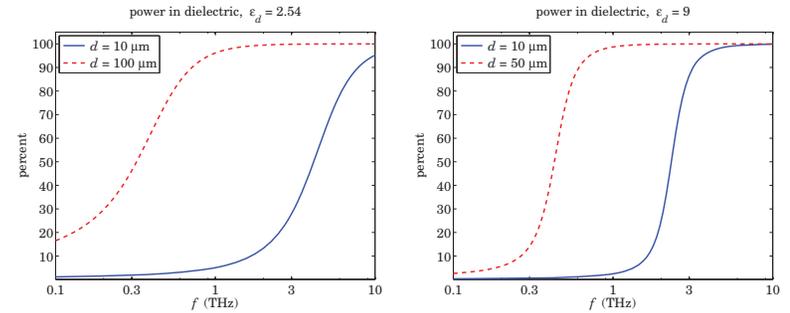


Fig. 10.19.9 Percentage of power in dielectric coating.

The attenuations in dB/cm were computed by the following code and plotted in Fig. 10.19.10.

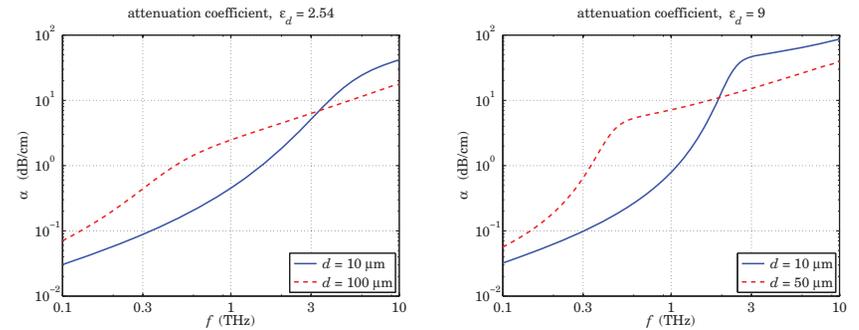


Fig. 10.19.10 Attenuation coefficient in dB/cm.

```

sigma = 5.8e7;           % copper conductivity
tand = 1e-3;            % loss tangent

at1 = goubatt(a,b1,ed,f,be1,sigma,tand);
at2 = goubatt(a,b2,ed,f,be2,sigma,tand);

dbcm = 20*log10(exp(1))/100; % convert to dB/cm
figure; loglog(ff, at1*dbcm,'b-', ff, at2*dbcm,'r--')
    
```

The attenuations are high, but still acceptable for propagation over short distances of the order of centimeters. The above frequency plots do not necessarily imply single-mode operation, because frequency range may exceed the cutoff frequencies of the lowest TM mode. The calculated cutoff frequencies with the help of the function, `gcut`, are as follows,

	$\epsilon_d = 2.54$	$\epsilon_d = 9$
thickness d_1	$f_c = 12.0776$ THz	$f_c = 5.2990$ THz
thickness d_2	$f_c = 1.2008$ THz	$f_c = 1.0577$ THz

Single-mode operation is below these frequencies. They were computed by the following MATLAB code in which Eq. (10.20.6) was used as the initial search point for `fzero`,

```

b1 = a + d1;
f10 = c0/2/sqrt(ed-1)/d1;
fc1 = fzero(@(f) gcut(a,b1,ed,f), f10)

b2 = a + d2;
f20 = c0/2/sqrt(ed-1)/d2;
fc2 = fzero(@(f) gcut(a,b2,ed,f), f20)
    
```

For the remainder of this example, let us fix the frequency to $f = 10^{12}$ Hz, or, 1 THz. Figures 10.19.11 and 10.19.12 show the dependence of the wavenumber β and phase velocity v_{ph} on the coating thickness. As the thickness increases, and more of the wave resides in the coating, the phase velocity tends to that of the dielectric, i.e., $c_0/\sqrt{\epsilon_d}$.

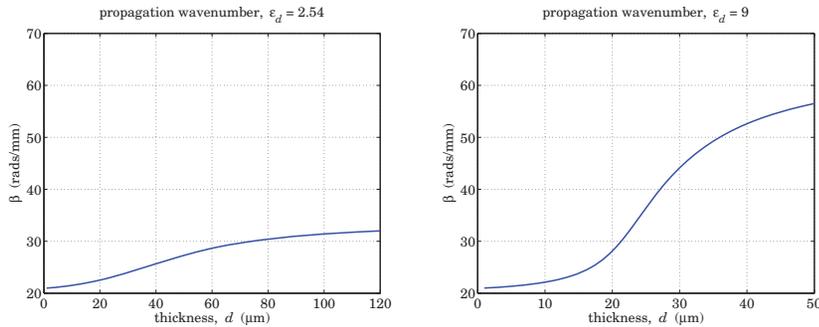


Fig. 10.19.11 Propagation wavenumbers vs. thickness at 1 THz.

The MATLAB code for generating these graphs is as follows. The power residing in the coating and the attenuations are also computed and plotted in Figs. 10.19.14 and 10.19.13. The function `goubau` is vectorized in the frequency variable f , but not in the thickness b . Therefore, the solution for each thickness must be determined with a loop,

```

ff0 = 1; f0 = 1e12*ff0; w = 2*pi*f0;
sigma = 5.8e7; tand = 1e-3;

if ed==2.54, dd = linspace(1,120,120); end % dd in microns
if ed==9.00, dd = linspace(1,50,50); end
    
```

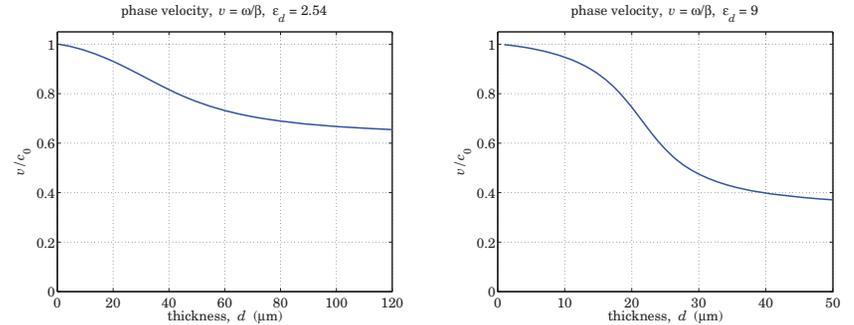


Fig. 10.19.12 Phase velocity vs. thickness at 1 THz.

```

d = dd*1e-6; b = a+d; % d,b in meters

for i=1:length(d)
[be(i),~,~,~,Pd(i)] = goubau(a,b(i),ed,f0,rel);
[atot(i),ac(i),ad(i)] = goubatt(a,b(i),ed,f0,be(i),sigma,tand);
end

v = w./be/c0; % v in units of c0
cm = 20*log10(exp(1))/100; % conversion to dB/cm

figure; plot(dd, be/1000, 'b-') % be in units of rads/mm
figure; plot(dd, v, 'b-')
figure; plot(dd,100*Pd, 'r-')
figure; plot(dd,atot*cm, 'b-', dd,ac*cm, 'g--', dd,ad*cm, 'r--')
    
```

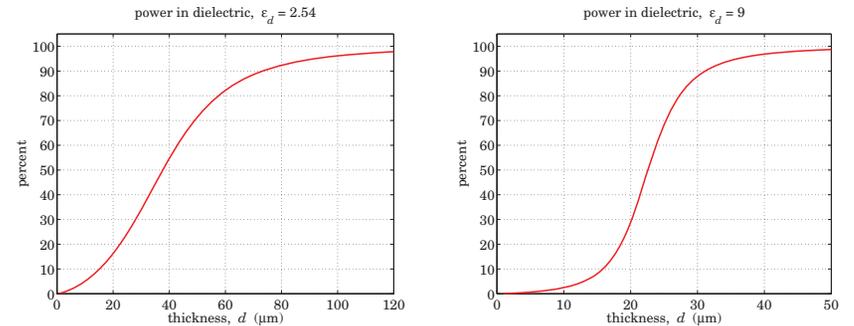


Fig. 10.19.13 Power in dielectric vs. thickness at 1 THz.

The thickness range for each ϵ_d was determined by finding the maximum thickness from the cutoff condition (10.19.27). For the given a, f , the MATLAB function `gcut` can be used to find the cutoff width, as well as to plot the cutoff function (10.19.31) versus b , as shown in Fig. 10.19.15,

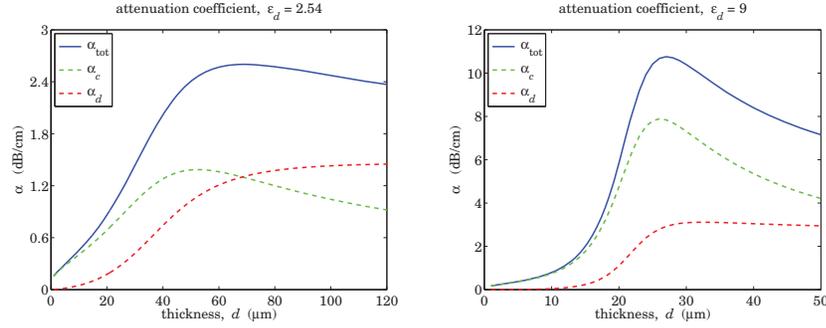


Fig. 10.19.14 Attenuation vs. thickness at 1 THz.

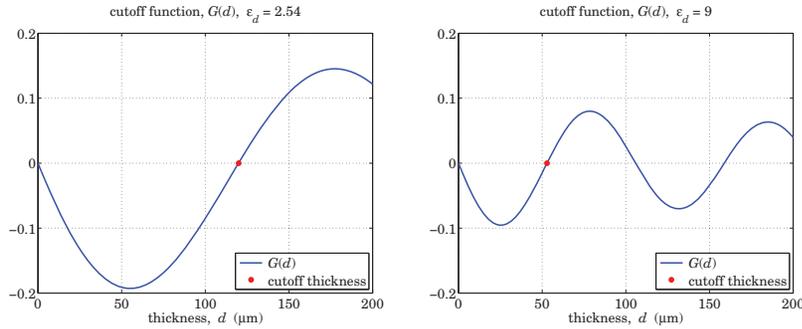


Fig. 10.19.15 Cutoff thickness at 1 THz.

$$G(b) = J_0(h_0 b) Y_0(h_0 a) - J_0(h_0 a) Y_0(h_0 b) \quad (10.19.31)$$

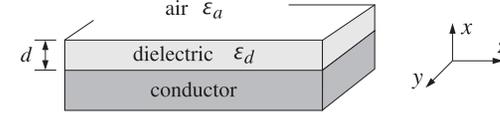
The computed values of the cutoff thicknesses were $d_c = 52.88 \mu\text{m}$ for $\epsilon_d = 2.54$, and $d_c = 119.88 \mu\text{m}$ for $\epsilon_d = 9$. The MATLAB code is given below,

```
f0 = 1e12; a=100e-6;
d0 = c0/2/sqrt(ed-1)/f0; % see Eq. (10.20.6)
b0 = a+d0; % initial search point for fzero
dc = fzero(@(b) gcut(a,b,ed,f0), b0) - a; % cutoff thickness
dd = linspace(0,200,201); d = dd * 1e-6; % dd in microns, d in meters
b = a + d;
figure; plot(dd,gcut(a,b,ed,f0),'b-', 'dc*1e6, 0, 'r.');
```

In conclusion, for THz applications, one can find a range of parameters a, b, ϵ_d of the Goubau line that allows single-mode operation, high lateral confinement, acceptable attenuation, and low dispersion. \square

10.20 Planar Limit of the Goubau Line

The planar limit of the Goubau line, shown below, consists of a planar conductor with a thin dielectric coating of thickness, say, d . It may be thought of as the limit of the Goubau line as the radii a, b tend to infinity, such that the difference, $d = b - a$, remains finite. It was originally studied by Attwood [988] as a simplified version of the Goubau line. The conductor is assumed to be perfect and the dielectric, lossless.



To see how this evolves from the Goubau solution, let us rewrite Eq. (10.19.11) in a normalized way by making the substitutions for the constants E_0, E_1 ,

$$E_0 \rightarrow \frac{E_0}{Z_0(hb)}, \quad E_1 \rightarrow \frac{E_0}{K_0(\gamma b)}$$

$a \leq \rho \leq b$	$b \leq \rho < \infty$
$E_z = E_0 \frac{Z_0(h\rho)}{Z_0(hb)}$	$E_z = E_0 \frac{K_0(\gamma\rho)}{K_0(\gamma b)}$
$E_\rho = E_0 \frac{j\beta}{h} \frac{Z_1(h\rho)}{Z_0(hb)}$	$E_\rho = -E_0 \frac{j\beta}{\gamma} \frac{K_1(\gamma\rho)}{K_0(\gamma b)}$
$H_\phi = E_0 \frac{jk_0\epsilon_d}{\eta_0 h} \frac{Z_1(h\rho)}{Z_0(hb)}$	$H_\phi = -E_0 \frac{jk_0\epsilon_d}{\eta_0 \gamma} \frac{K_1(\gamma\rho)}{K_0(\gamma b)}$
$h = \sqrt{k_0^2 \epsilon_d - \beta^2}$	$\gamma = \sqrt{\beta^2 - k_0^2 \epsilon_a}$

(10.20.1)

The boundary condition at $\rho = b$ is automatically satisfied for E_z , whereas that for H_ϕ is equivalent to the characteristic equation (10.19.14). Following the same limiting procedure as in Sec. 10.17, we set, $\rho = a + x$, and, $b = a + d$, and take the limit $a \rightarrow \infty$ with x, d kept fixed. In this limit, we may replace the Bessel functions by their asymptotic forms [1790], for $n = 0, 1$,

$$J_n(u) = \sqrt{\frac{2}{\pi u}} \cos\left(u - \frac{\pi}{4} - \frac{n\pi}{2}\right)$$

$$Y_n(u) = \sqrt{\frac{2}{\pi u}} \sin\left(u - \frac{\pi}{4} - \frac{n\pi}{2}\right) \quad (10.20.2)$$

$$K_n(u) = \sqrt{\frac{\pi}{2u}} e^{-u}$$

Using (10.20.2) and some trigonometric identities, we find the following asymptotic forms, and limits as $a \rightarrow \infty$,

$$\begin{aligned} \frac{Z_0(h\rho)}{Z_0(hb)} &= \sqrt{\frac{b}{\rho}} \cdot \frac{\sin(h(\rho-a))}{\sin(h(b-a))} = \sqrt{\frac{a+d}{a+x}} \cdot \frac{\sin(hx)}{\sin(hd)} \rightarrow \frac{\sin(hx)}{\sin(hd)} \\ \frac{Z_1(h\rho)}{Z_0(hb)} &= -\sqrt{\frac{b}{\rho}} \cdot \frac{\cos(h(\rho-a))}{\sin(h(b-a))} = -\sqrt{\frac{a+d}{a+x}} \cdot \frac{\cos(hx)}{\sin(hd)} \rightarrow -\frac{\cos(hx)}{\sin(hd)} \\ \frac{K_0(\gamma\rho)}{K_0(\gamma b)} &= \frac{K_1(\gamma\rho)}{K_0(\gamma b)} = \sqrt{\frac{b}{\rho}} \cdot e^{-\gamma(\rho-b)} = \sqrt{\frac{a+d}{a+x}} \cdot e^{-\gamma(x-d)} \rightarrow e^{-\gamma(x-d)} \end{aligned} \quad (10.20.3)$$

For example, we have for $n = 0, 1$,

$$\begin{aligned} Z_n(h\rho) &= J_n(h\rho) - \frac{J_0(ha)}{Y_0(ha)} Y_0(h\rho) \\ &= \sqrt{\frac{2}{\pi h\rho}} \left[\cos\left(h\rho - \frac{\pi}{4} - \frac{n\pi}{2}\right) - \frac{\cos(ha - \pi/4)}{\sin(ha - \pi/4)} \sin\left(h\rho - \frac{\pi}{4} - \frac{n\pi}{2}\right) \right] \\ &= \sqrt{\frac{2}{\pi h\rho}} \cdot \frac{-\sin(h(\rho-a) - n\pi/2)}{\sin(ha - \pi/4)} \end{aligned}$$

The characteristic equation (10.19.14) simplifies as follows, at $\rho = b$, or, $x = d$,

$$\frac{h}{\varepsilon_d} \frac{Z_0(hb)}{Z_1(hb)} = -\frac{\gamma}{\varepsilon_a} \frac{K_0(\gamma b)}{K_1(\gamma b)} \Rightarrow -\frac{h}{\varepsilon_d} \frac{\sin(hd)}{\cos(hd)} = -\frac{\gamma}{\varepsilon_a}, \quad \text{or,} \quad \boxed{\frac{h}{\varepsilon_d} \tan(hd) = \frac{\gamma}{\varepsilon_a}} \quad (10.20.4)$$

As shown by King-Wiltse [994], this simplified form of the characteristic equation can be used as a substitute of (10.19.14) at higher frequencies, near THz. Similarly, the field solutions (10.20.1) become,

$0 \leq x \leq d$	$d \leq x < \infty$
$E_z = E_0 \frac{\sin(hx)}{\sin(hd)}$	$E_z = E_0 e^{-\gamma(x-d)}$
$E_x = -E_0 \frac{j\beta}{h} \frac{\cos(hx)}{\sin(hd)}$	$E_x = -E_0 \frac{j\beta}{\gamma} e^{-\gamma(x-d)}$
$H_y = -E_0 \frac{jk_0 \varepsilon_d}{\eta_0 h} \frac{\cos(hx)}{\sin(hd)}$	$H_y = -E_0 \frac{jk_0 \varepsilon_a}{\eta_0 \gamma} e^{-\gamma(x-d)}$
$h = \sqrt{k_0^2 \varepsilon_d - \beta^2}$	$\gamma = \sqrt{\beta^2 - k_0^2 \varepsilon_a}$

(10.20.5)

where they must be multiplied by $e^{j\omega t - j\beta z}$. The tangential component E_z vanishes on the (assumed perfect) conductor at $x = 0$, and it matches the air side at $x = d$. The matching of H_y at $x = d$ is equivalent to Eq. (10.20.4).

As in the case of the Goubau line, the cutoff condition is $\gamma = 0$, which implies through (10.20.4) that $\tan(h_0 d) = 0$, $h_0 = k_0 \sqrt{\varepsilon_d - 1}$. The lowest cutoff corresponds to $h_0 d = \pi$, or, expressed as a frequency-thickness relation,

$$\boxed{fd = \frac{c_0}{2\sqrt{\varepsilon_d - 1}}} \quad (10.20.6)$$

The characteristic equation (10.20.4) can be solved recursively using the same iterative procedure as for the Goubau line, that is, by writing it in the following form,

$$\frac{\gamma}{h} = \frac{1}{\varepsilon_d} \tan(hd) \equiv F(\beta)$$

with $\varepsilon_a = 1$, and solving the left-hand side for β ,

$$\beta = k_0 \sqrt{\frac{1 + \varepsilon_d F^2(\beta)}{1 + F^2(\beta)}} \quad (10.20.7)$$

and turning it into the recursion, for $n = 0, 1, 2, \dots$,

$$\beta_{n+1} = r k_0 \sqrt{\frac{1 + \varepsilon_d F^2(\beta_n)}{1 + F^2(\beta_n)}} + (1-r)\beta_n \quad (10.20.8)$$

where r is a relaxation parameter, $0 < r \leq 1$. The following MATLAB function, `attw`, implements this iteration:

```
[be,g,h,N,E,pd] = attw(d,ed,f,r,to1,be0) % Attwood surface waveguide
[be,g,h,N,E,pd] = attw(d,ed,f,r,to1)   (be0 = 0.999*k0*sqrt(ed))
[be,g,h,N,E,pd] = attw(d,ed,f,r)       (to1 = 1e-10)
```

```
% d = coating thickness [meters]
% ed = relative dielectric constant of coating (ed>1)
% f = vector of frequencies [Hz]
% r = relaxation parameter (0 < r <= 1)
% to1 = computational tolerance, default to1 = 1e-10
% be0 = initializing vector, size(f), default be0 = 0.999*k0*sqrt(ed)
%
% be = propagation wavenumber [rads/m], size(f)
% ga = lateral decay constant [1/m], size(f)
% h = lateral wavenumber in dielectric coating [rads/m], size(f)
% N = number of iterations to converge (for all f)
% E = computational error of characteristic equation, size(f)
% pd = proportion of transmitted power in dielectric coating
```

The attenuation coefficient may be computed by,

$$\alpha = \frac{P'_{\text{loss}}}{2P_T} \quad (10.20.9)$$

where P'_{loss} is the power loss per unit conductor length, and P_T , the transmitted power.

The two parts of P'_{loss} due to the losses in the conductor and to the losses in the dielectric coating are given as follows, relative to a finite strip in the y -direction of width, $\Delta y = 1$ meter,

$$P'_{\text{loss}} = P'_c + P'_d = \frac{1}{2} R_s |H_y|^2 \Big|_{x=0} + \frac{1}{2} \omega \varepsilon_l \int_0^d [|E_z|^2 + |E_x|^2] dx \quad (10.20.10)$$

where, in terms of the loss-tangent of the dielectric, $\epsilon_l = \epsilon_0 \epsilon_l = \epsilon_0 \epsilon_d \tan \theta$. The transmitted power is obtained by integrating the z-component of the Poynting vector on a $dx \cdot \Delta y$ cross-sectional area of the coating and the air side,

$$P_T = \int_0^d \frac{1}{2} \operatorname{Re}[E_x H_y^*] dx + \int_d^\infty \frac{1}{2} \operatorname{Re}[E_x H_y^*] dx \quad (10.20.11)$$

Using the solutions (10.20.5), the indicated integrals can be done easily, resulting in,

$$P'_{\text{loss}} = \frac{|E_0|^2 k_0^2 \epsilon_d^2}{2 \eta_0^2 h^2 \sin^2(hd)} \left\{ R_s + \frac{\eta_0 \tan \theta}{2 k_0 \epsilon_d} \left(\beta^2 \left[d + \frac{\sin(2hd)}{2h} \right] + h^2 \left[d - \frac{\sin(2hd)}{2h} \right] \right) \right\}$$

$$P_T = \frac{|E_0|^2 \beta k_0 \epsilon_d^2}{4 \eta_0 h^2 \sin^2(hd)} \left\{ \frac{1}{\epsilon_d} \left[d + \frac{\sin(2hd)}{2h} \right] + \frac{\cos^2(hd)}{y} \right\}$$

where we used Eq. (10.20.4) and set $\epsilon_a = 1$. Dividing, we obtain the total attenuation coefficient, with the R_s term representing the conductor losses, and the $\tan \theta$ term, the dielectric ones,

$$\alpha = \frac{k_0}{\beta \eta_0} \cdot \frac{R_s + \frac{\eta_0 \tan \theta}{2 k_0 \epsilon_d} \left(\beta^2 \left[d + \frac{\sin(2hd)}{2h} \right] + h^2 \left[d - \frac{\sin(2hd)}{2h} \right] \right)}{\frac{1}{\epsilon_d} \left[d + \frac{\sin(2hd)}{2h} \right] + \frac{\cos^2(hd)}{y}} \quad (10.20.12)$$

It also follows from the expression for P_T that the proportion of the transmitted power within the dielectric is given by,

$$p_d = \frac{\frac{1}{\epsilon_d} \left[d + \frac{\sin(2hd)}{2h} \right]}{\frac{1}{\epsilon_d} \left[d + \frac{\sin(2hd)}{2h} \right] + \frac{\cos^2(hd)}{y}} = \text{power in dielectric} \quad (10.20.13)$$

The following MATLAB function, attwatt, implements Eqs. (10.20.12) and (10.20.13):

```
[atot,ac,ad,pd] = attwatt(d,ed,f,be,sigma,tand)

% d = coating thickness [meters]
% ed = relative dielectric constant of coating (ed>1)
% f = vector of frequencies [Hz]
% be = propagation wavenumber [rads/m], same size as f, obtained from GOUBAU
% sigma = conductivity of inner conductor [siemens/m]
% tand = loss tangent of dielectric coating, scalar, or, size(f)
%
% atot = total attenuation coefficient [nepers/m], size(f)
% ac = attenuation due to conductor [nepers/m], size(f)
% ad = attenuation due to dielectric [nepers/m], size(f)
% pd = proportion of transmitted power in dielectric coating
%
% conversions: dB/m = 8.68589 * atot = 20*log10(exp(1)) * atot
%              dB/100ft = 8.68589 * 30.48 * atot
```

Example 10.20.1: This example is from [988] and demonstrates that the planar version behaves very similarly to the Goubau line with respect to the dependence of the wavenumber β and attenuation α on frequency and coating thickness.

Consider a copper planar conductor with conductivity $\sigma = 5.8 \times 10^7$ siemens/m, covered with a dielectric coating of permittivity $\epsilon_d = 4$. Three coating thicknesses are compared, $d_1 = 1$ mm, $d_2 = 5$ mm, and $d_3 = 10$ mm, over the frequency range, $0.3 \leq f \leq 30$ GHz. Figs. 10.20.1 and 10.20.2 show the propagation and lateral wavenumbers β, γ , as well as the phase velocities, $v_{\text{ph}} = \omega/\beta$, and percentage of power in the dielectric for the three thicknesses.

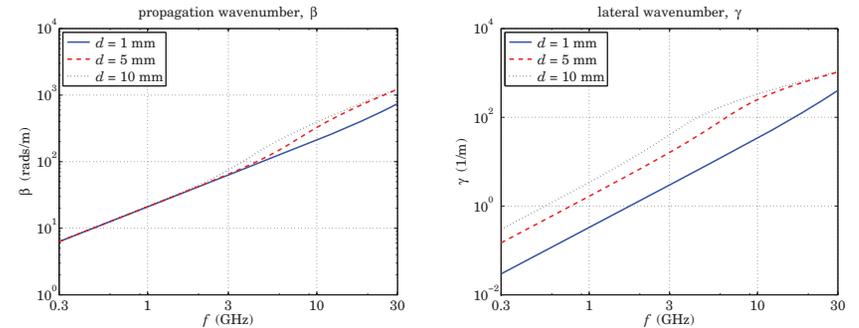


Fig. 10.20.1 Propagation and lateral wavenumbers, β, γ , vs. frequency.

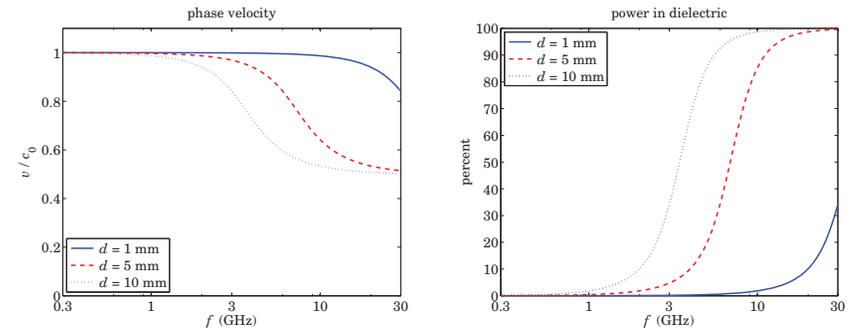


Fig. 10.20.2 Phase velocities and power in dielectric.

As in the Goubau line case, the wavenumber β switches gradually from near $\beta = k_0$ to near $\beta = k_0 \sqrt{\epsilon_d}$. As the coating thickness and frequency increase, the power confinement in the coating increases, while the field strength outside the coating decreases with distance, as measured by the value of the decay constant γ^{-1} . The following MATLAB code generates the graphs,

```
c0 = 299792458;
ed = 4;
```

```

d1 = 1e-3; d2 = 5e-3; d3 = 10e-3; % meters

ff = 3*logspace(-1,1,100); % 0.3 <= ff <= 30 GHz
f = ff * 1e9; % Hz

rel = 0.01; % relaxation parameter

[be1,ga1,h1,N1,E1,pd1] = attw(d1,ed,f,rel);
[be2,ga2,h2,N2,E2,pd2] = attw(d2,ed,f,rel);
[be3,ga3,h3,N3,E3,pd3] = attw(d3,ed,f,rel);

v1 = w./be1/c0; v2 = w./be2/c0; v3 = w./be3/c0;

figure; loglog(ff,be1,'b-', ff,be2,'r--', ff,be3,'k:');
figure; loglog(ff,ga1,'b-', ff,ga2,'r--', ff,ga3,'k:');
figure; semilogx(ff,v1,'b-', ff,v2,'r--',ff,v3,'b:');
figure; semilogx(ff,100*pd1,'b-', ff,100*pd2,'r--', ff,100*pd3,'k:');

```

The attenuations in the conductor and the dielectric are shown in Fig. 10.20.3. The loss tangent in the dielectric was assumed in [988] to arise from a very small conductivity, i.e.,

$$\tan \theta = \frac{\sigma_d}{\omega \epsilon_d} = \frac{\sigma_d}{\omega \epsilon_0 \epsilon_d}, \quad \sigma_d = 6.6667 \times 10^{-4} \text{ siemens/m}$$

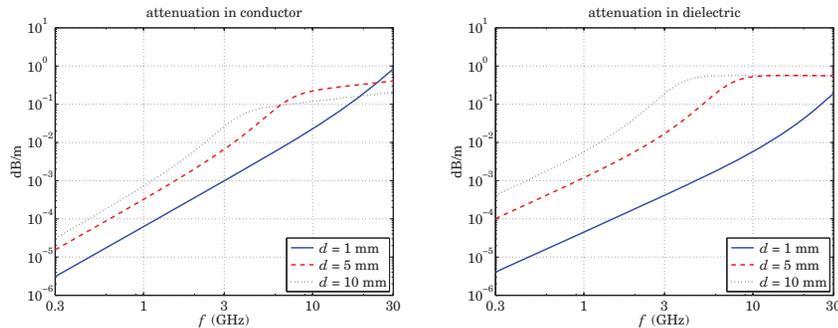


Fig. 10.20.3 Attenuations in conductor and dielectric coating.

The MATLAB code generating the graphs was,

```

ep0 = 8.854187817e-12; % NIST value, vacuum permittivity
sigma = 5.8e7; % copper conductor
sigd = 6.6667e-4; % dielectric

w = 2*pi*f;
tand = sigd./w/ep0/ed;

[at1,ac1,ad1] = attwatt(d1,ed,f,be1,sigma,tand);
[at2,ac2,ad2] = attwatt(d2,ed,f,be2,sigma,tand);
[at3,ac3,ad3] = attwatt(d3,ed,f,be3,sigma,tand);

db = 8.68589; % convert to dB/m

```

```

figure; loglog(ff,db*ac1,'b-', ff,db*ac2,'r--', ff,db*ac3,'k:');
figure; loglog(ff,db*ad1,'b-', ff,db*ad2,'r--', ff,db*ad3,'k:');

```

As expected, the attenuation in the dielectric increases with thickness and frequency. □

10.21 Problems

10.1 Prove Eq. (10.3.8). To do so, introduce the following polar forms,

$$-\epsilon_f = \epsilon_R + j\epsilon_I = |\epsilon_f|e^{j\theta}, \quad -\epsilon_f - \epsilon_c = \epsilon_R - \epsilon_c + j\epsilon_I = |\epsilon_f + \epsilon_c|e^{j\phi}$$

Then, using Eqs. (10.2.6) and some trig identities, show that Eq. (10.3.8) reduces to the equivalent condition,

$$|\epsilon_f|(\sin \theta + \sin(\phi - \theta)) + \epsilon_c(\sin \theta - \sin \phi) = \epsilon_I \left(1 + \frac{\epsilon_c}{|\epsilon_f|}\right)$$

Then, verify this condition by first proving the relationships,

$$\sin(\phi - \theta) = \frac{\epsilon_c \epsilon_I}{|\epsilon_f| \cdot |\epsilon_f + \epsilon_c|}, \quad \sin \theta = \frac{\epsilon_I}{|\epsilon_f|}, \quad \sin \phi = \frac{\epsilon_I}{|\epsilon_f + \epsilon_c|}$$

10.2 Prove Eqs. (10.3.9)–(10.3.11) and thereby prove the equality $v_g = v_{en}$ for the lossless case of surface plasmons along a single metal-dielectric interface.

10.3 Using some hyperbolic trigonometric identities, show that Eq. (10.1.6) for the transverse magnetic field in a plasmonic waveguide can be written in the following equivalent forms, which are commonly found in the literature,

$$H_y(x) = H_1 \cdot \begin{cases} \cosh(y(x-a)) - \frac{p_c \alpha_c}{y} \sinh(y(x-a)), & |x| \leq a \\ e^{-\alpha_c(x-a)}, & x \geq a \\ [\cosh(2ya) + \frac{p_c \alpha_c}{y} \sinh(2ya)] e^{\alpha_c(x+a)}, & x \leq -a \end{cases} \quad (10.21.1)$$

and

$$H_y(x) = H_2 \cdot \begin{cases} \cosh(y(x+a)) + \frac{p_s \alpha_s}{y} \sinh(y(x+a)), & |x| \leq a \\ [\cosh(2ya) + \frac{p_s \alpha_s}{y} \sinh(2ya)] e^{-\alpha_c(x-a)}, & x \geq a \\ e^{\alpha_s(x+a)}, & x \leq -a \end{cases} \quad (10.21.2)$$

and show that the new constants H_1, H_2 are related to H_0 of Eq. (10.1.6) by

$$H_1 = H_0 \cosh(ya + \psi), \quad H_2 = H_0 \cosh(ya - \psi)$$

10.4 Consider a symmetric MDM plasmonic waveguide with film width $2a$ and permittivities ϵ_f and $\epsilon_c = \epsilon_s$, and assume lossless media so that ϵ_f is real positive and ϵ_c , real negative. We saw at the end of Sec. 10.4 that complex modes can exist in the lossless case that are highly damped and carry no net power in the propagation direction. Consider a TM_0 complex mode with $\beta = \beta_R - j\beta_I$, $\gamma = \gamma_R + j\gamma_I$, and $\alpha_c = \alpha_R + j\alpha_I$, satisfying the characteristic equation,

$\tanh(\gamma a) = -p_c \alpha_c / \gamma$, $p_c = \epsilon_f / \epsilon_c$. The magnetic field is given by Eq. (10.1.5) with $\psi = 0$. By integrating the z -component of the Poynting vector over the transverse x -direction, show that up to an overall positive constant, the powers flowing within the dielectric film and metal sides are given by

$$P_f = \frac{\beta_R}{\epsilon_f} \left[\frac{\sinh(2\gamma_R a)}{2\gamma_R} + \frac{\sin(2\gamma_I a)}{2\gamma_I} \right], \quad 2P_c = \frac{\beta_R}{\epsilon_c} \left[\frac{\cosh(2\gamma_R a) + \cos(2\gamma_I a)}{2\alpha_R} \right]$$

Using the characteristic equation, prove that $P_f = -2P_c$, so that the net power is zero, $P = P_f + 2P_c = 0$. *Hint:* Prove and use the following two results, $\alpha_R \alpha_I = \gamma_R \gamma_I$, and

$$\tanh((\gamma_R + j\gamma_I) a) = \frac{\sinh(2\gamma_R a) + j \sin(2\gamma_I a)}{\cosh(2\gamma_R a) + \cos(2\gamma_I a)}$$

- 10.5 Consider a symmetric MDM plasmonic waveguide with film width $2a$ and permittivities ϵ_f for the dielectric film and $\epsilon_c = \epsilon_s$ for the metal sides. Let β, γ, α_c be the propagation parameters for the corresponding symmetric solution satisfying the characteristic equation, $\gamma \tanh(\gamma a) = -p_c \alpha_c$, where $p_c = \epsilon_f / \epsilon_c$. Show that if ϵ_c is changed by a small amount $\Delta\epsilon_c$, that is, $\epsilon_c \rightarrow \epsilon_c + \Delta\epsilon_c$, then the propagation wavenumber changes from β to $\beta + \Delta\beta$, where

$$\Delta\beta = \frac{\gamma}{\beta} \cdot \Delta\gamma, \quad \Delta\gamma = \frac{k_0 p_c \left(\frac{\gamma}{k_0} \right) \left(\frac{\alpha_c^2}{k_0^2 \epsilon_c} + \frac{1}{2} \right)}{(k_0 a) \left(\frac{\gamma^2 \alpha_c}{k_0^3} \right) \left(1 - \frac{p_c^2 \alpha_c^2}{\gamma^2} \right) - p_c (\epsilon_f - \epsilon_c)} \cdot \Delta\epsilon_c \quad (10.21.3)$$

- 10.6 *Computer Experiment: Anomalous Complex Modes.* This problem is based on [952]. Consider an MDM guide that has $\epsilon_c = \epsilon_s = -10$, $\epsilon_f = 1$, and film width $2a = \lambda_0/4$. It corresponds to a lossless case, and therefore the modes will come in pairs, $\pm\beta_R - j\beta_I$, and (except for the even $q = 0$ mode) carry no net power. The following two tables show the computed even and odd modes, where the index q and the \pm signs refer to Eq. (10.9.2), and even and odd correspond to the index $m = 0, 1$ in Eq. (10.9.2).

q	β/k_0 (even modes)	γ/k_0	α_c/k_0
0	1.21340	0.68728j	3.38709
1	-	0.07361 - 3.87013j	0.07127 - 3.99720j
1	+	-0.07361 - 3.87013j	0.07127 + 3.99720j
2	-	0.11711 - 7.93691j	0.11620 - 7.99965j
2	+	-0.11711 - 7.93691j	0.11620 + 7.99965j
3	-	0.12317 - 11.95816j	0.12274 - 11.99990j
3	+	-0.12317 - 11.95816j	0.12274 + 11.99990j

q	β/k_0 (odd modes)	γ/k_0	α_c/k_0
0	-1.50304j	1.80531j	2.78224
1	-	0.10788 - 5.91526j	0.10637 - 5.99917j
1	+	-0.10788 - 5.91526j	0.10637 + 5.99917j
2	-	0.12109 - 9.94970j	0.12048 - 9.99982j
2	+	-0.12109 - 9.94970j	0.12048 + 9.99982j
3	-	0.12441 - 13.96418j	0.12409 - 13.99993j
3	+	-0.12441 - 13.96418j	0.12409 + 13.99993j

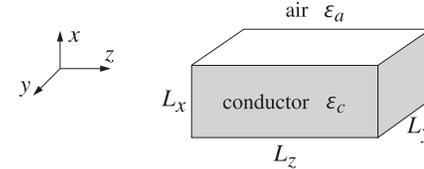
- a. Following Example 10.9.1, write MATLAB code that verifies the entries in the above tables. In addition, calculate the corresponding error norms of the two tables.

- b. For all cases, calculate the net power for each mode, as well as the amount of power and its direction flowing in the metal and the dielectric media. Note that the $q = 0$ even mode is the fundamental G-SPP mode. The $q = 0$ odd mode is truly evanescent in the sense that it has $\beta_R = 0$ and $\beta_I = -\text{Im}(\beta) > 0$, and it is laterally confined since α_c is positive real. This mode carries no power in any of the three media—explain why.

- 10.7 Verify the equality Eq. (10.16.9). As a preliminary step, use the definition, $\gamma_c^2 = k_0^2 \epsilon_c - \beta^2 = k_0^2 (\epsilon_R - j\epsilon_I) - (\beta_R - j\beta_I)^2$, to show the following relationship,

$$2\beta_I \text{Re}[\beta \epsilon_c^*] - |\beta|^2 \epsilon_I = \epsilon_c \text{Im}[\gamma_c^2] + \epsilon_I \gamma_c^2$$

- 10.8 The Zenneck surface wave was discussed in Sec. 10.17 with field solutions given by Eqs. (10.17.3). Consider a rectangular volume, $L_x L_y L_z$, within the conductor whose top side coincides with the interface with the air side, as shown below.



You may assume $L_y = 1$ meter. Following a similar discussion as that carried out for the Sommerfeld wire, show that the power flowing into this volume from all sides is given by,

$$\mathcal{P}_{\text{in}} = \omega \frac{1}{2} \epsilon_0 |E_0|^2 (1 - e^{-2\beta_I L_z}) (1 - e^{-2\alpha_c L_x}) \left[\frac{\text{Re}[\gamma_c \epsilon_c^*]}{2\beta_I \cdot |\gamma_c|^2} + \frac{\text{Re}[\beta \epsilon_c^*]}{2\alpha_c \cdot |\gamma_c|^2} \right]$$

where $\alpha_c = -\text{Im}(\gamma_c) > 0$. Similarly, show that the power dissipated into Ohmic losses within the $L_x L_y L_z$ volume is given by,

$$\mathcal{P}_{\text{loss}} = \omega \frac{1}{2} \epsilon_0 |E_0|^2 (1 - e^{-2\beta_I L_z}) (1 - e^{-2\alpha_c L_x}) \left[\frac{\epsilon_I |\beta|^2 + \epsilon_I |\gamma_c|^2}{2\beta_I \cdot 2\alpha_c \cdot |\gamma_c|^2} \right]$$

where $\epsilon_I = -\text{Im}(\epsilon_c)$. Finally, using the relationship $\gamma_c^2 = k_0^2 \epsilon_c - \beta^2$, show that $\mathcal{P}_{\text{in}} = \mathcal{P}_{\text{loss}}$.

- 10.9 The planar limit of the Goubau line discussed in Sec. 10.20 is obtained in the limit $a, b \rightarrow \infty$, such that the thickness, $d = b - a$, is kept constant. Show that the attenuation coefficient (10.20.12) of the planar case is obtained as the limit of the Goubau line attenuation (10.19.24). To do so, use the asymptotic forms (10.20.2) inside the integrands of Eqs. (10.19.22) to show the following limiting forms for the U, V, W ,

$$U \approx \frac{1}{a} \cdot \left[d + \frac{\sin(2hd)}{2h} \right], \quad V \approx \frac{1}{a} \cdot \left[d - \frac{\sin(2hd)}{2h} \right], \quad W \approx \frac{b}{a^2} \cdot \frac{\cos^2(hd)}{y}$$

Then, use these into (10.19.24) and apply the planar limit.