LECTURE NOTES

OF

ANTENNAS

Mário G. Silveirinha

Lisboa, May 2020

Index

1. Radiating systems

1.1 Fields radiated by a generic current distribution

Consider a generic source (antenna) described by the electric current density $\underline{\mathbf{j}}(\mathbf{r}')$. The electromagnetic fields emitted by the antenna are determined by a vector potential $\underline{\mathbf{A}}(\mathbf{r}')$:

$$\boxed{\underline{\mathbf{H}} = \frac{1}{\mu_0}\nabla\times\underline{\mathbf{A}}}, \qquad \boxed{\underline{\mathbf{E}} = \frac{1}{j\omega\varepsilon_0}\left(\nabla\times\underline{\mathbf{H}} - \underline{\mathbf{j}}\right)}. \tag{1.1}$$

The vector potential is given by:

$$\boxed{\underline{\mathbf{A}}(\mathbf{r}) = \mu_0\iiint dV'\,\underline{\mathbf{j}}(\mathbf{r}')\frac{e^{-jk_0|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}}. \tag{1.2}$$

The function $\Phi_0 = \dfrac{e^{-jk_0|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|}$ is the fundamental solution of the Helmholtz's equation:

$$\nabla^2\Phi_0 + k_0^2\Phi_0 = -\delta(\mathbf{r}-\mathbf{r}'). \tag{1.3}$$

Hence, the vector potential satisfies:

$$\nabla^2\underline{\mathbf{A}} + k_0^2\underline{\mathbf{A}} = -\mu_0\underline{\mathbf{j}}. \tag{1.4}$$

Using the vector identity $\nabla\times\nabla\times = \nabla\nabla\cdot - \nabla^2$, the electric field can be written explicitly in terms of the vector potential as follows:

$$\underline{\mathbf{E}} = \frac{1}{j\omega\varepsilon_0\mu_0}\left(\nabla\nabla\cdot\underline{\mathbf{A}} + k_0^2\underline{\mathbf{A}}\right) - \frac{1}{j\omega\varepsilon_0}\underline{\mathbf{j}}. \tag{1.5}$$

1.2 Far-field region

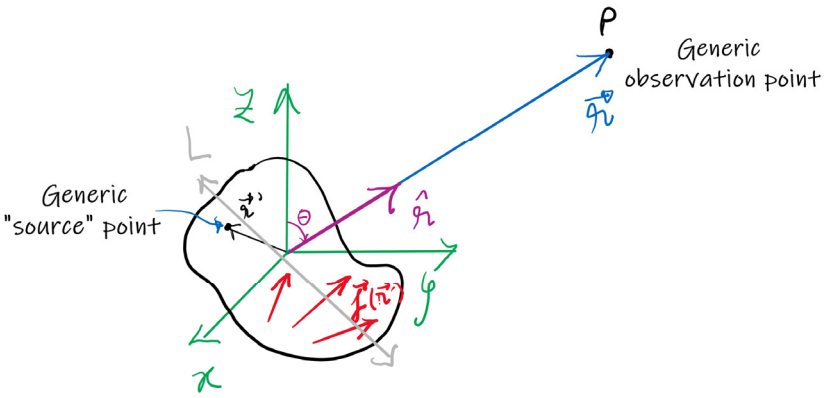
Suppose that the antenna is localized near the origin and is confined within some region with largest dimension L, so that $r' < L$. For observation points far from the source ($r >> L$) one can write:

$$\left|\mathbf{r}-\mathbf{r}'\right|^2 = r^2 + r'^2 - 2\mathbf{r}\cdot\mathbf{r}' \approx r^2\left(1-2\frac{1}{r^2}\mathbf{r}\cdot\mathbf{r}'\right). \tag{1.6}$$

Note that the term $r'^2/r^2 \le L^2/r^2$ is negligible as compared to the other two terms. Using the Taylor series $\sqrt{1+x} \approx 1+x/2$, one finds that:

$$\left|\mathbf{r}-\mathbf{r}'\right| \approx r\left(1-\frac{1}{r^2}\mathbf{r}\cdot\mathbf{r}'\right) = r - \hat{\mathbf{r}}\cdot\mathbf{r}'. \tag{1.7}$$

Here $\hat{\mathbf{r}} = \mathbf{r}/r$ is the observation direction, which only depends on (θ,φ): $\hat{\mathbf{r}} = \cos\varphi\sin\theta\hat{\mathbf{x}} + \sin\varphi\sin\theta\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}$.



For $r >> L$, one can use the approximations:

$$\frac{1}{\left|\mathbf{r}-\mathbf{r}'\right|} \approx \frac{1}{r}, \qquad e^{-jk_0\left|\mathbf{r}-\mathbf{r}'\right|} \approx e^{-jk_0 r}e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'}. \tag{1.8}$$

The term $\dfrac{1}{\left|\mathbf{r}-\mathbf{r}'\right|}$ is evaluated with a zero-order approximation ($\left|\mathbf{r}-\mathbf{r}'\right| \approx r$), whereas $e^{-jk_0\left|\mathbf{r}-\mathbf{r}'\right|}$ is evaluated with the 1st-order approximation ($\left|\mathbf{r}-\mathbf{r}'\right| \approx r - \hat{\mathbf{r}}\cdot\mathbf{r}'$). The reason is that due to the periodicity of the complex exponential, the phase corrections are negligible only when they are

much less than 2π. If one writes $\phi = k_0 \left|\mathbf{r}-\mathbf{r}'\right| = \underbrace{k_0 r}_{\text{zero-order}} \underbrace{-k_0 \hat{\mathbf{r}}\cdot\mathbf{r}'}_{\text{1st-order}} + \underbrace{\delta\phi}_{\text{higher-order}}$, the 1st-order term is much smaller than the zero-order term, but it is not necessarily much smaller than 2π, and hence it cannot be neglected. It can be show that the higher order corrections are negligible ($\left|\delta\phi\right| << 2\pi$) only if $r > 2L^2/\lambda_0$.

Substituting Eqs. (1.8) into Eq. (1.2) one finds that in the far-field region the vector potential is given by:

$$\underline{\mathbf{A}}(\mathbf{r})\Big|_{\text{far-field}} \approx \mu_0 \frac{e^{-jk_0 r}}{4\pi r} \underbrace{\iiint dV' \underline{\mathbf{j}}(\mathbf{r}') e^{+jk_0 \hat{\mathbf{r}}\cdot\mathbf{r}'}}_{\mathbf{f}(\hat{\mathbf{r}})} = \mu_0 \frac{e^{-jk_0 r}}{4\pi r} \mathbf{f}(\hat{\mathbf{r}}). \tag{1.9}$$

The function $\mathbf{f}(\hat{\mathbf{r}})$ is essentially a Fourier transform of the current distribution. This function depends exclusively on the direction of observation, i.e., $\mathbf{f} = \mathbf{f}(\theta,\phi)$, and thereby determines the *directional* properties of the far-field. The radial dependence of the field is determined uniquely by the factor $\frac{e^{-jk_0 r}}{4\pi r}$, which describes a spherical wave.

The electromagnetic fields in the far-field are obtained by substituting Eq. (1.9) in Eq. (1.1). For example,

$$\begin{aligned} \underline{\mathbf{H}}(\mathbf{r})\Big|_{\text{far-field}} &\approx \nabla \times \left\{ \frac{e^{-jk_0 r}}{4\pi r} \mathbf{f}(\hat{\mathbf{r}}) \right\} \\ &= \nabla \left\{ \frac{e^{-jk_0 r}}{4\pi r} \right\} \times \mathbf{f}(\hat{\mathbf{r}}) + \frac{e^{-jk_0 r}}{4\pi r} \nabla \times \mathbf{f} \end{aligned} \tag{1.10}$$

In the second identity, we used $\nabla \times \{g\,\mathbf{f}\} = \nabla g \times \mathbf{f} + g \nabla \times \mathbf{f}$. To proceed, we note that in spherical coordinates:

$$\begin{aligned} \nabla \times \mathbf{f} &= \frac{1}{r^2 \sin\theta} \begin{vmatrix} \hat{\mathbf{r}} & r\hat{\boldsymbol{\theta}} & r\sin\theta\hat{\boldsymbol{\varphi}} \\ \partial_r & \partial_\theta & \partial_\varphi \\ f_r & r f_\theta & r\sin\theta f_\varphi \end{vmatrix} = \frac{1}{r^2 \sin\theta} \begin{vmatrix} \hat{\mathbf{r}} & r\hat{\boldsymbol{\theta}} & r\sin\theta\hat{\boldsymbol{\varphi}} \\ 0 & \partial_\theta & \partial_\varphi \\ f_r & r f_\theta & r\sin\theta f_\varphi \end{vmatrix} \\ &= \frac{1}{r^2 \sin\theta} \hat{\mathbf{r}} \left(\partial_\theta \left(r\sin\theta f_\varphi \right) - \partial_\varphi \left(r f_\theta \right) \right) + \ldots = o\left(\frac{1}{r}\right) \end{aligned} \tag{1.11}$$

In the second identity, we used $\partial / \partial r = 0$ because $\mathbf{f} = \mathbf{f}(\theta, \phi)$. By evaluating the determinant in the second identity one can check that $\nabla \times \mathbf{f}$ decays as $1/r$ as $r \to \infty$. A term that decays as $1/r^n$ is denoted by $o\left(\frac{1}{r^n}\right)$. Using this result one finds that:

$$\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}} \approx \nabla\left\{\frac{e^{-jk_0 r}}{4\pi r}\right\} \times \mathbf{f}(\hat{\mathbf{r}}) + o\left(\frac{1}{r^2}\right). \tag{1.12}$$

From the formula of the gradient in spherical coordinates: $\nabla\left\{\frac{e^{-jk_0 r}}{4\pi r}\right\} = \partial_r\left\{\frac{e^{-jk_0 r}}{4\pi r}\right\}\hat{\mathbf{r}} = -jk_0\left(1 + \frac{1}{jk_0 r}\right)\frac{e^{-jk_0 r}}{4\pi r}\hat{\mathbf{r}}$. From here one gets:

$$\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}} \approx -jk_0 \hat{\mathbf{r}} \times \mathbf{f}(\hat{\mathbf{r}}) \frac{e^{-jk_0 r}}{4\pi r} + o\left(\frac{1}{r^2}\right). \tag{1.13}$$

The correction term falls-off as $o\left(\frac{1}{r^2}\right)$ and is negligible when $k_0 r >> 1$, or equivalently when $r >> \lambda_0$.

Following a similar analysis, it can be shown that the electric field $\underline{\mathbf{E}} = \frac{1}{j\omega\varepsilon_0}\nabla \times \underline{\mathbf{H}}$ in the far-field region satisfies:

$$\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}} \approx \frac{1}{j\omega\varepsilon_0}(-jk_0)^2 \hat{\mathbf{r}} \times \left(\hat{\mathbf{r}} \times \mathbf{f}(\hat{\mathbf{r}})\right) \frac{e^{-jk_0 r}}{4\pi r} + o\left(\frac{1}{r^2}\right). \tag{1.14}$$

It is convenient to introduce the vector:

$$\boxed{\mathbf{h}_e(\hat{\mathbf{r}}) \equiv \frac{1}{\underline{I}(0)}\hat{\mathbf{r}} \times \left(\hat{\mathbf{r}} \times \mathbf{f}(\hat{\mathbf{r}})\right) = \frac{1}{\underline{I}(0)}\left[\left(\iiint dV' \underline{\mathbf{j}}(\mathbf{r}') e^{+jk_0 \hat{\mathbf{r}} \cdot \mathbf{r}'}\right) \times \hat{\mathbf{r}}\right] \times \hat{\mathbf{r}}}. \tag{1.15}$$

Here, $\underline{I}(0)$ is the feeding current of the antenna. The current distribution $\underline{\mathbf{j}}$ includes the polarization and conduction currents on the materials that form the antenna. The complex vector $\mathbf{h}_e(\hat{\mathbf{r}})$ has units of length and is known as the ***antenna effective height***. The effective height only depends on the direction of observation $\mathbf{h}_e = \mathbf{h}_e(\theta, \varphi)$. Furthermore, from the properties of

the vector product, it is clear that $\mathbf{h}_e(\hat{\mathbf{r}})\cdot\hat{\mathbf{r}}=0$. Hence, the effective height is a vector of the form:

$$\mathbf{h}_e(\hat{\mathbf{r}})\equiv h_{e,\theta}\hat{\boldsymbol{\theta}}+h_{e,\varphi}\hat{\boldsymbol{\varphi}}, \tag{1.16}$$

with $h_{e,\theta}=h_{e,\theta}(\theta,\varphi)$ and $h_{e,\varphi}=h_{e,\varphi}(\theta,\varphi)$.

The electric field in the far-field region can be expressed in terms of the effective height:

$$\boxed{\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}}\approx\eta_0 jk_0\underline{I}(0)\mathbf{h}_e(\theta,\varphi)\frac{e^{-jk_0r}}{4\pi r}}. \tag{1.17}$$

This result is completely general and applies to any antenna. The directional properties of the radiated field are completely determined by the effective height. The far-field region is defined by the three conditions used in the derivation:

$$\boxed{r>>L,\qquad r>>\lambda_0,\qquad r>2L^2/\lambda_0}. \tag{1.18}$$

Furthermore, from Eqs. (1.13)-(1.14) it is clear that:

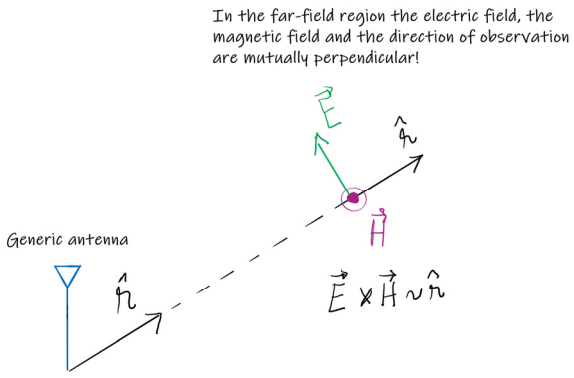
$$\boxed{\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}}\cdot\hat{\mathbf{r}}=0=\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}}\cdot\hat{\mathbf{r}}}. \tag{1.19}$$

This means that the radiated fields are transverse (perpendicular) to the direction of observation $\hat{\mathbf{r}}$ (direction of propagation). Thus, in spherical coordinates they are of the form $\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}}=\underline{E}_\theta\hat{\boldsymbol{\theta}}+\underline{E}_\varphi\hat{\boldsymbol{\varphi}}$ and $\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}}=\underline{H}_\theta\hat{\boldsymbol{\theta}}+\underline{H}_\varphi\hat{\boldsymbol{\varphi}}$.

Comparing Eqs. (1.13)-(1.14) one sees that $\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}}=\eta_0\,\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}}\times\hat{\mathbf{r}}$. Calculating the vector product of both sides of this equation with $\hat{\mathbf{r}}$, one gets $\hat{\mathbf{r}}\times\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}}=\eta_0\hat{\mathbf{r}}\times\left(\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}}\times\hat{\mathbf{r}}\right)=\eta_0\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}}$ [1]. From here it follows that:

$$\boxed{\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}}=\eta_0\,\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}}\times\hat{\mathbf{r}}},\qquad\boxed{\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}}=\frac{1}{\eta_0}\hat{\mathbf{r}}\times\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}}}. \tag{1.20}$$

[1] Note that $\mathbf{A}\times(\mathbf{B}\times\mathbf{C})=\mathbf{B}(\mathbf{A}\cdot\mathbf{C})-\mathbf{C}(\mathbf{A}\cdot\mathbf{B})$.



This shows that the electric and magnetic fields are also perpendicular to each other. Thus, in the far-field region the wave is transverse electromagnetic (TEM). Furthermore, the field amplitudes in the far-field are related by the free-space impedance:

$$\boxed{\left\|\underline{\mathbf{E}}\right\|_{\text{far-field}} = \eta_0 \left\|\underline{\mathbf{H}}\right\|_{\text{far-field}}}. \tag{1.21}$$

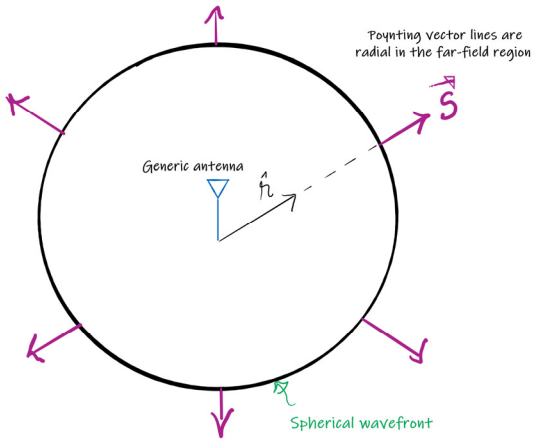
Note that the electromagnetic fields of a plane wave that propagates in free-space are linked precisely as in Eqs. (1.20)-(1.21), with the direction of propagation $\hat{\mathbf{d}}$ in the place of the direction of observation $\hat{\mathbf{r}}$.

1.3 Radiation intensity

The (time-averaged) Poynting vector is $\mathbf{S}_{\text{av}} = \frac{1}{2}\text{Re}\left\{\underline{\mathbf{E}} \times \underline{\mathbf{H}}^*\right\}$. In the far-field region, it is of the form $\mathbf{S}_{\text{av}}\big|_{\text{far-field}} = \frac{1}{2}\text{Re}\left\{\underline{\mathbf{E}} \times \left(\frac{1}{\eta_0}\hat{\mathbf{r}} \times \underline{\mathbf{E}}\right)^*\right\}$. Using the vector identity in the footnote 1 and Eqs. (1.19) and (1.21) one finds that the Poynting vector is oriented along the radial direction:

$$\mathbf{S}_{\text{av}}\big|_{\text{far-field}} = \frac{\left|\underline{\mathbf{E}}\right|^2}{2\eta_0}\hat{\mathbf{r}} = \frac{\eta_0 \left|\underline{\mathbf{H}}\right|^2}{2}\hat{\mathbf{r}}. \tag{1.22}$$

Thus, in the far-field the energy flows along the radial direction. This confirms that the wave is spherical.



The radiation intensity of the antenna (power radiated per solid angle) is given by:

$$\boxed{U = S_{\text{av}} r^2 \Big|_{\text{far-field}} = \frac{|\underline{\mathbf{E}}|^2}{2\eta_0} r^2 \Big|_{\text{far-field}}}. \tag{1.23}$$

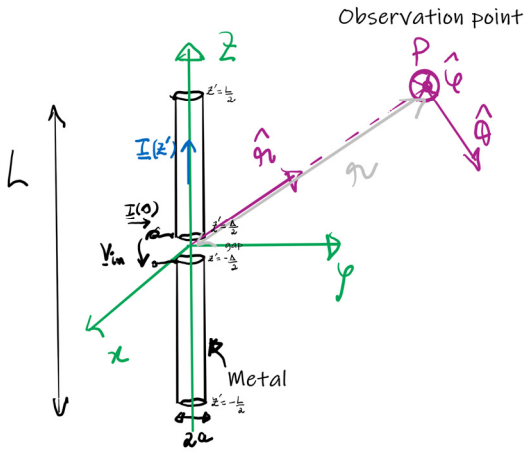
It can be expressed in terms of the effective height as $U = \frac{1}{2\eta_0} \left| \eta_0 j k_0 \underline{I}(0) \mathbf{h}_e \frac{e^{-jk_0 r}}{4\pi r} \right|^2 r^2$, which is the same as:

$$\boxed{U = \frac{\eta_0}{32\pi^2} k_0^2 \left|\underline{I}(0)\right|^2 |\mathbf{h}_e|^2 = \frac{\eta_0}{8} \left|\underline{I}(0)\right|^2 \frac{|\mathbf{h}_e|^2}{\lambda_0^2}}. \tag{1.24}$$

2. Linear antennas

2.1 The dipole antenna

The dipole antenna consists of a metallic cylinder fed through its center point by some generator or transmission line. The voltage at the antenna terminals is $\underline{V}_{in}$ and the feeding current is $\underline{I}(0)$. The dipole length is denoted by L and the radius of the cylindrical wire by a.



It is assumed that $L >> a$, which corresponds to the so-called "thin wire approximation". In such case, one can assume that the current density is oriented along the axis of the dipole: $\underline{\mathbf{j}} \approx j_z \hat{\mathbf{z}}$. For a very good conductor, the current density is confined near the surface of the metallic wire. The total current flowing along z is $\underline{I}(z) \approx \iint j_z(x,y,z)dxdy$, with the integral taken over a cross-section ($z = const.$) of the wire.

As discussed in Chapter 1, the fields radiated by an antenna are fully determined by the vector potential [Eq. (1.2)], which depends on the current $\underline{I}(z)$. The problem is that the current distribution is usually unknown. The calculation of the currents induced on the materials that form the antenna is in general a very complicated problem, and requires the use of numerical

methods. In the next sub-section, we show that for a dipole antenna $\underline{I}(z)$ can be obtained from the solution of an *integral equation*.

2.2 Hallén's integral equation

The idea to find the unknown current distribution is to explore the fact that the fields radiated by the current must satisfy some constraints on the dipole surface.

To develop this idea, let us first evaluate the electric field radiated by $\underline{I}(z)$. It is clear that $\underline{\mathbf{A}} = \underline{A}_z \hat{\mathbf{z}}$ because the current density is directed along *z*. Hence, using Eq. (1.5) one finds that the *z*-component of the electric field away from the source region is given by:

$$\underline{E}_z = \frac{c^2}{j\omega}\left(\frac{\partial^2 \underline{A}_z}{\partial z^2} + k_0^2 \underline{A}_z\right). \tag{2.1}$$

Let us evaluate $\underline{E}_z$ over the *z*-axis ($\mathbf{r} = (0,0,z)$). The metal is assumed to be a perfect electric conductor (PEC), and the dipole is supposed to be a hollow cylindrical tube, so that the current is completely confined to the surface ($x'^2 + y'^2 = a^2$). Then, for $\mathbf{r} = (0,0,z)$, we can write:

$$\begin{aligned} |\mathbf{r} - \mathbf{r}'|^2 &= (x - x')^2 + (y - y')^2 + (z - z')^2 \\ &= a^2 + (z - z')^2 \end{aligned}. \tag{2.2}$$

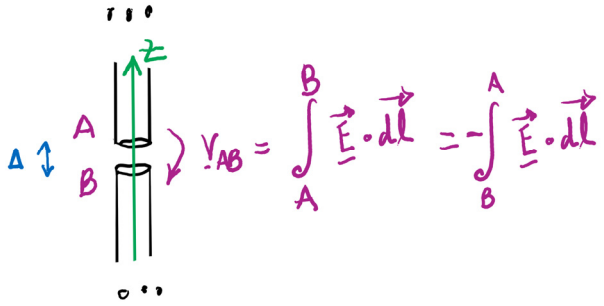
From Eq. (1.2), one finds that along the *z*-axis:

$$\begin{aligned} \underline{A}_z &= \mu_0 \iiint dV' \underline{j}_z(\mathbf{r}') \frac{e^{-jk_0\sqrt{a^2+(z-z')^2}}}{4\pi\sqrt{a^2+(z-z')^2}} \\ &= \mu_0 \int dz' \underline{I}(z') \frac{e^{-jk_0\sqrt{a^2+(z-z')^2}}}{4\pi\sqrt{a^2+(z-z')^2}} \end{aligned} \quad \text{(over the } z\text{-axis)}. \tag{2.3}$$

From Eqs. (2.1) and (2.3), it should be clear that $\underline{E}_z$ is fully determined by the current on the antenna.

The dipole surface is defined by $-\Delta/2 \le |z| \le L/2$, where L is the length of the dipole. Here, Δ is the generator gap width. Now the key point is that for a perfectly conducting metal, $\underline{E}_z$ is required to vanish on the surface of the dipole. Due to this reason, $\underline{E}_z$ will also be (approximately) zero over the axis of the hollow metallic tube:

$$\underline{E}_z = 0, \qquad \Delta/2 \le |z| \le L/2. \tag{2.4}$$



On the other hand, the voltage at the antenna terminals is $\underline{V}_{in} = -\int_{-\Delta/2}^{+\Delta/2} \mathbf{E}\cdot\mathbf{dl} = -\int_{-\Delta/2}^{+\Delta/2} E_z dz$. Assuming that the gap width is very small compared to L and that the electric field is approximately constant in the gap, one obtains $\underline{V}_{in} \approx -E_z\Delta$. Hence, over the dipole region of the z-axis one has:

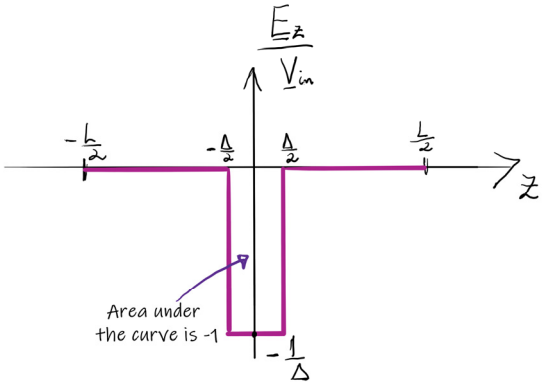
$$\underline{E}_z = \begin{cases} 0, & \Delta/2 \le |z| \le L/2 \\ -\dfrac{\underline{V}_{in}}{\Delta}, & |z| \le \Delta/2 \end{cases} \tag{2.5}$$

For an infinitesimal gap $\Delta \to 0^+$, the electric field diverges to infinity at the origin (generator) and vanishes elsewhere ($0^+ \le |z| \le L/2$). Therefore, it can be approximated by a delta-function:

$$\underline{E}_z \approx -\underline{V}_{in}\delta(z), \qquad \text{(region } |z| \le L/2 \text{ of the } z\text{-axis)}. \tag{2.6}$$

Combining the above result with Eq. (2.1), we find that:

$$\frac{d^2\underline{A}_z}{dz^2} + k_0^2\underline{A}_z = -\underline{V}_{in}\frac{j\omega}{c^2}\delta(z). \tag{2.7}$$



This differential equation can now be solved for $\underline{A}_z$. It is clear that for $z \neq 0$, $\underline{A}_z$ is a solution of the homogeneous equation $\frac{d^2 \underline{A}_z}{dz^2} + k_0^2 \underline{A}_z = 0$, and thereby is a linear combination of $\cos(k_0 z)$ and $\sin(k_0 z)$:

$$\underline{A}_z = \begin{cases} C_1 \cos(k_0 z) + C_2 \sin(k_0 z), & z > 0 \\ C_1' \cos(k_0 z) + C_2' \sin(k_0 z), & z < 0 \end{cases}. \tag{2.8}$$

By symmetry, the current, and thereby also $\underline{A}_z$, must be an even function of z. This requires that $C_1 = C_1'$ and $C_2 = -C_2'$. This implies that:

$$\underline{A}_z = C_1 \cos(k_0 z) + C_2 \sin(k_0 |z|). \tag{2.9}$$

Substituting the above formula into Eq. (2.7) and taking into account that $\frac{d^2}{dz^2} \sin(k_0 |z|) = k_0 \frac{d}{dz} (\cos(k_0 z) \operatorname{sgn} z) = -k_0^2 \sin(k_0 |z|) + 2k_0 \delta(z)$, it is found that $2k_0 C_2 = -\underline{V}_{in} \frac{j\omega}{c^2}$, or equivalently $C_2 = -j \frac{\underline{V}_{in}}{2c}$. This shows that the vector potential along the dipole axis is:

$$\underline{A}_z = C_1 \cos(k_0 z) - j \frac{\underline{V}_{in}}{2c} \sin(k_0 |z|). \tag{2.10}$$

Combining this result with Eq. (2.3), we obtain the so-called Hallén's integral equation (here $C = C_1 / \mu_0$ is some unknown constant):

$$\boxed{\underline{C}\cos(k_0 z) - j\frac{\underline{V}_{in}}{2\eta_0}\sin(k_0|z|) = \int_{-L/2}^{+L/2} dz'\,\underline{I}(z')\frac{e^{-jk_0\sqrt{a^2+(z-z')^2}}}{4\pi\sqrt{a^2+(z-z')^2}}}. \tag{2.11}$$

The unknowns of the equation are $\underline{C}$ (a complex number) and $\underline{I}(z')$ (a complex function). The current is subject to the physical constraint

$$\underline{I}(\pm L/2) = 0, \tag{2.12}$$

i.e., it must vanish at the dipole ends. The Hallén equation must be satisfied for any $|z| \le L/2$.

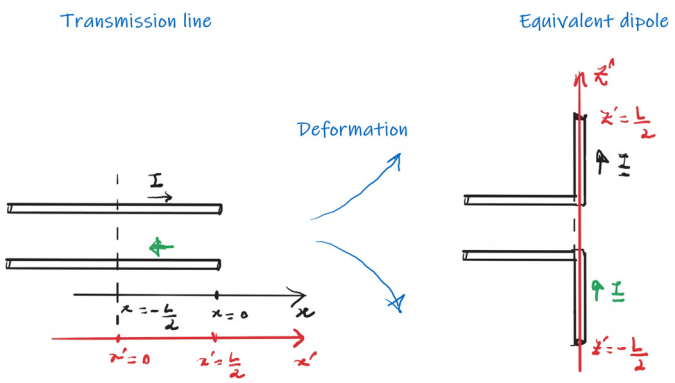
In general, an "integral equation" is an equation of the form:

$$g(z) = \int dz' K(z,z') f(z'), \tag{2.13}$$

with $g(z)$ some given function and $f(z')$ the unknown of the equation. The term $K(z,z')$ is known as the *kernel* of the equation. Evidently, for the Hallén's equation the kernel is $K(z,z') = \dfrac{e^{-jk_0\sqrt{a^2+(z-z')^2}}}{4\pi\sqrt{a^2+(z-z')^2}}$. The numerical solution of the Hallén's equation will be discussed in detail in the context of the computational work.

2.3 Sinusoidal current approximation

Fortunately, it is possible to obtain an approximate formula for the current $\underline{I}(z')$ based on a transmission line analogy. The idea is the following. Consider a transmission line terminated in open-circuit. The current in the line is a superposition of two counter-propagating waves: $\underline{I}_{\text{line}}(x) = \underline{I}^+(x) + \underline{I}^-(x) = \underline{I}_0^+ e^{-j\beta x} + \underline{I}_0^- e^{+j\beta x}$. Let us pick the origin of the x-axis to be at the end of the line. Then, $\underline{I}(0) = 0$, i.e., $\underline{I}_0^+ = -\underline{I}_0^-$. This implies that $\underline{I}_{\text{line}}(x) = -\underline{I}_m \sin\beta x$ with $\underline{I}_m = 2j\underline{I}_0^+$.



Now the key observation is that a dipole antenna can be regarded as a deformed transmission line, with the two conductors bent by 90º, as illustrated in the figure. The sinusoidal current approximation is based on the hypothesis that after the deformation of the line the current distribution remains unchanged. To explore the consequences of this hypothesis, it is convenient to consider a primed reference axis such that $x'=0$ is the point where the line is bent. It is clear that $x'=x+L/2$, so that $\underline{I}_{\text{line}}\left(x'\right)=\underline{I}_m\sin\beta\left(\frac{L}{2}-x'\right)$.

The bending transforms the horizontal positive x'-axis into the vertical positive z'-axis ($x'\to z'$, $\underline{I}_{\text{line}}\to\underline{I}$). Hence, for $z'>0$, the current on the dipole is given by $\underline{I}\left(z'\right)=\underline{I}_m\sin k_0\left(\frac{L}{2}-z'\right)$. We replaced $\beta\to k_0$, because the two conductors of the equivalent line are surrounded by air. In a transmission line the two conductors transport anti-paralllel currents (with opposite signs). In contrast, after the deformation the currents in the two conductors (the arms of the dipole) become parallel (have the same sign). This means that $\underline{I}\left(z'\right)$ must be an even function of z' so that:

$$\boxed{\underline{I}\left(z'\right)=\underline{I}_m\sin k_0\left(\frac{L}{2}-\left|z'\right|\right)},\qquad\text{(sinusoidal current approximation).}\qquad(2.14)$$

For a sufficiently long dipole, $\underline{I}_m$ represents the peak (maximum) current: $\max\left|\underline{I}(z')\right| = \left|\underline{I}_m\right|$. For a dipole with length inferior to half-wavelength ($L < \lambda_0/2$ or equivalently $k_0 L/2 < \pi/2$), the peak current is $\max\left|\underline{I}(z')\right| = \left|\underline{I}_m\right| \sin\dfrac{k_0 L}{2}$.

2.4 Far-field of the dipole

The far-field of the dipole antenna is determined by the effective height [Eq. (1.15)]. Using $\underline{\mathbf{j}} \approx j_z \hat{\mathbf{z}}$, one gets:

$$\begin{aligned}\mathbf{h}_e &= \frac{1}{\underline{I}(0)}\left(\iiint dV' j_z(\mathbf{r}') e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'}\right)\left(\hat{\mathbf{z}}\times\hat{\mathbf{r}}\right)\times\hat{\mathbf{r}} \\ &\approx \frac{1}{\underline{I}(0)}\left(\int_{-L/2}^{L/2} dz' e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'}\Big|_{x'=y'=0} \iint dx'dy' j_z(\mathbf{r}')\right)\left(\hat{\mathbf{z}}\times\hat{\mathbf{r}}\right)\times\hat{\mathbf{r}} \\ &= \frac{1}{\underline{I}(0)}\left(\int_{-L/2}^{L/2} dz' e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'}\Big|_{x'=y'=0} \underline{I}(z')\right)\left(\hat{\mathbf{z}}\times\hat{\mathbf{r}}\right)\times\hat{\mathbf{r}}\end{aligned} \tag{2.15}$$

Taking into account that $\left(\hat{\mathbf{z}}\times\hat{\mathbf{r}}\right)\times\hat{\mathbf{r}} = \sin\theta\hat{\boldsymbol{\varphi}}\times\hat{\mathbf{r}} = \sin\theta\hat{\boldsymbol{\theta}}$ and that $\hat{\mathbf{r}}\cdot\mathbf{r}' = \hat{\mathbf{r}}\cdot z'\hat{\mathbf{z}} = z'\cos\theta$, one gets:

$$\mathbf{h}_e = h_e\hat{\boldsymbol{\theta}}, \quad \text{with} \quad \boxed{h_e = \sin\theta\int_{-L/2}^{L/2} dz' \frac{\underline{I}(z')}{\underline{I}(0)} e^{+jk_0 z'\cos\theta}}. \tag{2.16}$$

Thus, the dipole electromagnetic fields in the far-field zone are given by:

$$\begin{aligned}\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}} &\approx \underline{E}_\theta\hat{\boldsymbol{\theta}} \\ \underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}} &\approx \underline{H}_\varphi\hat{\boldsymbol{\varphi}}\end{aligned}, \qquad \text{with} \quad \boxed{\begin{aligned}\underline{E}_\theta &\approx \eta_0 jk_0\underline{I}(0)h_e\frac{e^{-jk_0 r}}{4\pi r} \\ \underline{H}_\varphi &\approx \underline{E}_\theta/\eta_0\end{aligned}}. \tag{2.17}$$

2.5 Effective height with the sinusoidal current approximation

Within the sinusoidal current approximation [Eq. (2.14)], the (scalar) effective height is given by:

$$h_e(\theta) = \frac{\sin\theta}{\underline{I}(0)}\int_{-L/2}^{L/2} dz'\underline{I}_m \sin k_0\left(\frac{L}{2} - |z'|\right) e^{+jk_0 z'\cos\theta}. \tag{2.18}$$

with the feeding current

$$\underline{I}(0) = \underline{I}_m \sin\frac{k_0 L}{2}. \tag{2.19}$$

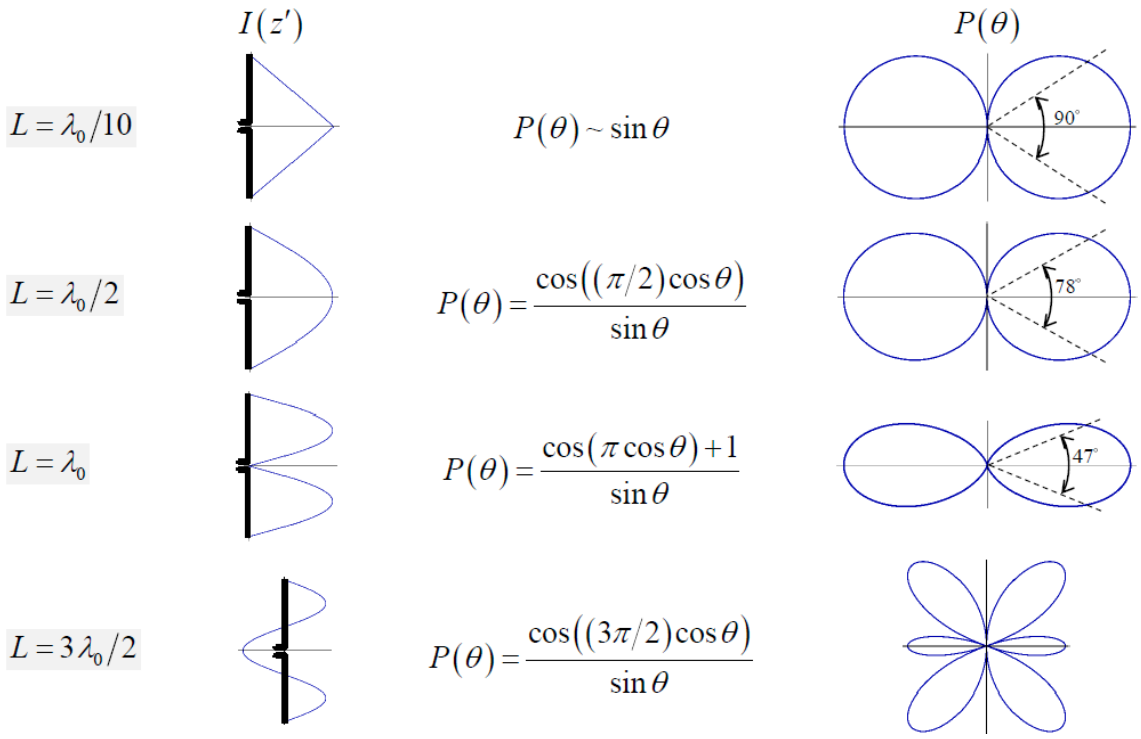
The integral can be evaluated explicitly yielding the closed-form result:

$$\boxed{h_e(\theta) = \frac{2\underline{I}_m}{\underline{I}(0)k_0} P(\theta)}, \quad \text{with} \quad \boxed{P(\theta) = \frac{\cos\left(\frac{k_0 L}{2}\cos\theta\right) - \cos\left(\frac{k_0 L}{2}\right)}{\sin\theta}}. \tag{2.20}$$

The function $P(\theta)$ is known as the ***pattern factor***. The dipole radiation intensity [Eq. (1.24)] can be expressed in terms of the pattern factor as:

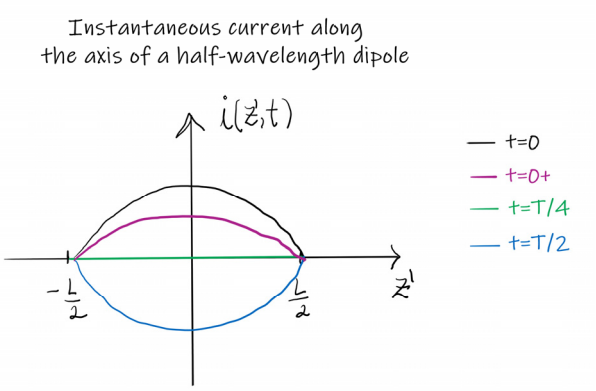
$$\boxed{U = \frac{\eta_0}{8\pi^2}\left|\underline{I}_m\right|^2 \left|P(\theta)\right|^2}. \tag{2.21}$$

Clearly, the pattern factor determines the directional properties of the antenna (radiation pattern). The pattern factor depends uniquely on the height of the dipole L relative to the free-space wavelength. A sketch of the current profile (snapshot in time) and of the pattern factor for different values of L/λ_0 is shown in the figure below. As seen, the radiation pattern becomes more directive as the the dipole length increases for $L/\lambda_0 \le 1.0$. In particular, the half-power beamwidth (HPBW) decreases from 90º (for a short-dipole) down to 78º for $L/\lambda_0 = 0.5$, and to 47º for L/λ_0. In contrast, for $L/\lambda_0 = 1.5$ the radiation pattern is formed by multiple lobes. Thus, increasing the length of the dipole beyond some threshold value does not lead to a more directive beam.



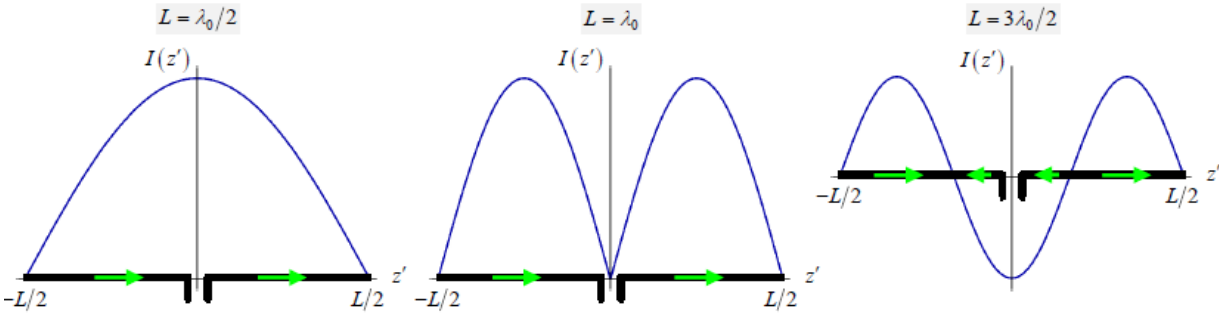
The formation of multiple lobes occurs because when $L/\lambda_0 > 1.0$ the sign of the instantaneous current varies along the dipole antenna. Note that the instantaneous current is:

$$i(z',t) = \mathrm{Re}\left\{\underline{I}(z')e^{j\omega t}\right\} = I_m \sin\left(k_0\left(\frac{L}{2} - |z'|\right)\right)\cos(\omega t). \tag{2.22}$$



The sinusoidal current approximation was used and it is assumed that I_m is real-valued. For $L/\lambda_0 < 1.0$ the sign of $i(z',t)$ is independent of z' for a fixed t (see the figure for the case of the half-wavelength dipole). In such conditions, all the antenna sections radiate in phase and hence there is a constructive interference of the radiated fields. In contrast, when $L/\lambda_0 > 1.0$ the

sections of the antenna where $i(z',t)>0$ radiate fields in opposition of phase with respect to the sections of the antenna where $i(z',t)<0$. This leads to a destructive interference of the radiated fields and to the formation of nulls and multiple lobes in the radiation pattern. The multiple lobes appear for $L/\lambda_0 > 1.2$.



2.6 *Radiated power and radiation resistance*

The power radiated by a dipole antenna can be found by integrating the radiation intensity [Eq. (2.21)] over all the solid angles:

$$P_{rad} = \iint U d\Omega = \frac{\eta_0}{8\pi^2}\left|\underline{I}_m\right|^2 \iint \left|P(\theta)\right|^2 d\Omega \,. \tag{2.23}$$

Using $\iint \left|P(\theta)\right|^2 d\Omega = \iint \left|P(\theta)\right|^2 \underbrace{\sin\theta d\theta d\varphi}_{d\Omega} = 2\pi \times \int_0^{\pi} \left|P(\theta)\right|^2 \sin\theta d\theta$, one gets:

$$P_{rad} = \iint U d\Omega = \frac{\eta_0}{4\pi}\left|\underline{I}_m\right|^2 \int_0^{\pi} \left|P(\theta)\right|^2 \sin\theta d\theta \,. \tag{2.24}$$

The radiation resistance of a dipole antenna $R_{rad} = 2P_{rad} / \left|\underline{I}(0)\right|^2$ is given by (see Eq. (2.19)):

$$R_{rad} = \frac{\eta_0}{2\pi} \frac{1}{\sin^2\left(\dfrac{k_0 L}{2}\right)} \int_0^{\pi} \left|P(\theta)\right|^2 \sin\theta d\theta \,. \tag{2.25}$$

For a half-wavelength dipole ($L=\lambda_0/2$, i.e., $k_0L=\pi$) one finds that $R_{rad}=\frac{\eta_0}{2\pi}\int_0^{\pi}\left|P(\theta)\right|^2\sin\theta d\theta\approx 60\int_0^{\pi}\left|P(\theta)\right|^2\sin\theta d\theta$ with $P(\theta)\big|_{\lambda/2}=\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta}$. Doing the integration numerically:

$$\left(R_{rad}\right)_{\lambda/2}\approx 73.1\Omega\,. \tag{2.26}$$

Proceeding similarly, it is possible to show that for $L=\lambda_0$ and $L=3\lambda_0/2$ one has:

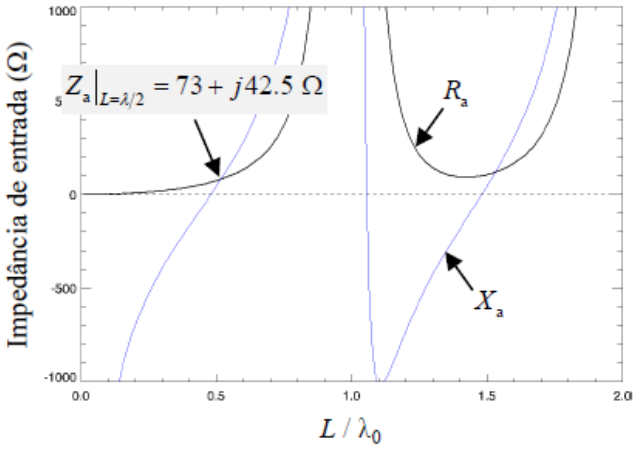
$$\left(R_{rad}\right)_{\lambda}\approx\infty, \qquad \left(R_{rad}\right)_{3\lambda/2}\approx 105.3\Omega \tag{2.27}$$

2.7 Input impedance

The input impedance of the dipole antenna is defined as

$$\underline{Z}_{in}\equiv R_{in}+jX_{in}=\frac{\underline{V}_{in}}{\underline{I}(0)}\,. \tag{2.28}$$

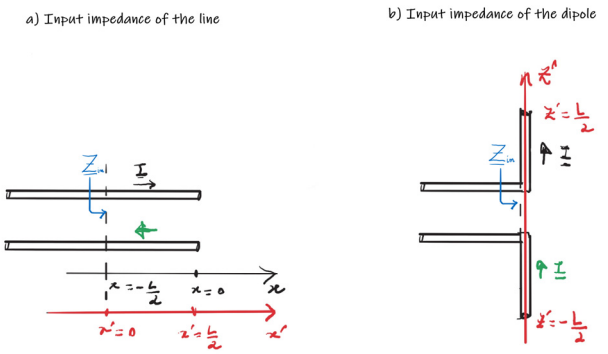
(the input impedance can also be denoted as $\underline{Z}_a\equiv R_a+jX_a$). The input impedance of the dipole antenna can be rigorously calculated from the solution of the Hallén's integral equation. The variation of the resistance and reactance with L/λ_0 is depicted in the figure below:



L	R_{rad}
dl	≈ 0
$\lambda_0/2$	73
λ_0	$\approx\infty$
$3\lambda_0/2$	105.3

In general, R_{in} depends on the radiation resistance and on the loss resistance: $R_{in}=R_{rad}+R_L$. For a lossless antenna, one has $R_{in}=R_{rad}$. This means that the input resistance of a dipole antenna can be estimated using Eq. (2.25) [see also Eqs. (2.26)-(2.27)]. For a short-

dipole ($L/\lambda_0 << 1$) the input resistance is similar to that of a Hertz dipole, and hence $R_{rad}\big|_{\text{short-dipole}} \approx 0$.



The variation of the reactance X_{in} with L/λ_0 can be justified using the transmission line analogy of Sect. 2.3. In fact, if one supposes that the bending of the line conductors does not significantly influence the input reactance it follows that:

$$X_{in}\Big|_{\substack{\text{dipole}\\ \text{length L}}} \approx X_{in}\Big|_{\substack{\text{line}\\ \text{length } l=L/2}} \tag{2.29}$$

where $X_{in}\big|_{\text{line}} = \text{Im}\left\{Z_{in}\big|_{\text{line}}\right\}$ is the reactance of the equivalent transmission line with length $l = L/2$. For a transmission line terminated in open-circuit, one has $Z_{in}\big|_{\text{line}} = -jZ_0\cot(\beta l) = -jZ_0\cot(k_0 L/2)$, where Z_0 is the characteristic impedance of the line. Thus, the analogy yields:

$$X_{in}\Big|_{\substack{\text{dipole}\\ \text{length L}}} \approx -Z_0\cot(k_0 L/2). \tag{2.30}$$

The formula predicts that the antenna is resonant ($X_{in}\big|_{\text{dipole}} = 0$) when $\frac{k_0 L}{2} = \frac{\pi}{2}, \frac{3\pi}{2}, \ldots$ (zero of the cotangent function), or equivalently $L = 0.5\lambda_0, 1.5\lambda_0, \ldots$ This prediction is in good agreement with the exact $X_{in}\big|_{\text{dipole}}$ represented in the prevous figure. The first resonance of the antenna occurs for a height slightly (a few percent) shorter than $L = 0.5\lambda_0$. Thus, when $L \approx 0.5\lambda_0$ the dipole is resonant and hence can be directly fed by a transmission line, ideally with a

characteristic impedance near $\left(R_{rad}\right)_{\lambda/2} \approx 73.1\Omega$. An antenna with $L \approx 1.5\lambda_0$ is also resonant, but it is not interesting for applications because of its multiple-lobe radiation pattern. It is worth noting that an antenna with $L \approx 1.0\lambda_0$ has an extremely large input impedance (typically on the order of $\mathrm{M}\Omega$) and hence it is a rather inefficient radiator, as it is inevitably strongly mismatched from the feeding line.

For a short dipole the input reactance X_{in} is very large and negative. Due to this reason it is difficult to design a transformer that can match a short dipole to a typical feed. The property $X_{in} < 0$ shows that a short dipole behaves as a capacitor, consistent with the fact that its near-field is predominantly electric.

2.8 The half-wavelength dipole

As discussed in the previous section, the half-wavelength dipole is particularly interesting for applications. For a half-wavelength dipole $L = \lambda_0 / 2$ or equivalently $k_0 L = \pi$. This leads to:

$$\underline{I}(0)\big|_{\lambda/2} = \underline{I}_m . \tag{2.31}$$

$$h_e(\theta)\big|_{\lambda/2} = \frac{2}{k_0} P(\theta), \qquad \text{with} \quad P(\theta)\big|_{\lambda/2} = \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \quad . \tag{2.32}$$

In particular, the far-field of the half-wavelength dipole is given by:

$$\underline{E}_\theta\big|_{\lambda/2} \approx \eta_0 j \underline{I}(0) \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \frac{e^{-jk_0 r}}{2\pi r} . \tag{2.33}$$

The electric field intensity can be written as:

$$\boxed{\left|\underline{E}\right|_{\lambda/2} = \frac{\eta_0}{2\pi r}\left|\underline{I}(0)\right| \left|\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta}\right| \approx \frac{60}{r}\left|\underline{I}(0)\right| \left|\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta}\right|} . \tag{2.34}$$

The directive gain of the dipole antenna is given by:

$$g = \frac{U}{U_{iso}} = \frac{U}{P_{rad}/(4\pi)} = \frac{4\pi U}{\frac{1}{2}R_{rad}\left|I(0)\right|^2}. \tag{2.35}$$

Hence, using Eq. (2.21)

$$g = \frac{4\pi \frac{\eta_0}{8\pi^2}\left|\underline{I}_m\right|^2 \left|P(\theta)\right|^2}{\frac{1}{2}R_{rad}\left|I(0)\right|^2} \overset{L=\lambda/2}{=} \frac{\eta_0}{\pi R_{rad}}\left|P(\theta)\right|^2 = \frac{120}{73.1}\left|P(\theta)\right|^2. \tag{2.36}$$

This shows that:

$$\boxed{g\big|_{L=\lambda/2} = 1.64\left|\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta}\right|^2}. \tag{2.37}$$

The directive gain has rotation symmetry about the *z*-axis, i.e. about the dipole axis. Furthermore, the radiation pattern has mirror-symmetry with respect to the *xoy* plane. The radiation maximum occurs for $\theta = 90^\circ$, so that the directivity is:

$$\boxed{D_{\lambda/2} = \max g\big|_{L=\lambda/2} = 1.64}. \tag{2.38}$$

Thus, the half-wavelength dipole is slightly more directive than the Hertz dipole ($D_{Hertz} = 1.5$). The half-power beamwidth of the $L = \lambda_0/2$ dipole is 78º.

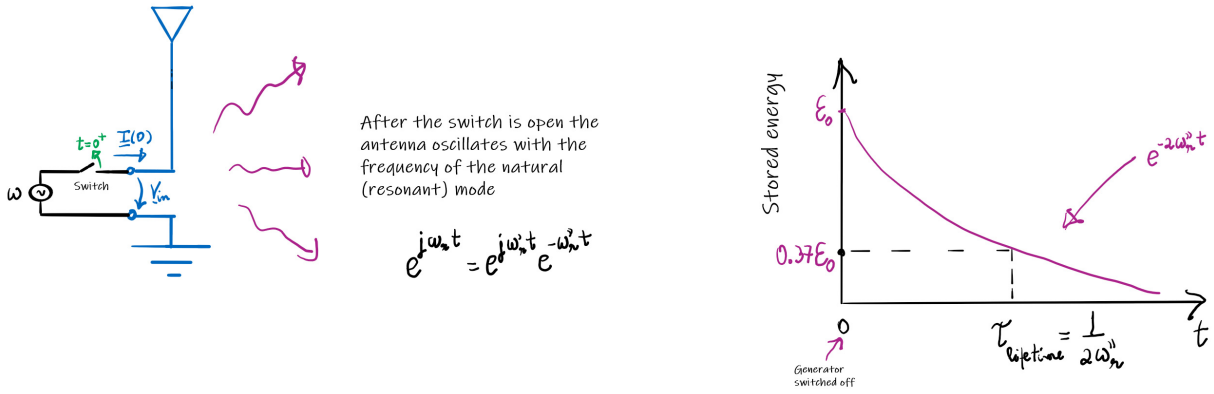
2.9 Further discussion about the resonant condition

The feeding current of an antenna ($\underline{I}(0)$) may be regarded as the response to the input voltage ($\underline{V}_{in}$). The two quantities are related through the input impedance:

$$\underline{I}(0) = \frac{\underline{V}_{in}}{\underline{Z}_{in}(\omega)}. \tag{2.39}$$

Thus, $\dfrac{1}{\underline{Z}_{in}(\omega)}$ may be regarded as the transfer function of a system with input $\underline{V}_{in}$ and output $\underline{I}(0)$. Is it possible to have a response ($\underline{I}(0) \neq 0$) without an input ($\underline{V}_{in} = 0$)? Clearly, this may

be feasible only for some (resonant) frequency (ω_r) that ensures that $\underline{Z}_{in}(\omega_r)=0$, so that the transfer function has a pole. In such conditions, the antenna is said to be resonant at the frequency ω_r because it can emit energy (initially stored in the system) even though the excitation voltage is identical to zero ($\underline{V}_{in}=0$).



An antenna supports multiple *natural modes*, i.e., multiple resonances. When the generator that feeds the antenna is switched off, let us say at $t=0$, there is still energy in the near-field of the antenna. For $t>0$, the antenna will oscillate freely (natural mode) and radiate away all stored energy with frequency ω_r.

In general, both the input resistance and the reactance depend on frequency $\underline{Z}_{in}\equiv R_{in}(\omega)+jX_{in}(\omega)$. Let us consider a real-valued frequency ω_0 for which $X_{in}(\omega_0)=0$. Let us show that the system is resonant near ω_0. To do this, we suppose that the variations of $R_{in}(\omega)$ near ω_0 are negligible so that $R_{in}(\omega)\approx R_{in}(\omega_0)$. Furthermore, the reactance is expanded in a Taylor series around ω_0: $X_{in}(\omega)\approx X_{in}(\omega_0)+(\omega-\omega_0)X'_{in}(\omega_0)=(\omega-\omega_0)X'_{in}(\omega_0)$. These approximations yield:

$$\underline{Z}_{in}\approx R_{in}(\omega_0)+j(\omega-\omega_0)X'_{in}(\omega_0). \tag{2.40}$$

As previously noted, it is possible to have oscillations in the antenna without an input when $\underline{Z}_{in}(\omega_r)=0$. Solving this equation with respect to ω_r one finds that:

$$\omega_r = \omega_0 + j\frac{R_{in}(\omega_0)}{X'_{in}(\omega_0)}. \tag{2.41}$$

The resonance frequency is *complex-valued* $\omega_r = \omega'_r + j\omega''_r$. This means that the time variation of the natural (free) oscillations (without any excitation) is of the form $e^{j\omega_r t} = e^{j\omega'_r t}e^{-\omega''_r t}$. The real part of the resonance frequency ($\omega'_r = \omega_0$) determines the time period of the oscillations ($T = 2\pi/\omega'_r$). The imaginary part ($\omega''_r = \frac{R_{in}(\omega_0)}{X'_{in}(\omega_0)}$) determines the decay rate of the oscillations.

The oscillations are exponentially damped ($e^{-\omega''_r t} \to 0$) because they describe the relaxation of the energy initially stored in the system[1]. The duration (lifetime) of the natural oscillation is roughly $\tau_{\text{lifetime}} = \frac{1}{2\omega''_r}$.

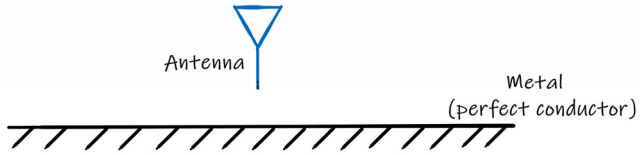
For example, suppose that a dipole antenna is excited by a generator with some oscillation frequency ω and that the generator is switched off at $t = 0$. For $t > 0$, the dipole antenna will support damped oscillations with a frequency ω_0 determined by $X_{in}(\omega_0) = 0$. This condition is satisfied when $k_0 L = \pi$ (half-wavelength dipole), which yields $\omega_0 = \pi c / L$. This discussion explains why a half-wavelength dipole antenna is an efficient radiator: $k_0 L = \pi$ corresponds to a natural frequency of the antenna, i.e., a frequency for which the antenna supports free oscillations without any excitation. When the excitation frequency matches the resonant frequency the oscillations can be rather strong, similar to any other resonant system.

2.10 Image method

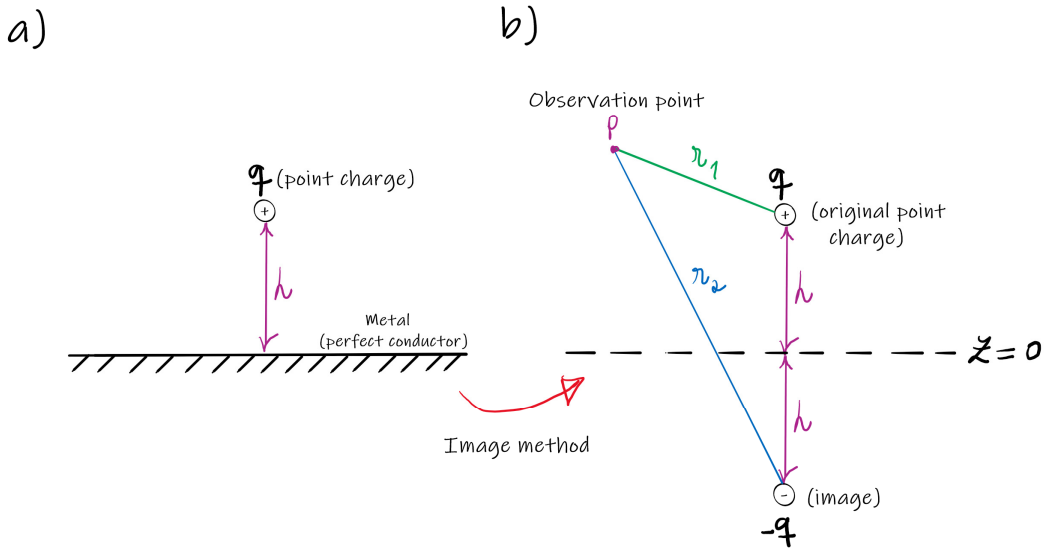
So far, we discussed the radiation of antennas in free-space. However, in practice the antennas are positioned near other objects. The image method is a powerful technique that enables

[1] The described analysis is meaningful only when $\omega''_r << \omega'_r$, so that the Taylor series expansion can be justified.

reducing a radiation problem where the emitter is in the vicinity of a metal (planar) surface to a radiation problem where the emitter is in free-space.



To begin with, we start with a simple electrostatic problem. Consider a point charge q positioned at a distance h from a metal surface (plane $z=0$). The metal is assumed to be a perfect conductor, so that the electric field component parallel (tangent) to the surface must vanish ($\mathbf{E}_{\tan}=0$), or equivalently the electrostatic potential ϕ must be a constant on the metal surface ($\phi|_{z=0}=const.$). The point charge attracts charges of opposite sign in the metal and induces a nontrivial charge distribution on the metal surface. The unknown charge distribution must ensure that the total electric potential (created by itself and by the point charge) satisfies $\phi|_{z=0}=const.$. Finding the induced charge distribution is not a simple problem!

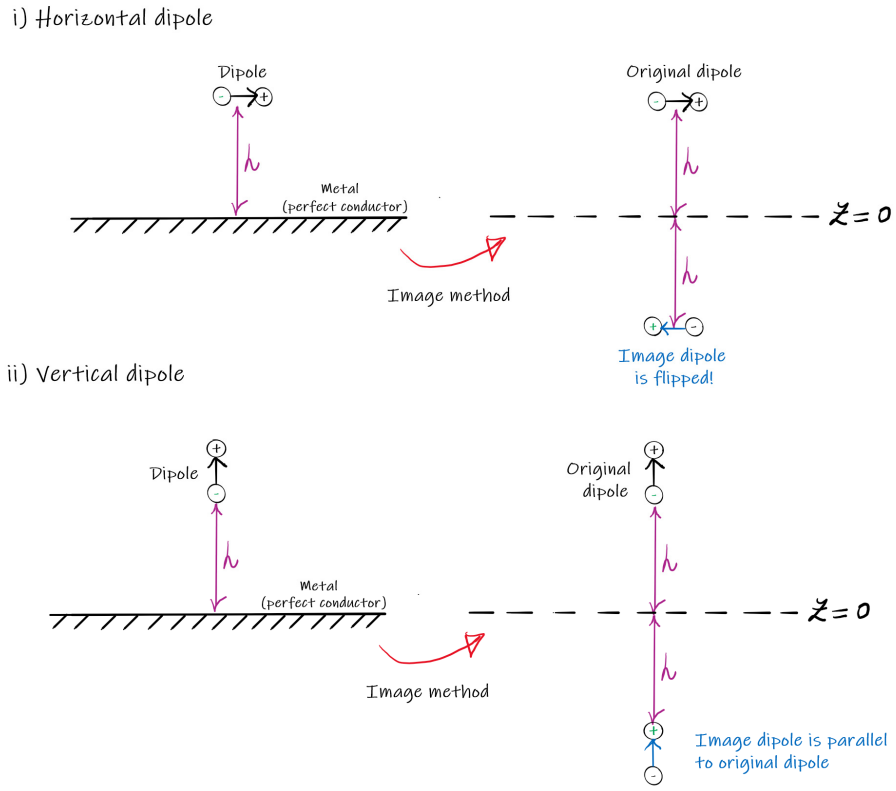


Fortunately, there is an easy way. We consider an auxiliary problem where the metal plane is removed and where a second (*image*) point charge $-q$ is placed at the mirror symmetric point

with respect to $z=0$. The potential created by the two point charges can be found trivially from Coulomb's theory:

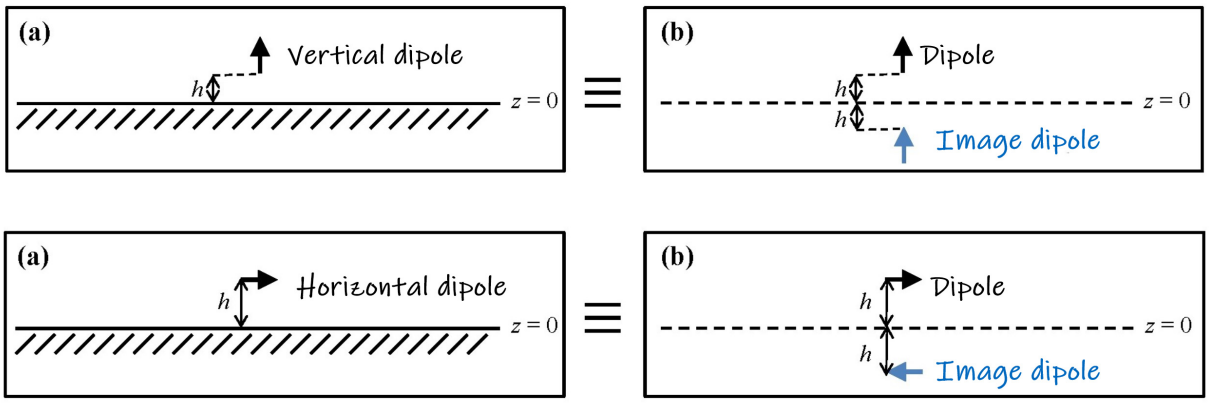
$$\phi\big|_P = \frac{q}{4\pi\varepsilon_0 r_1} + \frac{-q}{4\pi\varepsilon_0 r_2}. \tag{2.42}$$

For any point on the plane $z=0$, one has $r_1 = r_2$. This implies that $\phi\big|_{z=0} = 0$. Thus, the $z=0$ plane of the auxiliary problem is a plane of constant potential, and thereby it cannot be distinguished from a perfect electric conductor (PEC). In particular, the $z=0$ plane of the auxiliary problem can be replaced by a PEC without perturbing the fields in the region $z>0$. Such configuration corresponds precisely to the original problem. Thus, the solution of the original problem (with the metal plane) is coincident with the solution of the auxiliary problem [Eq. (2.42)] (without the metal plane but with an *image* charge).

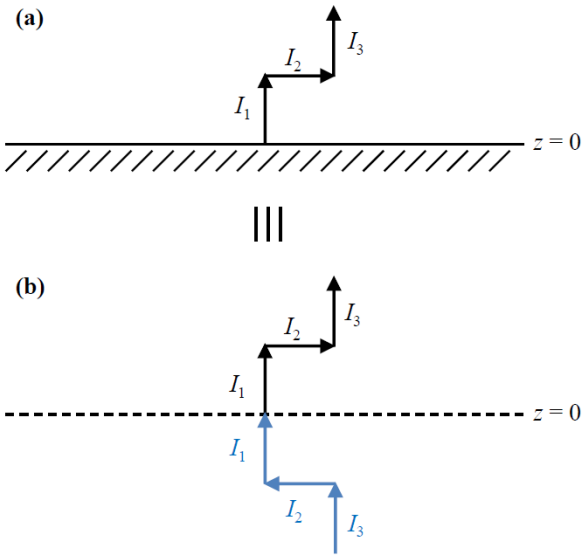


The same ideas can be generalized to solve more complex dynamical problems, for example to find the radiation from a (Hertz) dipole placed in the vicinity of a metal plate. There are two

cases of interest: i) horizontal dipole, ii) vertical dipole. To understand what is the image of a dipole, we recall that a dipole consists of two point charges with opposite signs separated by some distance. Thus, in analogy with the electrostatic case discussed previously, we find that the image of a vertical dipole is another vertical dipole oriented in the same direction, whereas the image of a horizontal dipole is an anti-parallel horizontal dipole (see the figure). The radiation from the dipole in the vicinity of a metal plate is coincident (in the half-space $z>0$) with the field emitted by the dipole and by its image in free-space. Note that the field emitted by two Hertz dipoles can be easily found using the superposition principle.

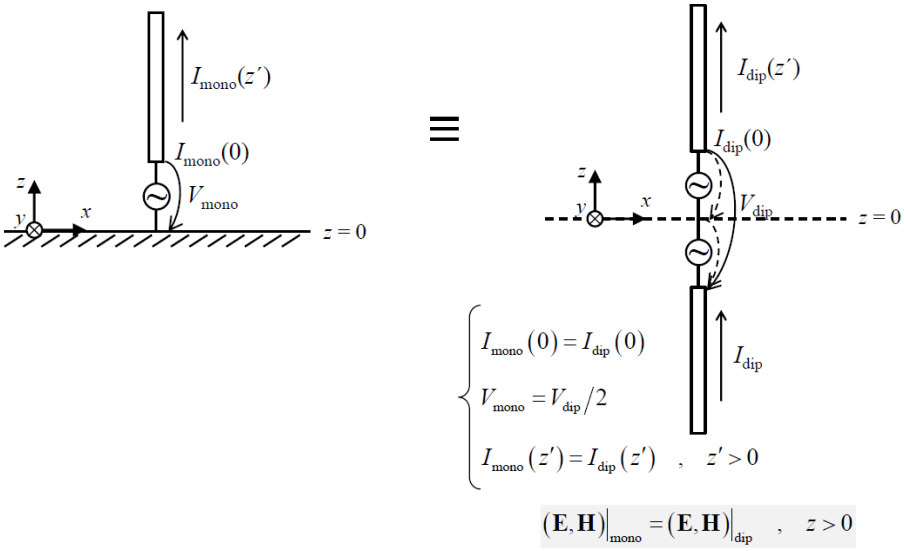


The same idea can be applied to more complex current distributions, as an arbitrary current distribution can always be regarded as a collection of Hertz dipoles. This concept is illustrated in the figure below.



2.11 The monopole antenna

The monopole antenna consists of a cylindrical conductor with height L_{mono} above a metal-ground plane (figure below, left). The antenna is excited by a voltage generator ($\underline{V}_{\text{mono}}$), as illustrated in the figure. From the image method, the monopole antenna can be reduced to an equivalent dipole antenna with height $L_{\text{dip}} = 2L_{\text{mono}}$ (figure below, right). The fields radiated by the two antennas are identical in the region $z > 0$. Furthermore, the current distributions ($\underline{I}(z')\big|_{\text{mono}} = \underline{I}(z')\big|_{\text{dip}}$ for $z' > 0$) and the feeding currents are also identical ($\underline{I}(0)\big|_{\text{mono}} = \underline{I}(0)\big|_{\text{dip}}$). On the other, the input voltage is two times larger for the equivalent dipole antenna: $\underline{V}_{\text{dip}} = 2\underline{V}_{\text{mono}}$.



Due to these properties the input impedance of the monopole antenna is half of the input impedance of the equivalent dipole:

$$\underline{Z}_{in}\big|_{\text{mono}}=\frac{\underline{V}_{\text{mono}}}{\underline{I}(0)\big|_{\text{mono}}}=\frac{\underline{V}_{\text{dip}}/2}{\underline{I}(0)\big|_{\text{dip}}}=\frac{1}{2}\underline{Z}_{in}\big|_{\text{dip}}. \tag{2.43}$$

For example, for $L_{\text{mono}}=\lambda_0/4$ one has $\underline{Z}_{in}\big|_{\substack{\text{mono}\\L_{\text{mono}}=\lambda_0/4}}=\frac{1}{2}\underline{Z}_{in}\big|_{\substack{\text{dip}\\L_{\text{dip}}=\lambda_0/2}}\approx\frac{73.1}{2}\approx 36.5\Omega$. Thus, a monopole antenna with height $L_{\text{mono}}\approx\lambda_0/4$ is resonant.

On the other hand, since the radiated fields are identical in the half-space $z>0$, the radiation intensity of the two antennas are identical for $z>0$:

$$U\big|_{\text{mono}}=\begin{cases}U\big|_{\text{dip}}, & 0<\theta<90^{\text{o}}\\ 0, & 90^{\text{o}}<\theta<180^{\text{o}}\end{cases}. \tag{2.44}$$

Since the monopole does not radiate to the lower-half space, it emits only half of the power of the equivalent dipole:

$$P_{rad}\big|_{\text{mono}}=\frac{1}{2}P_{rad}\Big|_{\text{dip}}. \tag{2.45}$$

Combining the two previous equations, one sees that the directive gain $g = 4\pi U / P_{rad}$ satisfies:

$$g\big|_{\text{mono}} = \begin{cases} \dfrac{4\pi U\big|_{\text{mono}}}{P_{rad}\big|_{\text{mono}}} = \dfrac{4\pi U\big|_{\text{dip}}}{P_{rad}\big|_{\text{dip}} / 2} = 2\, g\big|_{\text{dip}}, & \text{for } 0<\theta<90^{\text{o}} \\ 0, & \text{for } 90^{\text{o}}<\theta<180^{\text{o}} \end{cases} . \tag{2.46}$$

Thus, the directivity of the monopole satisfies:

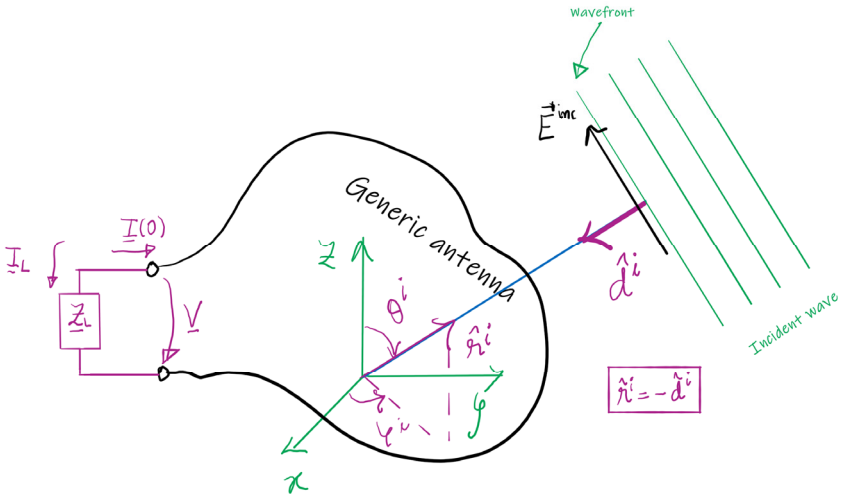
$$D\big|_{\text{mono}} = 2\, D\big|_{\text{dip}} . \tag{2.47}$$

For example, for $L_{\text{mono}} = \lambda_0 / 4$ the directivity of the monopole is $D\big|_{\text{mono}} = 2\, D\big|_{\text{dip},\, L_{\text{dip}}=\lambda_0/2} = 2 \times 1.64 = 3.28$. The monopole is more directive than the equivalent dipole because all the radiated energy is channeled to the $z > 0$ half-space.

3. Receiving antenna

3.1 The receiving antenna

So far, our study was focused on radiation problems where the antenna is regarded as an emitter. An excitation with a generator gives rise to the radiation of an electromagnetic wave that propagates in space. However, an antenna can also be used as "receiver" of some incident electromagnetic wave. The incident wave induces an electric signal at the antenna terminals, which are typically connected to a load $\underline{Z}_L = R_L + jX_L$.



Usually, in typical systems the incident wave can be approximated by a plane wave. A generic plane wave propagating in free-space has an electric field of the form

$$\underline{\mathbf{E}}^{\text{inc}} = \underline{\mathbf{E}}_0^{\text{inc}} e^{-jk_0 \mathbf{r}\cdot\hat{\mathbf{d}}^i}. \tag{3.1}$$

Here, $k_0 = \omega / c$ is the free-space wave number, $\underline{\mathbf{E}}_0^{\text{inc}}$ is a constant (complex) vector (which gives the electric field at the origin; the antenna terminals are supposed to be centered at the origin), $\mathbf{r} = (x, y, z)$ is the position vector, and $\hat{\mathbf{d}}^i$ is a unit vector that determines the direction of propagation of the wave.

The vector $\underline{\mathbf{E}}_0^{\text{inc}}$ is constrained by the condition $\underline{\mathbf{E}}_0^{\text{inc}} \cdot \hat{\mathbf{d}}^i = 0$, which guarantees that the electric field is orthogonal to the direction of propagation. Furthermore, the vector $\underline{\mathbf{E}}_0^{\text{inc}}$ determines the polarization state of the incident wave (e.g., circular, elliptical or linear polarization). The direction of arrival of the wave, as seen by the antenna, is

$$\hat{\mathbf{r}}^i = -\hat{\mathbf{d}}^i . \tag{3.2}$$

Since $\underline{\mathbf{E}}_0^{\text{inc}} \cdot \hat{\mathbf{r}}^i = -\underline{\mathbf{E}}_0^{\text{inc}} \cdot \hat{\mathbf{d}}^i = 0$ the vector $\underline{\mathbf{E}}_0^{\text{inc}}$ can be written in terms of the vectors $\hat{\boldsymbol{\theta}}^i, \hat{\boldsymbol{\varphi}}^i$ associated with the considered spherical coordinate system:

$$\underline{\mathbf{E}}_0^{\text{inc}} = \underline{E}_1^i \hat{\boldsymbol{\theta}}^i + \underline{E}_2^i \hat{\boldsymbol{\varphi}}^i . \tag{3.3}$$

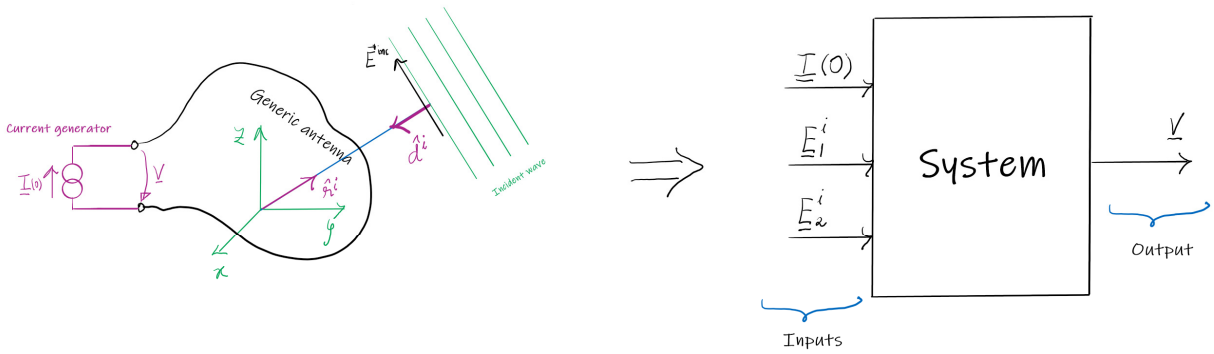
The incident electromagnetic wave induces currents on the materials that form the antenna. The antenna response depends on $\hat{\mathbf{r}}^i$ and $\underline{\mathbf{E}}_0^{\text{inc}}$, i.e., it depends on the direction of arrival of the incident wave, on the field amplitude, and on the polarization state.

In practice, one is mostly interested in understanding how the incident wave affects the voltage and current at the antenna terminals and finding the power delivered to the load.

3.2 Simultaneous excitation

To understand how an antenna reacts to an incident wave it is useful to consider the general scenario of a simultaneous excitation by a feeding current ($\underline{I}(0)$, enforced by some current generator) and by an incident plane wave ($\underline{\mathbf{E}}^{\text{inc}}$), so that the antenna is operated simultaneously as an emitter and a receiver.

Simultaneous excitation of the antenna



Thus, $\underline{I}(0)$ and $\underline{\mathbf{E}}_0^{\text{inc}}$ may be regarded as the inputs of the system, i.e., the parameters that determine the system response, and any other quantities of interest (e.g., the electromagnetic fields at some point of space) are ouputs. In particular, the voltage $\underline{V}$ at the antenna terminals can be seen as a function of $\underline{I}(0)$ and $\underline{\mathbf{E}}_0^{\text{inc}} = \underline{E}_1^i \hat{\boldsymbol{\theta}}^i + \underline{E}_2^i \hat{\boldsymbol{\varphi}}^i$:

$$\underline{V} = \underline{V}\left(\underline{I}(0), \underline{E}_1^i, \underline{E}_2^i\right). \tag{3.4}$$

Now the key point is that the Maxwell's equations are linear (it is also assumed that all the materials that form the antenna are linear). Due to this reason, the input-output relations in the frequency domain are forcibly linear, independent of the complexity of the antenna! Thus, it is possible to write:

$$\underline{V} = \underline{Z}_a \underline{I}(0) + h_1^r \underline{E}_1^i + h_2^r \underline{E}_2^i. \tag{3.5}$$

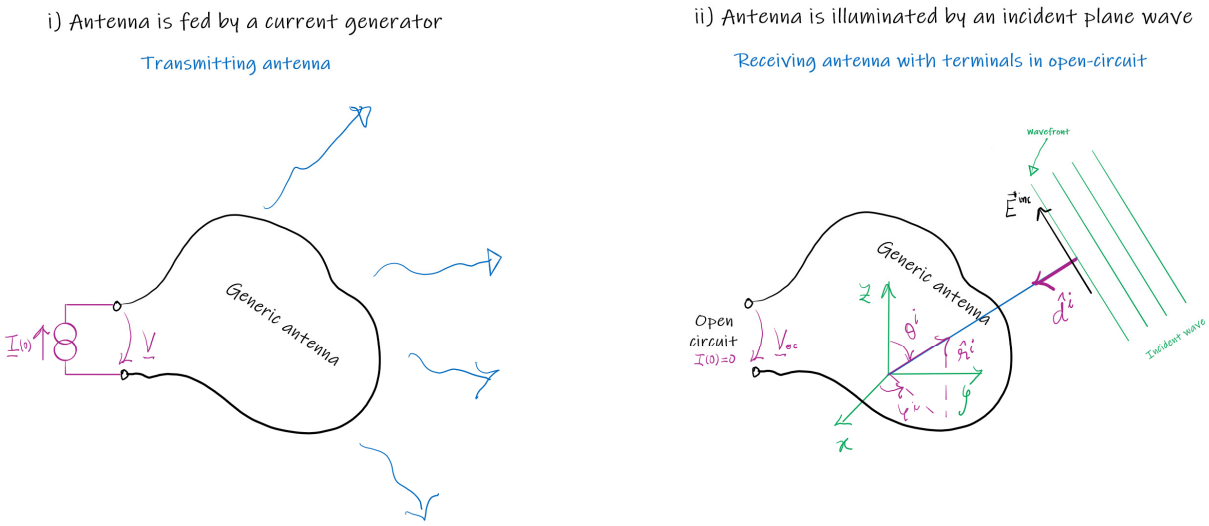
for some (unknown) constants $\underline{Z}_a, h_1^r, h_2^r$ that may depend on frequency and on the antenna. It is convenient to introduce the vector:

$$\mathbf{h}_e^r = h_1^r \hat{\boldsymbol{\theta}}^i + h_2^r \hat{\boldsymbol{\varphi}}^i. \tag{3.6}$$

The vector $\mathbf{h}_e^r$ is known as the *effective length* of the receiving antenna[1] and has units of length [m]. Note that $\mathbf{h}_e^r \cdot \hat{\mathbf{r}}^i = 0$. The voltage at the antenna terminals can be written as:

$$\boxed{\underline{V} = \underline{Z}_a \underline{I}(0) + \mathbf{h}_e^r \cdot \underline{\mathbf{E}}_0^{\text{inc}}}. \tag{3.7}$$

The formula is valid for arbitrary excitations and in particular we can consider two special cases: (i) the antenna is fed only by the current generator ($\underline{\mathbf{E}}_0^{\text{inc}} = 0$) and (ii) the current generator is turned off ($\underline{I}(0) = 0$) so that the antenna terminals are in open-circuit.



In case (i) it is obvious that the antenna is operated in transmitting mode. In this situation $\underline{V} = \underline{Z}_a \underline{I}(0)$, so that $\underline{Z}_a = \underline{V} / \underline{I}(0)$, which is nothing but the input impedance of the transmitting antenna studied in the earlier chapters:

$$\underline{Z}_a = \underline{Z}_{in} = \text{input impedance of the transmitting antenna}. \tag{3.8}$$

In case (ii), it is possible to identify $\underline{V} = \mathbf{h}_e^r \cdot \underline{\mathbf{E}}_0^{\text{inc}}$ as the voltage $\underline{V}_{oc}$ induced by the incident wave at the antenna terminals when they are in open-circuit:

$$\underline{V}_{oc} = \mathbf{h}_e^r \cdot \underline{\mathbf{E}}_0^{\text{inc}}. \tag{3.9}$$

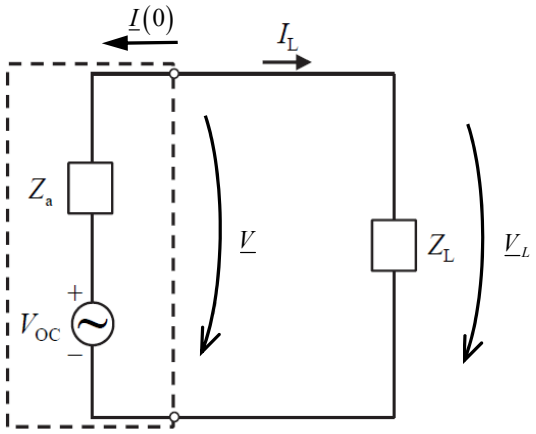
Thus, the general input-output relation (3.7) can be written as:

[1] In this definition it is implicit that the origin is coincident with the antenna terminals, so that the incident field is referred to the antenna terminals.

$$\boxed{\underline{V} = \underline{Z}_a \underline{I}(0) + \underline{V}_{oc}}. \tag{3.10}$$

3.3 Equivalent circuit of the receiving antenna

Consider again the scenario of Sect. 3.1, where the antenna terminals are connected to a load $\underline{Z}_L = R_L + jX_L$, rather than to a current generator. Due to the presence of the load the current entering into the antenna terminals $\underline{I}(0)$ can be nontrivial. Evidently, the antenna is indifferent to how the current $\underline{I}(0)$ is "generated". Its response is fully determined by the values of $\underline{I}(0)$ and of $\underline{\mathbf{E}}_0^{\text{inc}}$, independent of the internal structure of the (lumped) element connected to its terminals (a generator or a load). This means that Eq. (3.10) remains valid when the current generator is replaced by a load (see the figure in page 31). Based on this observation and on Eq. (3.10), we obtain the following equivalent circuit for the receiving antenna:



The current $\underline{I}_L = -\underline{I}(0)$ is the current flowing on the load. The equivalent circuit, despite its simplicity, characterizes fully (and exactly) the response of the antenna from the point of view of the load. The receiving antenna is equivalent to a generator with open-circuit voltage $\underline{V}_{oc} = \mathbf{h}_e^r \cdot \underline{\mathbf{E}}_0^{\text{inc}}$ and with internal impedance $\underline{Z}_a$. As previously mentioned, the internal impedance $\underline{Z}_a$ is nothing more than the impedance of the antenna operated as a transmitter. The voltage $\underline{V}_{oc}$

is proportional to the incident electric field. In general $\underline{V}_{oc}$ depends on the polarization of the incoming wave and depends on the direction of arrival of the wave $\underline{V}_{oc} = \underline{V}_{oc}\left(\hat{\mathbf{r}}^i\right) = \underline{V}_{oc}\left(\theta^i, \varphi^i\right)$.

3.4 Power delivered to a load and impedance matching

The power delivered to a load can be easily found using the equivalent circuit of the receiving antenna:

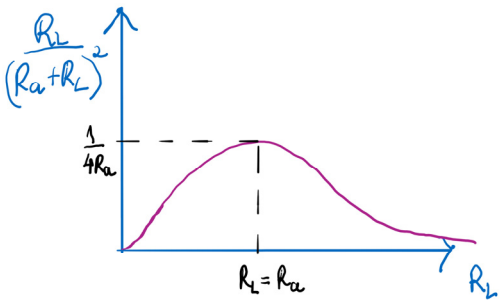
$$P_r = \frac{1}{2}\mathrm{Re}\left\{\underline{V}_L \underline{I}_L^*\right\} = \frac{1}{2}\mathrm{Re}\left\{\underline{Z}_L \left|\underline{I}_L\right|^2\right\} = \frac{1}{2} R_L \left|\underline{I}_L\right|^2 . \tag{3.11}$$

Since the current flowing on the load is $\underline{I}_L = \dfrac{\underline{V}_{oc}}{\underline{Z}_a + \underline{Z}_L}$, one gets:

$$P_r = \frac{1}{2}\frac{R_L}{\left|\underline{Z}_a + \underline{Z}_L\right|^2}\left|\underline{V}_{oc}\right|^2 = \frac{1}{2}\frac{R_L}{\left(R_a + R_L\right)^2 + \left(X_a + X_L\right)^2}\left|\underline{V}_{oc}\right|^2 . \tag{3.12}$$

We used $\underline{Z}_L = R_L + jX_L$ and $\underline{Z}_a = R_a + jX_a$.

For a given incident wave and antenna, the values of $\left|\underline{V}_{oc}\right|$ and $\underline{Z}_a = R_a + jX_a$ are fixed. In such a case, the power extracted from the incident field depends on the load: $P_r = P_r\left(R_L, X_L\right)$. The optimal load is the one that maximizes P_r. Evidently, to maximize P_r we need to ensure that (i) $X_a + X_L = 0$ and (ii) that $\dfrac{R_L}{\left(R_a + R_L\right)^2}$ is maximal. The plot of the function $\dfrac{R_L}{\left(R_a + R_L\right)^2}$ as a function of R_L is illustrated in the figure below.



The maximum is reached when $R_L = R_a$. The conjunction of (i) and (ii) shows that the optimal load impedance is a conjugated matched load:

$$\boxed{\left(\underline{Z}_L\right)_{\text{opt}} = \underline{Z}_a^* = R_a - jX_a}. \tag{3.13}$$

The power delivered to the optimal load is:

$$\left(P_r\right)_{\text{opt}} = \frac{\left|\underline{V}_{oc}\right|^2}{8R_a}. \tag{3.14}$$

Note that $\left(P_r\right)_{\text{opt}}$ is the maximum power that the antenna can extract from a given incident wave. Typically the wave that arrives at an antenna can be a rather weak signal and thus it is essential to guarantee a good impedance matching (if necessary using an impedance transformer).

The power delivered to a generic load can be written in terms of $\left(P_r\right)_{\text{opt}}$ as follows:

$$\boxed{P_r = C_i\left(P_r\right)_{\text{opt}}}. \tag{3.15}$$

where

$$\boxed{C_i = \frac{4R_aR_L}{\left(R_a + R_L\right)^2 + \left(X_a + X_L\right)^2}}. \tag{3.16}$$

is the *impedance matching coefficient*. The impedance matching coefficient satisfies:

$$0 \le C_i \le 1. \tag{3.17}$$

The maximum $C_i = 1$ is reached for a matched load ($\left(\underline{Z}_L\right)_{\text{opt}} = \underline{Z}_a^*$) and the minimum $C_i = 0$ for any purely reactive load.

3.5 Polarization matching, available power and effective area

Let us now analyze in greater detail the formula for the open-circuit voltage $\underline{V}_{oc} = \mathbf{h}_e^r \cdot \underline{\mathbf{E}}_0^{\text{inc}}$. It is clear from the properties of the inner product (Schwarz inequality):

$$\left|\underline{V}_{oc}\right| \le \left|\mathbf{h}_e^r\right|\left|\underline{\mathbf{E}}_0^{\text{inc}}\right|. \tag{3.18}$$

Note that for a generic complex vector $\mathbf{A}$ one has $|\mathbf{A}|^2 = \mathbf{A}\cdot\mathbf{A}^* = |A_x|^2 + |A_y|^2 + |A_z|^2$. The Schwarz inequality becomes an identity only when $\mathbf{h}_e^r \sim \underline{\mathbf{E}}_0^{\text{inc},*}$. Here the "tilda" symbol means that $\mathbf{h}_e^r$ is proportional to $\underline{\mathbf{E}}_0^{\text{inc},*}$. Notice that the electric field needs to be conjugated[1].

It is convenient to introduce the coefficient:

$$\boxed{C_p = \frac{\left|\mathbf{h}_e^r \cdot \underline{\mathbf{E}}_0^{\text{inc}}\right|^2}{\left|\mathbf{h}_e^r\right|^2 \left|\underline{\mathbf{E}}_0^{\text{inc}}\right|^2}}. \tag{3.19}$$

From the previous discussion, it is evident that $0 \le C_p \le 1$, and that

$$\left|\underline{V}_{oc}\right| = \left|\mathbf{h}_e^r\right|\left|\underline{\mathbf{E}}_0^{\text{inc}}\right|\sqrt{C_p}\,. \tag{3.20}$$

The coefficient C_p is totally insensitive to the intensity of the incident field (value of $\left|\underline{\mathbf{E}}_0^{\text{inc}}\right|$). However, it depends on the polarization state of the wave, which is determined by the "orientation" of the complex vector $\underline{\mathbf{E}}_0^{\text{inc}}$. Due to this reason C_p is known as the "polarization matching coefficient". When $C_p = 1$ the incoming wave is *polarization matched* to the antenna. The condition $C_p = 1$ is satisfied when:

$$\underline{\mathbf{E}}_0^{\text{inc}} \sim \mathbf{h}_e^{r,*}, \qquad \text{(polarization matching condition: } C_p = 1\text{)}. \tag{3.21}$$

The coefficient C_p describes how the wave polarization affects the received power. Indeed, from Eqs. (3.14)-(3.15), it is possible to write:

$$P_r = C_i C_p \frac{\left|\mathbf{h}_e^r\right|^2 \left|\underline{\mathbf{E}}_0^{\text{inc}}\right|^2}{8R_a}\,. \tag{3.22}$$

---

[1] When $\mathbf{h}_e^r = c\,\underline{\mathbf{E}}_0^{\text{inc},*}$ with c some constant, one has $\underline{V}_{oc} = \mathbf{h}_e^r \cdot \underline{\mathbf{E}}_0^{\text{inc}} = c\left|\underline{\mathbf{E}}_0^{\text{inc}}\right|^2$. Hence, $\left|\underline{V}_{oc}\right| = |c|\left|\underline{\mathbf{E}}_0^{\text{inc}}\right|\left|\underline{\mathbf{E}}_0^{\text{inc}}\right| = \left|\mathbf{h}_e^r\right|\left|\underline{\mathbf{E}}_0^{\text{inc}}\right|$, which confirms that the Schwarz inequality becomes an identity.

The received power can also be expressed in terms of the Poynting vector intensity $S^{\text{inc}} = \dfrac{\left|\underline{\mathbf{E}}_0^{\text{inc}}\right|^2}{2\eta_0}$ as follows:

$$\boxed{P_r = C_i C_p A_{ef} S^{\text{inc}}}, \tag{3.23}$$

where by definition

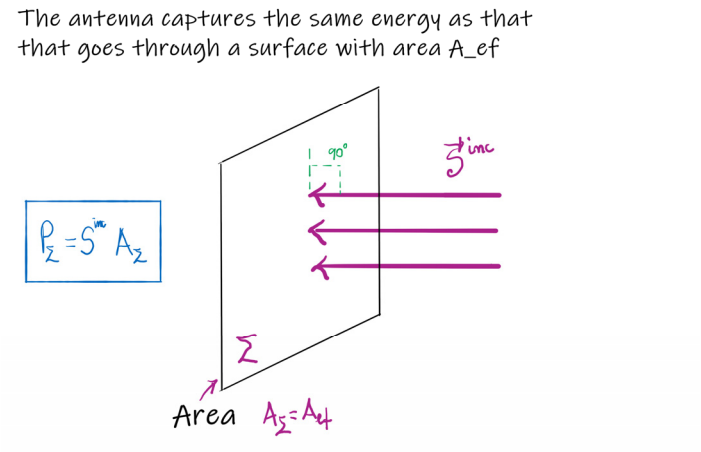
$$\boxed{A_{ef} = \frac{\eta_0 \left|\mathbf{h}_e^r\right|^2}{4R_a}} \tag{3.24}$$

is the effective area of the receiving antenna. Note that the effective area depends exclusively on the antenna parameters and is independent of the incident wave. Equation (3.23) shows that for a given intensity of the incident wave (S^{inc} fixed) the maximum power that can be collected by the antenna is

$$\boxed{P_a = A_{ef} S^{\text{inc}}}. \tag{3.25}$$

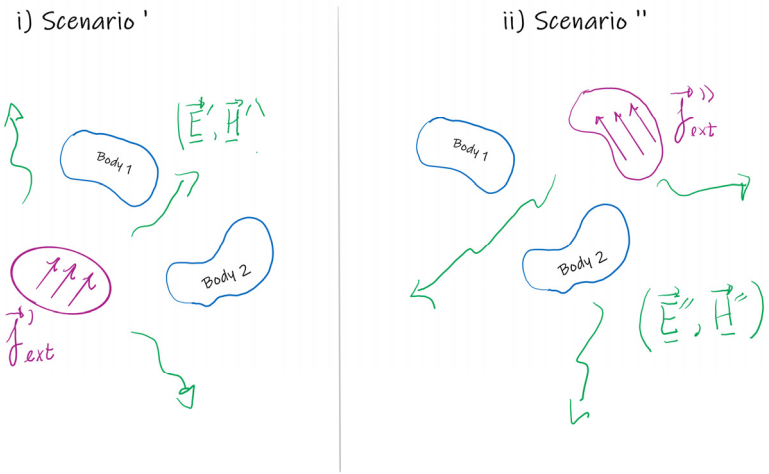
Due to this reason P_a is known as the *available power*. The received power is coincident with the available power when $C_i = 1$ (load is matched to the antenna) and when $C_p = 1$ (polarization of the wave is matched to the antenna). In other words, the received power is coincident with the available power when there is both polarization and impedance matching.

The parameter A_{ef} has units of area. It may be heuristically understood as the "area" from which the antenna can collect energy from. Recall that the power that goes through some surface Σ with area A_Σ is given by $P_\Sigma = \iint_\Sigma \mathbf{S}^{\text{inc}} \cdot \hat{\mathbf{n}}\, ds = S^{\text{inc}} A_\Sigma$. The effective area can be rather different from the physical area of the antenna, especially for electrically small antennas. For electrically large antennas the physical area and effective area are typically of the same order of magnitude.



3.6 Reciprocity theorem

One of the most remarkable results of antenna theory is that the properties of an antenna in receiving mode are strictly determined by the properties of the same antenna in transmitting mode. For example, if an antenna radiates efficiently in a certain direction of space then it also captures efficiently the radiation that arrives along the same direction. The theoretical basis of this fundamental result is the "reciprocity theorem" which is the topic of discussion of the present section.



The theorem considers an arbitrary physical environment, under two different excitations $\mathbf{j}'_{ext}$ and $\mathbf{j}''_{ext}$. The properties of the environment (which may be formed by arbitrarily shaped

dielectric bodies or conductors) are described by the space varying permittivity $\varepsilon = \varepsilon(x,y,z)$ and space varying permeability $\mu = \mu(x,y,z)$. The fields $(\underline{\mathbf{E}}', \underline{\mathbf{H}}')$ radiated by the (external) current density $\mathbf{j}'_{ext}$ satisfy the frequency domain Maxwell's equations:

$$\nabla \times \underline{\mathbf{E}}' = -j\omega\mu\underline{\mathbf{H}}', \qquad \nabla \times \underline{\mathbf{H}}' = \mathbf{j}'_{ext} + j\omega\varepsilon\underline{\mathbf{E}}' \tag{3.26a}$$

Similarly, the fields $(\underline{\mathbf{E}}'', \underline{\mathbf{H}}'')$ radiated in the second scenario with the excitation $\mathbf{j}''_{ext}$ satisfy:

$$\nabla \times \underline{\mathbf{E}}'' = -j\omega\mu\underline{\mathbf{H}}'', \qquad \nabla \times \underline{\mathbf{H}}'' = \mathbf{j}''_{ext} + j\omega\varepsilon\underline{\mathbf{E}}'' \tag{3.26b}$$

The reciprocity theorem in differential form establishes that the two field distributions (associated with different excitations in the same physical environment) are constrained by:

$$\boxed{\nabla \cdot \left\{ \underline{\mathbf{E}}' \times \underline{\mathbf{H}}'' - \underline{\mathbf{E}}'' \times \underline{\mathbf{H}}' \right\} = -\underline{\mathbf{E}}' \cdot \mathbf{j}''_{ext} + \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext}}. \tag{3.27}$$

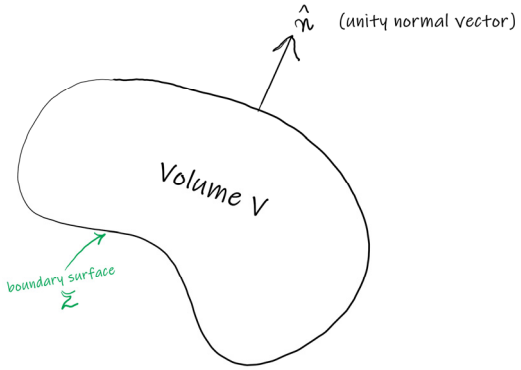
The proof of the theorem is relatively straightforward. Using the vector identity $\nabla \cdot \{\mathbf{A} \times \mathbf{B}\} = -\mathbf{A} \cdot \nabla \times \mathbf{B} + \mathbf{B} \cdot \nabla \times \mathbf{A}$ and the Maxwell equations (3.26) one gets:

$$\begin{aligned} \nabla \cdot \left\{ \underline{\mathbf{E}}' \times \underline{\mathbf{H}}'' \right\} &= -\underline{\mathbf{E}}' \cdot \nabla \times \underline{\mathbf{H}}'' + \underline{\mathbf{H}}'' \cdot \nabla \times \underline{\mathbf{E}}' \\ &= -\underline{\mathbf{E}}' \cdot \left(\mathbf{j}''_{ext} + j\omega\varepsilon\underline{\mathbf{E}}'' \right) + \underline{\mathbf{H}}'' \cdot \left(-j\omega\mu\underline{\mathbf{H}}' \right). \\ &= -\underline{\mathbf{E}}' \cdot \mathbf{j}''_{ext} - j\omega\varepsilon\underline{\mathbf{E}}' \cdot \underline{\mathbf{E}}'' - j\omega\mu\underline{\mathbf{H}}' \cdot \underline{\mathbf{H}}'' \end{aligned} \tag{3.28}$$

By interchanging the symbols ' and " one readily sees that $\nabla \cdot \left\{ \underline{\mathbf{E}}'' \times \underline{\mathbf{H}}' \right\} = -\underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext} - j\omega\varepsilon\underline{\mathbf{E}}' \cdot \underline{\mathbf{E}}'' - j\omega\mu\underline{\mathbf{H}}' \cdot \underline{\mathbf{H}}''$. Notice that the last two terms on the right-hand side of the equation are unchanged when the symbols ' and " are interchanged. Subtracting the two equations member by member one obtains Eq. (3.27), which proves the theorem.

A more enlightening version of the theorem is obtained by integrating both sides of Eq. (3.27) over some arbitrarily shaped volume *V*. From Gauss' theorem, the integral of the divergence of a vector field is identical to the flux over the boundary surface: $\iiint_V \nabla \cdot \mathbf{F} = \iint_\Sigma \mathbf{F} \cdot \hat{\mathbf{n}}\, ds$. This property yields the reciprocity theorem in integral form:

$$\iint_\Sigma ds\hat{\mathbf{n}} \cdot \left\{ \underline{\mathbf{E}}' \times \underline{\mathbf{H}}'' - \underline{\mathbf{E}}'' \times \underline{\mathbf{H}}' \right\} = \iiint_V dV \left(-\underline{\mathbf{E}}' \cdot \mathbf{j}''_{ext} + \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext} \right). \tag{3.29}$$



In the context of antenna theory, it is especially interesting to take the volume V as the entire space. In such a case, the contribution from the surface integral vanishes. To prove this, take V as a spherical surface of radius R centered at the origin of the coordinate system. In this case, the unity normal vector is the radial vector $\hat{\mathbf{n}} = \hat{\mathbf{r}}$. For a sufficiently large R, the fields can be approximated by the far-field formulas [Eq. (1.20)]: $\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}} = \eta_0 \underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}} \times \hat{\mathbf{r}}$ and $\underline{\mathbf{H}}(\mathbf{r})\big|_{\text{far-field}} = \dfrac{1}{\eta_0}\hat{\mathbf{r}} \times \underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}}$. Hence, we have that:

$$\hat{\mathbf{n}} \cdot \left(\underline{\mathbf{E}}' \times \underline{\mathbf{H}}''\right)\big|_{\text{far-field}} = \hat{\mathbf{r}} \cdot \left(\underline{\mathbf{E}}' \times \left(\frac{1}{\eta_0}\hat{\mathbf{r}} \times \underline{\mathbf{E}}''\right)\right) = \frac{1}{\eta_0}\underline{\mathbf{E}}' \cdot \underline{\mathbf{E}}''. \tag{3.30}$$

The last identity is obtained using $\mathbf{A} \times (\mathbf{B} \times \mathbf{C}) = \mathbf{B}(\mathbf{A} \cdot \mathbf{C}) - \mathbf{C}(\mathbf{A} \cdot \mathbf{B})$ and taking into account that $\hat{\mathbf{r}} \cdot \underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}} = 0$. But interchanging the ' and " symbols does not change the right-hand side of the above equation, and hence it follows that $\hat{\mathbf{n}} \cdot \left(\underline{\mathbf{E}}' \times \underline{\mathbf{H}}'' - \underline{\mathbf{E}}'' \times \underline{\mathbf{H}}'\right)\big|_{\text{far-field}} = 0$ (the correction term is on the order of $1/R^3$, which approaches zero faster than the area of the spherical surface approaches infinity). This confirms that the surface integral vanishes when $R \to \infty$. Thereby, we proved that:

$$\boxed{\iiint\limits_{\text{all space}} dV \left(-\underline{\mathbf{E}}' \cdot \mathbf{j}''_{ext} + \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext}\right) = 0}. \tag{3.31}$$

This is a rather remarkable result: the fields radiated by two different and totally arbitrary current distributions in the same physical environment are not independent, and are required to satisfy

the above reciprocity constraint! Note that Eq. (3.31) is only applicable when the current distributions are localized in some bounded volume of space, so that the far-field approximation can be used.

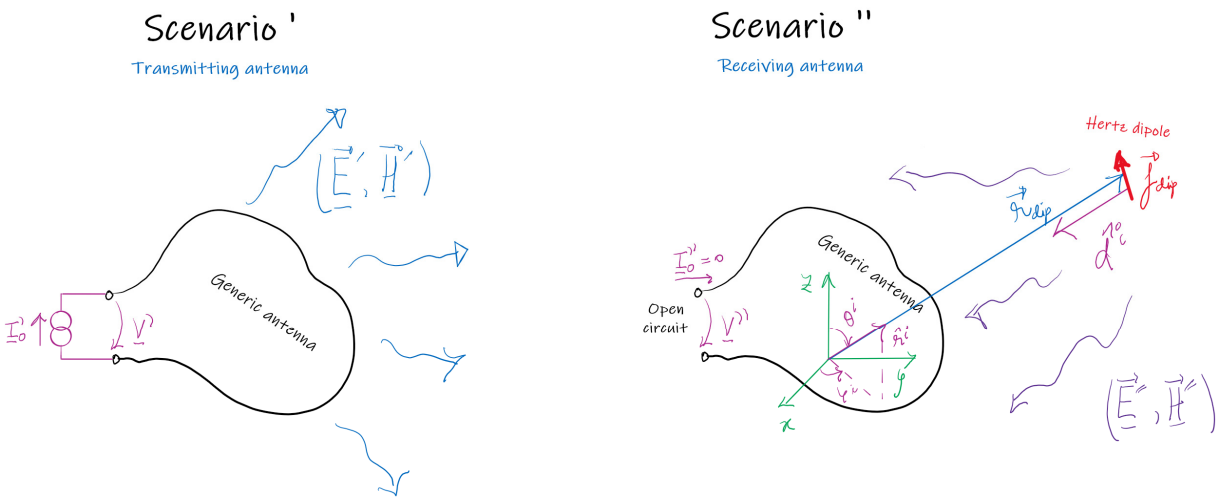
3.7 Application of the reciprocity theorem to antenna theory

We will now use the reciprocity theorem to prove one of the key results of antenna theory: the effective height of an antenna in transmitting mode is identical to the effective height of the same antenna in receiving mode:

$$\boxed{\mathbf{h}_e^r = \mathbf{h}_e} \tag{3.32}$$

We recall that $\mathbf{h}_e^r$ determines the voltage induced at the (receiving) antenna terminals when they are in open-circuit [see Eq. (3.9)], whereas $\mathbf{h}_e$ determines the far-field of the transmitting antenna [see Eq. (1.17)].

To prove that $\mathbf{h}_e^r = \mathbf{h}_e$ we apply the reciprocity theorem to the scenarios illustrated in the figure below.



The scenario on the left (' scenario) considers an arbitrary antenna fed by a current generator (transmitting antenna). The antenna radiates the fields $(\underline{\mathbf{E}}', \underline{\mathbf{H}}')$. The gap distance between the antenna terminals is supposed to be *dl* (very small compared to the wavelength). We suppose

that the antenna terminals are centered at the origin and that the current $\underline{I}'_0$ flows along z. Therefore, the external excitation in the first scenario is described by:

$$\mathbf{j}'_{ext} = \underline{I}'_0 dl\, \hat{\mathbf{z}}\delta(\mathbf{r}). \tag{3.33}$$

Note that $\mathbf{j}'_{ext}$ is the same as for an Hertz dipole (current distribution is constant over a length dl).

On the other hand, the second scenario (" scenario) depicts the same antenna with its terminals in open-circuit. The antenna is illuminated by the fields created by a Hertz dipole positioned at the point $\mathbf{r}_{\text{dip}}$. The Hertz dipole is in the far-field of the antenna and is modeled by a current distribution of the form:

$$\mathbf{j}''_{ext} = \mathbf{j}_{\text{dip}}\delta(\mathbf{r} - \mathbf{r}_{\text{dip}}). \tag{3.34}$$

Here, $\mathbf{j}_{\text{dip}}$ is an arbitrary vector that depends on the orientation of the Hertz dipole and on the feeding current. The total fields in this second scenario are $(\underline{\mathbf{E}}'', \underline{\mathbf{H}}'')$. They can be seen as the superposition of the primary field radiated by the Hertz dipole alone in free-space ($\underline{\mathbf{E}}''^{\text{inc}}$) plus the field back-scattered by the antenna ($\underline{\mathbf{E}}''^{\text{bs}}$), so that $\underline{\mathbf{E}}'' = \underline{\mathbf{E}}''^{\text{inc}} + \underline{\mathbf{E}}''^{\text{bs}}$.

Applying the reciprocity theorem [Eq. (3.31)] to the ' and " scenarios one finds that:

$$-\underline{\mathbf{E}}'(\mathbf{r}_{\text{dip}}) \cdot \mathbf{j}_{\text{dip}} + \underline{I}'_0 dl\, \underline{\mathbf{E}}''(0) \cdot \hat{\mathbf{z}} = 0. \tag{3.35}$$

The voltage induced at the antenna terminals in the second scenario is $\underline{V}'' = -\int_{-dl/2}^{dl/2} \mathbf{E}''\big|_{\text{gap}} \cdot \mathbf{dl} \approx -dl\, \underline{\mathbf{E}}''(0) \cdot \hat{\mathbf{z}}$. Taking into account that $\underline{V}''$ is nothing more than the open-circuit voltage at the antenna terminals ($\underline{V}'' = \underline{V}_{oc}$), we find that:

$$\underline{I}'_0 \underline{V}_{oc} = -\underline{\mathbf{E}}'(\mathbf{r}_{\text{dip}}) \cdot \mathbf{j}_{\text{dip}}. \tag{3.36}$$

Next, we note that since the Hertz dipole is in the far-field of the antenna one can evaluate $\underline{\mathbf{E}}'\left(\mathbf{r}_{\text{dip}}\right)$ using Eq. (1.17), i.e., $\underline{\mathbf{E}}'\left(\mathbf{r}_{\text{dip}}\right) \approx \eta_0 jk_0 \underline{I}'_0 \mathbf{h}_e\left(\hat{\mathbf{r}}^i\right)\dfrac{e^{-jk_0 r_{\text{dip}}}}{4\pi r_{\text{dip}}}$ with $r_{\text{dip}} = \left|\mathbf{r}_{\text{dip}}\right|$ the distance between the antenna terminals and the Hertz dipole. This implies that:

$$\underline{V}_{oc} = -\eta_0 jk_0 \mathbf{h}_e\left(\hat{\mathbf{r}}^i\right)\frac{e^{-jk_0 r_{\text{dip}}}}{4\pi r_{\text{dip}}} \cdot \mathbf{j}_{\text{dip}} \,. \tag{3.37}$$

Since $\mathbf{h}_e \cdot \hat{\mathbf{r}}^i = 0$ it is possible to write $\mathbf{h}_e \cdot \mathbf{j}_{\text{dip}} = \mathbf{h}_e \cdot \left[\mathbf{j}_{\text{dip}} - \hat{\mathbf{r}}^i\left(\hat{\mathbf{r}}^i \cdot \mathbf{j}_{\text{dip}}\right)\right] = \mathbf{h}_e \cdot \left[\hat{\mathbf{r}}^i \times \left(\mathbf{j}_{\text{dip}} \times \hat{\mathbf{r}}^i\right)\right]$. Hence, the voltage at the terminals of the receiving antenna is given by:

$$\underline{V}_{oc} = \mathbf{h}_e\left(\hat{\mathbf{r}}^i\right) \cdot \left[\eta_0 jk_0 \left(\mathbf{j}_{\text{dip}} \times \hat{\mathbf{r}}^i\right) \times \hat{\mathbf{r}}^i \frac{e^{-jk_0 r_{\text{dip}}}}{4\pi r_{\text{dip}}}\right]. \tag{3.38}$$

To proceed, we calculate the complex amplitude of the incident field $\underline{\mathbf{E}}''^{\text{inc}}$ radiated by the Hertz dipole ($\underline{\mathbf{E}}''^{\text{inc}}$) evaluated the receiving antenna terminals (i.e., at the origin). It is given by:

$$\begin{aligned} \underline{\mathbf{E}}''^{\text{inc}}_0 \equiv \underline{\mathbf{E}}''^{\text{inc}}\Big|_{\substack{\text{Hertz field} \\ \text{evaluated at the origin}}} &\approx \eta_0 jk_0 \left[\underline{I}^{\text{Hertz}}\left(0\right)\mathbf{h}_e^{\text{Hertz}}\left(-\hat{\mathbf{r}}^i\right)\right]\frac{e^{-jk_0 r_{\text{dip}}}}{4\pi r_{\text{dip}}} \\ &= \eta_0 jk_0 \left(\mathbf{j}_{\text{dip}} \times \hat{\mathbf{r}}^i\right) \times \hat{\mathbf{r}}^i \frac{e^{-jk_0 r_{\text{dip}}}}{4\pi r_{\text{dip}}} \end{aligned} \tag{3.39}$$

The identity in the first line follows from Eq. (1.17) and the identity in the second line from Eq. (1.15) evaluated with the current distribution of the Hertz dipole $\mathbf{j} = \mathbf{j}_{\text{dip}}\delta\left(\mathbf{r}\right)$ (referred to a coordinate system centered at the dipole position). In the above, $\mathbf{h}_e^{\text{Hertz}}$ is the effective length of the Hertz dipole and $\underline{I}^{\text{Hertz}}\left(0\right)$ is the feeding current. From Eq. (1.15), one can see that $\underline{I}^{\text{Hertz}}\left(0\right)\mathbf{h}_e^{\text{Hertz}}\left(\hat{\mathbf{r}}\right) = \left(\mathbf{j}_{\text{dip}} \times \hat{\mathbf{r}}\right) \times \hat{\mathbf{r}}$. The Hertz dipole effective length is evaluated along $-\hat{\mathbf{r}}^i$, as this is the direction along which the Hertz dipole sees the receiving antenna. Comparing Eqs. (3.38)-(3.39) it is clear that:

$$\underline{V}_{oc} = \mathbf{h}_e\left(\hat{\mathbf{r}}^i\right) \cdot \underline{\mathbf{E}}''^{\text{inc}}_0 \,. \tag{3.40}$$

This shows that the voltage induced on the terminals of the receiving antenna in scenario " is determined by the effective length of the transmitting antenna ($\mathbf{h}_e\left(\hat{\mathbf{r}}^i\right)$) and by the complex amplitude of the incident wave evaluated at the antenna terminals ($\underline{\mathbf{E}}_0''^{\text{inc}}$). Comparing with the definition of the effective length of the receiving antenna [Eq. (3.9)] and taking into account that the incident field is arbitrary, it follows that $\mathbf{h}_e^r\left(\hat{\mathbf{r}}^i\right) = \mathbf{h}_e\left(\hat{\mathbf{r}}^i\right)$ as we wanted to prove.

3.8 Relation between the effective area and power gain

Using $\mathbf{h}_e^r = \mathbf{h}_e$ it is possible to link the effective area of an antenna A_{ef} (which describes how well it captures power in receiving mode) with the power gain G (which describes how directive is the transmitting antenna with respect to an isotropic emitter).

The power gain is determined by:

$$G = \frac{4\pi U}{P_{in}} = \frac{4\pi U}{\frac{1}{2} R_a \left|\underline{I}(0)\right|^2}, \tag{3.41}$$

where U is the radiation intensity and R_a is the input resistance of the antenna ($R_a = R_{in} = \text{Re}\{\underline{Z}_{in}\}$). Thus, using Eq. (1.24) one finds:

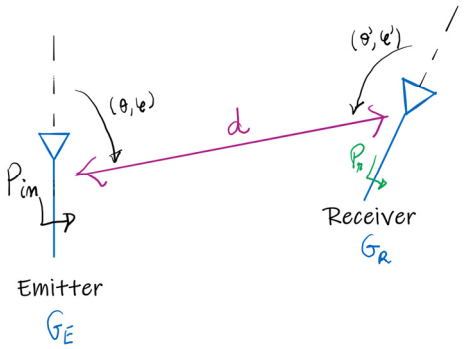
$$G = \pi\eta_0 \frac{\left|\mathbf{h}_e\right|^2}{R_a \lambda_0^2}. \tag{3.42}$$

Comparing the above formula with Eq. (3.24) and using the reciprocity result $\mathbf{h}_e^r = \mathbf{h}_e$, one obtains the universal relation (valid for any antenna) between the effective area and the power gain:

$$\boxed{A_{ef} = \frac{\lambda_0^2}{4\pi} G}. \tag{3.43}$$

3.9 *Friis formula*

Consider a scenario where the fields radiated by a generic emitter illuminate a generic receiving antenna. Using the theory developed so far, it is straightforward to relate the input power (P_{in}) accepted by the emitter with the power captured by the receiver (P_r) (see the figure). It is supposed that the two antennas are in the far-field of one another. The distance between the antennas is d.



The Poynting vector associated with the field radiated by the emitter is determined by the power gain of the antenna (G_E)

$$S^{\text{inc}} = \frac{P_{in}}{4\pi r^2} G_E(\theta,\varphi). \tag{3.44}$$

On the other end, the power captured by the receiving antenna is determined by its effective area $P_r = C_i C_p A_{ef,R}(\theta',\varphi') S^{\text{inc}}\big|_{r=d}$. This means that:

$$P_r = \frac{P_{in}}{4\pi d^2} G_E(\theta,\varphi) A_{ef,R}(\theta',\varphi') C_i C_p. \tag{3.45}$$

Using the universal relation between the power gain and the effective area the same result can be written as:

$$\boxed{P_r = P_{in}\left(\frac{\lambda_0}{4\pi d}\right)^2 G_E(\theta,\varphi) G_R(\theta',\varphi') C_i C_p = \frac{P_{in}}{(\lambda_0 d)^2} A_{ef,E}(\theta,\varphi) A_{ef,R}(\theta',\varphi') C_i C_p}. \tag{3.46}$$

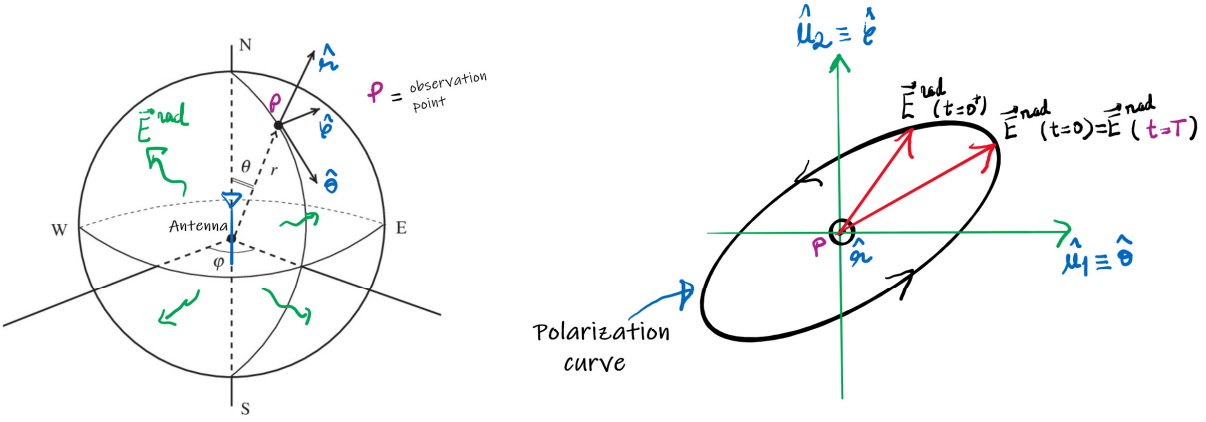
In the particular case of polarization and impedance matching, one finds that the power available at the terminals of the receiving antenna is $P_r \equiv P_a = P_{in}\left(\frac{\lambda_0}{4\pi d}\right)^2 G_E(\theta,\varphi) G_R(\theta',\varphi')$. This is known as the Friis formula.

3.10 Antenna polarization

Consider a generic antenna operated as a transmitter. The far-field of the antenna is determined by its effective length [see Eq. (1.17)]. In particular, the field radiated by the antenna along the direction $\hat{\mathbf{r}}$ is proportional to the effective length: $\underline{\mathbf{E}}^{\text{rad}} \sim \mathbf{h}_e(\hat{\mathbf{r}})$. Evidently, the instantaneous field is:

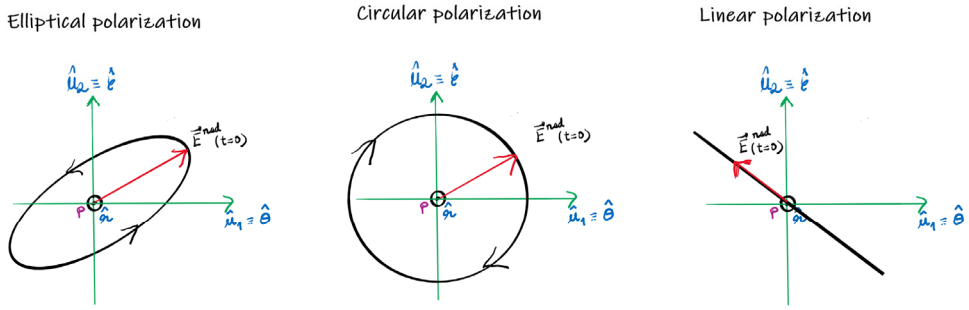
$$\mathbf{E}^{\text{rad}}(t) = \text{Re}\left\{\underline{\mathbf{E}}^{\text{rad}} e^{j\omega t}\right\} \sim \text{Re}\left\{\mathbf{h}_e(\hat{\mathbf{r}}) e^{j\omega t}\right\}. \tag{3.47}$$

(For simplicity in the last identity it was assumed that the proportionality constant that relates $\underline{\mathbf{E}}^{\text{rad}}$ and $\mathbf{h}_e(\hat{\mathbf{r}})$ is real-valued). As discussed in Sect. 1.2, the effective length is a vector of the form $\mathbf{h}_e(\hat{\mathbf{r}}) \equiv h_{e,\theta}\hat{\boldsymbol{\theta}} + h_{e,\varphi}\hat{\boldsymbol{\varphi}}$ because $\mathbf{h}_e(\hat{\mathbf{r}}) \cdot \hat{\mathbf{r}} = 0$. This guarantees that the electric field is orthogonal to the direction of the propagation ($\hat{\mathbf{r}}$).



The closed curve determined by the time evolution of $\mathbf{E}^{\text{rad}}(t)$ in a full time cycle (with period $T = 2\pi/\omega$) is known as the *polarization curve*. The polarization curve is perpendicular to the

direction of propagation of the wave. Because of the orthogonality condition $\mathbf{h}_e(\hat{\mathbf{r}})\cdot\hat{\mathbf{r}}=0$ the polarization curve is generically an ellipse, with the particular (degenerate) cases of a circumference and of a line segment. The polarization of the antenna along the direction $\hat{\mathbf{r}}$ is said to be elliptical (circular, linear) when the corresponding polarization curve is an ellipse (circumference, line segment). Note that the antenna polarization is direction dependent.



For the cases of elliptical and circular polarization, there are two different polarization states for the same polarization curve. Indeed, there are two possible trajectories for the electric field, corresponding respectively to clockwise and anticlockwise rotations. The sense of rotation of the electric field is determined using the "right-hand rule". This is done by imitating the trajectory of the electric field with the "right-hand" and by finding the corresponding orientation of the "thumb". When the "thumb" direction is coincident with the direction of propagation ($\hat{\mathbf{r}}$), the field is said to rotate to the "right". Otherwise, it rotates to the "left". For example, the leftmost panel of the above figure corresponds to a "right" elliptical polarization, whereas the middle panel to a left circular polarization.

Let us write the complex amplitude of the electric field as:

$$\underline{\mathbf{E}}=\underline{E}_1\hat{\mathbf{u}}_1+\underline{E}_2\hat{\mathbf{u}}_2\,. \tag{3.48}$$

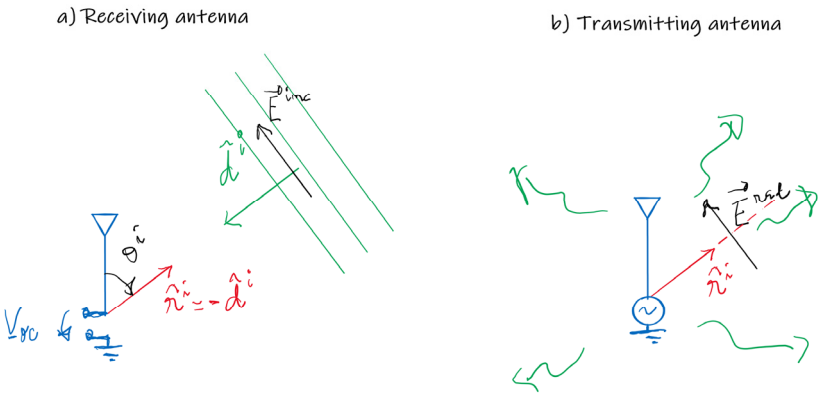
with $\hat{\mathbf{u}}_1\cdot\hat{\mathbf{u}}_2=0$. It is assumed that the direction of propagation of the wave is determined by $\hat{\mathbf{d}}=\hat{\mathbf{u}}_1\times\hat{\mathbf{u}}_2$. For antenna problems, one typically has $\hat{\mathbf{u}}_1=\hat{\boldsymbol{\theta}}$, $\hat{\mathbf{u}}_2=\hat{\boldsymbol{\varphi}}$ and $\hat{\mathbf{d}}=\hat{\mathbf{r}}$. Then, it is easy to prove the following:

- When $|\underline{E}_1| = |\underline{E}_2|$ *and* $\arg \underline{E}_1 = \arg \underline{E}_2 \pm 90^\text{o}$ the polarization state is circular.
- When $\underline{E}_1 = 0$ *or* $\underline{E}_2 = 0$ *or* $\arg \underline{E}_1 = \arg \underline{E}_2$ *or* $\arg \underline{E}_1 = \arg \underline{E}_2 + 180^\text{o}$ the polarization state is linear.
- If none of the above holds true, then the polarization state is elliptical.

Furthermore, in the case of circular polarization it can be shown that when $\arg \underline{E}_2 = \arg \underline{E}_1 + 90^\text{o}$ the wave is left circularly polarized (LCP), whereas when $\arg \underline{E}_2 = \arg \underline{E}_1 - 90^\text{o}$ the wave is right circularly polarized (RCP). It is underlined that this result holds true only when the direction of propagation is $\hat{\mathbf{d}} = \hat{\mathbf{u}}_1 \times \hat{\mathbf{u}}_2$. A general and simple way to find the sense of rotation is to determine the time evolution (trajectory) of the electric field on the polarization curve, e.g., by calculating $\mathbf{E}(t=0)$ and $\mathbf{E}(t=0^+)$.

3.11 Optimal polarization for the incident wave

Consider a receiving antenna with terminals in open-circuit illuminated by a plane wave (left panel of the figure below). As previously discussed, the voltage induced at the antenna terminals is $\underline{V}_{oc} = \mathbf{h}_e \cdot \underline{\mathbf{E}}_0^{\text{inc}}$ with $\mathbf{h}_e = \mathbf{h}_e^r$ the effective length of the antenna and $\underline{\mathbf{E}}_0^{\text{inc}}$ the complex amplitude of the incident field evaluated at the antenna terminals. As proven in Sect. 3.5, the open-circuit voltage depends on the polarization of the incident wave, and for a fixed intensity of the energy flux of the incoming wave (fixed value of S^{inc}) the amplitude $|\underline{V}_{oc}|$ is maximized when the polarization of the incident field is such that $\underline{\mathbf{E}}_0^{\text{inc}} \sim \mathbf{h}_e^*$ [Eq. (3.21)]. In this case, the incident wave is polarization matched to the antenna (optimal polarization).



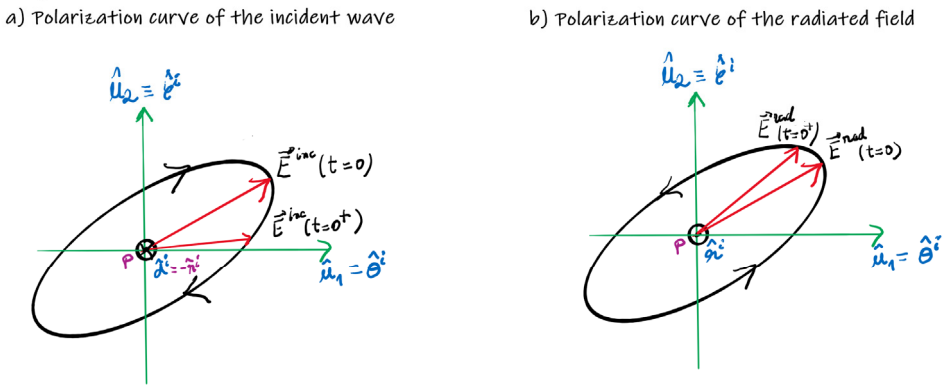
It is interesting to analyze the polarization curve of the incident field for the optimal case. The instantaneous incident field is such that:

$$\mathbf{E}^{\text{inc}}(t) = \text{Re}\left\{\underline{\mathbf{E}}_0^{\text{inc}} e^{j\omega t}\right\} \sim \text{Re}\left\{\mathbf{h}_e^* e^{j\omega t}\right\}. \tag{3.49}$$

Taking into account that $\text{Re}\{\mathbf{A}\} = \text{Re}\{\mathbf{A}^*\}$ for a generic vector $\mathbf{A}$, one finds with the help of Eq. (3.47) that $\mathbf{E}^{\text{inc}}(t) \sim \text{Re}\left\{\mathbf{h}_e e^{-j\omega t}\right\} \sim \mathbf{E}^{\text{rad}}(-t)$, i.e., the optimal incident field is related to the field radiated by antenna in transmitting mode as:

$$\mathbf{E}^{\text{inc}}(t) \sim \mathbf{E}^{\text{rad}}(-t). \tag{3.50}$$

Thus, the polarization curve of the (optimal) incident wave is identical to the polarization curve of the transmitting antenna, apart from some irrelevant scale factor. However, the trajectory of the electric field is different in the two cases because the time is flipped $t \to -t$ (see an example for in the figure below).

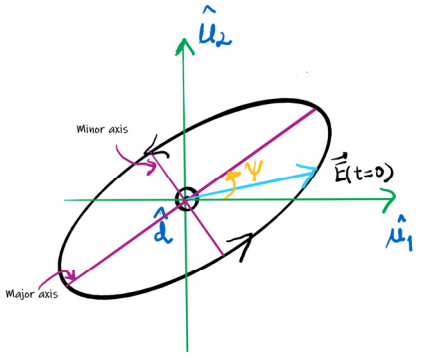


Even though the absolute physical sense of rotation of the electric field is different in two cases, the state of polarization of the two waves ($\mathbf{E}^{\text{inc}}$ and $\mathbf{E}^{\text{rad}}$) is identical! For example, in the case of the figure both curves are associated with a right-elliptical polarization. The reason is that the direction of propagation is also flipped because $\hat{\mathbf{d}}^i$ (the direction of the propagation of the incoming wave) is related to the direction of observation as $\hat{\mathbf{d}}^i = -\hat{\mathbf{r}}^i$. Thus, it follows that the optimal polarization of the incident wave is the *same* as the antenna polarization! For example, if an antenna radiates a left circularly polarized wave in some direction of space and then the optimal polarization for the incoming wave when the antenna is used as a receiver is also a left circular polarization.

It is somewhat intriguing that the absolute sense of rotation of the optimal polarization $\mathbf{E}^{\text{inc}}$ and that of $\mathbf{E}^{\text{rad}}$ are different. The physical justification is rooted in the invariance of the Maxwell's equations under a time reversal $t \to -t$ (see the problem set 1). Specifically, if some field distribution satisfies the Maxwell's equations, then the time-reversed field distribution also does. Under a time reversal the radiated fields are returned back to the antenna and thereby are associated with the optimal polarization state ($\mathbf{E}^{\text{inc}}(t) \sim \mathbf{E}^{\text{rad}}(-t)$)!

3.12 The Poincaré sphere

The polarization state of a wave can be geometrically represented on a sphere of unity radius, known as the Poincaré sphere. The coordinates of the polarization state on the sphere are found from two parameters AR and ψ determined by the polarization curve.



As previously discussed, the polarization curve is generically an ellipse and lies on the plane generated by two vectors $\hat{\mathbf{u}}_1$ and $\hat{\mathbf{u}}_2$ perpendicular to the direction of propagation. By definition, $0 \le \psi < 180^{\circ}$ is the tilt angle determined by the major axis of the ellipse and some reference axis (typically $\hat{\mathbf{u}}_1$). On the other hand, AR is the axial ratio given by the quotient of the principal axes of the polarization ellipse:

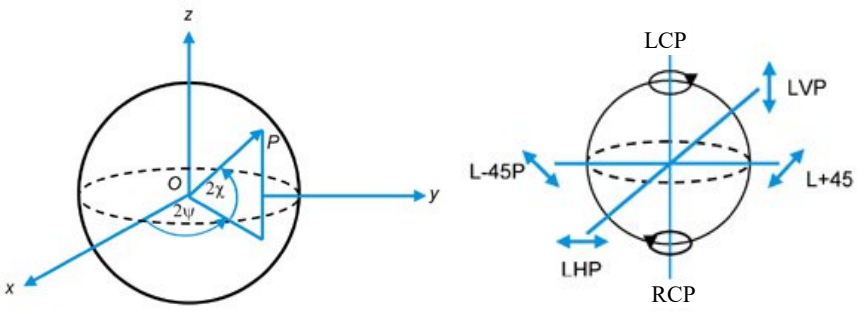
$$\text{AR} = \frac{\text{major axis}}{\text{minor axis}} \ge 1. \tag{3.51}$$

We introduce an angle $-45^{\circ} \le \chi \le 45^{\circ}$ defined by:

$$\cot \chi = \pm \text{AR}, \tag{3.52}$$

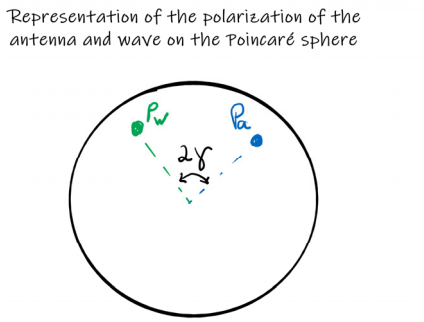
with the "+" sign ("-" sign) picked when the sense of rotation of the polarization curve is to the "left" ("right"). The coordinates of the polarization state on the Poincaré sphere are (see the figure for a geometrical interpretation):

$$\text{pol. state} \rightarrow \left(\cos 2\psi \cos 2\chi, \sin 2\psi \cos 2\chi, \sin 2\chi\right). \tag{3.53}$$



The poles of the Poincaré sphere correspond to the two circularly polarized states (LCP and RCP) and the equator of the sphere is associated with the different linearly polarized states.

The "polarization matching coefficient" C_p has an interesting geometrical interpretation in the Poincaré sphere.



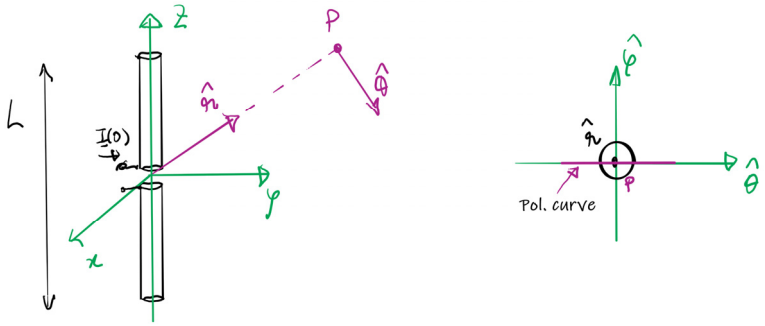
Specifically, suppose that both the polarization states of the antenna ($\mathbf{P}_a$) and of the wave ($\mathbf{P}_w$) are represented on the Poincaré sphere, as shown above. Let 2γ be the angle defined by $\mathbf{P}_a$ and $\mathbf{P}_w$. Then, it can be shown that:

$$C_p = \cos^2 \gamma \,. \tag{3.54}$$

Thus, using the Poincaré sphere one can visualize how well is the polarization of the incoming wave matched to the antenna. The polarization matching is good when the geometrical distance between $\mathbf{P}_a$ and $\mathbf{P}_w$ is small. The optimal polarization corresponds to the case $\mathbf{P}_a = \mathbf{P}_w$ ($C_p = 1, \;\; 2\gamma = 0^o$). The worst "polarization" scenario is when $\mathbf{P}_a$ and $\mathbf{P}_w$ are antipodal points in the Poincaré sphere ($C_p = 0, \;\; 2\gamma = 180^o$).

3.13 Examples with the dipole antenna

Let us consider the particular case of a dipole antenna. The effective length of the antenna is of the form (see Sect. 2.4) $\mathbf{h}_e = h_e \hat{\boldsymbol{\theta}}$. This implies that the dipole antenna radiates a linearly polarized wave for any direction of the observation. Indeed, in the far-field the electric field oscillates exclusively along the $\hat{\boldsymbol{\theta}}$ direction, and thereby the polarization curve is a line segment.



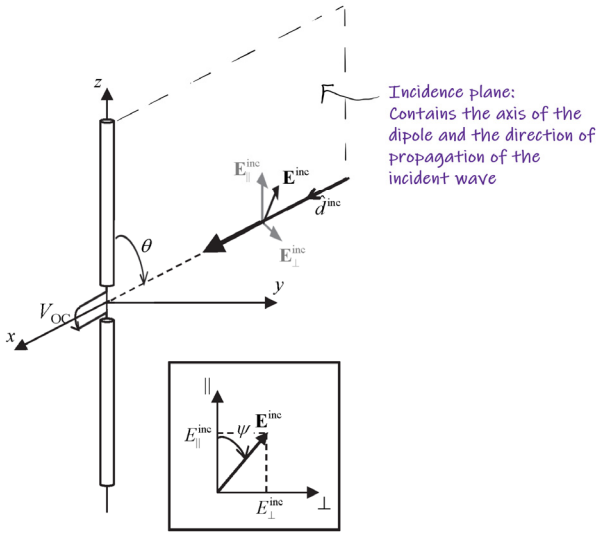
This means that the optimal polarization for an incoming wave that illuminates the antenna is also linearly polarized and must be directed along $\hat{\boldsymbol{\theta}}$. In fact, the polarization matching coefficient for a dipole antenna is:

$$C_p = \frac{\left|\hat{\boldsymbol{\theta}}^i \cdot \underline{\mathbf{E}}^{\text{inc}}\right|^2}{\left|\underline{\mathbf{E}}^{\text{inc}}\right|^2} \qquad \text{(dipole antenna).} \tag{3.55}$$

When the incoming plane wave is linearly polarized the coefficient C_p can be written as:

$$C_p = \cos^2\psi \qquad \text{(dipole antenna + linear pol.).} \tag{3.56}$$

where ψ is the angle between the incident electric field and the so-called *incidence plane* (see the figure below).



The incidence plane is by definition the plane generated by the antenna axis (z-axis) and the direction of arrival of the incoming wave ($\hat{\mathbf{d}}^i$). Only the component of the incident electric field parallel to the incident plane (the $\| = \hat{\boldsymbol{\theta}}^i$ component) interacts with the antenna. The component perpendicular to the incidence plane ($\perp = \hat{\boldsymbol{\varphi}}^i$) is orthogonal to the antenna axis, and hence cannot induce a current on the dipole.

When the incoming wave is circularly polarized, the incident field is of the form (see Sect. 3.10):

$$\underline{\mathbf{E}}_0^{\text{inc}} = E_0 \left(\hat{\boldsymbol{\theta}}^i \pm j \hat{\boldsymbol{\varphi}}^i \right). \tag{3.57}$$

In this case, the polarization matching coefficient is:

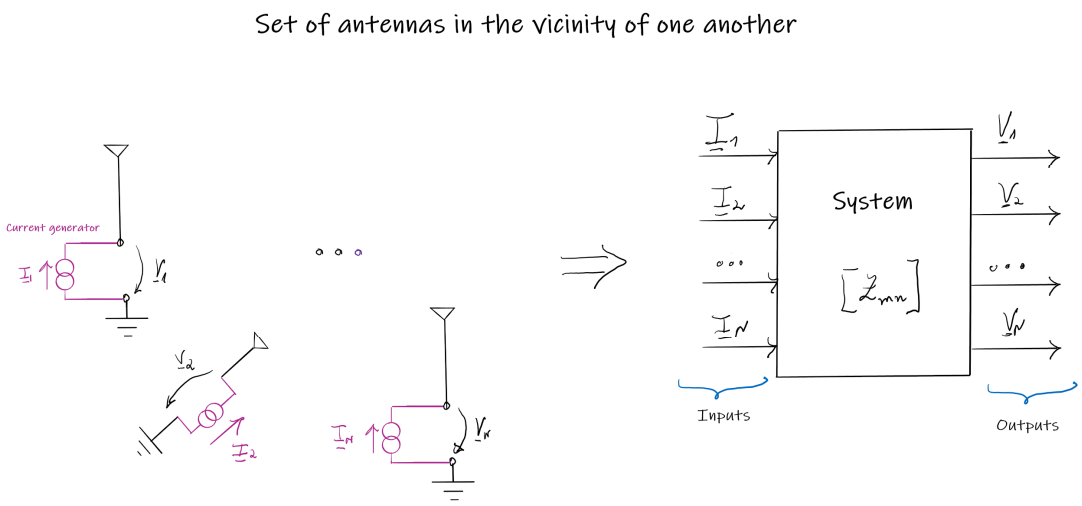
$$C_p = \frac{\left| E_0 \right|^2}{\left| E_0 \left(\hat{\boldsymbol{\theta}}^i \pm j \hat{\boldsymbol{\varphi}}^i \right) \right|^2} = \frac{1}{\left| 1 \right|^2 + \left| \pm j \right|^2} = \frac{1}{2} \qquad \text{(circular pol.).} \tag{3.58}$$

Thus, as expected, the antenna is only sensitive to half of the incident energy flux. The other half is associated with an electric field oriented along $\hat{\boldsymbol{\varphi}}^i$, which as explained above does not interact with the antenna.

4. Mutual and self impedances

4.1 Antenna coupling

In antenna systems it is often useful to place different radiating elements in the vicinity of one another. The most relevant examples are the antenna arrays where many identical radiating elements are excited simultaneously to obtain a tailored radiation pattern with some desired properties. In such systems, the typical distance between the radiating elements is on the order of half-wavelength, which corresponds to the near-field zone where the fields are strongly reactive and very intense. This coupling between the antennas can modify their radiation properties compared to the situation where an antenna stands alone in free-space.



The electromagnetic fields in a system formed by several antennas (n=1,2,…, N) are completely determined by the currents at the feeding points $(\underline{I}_1(0),\underline{I}_2(0),...,\underline{I}_N(0))$. In particular, the voltages at the antenna terminals $(\underline{V}_1,\underline{V}_2,...,\underline{V}_N)$ are functions of the currents. Thus, we have a multiple input-output system, as illustrated in the figure. The input-output relations are formally (we drop the label (0) to simplify the notations):

$$\begin{aligned}
&\underline{V}_1 = \underline{V}_1\left(\underline{I}_1, \underline{I}_2, ..., \underline{I}_N\right) \\
&\underline{V}_2 = \underline{V}_2\left(\underline{I}_1, \underline{I}_2, ..., \underline{I}_N\right) \\
&... \\
&\underline{V}_N = \underline{V}_N\left(\underline{I}_1, \underline{I}_2, ..., \underline{I}_N\right)
\end{aligned} \tag{4.1}$$

Now, because of the linearity of the Maxwell's equations and of the materials that form the antennas, the input-output relations in the frequency domain are forcibly linear. This means that there are constants $\underline{Z}_{mn}$ (m,n=1,…,N) such that:

$$\begin{aligned}
&\underline{V}_1 = \underline{Z}_{11}\underline{I}_1 + \underline{Z}_{12}\underline{I}_2 + ... + \underline{Z}_{1N}\underline{I}_N \\
&\underline{V}_2 = \underline{Z}_{21}\underline{I}_1 + \underline{Z}_{22}\underline{I}_2 + ... + \underline{Z}_{2N}\underline{I}_N \\
&... \\
&\underline{V}_N = \underline{Z}_{N1}\underline{I}_1 + \underline{Z}_{N2}\underline{I}_2 + ... + \underline{Z}_{NN}\underline{I}_N
\end{aligned} \tag{4.2}$$

The constants $\underline{Z}_{mn}$ dependent on the antenna and vary with frequency. It is evident that $\underline{Z}_{mn}$ has units of impedance. The set of equations (4.2) can be written in a compact way in a matrix form as $\mathbf{V} = \mathbf{Z} \cdot \mathbf{I}$ where $\mathbf{V}, \mathbf{I}$ are the column vectors with the voltages and currents at the antenna terminals, respectively, and $\mathbf{Z}$ is the so-called impedance matrix:

$$\underbrace{\begin{pmatrix} \underline{V}_1 \\ \underline{V}_2 \\ ... \\ \underline{V}_N \end{pmatrix}}_{\mathbf{V}} = \underbrace{\begin{pmatrix} \underline{Z}_{11} & \underline{Z}_{12} & ... & \underline{Z}_{N1} \\ \underline{Z}_{21} & \underline{Z}_{22} & ... & \underline{Z}_{12} \\ ... & ... & ... & ... \\ \underline{Z}_{N1} & \underline{Z}_{N1} & ... & \underline{Z}_{NN} \end{pmatrix}}_{\mathbf{Z}} \underbrace{\begin{pmatrix} \underline{I}_1 \\ \underline{I}_2 \\ ... \\ \underline{I}_N \end{pmatrix}}_{\mathbf{I}} \tag{4.3}$$

The diagonal elements of the impedance matrix ($\underline{Z}_{11}, \underline{Z}_{22}, ...$) are known as the self-impedances, whereas the off-diagonal terms are the mutual impedances.

The input impedance seen by the generator that feeds the m-th antenna is:

$$\left(\underline{Z}_{in}\right)_{\substack{\text{m-th} \\ \text{antenna}}} \equiv \frac{\underline{V}_m}{\underline{I}_m} \tag{4.4}$$

For clarity, let us focus on the case of two antennas (N=2), for which it is possible to write:

$$\begin{aligned}\left(\underline{Z}_{in}\right)_1 &\equiv \underline{Z}_{11} + \underline{Z}_{12}\frac{\underline{I}_2}{\underline{I}_1} \\ \left(\underline{Z}_{in}\right)_2 &\equiv \underline{Z}_{21}\frac{\underline{I}_1}{\underline{I}_2} + \underline{Z}_{22}\end{aligned}. \tag{4.5}$$

The impedance matrix is independent of the excitation and is exclusively determined by the antennas and their relative orientation and distance in space. In contrast, the previous formula shows that input impedance cannot be expressed alone only through the impedance matrix but that it also depends on the *point of operation* of the system determined by the current ratio $\underline{I}_1 / \underline{I}_2$:

$$\left(\underline{Z}_{in}\right)_i \equiv \text{function}_i\left(\mathbf{Z}, \frac{\underline{I}_1}{\underline{I}_2}\right). \tag{4.6}$$

Due to this reason $\left(\underline{Z}_{in}\right)_i$ is a "driving point impedance". The design of a matching network that connects a generator to a given antenna requires the knowledge of the impedance matrix and of the intended point of operation determined by the current ratio $\underline{I}_1 / \underline{I}_2$, i.e., of the driving point impedance.

From the definition of the impedance matrix, one can see that the self-impedance of the *n*-th antenna is the input impedance seen at the antenna terminals when the remaining antennas are not excited, (i.e., when their terminals are in open-circuit). For example, for the case of two antennas, one has:

$$\underline{Z}_{11} = \left.\frac{\underline{V}_1}{\underline{I}_1}\right|_{\underline{I}_2=0} = \left(\underline{Z}_{in}\right)_1\Big|_{\substack{\text{input impedance when}\\ \text{only the 1st}\\ \text{antenna is fed}}} \tag{4.7}$$

Strictly speaking $\underline{Z}_{11}$ is not the antenna impedance $\underline{Z}_{a1}$ discussed in earlier chapters. Indeed, $\underline{Z}_{a1}$ is the input impedance when the antenna stands *alone* in free-space, whereas $\underline{Z}_{11}$ is the input impedance when the antenna is in vicinity of the other antennas, with the terminals of the other antennas in open-circuit. Evidently, $\underline{Z}_{a1}$ and $\underline{Z}_{11}$ become identical only when the relative

distance between the antennas approaches infinity. In practice, in most cases, the difference between $\underline{Z}_{a1}$ and $\underline{Z}_{11}$ is negligible. Thereby, it is common to use the approximation:

$$\underline{Z}_{nn} \approx \underline{Z}_{a,n} \equiv \text{impedance of the n-th antenna alone in free-space}. \tag{4.8}$$

The currents induced on the materials that form the antennas are also linear functions of the feeding currents $\underline{I}_1, \underline{I}_2, ..., \underline{I}_N$. This means that the antenna coupling can change the radiation properties of each individual antenna. In particular, it can change the far-field properties of each antenna determined by the effective length $\mathbf{h}_e$.

For example, analogous to the antenna input impedance, the effective length of a generic antenna depends on the antenna parameters and on the point of operation determined by the feeding currents:

$$\left(\mathbf{h}_e\right)_{\substack{\text{n-th}\\ \text{antenna}}} = \text{function}_n\left(\text{antenna parameters, feeding currents}\right). \tag{4.9}$$

For example, for two antennas, the "driving" effective length of the first antenna is

$$\left(\mathbf{h}_e\right)_{\text{antenna 1}} = \text{function}\left(\text{antenna 1,antenna 2, } \underline{I}_2 / \underline{I}_1\right). \tag{4.10}$$

Fortunately, it is often possible to neglect the effect of the other antennas and of the point of operation on the effective length $\mathbf{h}_e$ of a given antenna. Thus, in most antenna systems one can assume that:

$$\left(\mathbf{h}_e\right)_{\substack{\text{n-th}\\ \text{antenna}}} \approx \left(\mathbf{h}_e\right)^{\substack{\text{n-th}\\ \text{antenna}\\ \text{alone in free-space}}}. \tag{4.11}$$

With this approximation, one can find the field radiated by a set of coupled antennas from their radiation properties alone in free-space and from the knowledge the feeding currents. We will develop further this point when we study antenna arrays.

4.2 Reciprocity constraint on the mutual impedances

The mutual impedances $\underline{Z}_{mn}$ ($m \neq n$) are constrained by reciprocity as follows:

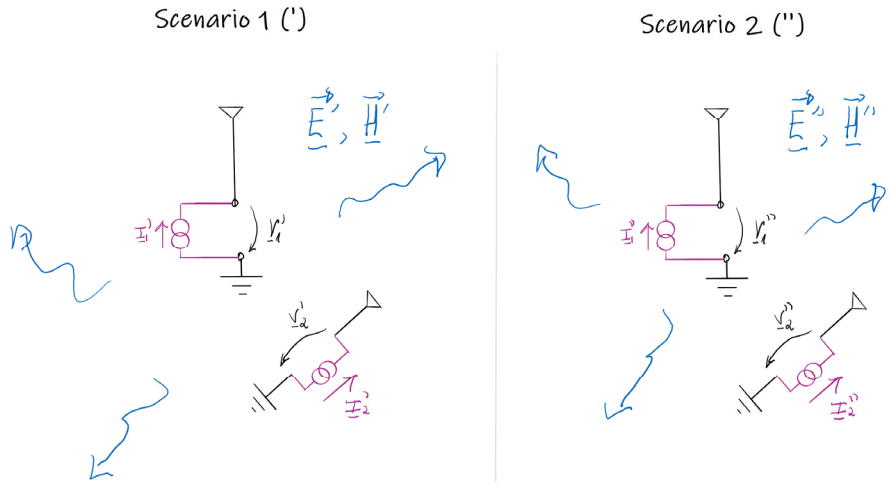
$$\underline{Z}_{mn} = \underline{Z}_{nm}. \tag{4.12}$$

In a matrix form, the same result is expressed as:

$$\boxed{\mathbf{Z} = \mathbf{Z}^T}, \tag{4.13}$$

where the superscript T represents the transpose of a matrix. This means that the impedance matrix is symmetric. For example, for a system with two antennas $\underline{Z}_{12} = \underline{Z}_{21}$.

Let us prove the enunciated result for the case $N=2$. To this end, we consider the two scenarios illustrated in the figure below. They correspond to different excitations of the same set of antennas. In the first scenario the excitation is determined by the one-primed current vector $\mathbf{I}' = \left(\underline{I}'_1 \quad \underline{I}'_2\right)^T$, and in the second scenario by the two-primed current vector $\mathbf{I}'' = \left(\underline{I}''_1 \quad \underline{I}''_2\right)^T$. The radiated fields in the first and second scenario are respectively $\left(\underline{\mathbf{E}}' \quad \underline{\mathbf{H}}'\right)$ and $\left(\underline{\mathbf{E}}'' \quad \underline{\mathbf{H}}''\right)$.



From the reciprocity theorem, we know that [Eq. (3.31)]

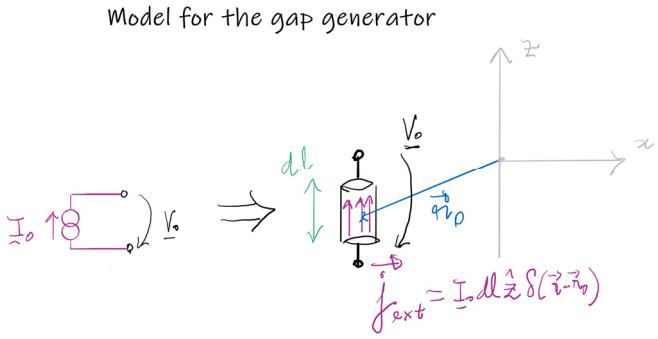
$$\iiint\limits_{\text{all space}} dV\, \underline{\mathbf{E}}' \cdot \mathbf{j}''_{ext} = \iiint\limits_{\text{all space}} dV\, \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext}\,. \tag{4.14}$$

The external currents in each scenario (the generators) are localized near the antenna terminals. Hence, one has:

$$\iiint\limits_{\text{all space}} dV\, \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext} = \iiint\limits_{\text{gap antenna 1}} dV\, \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext} + \iiint\limits_{\text{gap antenna 2}} dV\, \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext}\,. \tag{4.15}$$

Analogous to Sect. 3.7, it is assumed that the gap distance between the antenna terminals (dl) is very small compared to the wavelength and that the current is constant in the gap. This implies that the current density on the gap of the n-th antenna (n=1,2) is of the form $\mathbf{j}_{ext} = \underline{I}_n dl \hat{\mathbf{z}} \delta(\mathbf{r} - \mathbf{r}_{0n})$. Here, $\mathbf{r}_{0n}$ determines the coordinates of the gap of the n-th antenna and without loss of generality it is supposed that the gaps are oriented along $\hat{\mathbf{z}}$. Taking all these into account, one finds that:

$$\iiint\limits_{\text{gap antenna 1}} dV \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext} = dl \underline{\mathbf{E}}''(\mathbf{r}_1) \cdot \hat{\mathbf{z}} \underline{I}'_1 \approx -\underline{V}''_1 \underline{I}'_1 . \tag{4.16}$$



Similarly, one can show that $\iiint\limits_{\text{gap antenna 2}} dV \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext} \approx -\underline{V}''_2 \underline{I}'_2$. Therefore, it follows that:

$$\iiint\limits_{\text{all space}} dV \underline{\mathbf{E}}'' \cdot \mathbf{j}'_{ext} = -\underline{V}''_1 \underline{I}'_1 - \underline{V}''_2 \underline{I}'_2 = -\mathbf{I}'^T \cdot \mathbf{V}'' . \tag{4.17}$$

In the above $\mathbf{V}$ is a column vector with the double-primed voltages: $\mathbf{V}'' = \left(\underline{V}''_1 \quad \underline{V}''_2\right)^T$. Evidently, the same identity holds true when the single-primed and double-primed symbols are interchanged. Using this property in Eq. (4.14), we conclude that:

$$\mathbf{I}'^T \cdot \mathbf{V}'' = \mathbf{I}''^T \cdot \mathbf{V}' . \tag{4.18}$$

Independent of the current excitation, the voltages and currents are linked by the impedance matrix $\mathbf{V} = \mathbf{Z} \cdot \mathbf{I}$. This implies that:

$$\mathbf{I}'^T \cdot \mathbf{Z} \cdot \mathbf{I}'' = \mathbf{I}''^T \cdot \mathbf{Z} \cdot \mathbf{I}' . \tag{4.19}$$

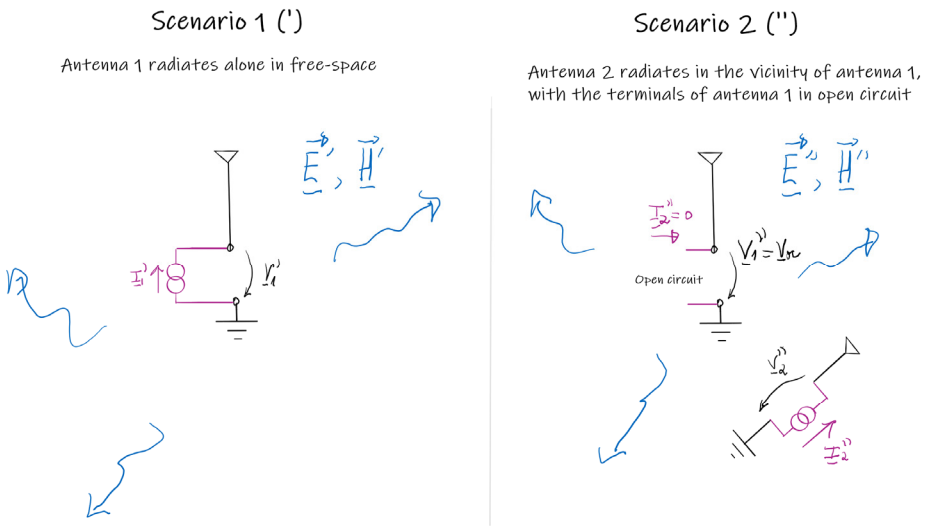
The right-hand side of the above equation is a scalar number and hence it is identical to its transpose: $\mathbf{I}''^{T}\cdot\mathbf{Z}\cdot\mathbf{I}'=\left[\mathbf{I}''^{T}\cdot\mathbf{Z}\cdot\mathbf{I}'\right]^{T}=\mathbf{I}'^{T}\cdot\mathbf{Z}^{T}\cdot\mathbf{I}''$. Thus, the reciprocity of the system implies that for any excitations $\mathbf{I}',\mathbf{I}''$ it is necessary that

$$\mathbf{I}'^{T}\cdot\left(\mathbf{Z}-\mathbf{Z}^{T}\right)\cdot\mathbf{I}''=0\,. \tag{4.20}$$

The identity can hold true for arbitrary excitations only if the matrix $\mathbf{Z}-\mathbf{Z}^{T}$ vanishes. This confirms that the impedance matrix is necessarily symmetric, as we wanted to show.

4.3 Mutual impedance with perturbation theory

In general, the impedance matrix needs to be determined using numerical methods, e.g., using a modified Halléns-type equation for the coupled antennas. However, as discussed in the following, it is possible to estimate the mutual impedances using with the parameters of the *uncoupled* antennas.To this end, we consider the two scenarios depicted in the figure below.



In the first scenario, the antenna 1 radiates alone in free-space. In the second scenario, the antenna 2 radiates in the vicinity of the first antenna, with the terminals of the antenna 1 in open-circuit. Evidently, we want to apply the reciprocity theorem to the considered scenarios. However, we need to do it it carefully. The reciprocity theorem assumes that the material structures in the two scenarios are the *same* and that only the external excitations are different.

That evidently it *not* the case in the scenarios under analysis, as the second antenna is absent in the first scenario!

We can circumvent this problem in a clever way. We will simply regard the second antenna as an "external current". This can be done by including in $\mathbf{j}''_{ext}$ (which usually describes only the current in the generator gaps) the polarization and conduction currents induced in the materials that form the antenna 2, $\mathbf{j}''_{\text{mat,ant2}}$. Thus, one can apply the reciprocity theorem in the usual way after the replacement $\mathbf{j}''_{ext} \to \mathbf{j}''_{ext} + \mathbf{j}''_{\text{mat,ant2}} \equiv \mathbf{j}''_2$, where $\mathbf{j}''_2$ gives the *total* current in the antenna 2 region. This yields:

$$\iiint\limits_{\text{all space}} dV \underline{\mathbf{E}}' \cdot \mathbf{j}''_2 = \iiint\limits_{\text{all space}} dV \underline{\mathbf{E}''} \cdot \mathbf{j}'_{ext} \,. \tag{4.21}$$

The term on the right-hand side can be evaluated explicitly as:

$$\iiint\limits_{\text{all space}} dV \underline{\mathbf{E}''} \cdot \mathbf{j}'_{ext} = \iiint\limits_{\text{gap 1}} dV \underline{\mathbf{E}''} \cdot \mathbf{j}'_{ext} = -\underline{I}'_1 \underline{V}''_1 \,. \tag{4.22}$$

The first identity uses the fact that in scenario 1 the external excitation is confined to the gap of the first antenna. The second identity is a consequence of Eq. (4.16). Since in the second scenario the first antenna is in open-circuit ($\underline{I}''_1 = 0$), the open-circuit voltage is given by $\underline{V}''_1 = Z_{11} \times 0 + Z_{12} \underline{I}''_2 = Z_{12} \underline{I}''_2$. Using this result in Eq. (4.22) and substituting the resulting formula into Eq. (4.21) it is found that:

$$Z_{12} = -\iiint\limits_{\text{antenna 2}} dV \frac{\underline{\underline{\mathbf{E}}}'}{\underline{I}'_1} \cdot \frac{\mathbf{j}''_2}{\underline{I}''_2} \,. \tag{4.23}$$

The parcel $\dfrac{\underline{\underline{\mathbf{E}}}'}{\underline{I}'_1}$ gives the field radiated by the antenna 1 alone in free-space normalized to the feeding current evaluated on the position of antenna 2. This term only depends on the fields radiated by the first antenna alone in free-space. On the other hand, the parcel $\dfrac{\mathbf{j}''_2}{\underline{I}''_2}$ determines the current distribution in antenna 2 normalized to the feeding current when it is operated as an

emitter in the vicinity of antenna 1, with the terminals of antenna 1 in open-circuit. Strictly speaking, $\frac{\mathbf{j}_2''}{\underline{I}_2''}$ can only be determined using numerical methods. However, to a first approximation (perturbation theory) we can assume that $\frac{\mathbf{j}_2''}{\underline{I}_2''}$ is coincident with the current distribution in antenna 2 when it radiates alone in free-space. This approximation yields the formula:

$$Z_{21} = Z_{12} \approx -\iiint_{\text{antenna 2}} dV \left.\frac{\underline{\underline{\mathbf{E}}}'}{\underline{I}_1'}\right|_{\substack{\text{field radiated by antenna 1}\\ \text{alone in free-space calculated over}\\ \text{antenna 2}}} \cdot \left.\frac{\mathbf{j}_2''}{\underline{I}_2''}\right|_{\substack{\text{current distribution induced on antenna 2}\\ \text{when it radiates alone in free-space}}} . \tag{4.24}$$

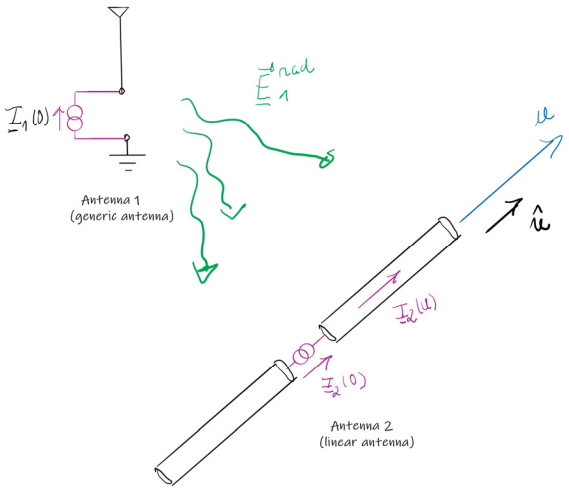
Note that the mutual impedances are written in terms of the current distribution and fields of the uncoupled antennas.

4.4 Application to linear antennas

Let us suppose now that antenna 2 is a linear antenna, e.g., a dipole antenna, so that the current distribution is approximately confined to a line segment oriented along the direction $\hat{\mathbf{u}}$. Then, the volume integral in Eq. (4.24) can be replaced by a line integral with the replacement $dV\,\mathbf{j}_2'' \to dl\,\underline{I}_2''\hat{\mathbf{u}}$. This yields the useful result:

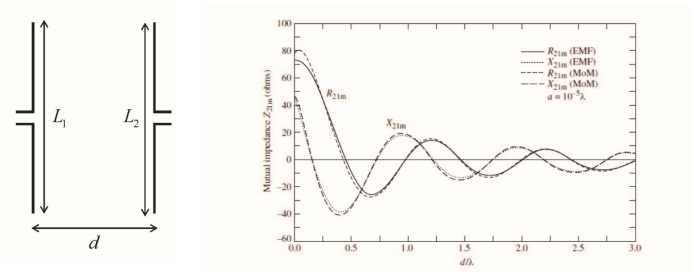
$$\boxed{Z_{21} = Z_{12} \approx -\int_{\substack{\text{antenna 2}\\ \text{axis}}} dl \left.\frac{\underline{\mathbf{E}}_1^{\text{rad}} \cdot \hat{\mathbf{u}}}{\underline{I}_1(0)}\right|_{\substack{\text{field radiated by antenna 1}\\ \text{alone in free-space calculated over}\\ \text{antenna 2 axis}}} \left.\frac{\underline{I}_2(u)}{\underline{I}_2(0)}\right|_{\substack{\text{current distribution induced on (the linear) antenna 2}\\ \text{when it radiates alone in free-space}}}} . \tag{4.25}$$

For simplicity, we dropped the primed and unprimed symbols. The integral is over the axis of the linear antenna (antenna 2). The variable *u* runs over the antenna 2 axis (see the figure).

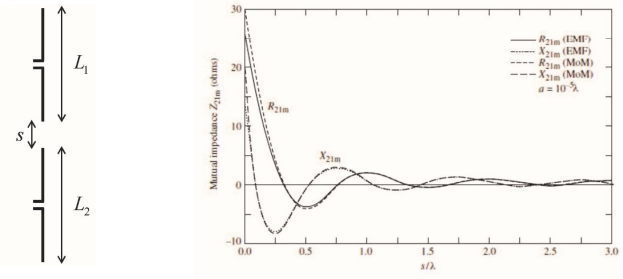


The derived formula can be used to calculate the mutual impedance of dipole antennas in different situations in a relatively straightforward (but cumbersome!) way. The results of such calculations for the case of side-by-side and collinear half-wavelength dipoles are shown in the figure below ($L_1 = L_2 = \lambda_0 / 2$; the mutual impedance is $\underline{Z}_{21} = \underline{Z}_{12} = R_m + jX_m$).

a) Side-by-side configuration



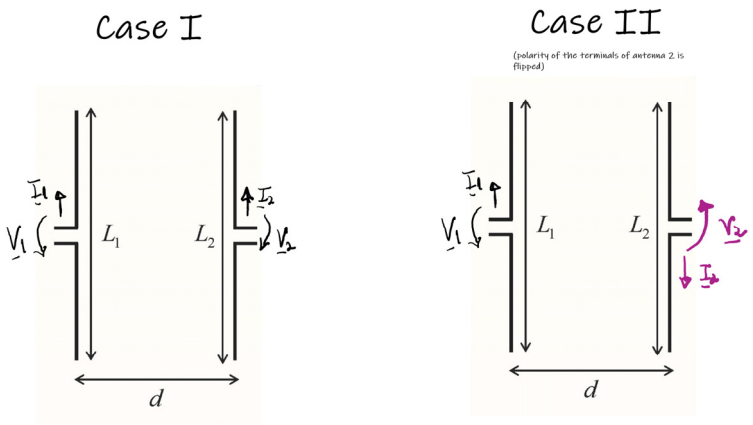
b) Collinear dipole configuration



As could be expected, the mutual coupling is stronger for a side-by-side configuration. The mutual coupling typically is weaker for larger separations between the dipoles, and it is

insignificant for distances much larger than the wavelength (especially for the collinear dipole configuration).

It is important to highlight that the real part of the mutual impedances can be either positive or negative! Specifically, if one writes $\underline{Z}_{21} = R_{21} + jX_{21}$ the signs of the resistance (R_{21}) and reactance (X_{21}) are not constrained in any manner. Furthermore, the value of $\underline{Z}_{21}$ depends on the relative polarity of the antenna terminals, i.e, on which terminals of the antennas are picked as the positive voltage terminals. For example, in the figure below the polarity of the terminals of antenna 2 is flipped in case II compared to case I. Due to this reason, it is simple to check that $\underline{Z}_{21}\big|_{\text{case I}} = -\underline{Z}_{21}\big|_{\text{case II}}$. The self-impedances are however identical in both cases: $\underline{Z}_{11}\big|_{\text{case I}} = \underline{Z}_{11}\big|_{\text{case II}}$ and $\underline{Z}_{22}\big|_{\text{case I}} = \underline{Z}_{22}\big|_{\text{case II}}$.



4.5 Accepted power

The power delivered to a set of antennas can be easily found from circuit theory. For the case of two antennas it is given by

$$P_{in,\text{tot}} = P_{in,1} + P_{in,2} = \frac{1}{2}\text{Re}\left\{\underline{V}_1\underline{I}_1^*\right\} + \frac{1}{2}\text{Re}\left\{\underline{V}_2\underline{I}_2^*\right\}. \tag{4.26}$$

This result can be written in matrix form as $P_{in,\text{tot}} = \frac{1}{2}\text{Re}\left\{\mathbf{I}^{T,*}\cdot\mathbf{V}\right\}$, where $\mathbf{V},\mathbf{I}$ are the column vectors with the voltages and currents at the antenna terminals. Using $\mathbf{V}=\mathbf{Z}\cdot\mathbf{I}$ it is found that:

$$P_{in,\text{tot}} = \frac{1}{2}\text{Re}\left\{\mathbf{I}^{T,*}\cdot\mathbf{Z}\cdot\mathbf{I}\right\}. \tag{4.27}$$

Taking into account that the impedance matrix is symmetric ($\mathbf{Z}=\mathbf{Z}^{T}$), it can be shown that the accepted power is determined only by the resistance matrix $\mathbf{R}=\text{Re}\left\{\mathbf{Z}\right\}$:

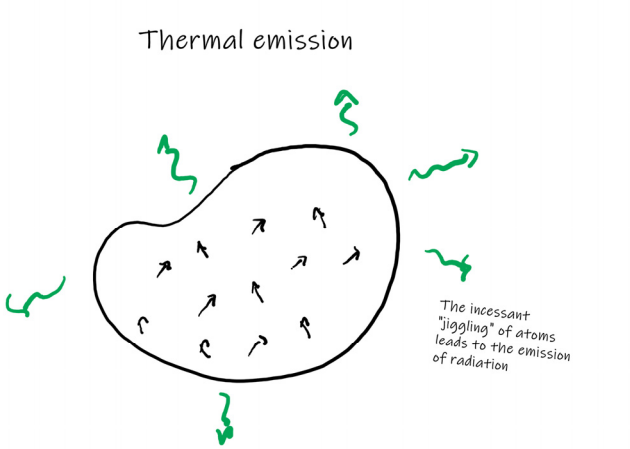
$$P_{in,\text{tot}} = \frac{1}{2}\mathbf{I}^{T,*}\cdot\mathbf{R}\cdot\mathbf{I}. \tag{4.28}$$

Evidently, for a passive system it is necessary that $P_{in,\text{tot}}\geq 0$ for any excitation. This is only possible if $\mathbf{R}$ is a positive definite matrix. For the case of two antennas, this requires that $R_{11}R_{22}-R_{12}^2>0$. Thus, the strength (but not the sign) of the mutual resistance is constrained by the value of the self-resistances. To conclude, we note that the accepted power is coincident with the total radiated power when the antennas are lossless.

5. Thermal noise

5.1 Thermal radiation

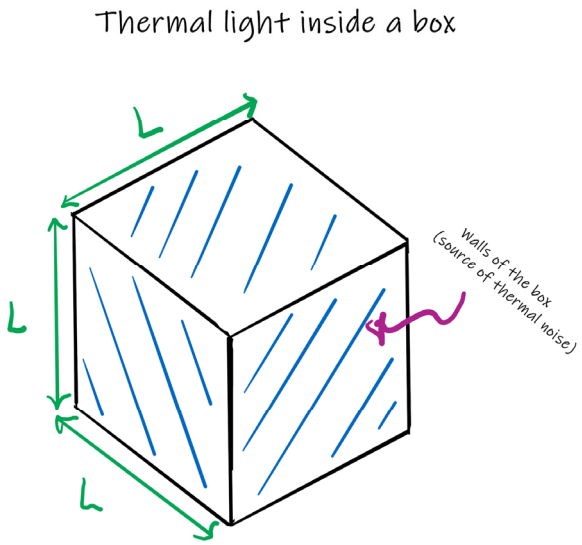
Microscopic systems are inherently characterized by some “agitation”. This agitation may be of quantum origin or due to thermal effects. In any of the cases, there is an incessant motion of matter (electrons, atoms, molecules, etc) at the nanoscale, and such vibrations lead to the emission of electromagnetic radiation. In other words, thermal effects cause a constant jiggling of the atoms, and these “fluctuations” lead to thermal emission. From the point of view of view of communication systems such radiation is “noise”. The level of noise determines the minimum power that one needs to deliver to a receiving antenna so that the incoming signal can be discriminated from the “thermal noise”. Thereby, it is essential to understand how to model thermal noise, what its origins are, and how to mitigate its effects.



5.2 Equipartition law

In order to characterize the radiation due to thermal fluctuations, we adopt a simple model where the system under study is a large box with dimensions $L \times L \times L$ filled with air. The walls of the box are made of some material, which is the source of thermal noise.

The radiation emitted by the walls of the box excites the natural modes of oscillation of the electromagnetic field. The natural modes are the oscillations supported by the field without an explicit external excitation.



The energy stored in each "mode" is determined by the "equipartition law". This fundamental principle essentially states that the probability of occupation of any state[1] of a system is the same. A detailed analysis beyond the scope of interest of this course (based on statistical physics concepts), shows that due to the equipartition law the thermal energy of any electromagnetic field mode is equal to (neglecting quantum corrections):

$$\mathcal{E}\big|_{\substack{\text{thermal energy}\\ \text{per mode}}} = k_B T \,. \tag{5.1}$$

Here, $k_B = 1.38\times10^{-23}\, J/K$ is the Boltzmann constant and T is the temperature in Kelvin.

Armed with this knowledge, next we want to determine the spectrum of the thermal energy inside the box. To do this, we need to find the electromagnetic modes, which is the topic of the next section.

[1] A "state" in this context should not be confused with a mode.

5.3 The natural modes

In general, the natural modes depend on *boundary conditions* at the box walls, and thereby on the material response. Here, to simplify the analysis and avoid considering any detailed material response, we will simply assume that the walls enforce the electromagnetic fields to satisfy *periodic* boundary conditions. The natural modes of the electromagnetic field inside such idealized box can be easily found from the modes in *free-space*. In free-space the modes are nothing but the usual plane wave solutions of Maxwell's equations. A generic plane wave is of the form:

$$\underline{\mathbf{E}} = \underline{\mathbf{E}}_0 e^{-j\mathbf{k}\cdot\mathbf{r}}, \tag{5.2}$$

where $\underline{\mathbf{E}}_0$ determines the polarization of the wave and $\mathbf{k} = k_x\hat{\mathbf{x}} + k_y\hat{\mathbf{y}} + k_z\hat{\mathbf{z}}$ is the wave vector. In free-space it satisfies:

$$\mathbf{k}\cdot\mathbf{k} = \left(\frac{\omega}{c}\right)^2. \tag{5.3}$$

One can construct the modal fields in the box from the fields in free-space simply imposing the periodic boundary conditions. For example, for the walls at $x = \pm L/2$ the fields must satisfy:

$$\underline{\mathbf{E}}\left(-L/2, y, z\right) = \underline{\mathbf{E}}\left(+L/2, y, z\right). \tag{5.4}$$

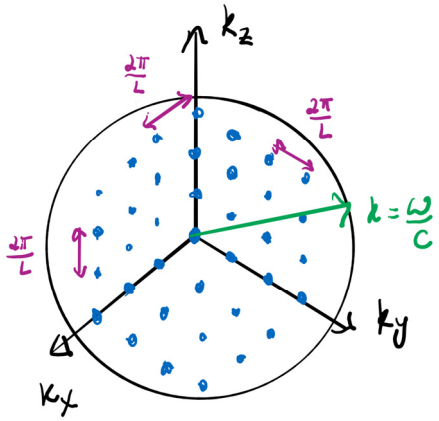
The above condition is equivalent to $\underline{\mathbf{E}}_0 e^{-j\mathbf{k}\cdot\mathbf{r}}\Big|_{\mathbf{r}=(-L/2,y,z)} = \underline{\mathbf{E}}_0 e^{-j\mathbf{k}\cdot\mathbf{r}}\Big|_{\mathbf{r}=(+L/2,y,z)}$, which leads to $e^{jk_xL} = 1$. Thereby, the allowed values for the *x*-component of the wave vector are $k_x = 2\pi n_x / L$, where $n_x = 0, \pm 1, \pm 2, \ldots$ is an integer number. Proceeding in the same way for the other two pairs of walls, it can be easily shown that the wave vector must be of the form:

$$\mathbf{k} = \frac{2\pi}{L} n_x\hat{\mathbf{x}} + \frac{2\pi}{L} n_y\hat{\mathbf{y}} + \frac{2\pi}{L} n_z\hat{\mathbf{z}}, \qquad n_x, n_y, n_z = 0, \pm 1, \pm 2, \ldots \tag{5.5}$$

Let us now find the distribution of modes in frequency. Specifically, let us count the number of modes N_ω, with an oscillation frequency $\tilde{\omega}$ less than ω. Evidently, from Eq. (5.3), the modes with $\tilde{\omega} < \omega$ have a wave vector $\mathbf{k}$ such that:

$$|\mathbf{k}| < \omega / c \,. \tag{5.6}$$

It is possible to visualize geometrically this condition representing each mode in the $\mathbf{k}-$space with a "point" (blue dot). The number of modes N_ω with oscillation frequency less than ω is determined by the number of points inside the sphere of radius $k = \omega / c$.



From Eq. (5.5), the distance between adjacent points in every direction of space is exactly $2\pi / L$. Hence, each point in the $\mathbf{k}-$space occupies the volume $(2\pi / L)^3$. The number of the points inside the sphere is roughly the volume of the sphere $\frac{4\pi}{3} k^3 = \frac{4\pi}{3}\left(\frac{\omega}{c}\right)^3$ divided by the volume occupied by each point. Thus, the number of modes N_ω is approximately:

$$N_\omega \approx 2 \times \frac{\frac{4\pi}{3}\left(\frac{\omega}{c}\right)^3}{\left(\frac{2\pi}{L}\right)^3} = \frac{1}{3\pi^2}\left(\frac{\omega}{c}\right)^3 L^3 \,. \tag{5.7}$$

The leading factor of 2 takes into account that there two independent polarization states of the electromagnetic field per mode. As seen, the number of modes is proportional to the volume of the box.

5.4 Spectrum of the thermal energy

With the equipartion law and the results of the previous section, we are ready to characterize the spectrum of the thermal energy inside the box. Specifically, let $E_\omega d\omega$ be the thermal energy associated with the radiation with a frequency in the small interval $[\omega, \omega + d\omega]$. The number of modes dN_ω in this interval is $dN_\omega = \frac{dN_\omega}{d\omega} d\omega = \frac{L^3\omega^2}{\pi^2 c^3} d\omega$. Thereby, from the equipartition law the thermal energy in the spectral range $[\omega, \omega + d\omega]$ is $E_\omega d\omega = k_B T dN_\omega = k_B T \frac{L^3\omega^2}{\pi^2 c^3} d\omega$. The energy volume density is by definition $W_\omega d\omega = \frac{1}{L^3} E_\omega d\omega$. It is given by:

$$W_\omega d\omega = k_B T \frac{\omega^2}{\pi^2 c^3} d\omega . \tag{5.8}$$

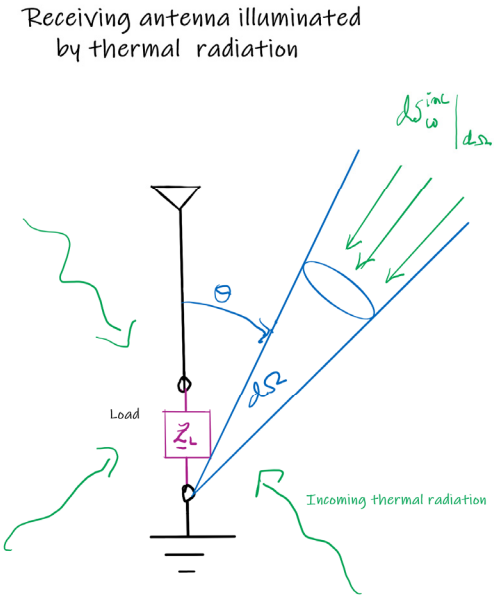
Note that the total energy per unit of volume in some spectral range $[\omega_1, \omega_2]$ is $\int_{\omega_1}^{\omega_2} W_\omega d\omega$. The total energy per unit of volume in the box is $\int_0^\infty W_\omega d\omega$ [1].

5.5 Thermal energy captured by an antenna

We are finally ready to apply the developed theory to antenna systems. Consider an antenna illuminated by the thermal radiation that propagates inside the "box". It is assumed that $L \to \infty$, so that the noise sources, i.e., the walls of the box are placed very far. We want to find the power

[1] This formula incorrectly predicts that the total energy inside the box is infinite. This happens due to a "ultraviolet catastrophe". The problem can be corrected by considering the quantization of the electromagnetic field energy in the box (Planck's theory), which leads to a correction in the formula $\mathcal{E} = k_B T$.

$dP_{r,\omega}$ captured by the antenna due to the incoming thermal radiation with frequency in the range $[\omega, \omega + d\omega]$. The antenna sees radiation (plane waves) arriving from all possible directions of space.



Let us first focus on the radiation that arrives inside some small solid angle $d\Omega$ oriented in the direction (θ, φ). From the theory of the receiving antenna the captured power is:

$$dP_{r,\omega}\Big|_{\substack{\text{solid angle} \\ d\Omega}} = C_i C_p A_{\text{ef}}(\theta,\varphi)\, dS_{\omega}^{\text{inc}}\Big|_{\substack{\text{solid angle} \\ d\Omega}}. \tag{5.9}$$

Here, $dS_{\omega}^{\text{inc}}\Big|_{\substack{\text{solid angle} \\ d\Omega}}$ is the intensity of the Poynting vector of the "plane waves" that arrive inside the solid angle $d\Omega$. The thermal radiation propagates isotropically. This implies, that the thermal energy density associated with the waves that propagate inside the solid angle $d\Omega$ is $W_{\omega} d\omega \frac{d\Omega}{4\pi}$, i.e., it corresponds to a fraction $\frac{d\Omega}{4\pi}$ of the total energy density. In free-space the Poynting vector is related to the energy density as (prove this result!):

$$S = Wc. \tag{5.10}$$

From here, one finds:

$$dS_{\omega}^{\text{inc}}\Big|_{\substack{\text{solid angle}\\ d\Omega}} = cW_{\omega}d\omega\frac{d\Omega}{4\pi} = k_B T\frac{\omega^2}{4\pi^3 c^2}d\Omega d\omega\,. \tag{5.11}$$

We can now substitute this formula into Eq. (5.9). The polarization matching coefficient is taken as $C_p = 1/2$ because all the polarization states are equiprobable (thermal radiation with the optimal polarization leads to $C_p = 1$ and thermal polarization with the "worst polarization" gives $C_p = 0$). It is found that:

$$dP_{r,\omega}\Big|_{\substack{\text{solid angle}\\ d\Omega}} = \frac{1}{2}\times C_i A_{\text{ef}}(\theta,\varphi)k_B T\frac{\omega^2}{4\pi^3 c^2}d\Omega d\omega\,. \tag{5.12}$$

Next, we integrate over all solid angles to take into account the radiation that arrives from all directions in the "sky". This gives:

$$\begin{aligned} dP_{r,\omega} &= \left(k_B T\frac{\omega^2}{8\pi^3 c^2}C_i \iint\limits_{\text{all solid angles}} A_{\text{ef}}(\theta,\varphi)d\Omega\right)d\omega \\ &= \left(k_B T\frac{\omega^2}{8\pi^3 c^2}C_i \iint\limits_{\text{all solid angles}} \frac{\lambda_0^2}{4\pi}G(\theta,\varphi)d\Omega\right)d\omega\,. \\ &= d\omega k_B T\frac{1}{8\pi^2}eC_i\left(\iint\limits_{\text{all solid angles}} g(\theta,\varphi)d\Omega\right) \end{aligned} \tag{5.13}$$

In the second identity we used the universal relation between the power gain and the effective area, and in the second identity we used $G = e\,g$, which relates the power gain and the directive gain through the efficiency. Note that here $dP_{r,\omega}$ is the power captured exclusively by the load, and does not include the power captured by the antenna body when the antenna is lossy.

From the definition of the directive gain of a transmitting antenna, one can write $g(\theta,\varphi) = \dfrac{4\pi U}{P_{rad}} = \dfrac{4\pi U}{\displaystyle\iint\limits_{\text{all solid angles}} U(\theta,\varphi)d\Omega}$. Thus, the integral of the directive gain over all solid angles is

$$\iint\limits_{\text{all solid angles}} g(\theta,\varphi)d\Omega = 4\pi\,. \tag{5.14}$$

Substituting this result into Eq. (5.13), we obtain the remarkably simple and universal (independent of the antenna) result:

$$dP_{r,\omega} = \frac{d\omega}{2\pi} k_B T \times C_i \times e \,. \qquad \text{(uniform “sky” temperature).} \qquad (5.15)$$

The power captured by the antenna load due to the incoming thermal radiation in the range $[\omega, \omega + d\omega]$ only depends on the temperature T of the noise sources, on the antenna efficiency and on the impedance matching coefficient. In particular, for a lossless antenna terminated with a matched load, the received power is:

$$dP_{r,\omega} = \frac{d\omega}{2\pi} k_B T \qquad \text{(matched load + lossless antenna).} \qquad (5.16)$$

Note that the power spectral density of thermal noise at the antenna terminals is “white”, i.e., $dP_{r,\omega} / d\omega = \frac{1}{2\pi} k_B T$ is independent of frequency.

5.6 Antenna noise temperature

In the previous section, it was assumed that the “sky” temperature is uniform. However, in practice an antenna captures energy radiated by different noise sources at different equivalent temperatures. One can take this into account by introducing a function $T_B(\theta, \varphi)$ that gives the equivalent temperature of the noise sources in the direction (θ, φ). This function is known as the *brightness temperature.* The only modification required in the analysis of the previous section is to replace $T \to T_B(\theta, \varphi)$ in Eq. (5.12). Then, integrating over all solid angles Eq. (5.13) becomes:

$$dP_{r,\omega} = \left(k_B \frac{\omega^2}{8\pi^3 c^2} C_i \iint\limits_{\text{all solid angles}} T_B(\theta, \varphi) A_{\text{ef}}(\theta, \varphi) d\Omega \right) d\omega \,. \qquad (5.17)$$

The *antenna noise temperature* T_A is defined as a weighted (by the effective area of the antenna) average of the brightness temperature:

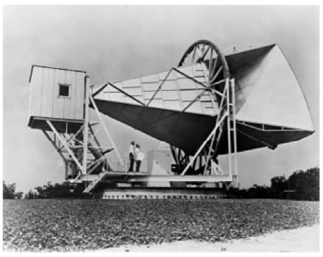
$$\boxed{T_A = \frac{\iint\limits_{\text{all solid angles}} T_B(\theta,\varphi)\, A_{\text{ef}}(\theta,\varphi)\, d\Omega}{\iint\limits_{\text{all solid angles}} A_{\text{ef}}(\theta,\varphi)\, d\Omega}}. \tag{5.18}$$

It is straightforward to check that Eq. (5.15) must be modified as

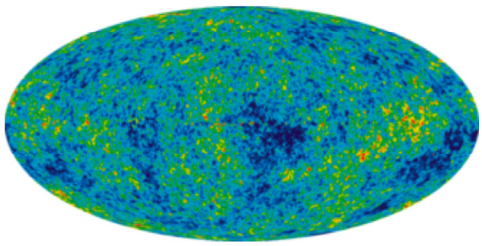
$$\boxed{dP_{r,\omega} = \frac{d\omega}{2\pi} k_B T_A \times C_i \times e}. \tag{5.19}$$

The antenna noise temperature depends on the environment wherein the antenna is placed. Desirably, the antenna should have a small gain along directions of space where the brightness temperature is larger, usually in the direction of the ground. Typical values for the brightness temperature for the microwave radiation spectrum are $T_B \approx 100-150K$ in the direction of the horizon (tangent to the Earth), $T_B \approx 5K$ in the zenith direction (perpendicular to the Earth), and $T_B \approx 300K$ in the direction of the ground (Earth surface).

The 6-meter horn antenna used by Penzias and Wilson



Map of the cosmic microwave background temperature (on average 2.7K)

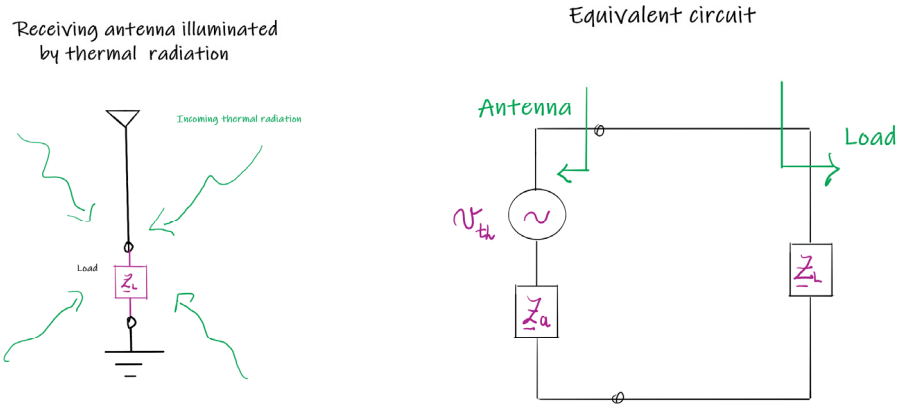


As curiosity, the characterization of the antenna noise is at the origin of a Physics Nobel prize. In 1964, Arno Penzias and Robert Wilson were experimenting with a 6 meter horn antenna that had been built to measure very faint radiowaves. To increase the sensitivity of their receiver they tried to eliminate all the recognizable noise sources and cooled down the receiver to liquid helium temperature. To their surprise, they found that a low, steady and mysterious noise persisted in their data. The residual background noise of 3.5K was 100 times more intense than what the recognizable noise sources could create. The background noise was independent of the direction along which the antenna was oriented and was present day and night with the same

intensity. They concluded that this noise, now known as “cosmic background radiation”, was coming from outside our galaxy. Today, this radiation is regarded as a fingerprint of the “Big Bang”, which supposedly released a tremendous blast of radiation that continues to arrive to us. In 1978, Penzias and Wilson were awarded the Physics Nobel Prize for their discovery.

5.7 Equivalent circuit

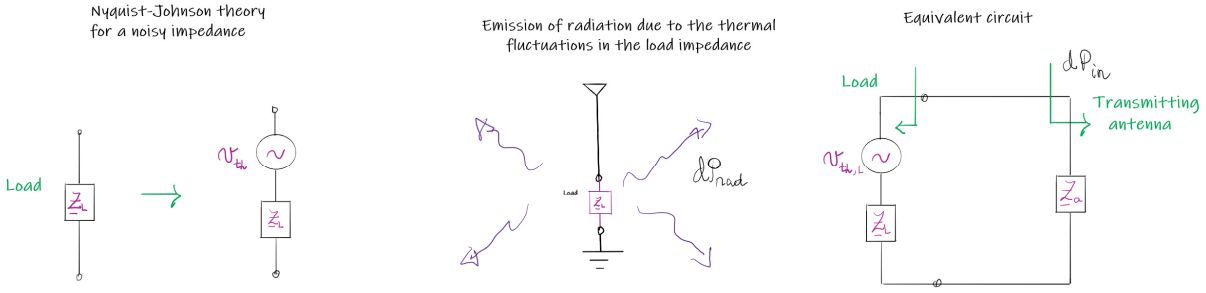
As shown in the figure below, one can introduce an equivalent circuit for the receiving antenna to model the effect of noise. The equivalent circuit is analogous to that of the receiving antenna, except that the generator amplitude is controlled by the noise, rather than by the amplitude of an incident plane wave.



A simple circuit analysis shows that the power delivered to the load in the equivalent circuit is $dP_{r,\omega} = \dfrac{\left\langle \left(v_{th,[\omega,\omega+d\omega]} \right)^2 \right\rangle}{4R_a} C_i$, with C_i the load impedance mismatch coefficient defined in the usual way, $\underline{Z}_a = R_a + jX_a$ is the antenna impedance, and $\left\langle \left(v_{th,[\omega,\omega+d\omega]} \right)^2 \right\rangle$ represents the expectation of $\left(v_{th,[\omega,\omega+d\omega]} \right)^2$ with $v_{th,[\omega,\omega+d\omega]}$ the part of the noise signal with spectrum in the range $[\omega, \omega + d\omega]$. Note that here v_{th} is an instantaneous noise voltage; for a time harmonic variation one has

$\left\langle v_{th}^2 \right\rangle = \left| \underline{V}_{th} \right|^2 / 2$. Comparing this result with Eq. (5.19), it is seen that the voltage of the equivalent noise generator is determined by:

$$\frac{\left\langle \left(v_{th,[\omega,\omega+d\omega]} \right)^2 \right\rangle}{d\omega/(2\pi)} = 4R_a k_B T_A \,. \tag{5.20}$$



It is interesting to note that load impedance ($\underline{Z}_L = R_L + jX_L$) is also subject to the influence of thermal noise. From the noise theory of Nyquist-Johnson, the effect of noise in the load can be modelled by an equivalent generator with amplitude:

$$\frac{\left\langle \left(v_{th,[\omega,\omega+d\omega]} \right)^2 \right\rangle_{\text{load}}}{d\omega/(2\pi)} = 4R_L k_B T \,. \tag{5.21}$$

where T is the physical temperature of the load. The noise in the load impedance gives rise to the *emission* of thermal radiation by the antenna. The emitted power can be found using the equivalent circuit for the transmitting antenna. Specifically, the input power at the terminals of the antenna is $dP_{in,\omega} = \dfrac{\left\langle v_{th,[\omega,\omega+d\omega]}^2 \right\rangle_{\text{load}}}{4R_L} C_i$. Thereby, from the Nyquist-Johnson result one finds that $dP_{in,\omega} = C_i k_B T \dfrac{d\omega}{2\pi}$. The power radiated $dP_{rad,\omega}$ (originated exclusively by the load) is related to the input power through the efficiency:

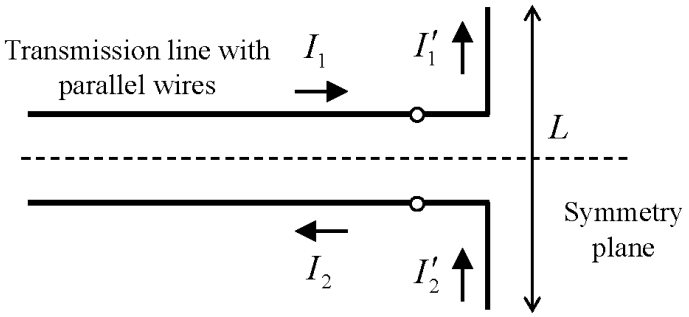
$$\boxed{dP_{rad,\omega} = k_B T \frac{d\omega}{2\pi} \times C_i \times e}. \tag{5.22}$$

Comparing this result with Eq. (5.19), one can see that when the load is in thermal equilibrium with the environment so that $T_B = T =$ physical temperature of the load , one has $dP_{rad,\omega} = dP_{r,\omega}$ so that the amount of received thermal energy is exactly identical to the amount of radiated thermal energy, as it should be!

6. Balancing devices

6.1 Balanced lines

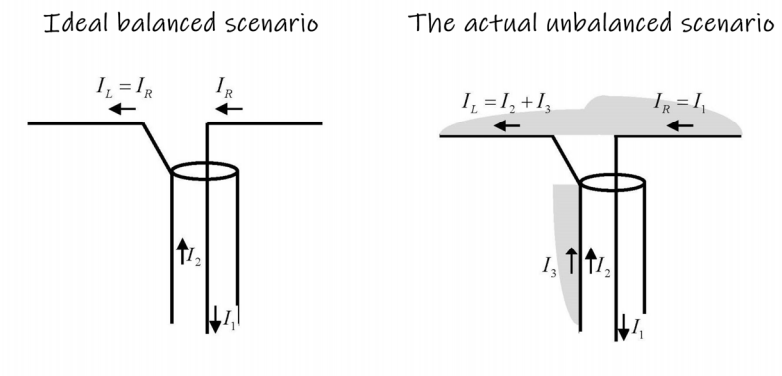
Antennas are fed by transmission lines. The line transports power from the generator to the antenna, and, desirably, should not contribute to the radiation pattern of the system. This can be enforced by guaranteeing that the two conductors that form the line transport symmetric currents ($\underline{I}_1 = \underline{I}_2$, with the currents oriented as shown below) so that the fields emitted by an individual conductor effectively cancel out the fields emitted by the other conductor. This is the normal mode of operation of transmission lines. When $\underline{I}_1 = \underline{I}_2$ the line is said to be balanced.



However, when a line is connected to an antenna the balancing condition $\underline{I}_1 = \underline{I}_2$ can be disrupted. This can lead to undesired effects such as an uncontrolled change of the input impedance or a corruption of the antenna radiation pattern.

To understand how such an effect can arise, let us first consider a configuration that guarantees that the line remains balanced. Specifically, suppose that a parallel wire transmission line is used to feed a dipole antenna, as illustrated in the figure. One can identify a symmetry plane equidistant from the two conductors and from the two arms of the dipole antenna. The mirror symmetry ensures that the currents remain identical when the line is connected to the antenna.

Indeed, by symmetry, the two-half spaces delimited by the symmetry plane are completely equivalent, and thereby if a certain current flows in the top conductor to the right, the same current will flow in the bottom conductor but to the left, so that the line is balanced[1].



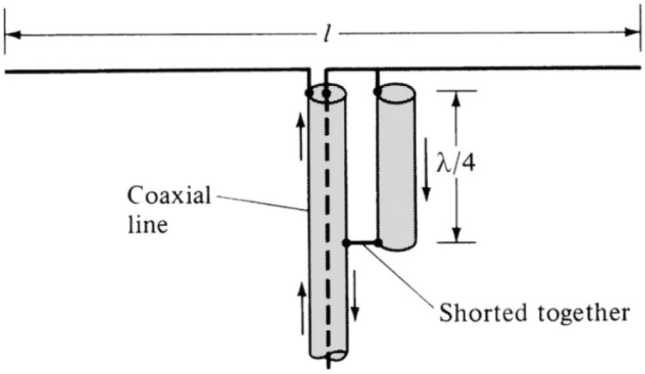
However, the condition $\underline{I}_1 = \underline{I}_2$ can be easily disrupted. To illustrate this consider the problem of feeding the same dipole antenna but with a coaxial cable. In the ideal balanced scenario the currents in the two arms of the dipole are identical. This requires that the current provided to the left arm of the dipole by the outer conductor of the coaxial cable is identical to the current provided to the right arm of the dipole by the inner conductor. However, in this system, there is no symmetry that guarantees such a behaviour. In fact, nothing prevents that $\underline{I}_L \neq \underline{I}_R$, as it is perfectly possible to induce a current on the outer side of the coaxial $\underline{I}_3$, leading to an excitation of the dipole antenna with unbalanced currents $\underline{I}_1 = \underline{I}_R \neq \underline{I}_L = \underline{I}_2 + \underline{I}_3$. In this situation the net current flowing in the outer conductor ($\underline{I}_2 + \underline{I}_3$) differs from the current flowing on the inner conductor $\underline{I}_1$ so that the line is unbalanced. The current $\underline{I}_3$ can itself radiate and corrupt the radiation pattern of a dipole antenna, leading to asymmetries and undesired sidelobes.

6.2 Baluns

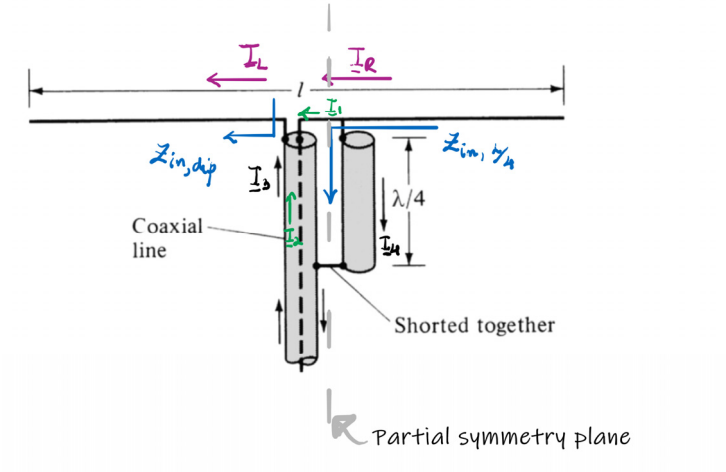
A “balun” is a device intended to enforce that a feeding line remains balanced after it is connected to an antenna. “Balun” is an acronym for “balanced to unbalanced”.

[1] This property is a simple consequence of the image method.

To illustrate the principle of operation of such devices, we analyze the so-called folded or quarter of wavelength balun represented below.



This balun eliminates the balancing problem that arises when a coaxial cable is connected to the dipole antenna. In short, the idea is to enforce some (partial) symmetry that can ensure that $\underline{I}_R \approx \underline{I}_L$. This is done by adding an "additional conductor" to the system, as represented in the figure above. The system is approximately symmetric with respect to the mid-plane that separates the coaxial from the additional conductor.



Effectively, the additional conductor forms a parallel wire transmission line with the outer side of the coaxial cable. Due to the partial symmetry of the system, one has $\underline{I}_3 \approx \underline{I}_4$ and $\underline{I}_1 \approx \underline{I}_2$. In particular, the fields emitted by the current flowing outside the coaxial cable are cancelled out by the fields radiated by the current flowing on the additional conductor. Thus, the radiation is mainly due to the arms of the dipole. Moreover, the current flowing on the right arm of the

dipole is now $\underline{I}_R = \underline{I}_1 + \underline{I}_4$ and is approximately identical to the current flowing on the left arm of the dipole $\underline{I}_L = \underline{I}_2 + \underline{I}_3$, guaranteeing a symmetric excitation of the dipole arms.

The rationale for the choice of the length of the additional conductor is the following. From the point of view of the feeding line, the load is formed by the "parallel" of the "dipole antenna" and of the "parallel wire transmission line" formed by the coaxial and by the additional conductor. Thus, the feeding line sees the load impedance:

$$\underline{Z}_{\text{load}} = \underline{Z}_{\text{in,dip}} \,||\, \underline{Z}_{\text{in},\frac{\lambda}{4}} . \tag{6.1}$$

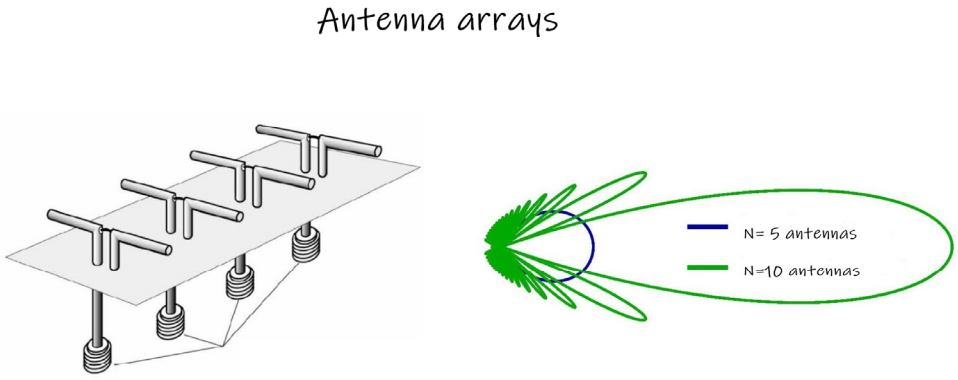
Here, $\underline{Z}_{\text{in,dip}}$ and $\underline{Z}_{\text{in},\frac{\lambda}{4}}$ are the impedances the dipole antenna and of the parallel wire line terminated in short-circuit. Ideally, we would like that $\underline{Z}_{\text{load}} = \underline{Z}_{\text{in,dip}}$. This can be enforced by choosing the length of the additional conductor to be a quarter of wavelength. Indeed, a quarter of wavelength transmission line transforms a short-circuit into an open circuit, so that $\underline{Z}_{\text{in},\frac{\lambda}{4}} = \infty$ and $\underline{Z}_{\text{load}} = \underline{Z}_{\text{in,dip}}$, as desired.

7. Antenna arrays

7.1 Introduction

For the purpose of point to point communications it is desirable to have directional antennas. This avoids wasting the emitted power along directions of space that are not part of the communication link. Usually, a single radiating element cannot provide the required gain. The solution for the problem is to use an "antenna array".

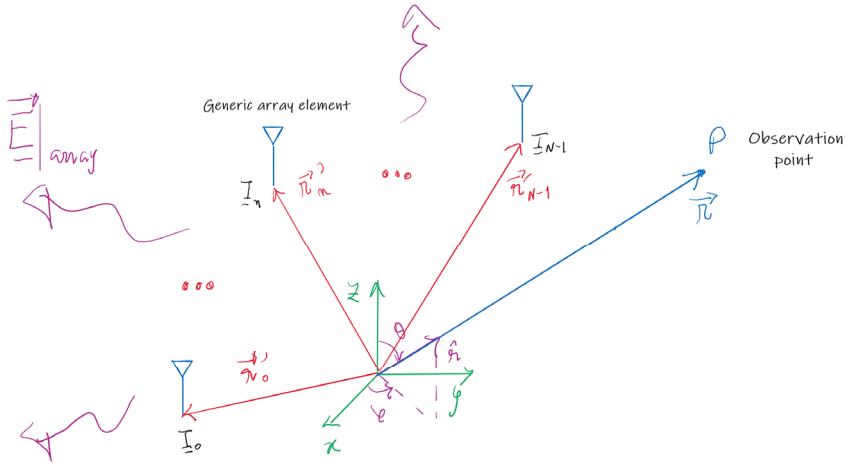
Antennas arrays are typically formed by several identical radiating elements. The amplitudes and phases of the feeding currents may be tailored to obtain a highly directive beam. In particular, it is possible to increase more and more the directivity of a radiating system by adding additional elements to the array.



Antenna arrays can also be used for "beam shaping", i.e., to synthesize a desired radiation pattern. For example, in cellular networks it is crucial to avoid the spillover of radiation to adjacent cells, and hence the radiation pattern should ideally cover uniformly the cell up to its boundary, and then drop off sharply to zero. Antenna arrays are the solution for this type of applications as they provide enough flexibility to tailor the beam almost arbitrarily.

7.2 Principle of pattern multiplication

We consider an array formed by N identical antennas. All the antennas have the same orientation, so that the antennas differ by one another by a simple *translation* in space. The n-th array element (n=0,1,…N-1) is centered at the point $\mathbf{r}_n'$ and is fed by the current $\underline{I}_n$.



We wish to obtain the electromagnetic fields in the far-field zone. The fields in the direction of observation $\hat{\mathbf{r}}$ are determined by the general formula derived in Chapter 1, repeated below for convenience:

$$\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}} = \eta_0 jk_0 \hat{\mathbf{r}} \times \left(\hat{\mathbf{r}} \times \mathbf{f}(\hat{\mathbf{r}})\right) \frac{e^{-jk_0 r}}{4\pi r}. \tag{7.1a}$$

$$\mathbf{f}(\hat{\mathbf{r}}) = \iiint dV' \underline{\mathbf{j}}(\mathbf{r}') e^{+jk_0 \hat{\mathbf{r}}\cdot\mathbf{r}'}. \tag{7.1b}$$

The integral in the definition of $\mathbf{f}(\hat{\mathbf{r}})$ is over the entire space. It is convenient to split the integral as a sum of parcels associated with the different antennas:

$$\begin{aligned} \mathbf{f}(\hat{\mathbf{r}}) &= \sum_{n=0}^{N-1} \iiint_{\text{volume n-th antenna}} \underline{\mathbf{j}}(\mathbf{r}')\big|_{\substack{\text{n-th}\\ \text{element}}} e^{+jk_0 \hat{\mathbf{r}}\cdot\mathbf{r}'} dV' \\ &= \sum_{n=0}^{N-1} e^{+jk_0 \hat{\mathbf{r}}\cdot\mathbf{r}_n'} \iiint_{\substack{\text{volume n-th antenna}\\ \text{shifted to the origin}}} \underline{\mathbf{j}}(\mathbf{R}')\big|_{\substack{\text{n-th}\\ \text{element}}} e^{+jk_0 \hat{\mathbf{r}}\cdot\mathbf{R}'} dV' \end{aligned}. \tag{7.2}$$

In the second identity, we introduced $\mathbf{R}' = \mathbf{r}' - \mathbf{r}_n'$ so that all the volume integrals are effectively shifted to the origin.

As discussed in Sect. 4.1, the current distribution $\underline{\mathbf{j}}(\mathbf{r}')$ induced on the antennas is a multilinear function of the feeding currents ($\underline{\mathbf{j}}(\mathbf{r}') = \text{multilinear function}\left(\underline{I}_0, \underline{I}_1, ..., \underline{I}_{N-1}\right)$). Nevertheless, it is typically a good approximation to neglect the antenna coupling so that $\underline{\mathbf{j}}\big|_{\substack{\text{n-th}\\ \text{element}}}$ induced at the *n*-th antenna element depends mainly on the feeding current of the same antenna: $\underline{\mathbf{j}}\big|_{\substack{\text{n-th}\\ \text{element}}} \approx \text{linear function}\left(\underline{I}_n\right)$. Furthermore, the profile of $\underline{\mathbf{j}}\big|_{\substack{\text{n-th}\\ \text{element}}}$ may be assumed to be roughly the same as when the antenna stands alone in free-space. Since all the antenna elements are identical, this approximation implies that $\underline{\mathbf{j}}\big|_{\substack{\text{n-th}\\ \text{element}}}(\mathbf{R}')/\underline{I}_n$ is the same for all the antennas. In fact, $\underline{\mathbf{j}}\big|_{\substack{\text{n-th}\\ \text{element}}}(\mathbf{R}')/\underline{I}_n$ can be identified with current induced at a generic array element centered at origin (and alone in free-space) normalized to the feeding current $\underline{\mathbf{j}}(\mathbf{r}')/\underline{I}(0)$.

From the previous discussion, it follows that the function $\mathbf{f}(\hat{\mathbf{r}})$ can be approximated by:

$$\mathbf{f}(\hat{\mathbf{r}}) \approx \sum_{n=0}^{N-1} e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}_n'}\underline{I}_n \iiint_{\substack{\text{generic array element}\\ \text{centered at the origin}}} \frac{\underline{\mathbf{j}}(\mathbf{r}')}{\underline{I}(0)} e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'} dV' . \tag{7.3}$$

Substituting this result into Eq. (7.1) it is found that:

$$\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}} = \left(\sum_{n=0}^{N-1} e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}_n'}\underline{I}_n\right)\eta_0 jk_0 \left(\frac{\mathbf{h}_e(\hat{\mathbf{r}})}{\underline{I}(0)}\right)\Bigg|_{\substack{\text{generic element}\\ \text{centered at the origin}\\ \text{alone in free-space}}} \frac{e^{-jk_0 r}}{4\pi r} . \tag{7.4}$$

where $\mathbf{h}_e(\hat{\mathbf{r}})$ is the effective length of a generic element of the array, defined in the usual way:

$$\mathbf{h}_e(\hat{\mathbf{r}}) = \left[\left(\iiint_{\substack{\text{generic array element}\\ \text{centered at the origin}}} \frac{\underline{\mathbf{j}}(\mathbf{r}')}{\underline{I}(0)} e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'} dV'\right)\times\hat{\mathbf{r}}\right]\times\hat{\mathbf{r}} . \tag{7.5}$$

The field radiated by the array [Eq. (7.4)] can be written in an enlightening way:

$$\boxed{\underline{\mathbf{E}}(\mathbf{r})\big|_{\substack{\text{far-field}\\ \text{array}}} = F(\hat{\mathbf{r}})\frac{\underline{\mathbf{E}}^{\text{rad}}(\mathbf{r})}{\underline{I}(0)}\Bigg|_{\substack{\text{single array element}\\ \text{positioned at the origin}}}} . \tag{7.6}$$

In the above, $\underline{\mathbf{E}}^{\text{rad}}(\mathbf{r})/\underline{I}(0)=\eta_0 jk_0\mathbf{h}_e(\hat{\mathbf{r}})\dfrac{e^{-jk_0r}}{4\pi r}$ is the field radiated by a single array element positioned at the origin normalized by the respective feeding current. The function $F(\hat{\mathbf{r}})$ is the so-called *array spatial factor*, and only depends on the feeding currents, on the position of the antennas, and on the direction of observation:

$$\boxed{F(\hat{\mathbf{r}})=\sum_{n=0}^{N-1}e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}_n'}\underline{I}_n}. \tag{7.7}$$

The formula (7.6) establishes that the fields emitted by the array are determined by the *multiplication* of the field emitted by a single element and the array spatial factor. This result is the foundation of array theory and is know as the *principle of pattern multiplication*.

Taking the absolute value of both sides of Eq. (7.6), one can write:

$$\left|\underline{\mathbf{E}}(\mathbf{r})\right|_{\substack{\text{far-field}\\ \text{array}}}=\left|F(\hat{\mathbf{r}})\right|\left.\frac{\left|\underline{\mathbf{E}}^{\text{rad}}(\mathbf{r})\right|}{\left|\underline{I}(0)\right|}\right|_{\substack{\text{single array element}\\ \text{positioned at the origin}}}. \tag{7.8}$$

The radiation intensity of the array is proportional to the square of the absolute value of the electric field: $U\sim\left|\underline{\mathbf{E}}(\mathbf{r})\right|^2$. This means that the (power) radiation pattern of the array is the product of the radiation pattern of a single element by $\left|F(\hat{\mathbf{r}})\right|^2$ (apart from an overall multiplying constant). In particular, the array spatial factor determines the radiation pattern of a hypothetical array of isotropic emitters:

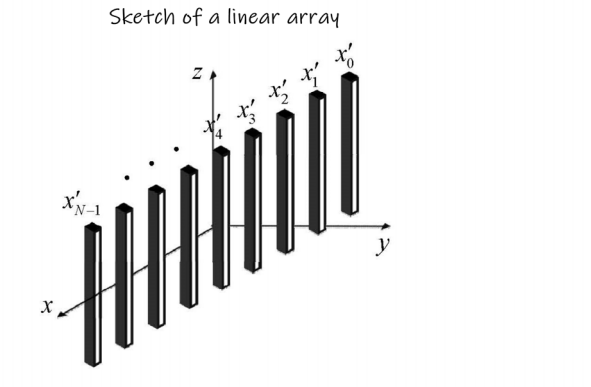
$$U_{\substack{\text{array of isotropic}\\ \text{emitters}}}\sim\left|F(\hat{\mathbf{r}})\right|^2. \tag{7.9}$$

For future reference, we note that the magnetic field emitted by the array is related to the electric field in the usual way: $\left.\underline{\mathbf{H}}(\mathbf{r})\right|_{\substack{\text{array}\\ \text{far-field}}}=\dfrac{1}{\eta_0}\hat{\mathbf{r}}\times\left.\underline{\mathbf{E}}(\mathbf{r})\right|_{\substack{\text{array}\\ \text{far-field}}}$.

7.3 Linear arrays

A "linear array" consists of an antenna array in which the elements are placed along some line: *the array axis*. For definiteness, we take the array axis to be the x axis. In this case, the coordinates of the array elements are of the form:

$$\mathbf{r}'_n = x'_n \hat{\mathbf{x}} \,. \tag{7.10}$$



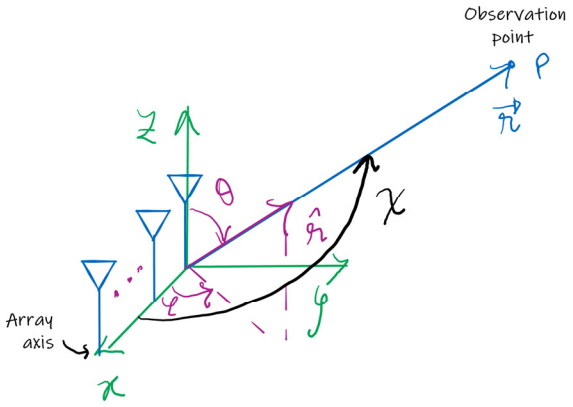
The array spatial factor can be written as:

$$\boxed{F = \sum_{n=0}^{N-1} \underline{I}_n e^{+jk_0 x'_n \cos\chi}}, \qquad \text{(linear array)} \tag{7.11}$$

where we introduced the angle χ defined by

$$\cos\chi = \hat{\mathbf{r}} \cdot \hat{\mathbf{x}} = \cos\varphi \sin\theta \,. \tag{7.12}$$

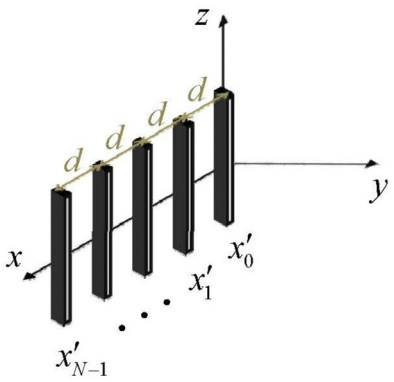
In the second identity, we used $\hat{\mathbf{r}} = \cos\varphi\sin\theta\hat{\mathbf{x}} + \sin\varphi\sin\theta\hat{\mathbf{y}} + \cos\theta\hat{\mathbf{z}}$. The parameter χ has a simple geometrical interpretation: it is the angle between the observation direction and the array axis (the x-axis). The angle χ varies in the range $0 \le \chi \le 180^{\circ}$. The array spatial factor depends on the direction of observation exclusively through χ. Due to this reason the radiation pattern of the array $|F|$ has symmetry of revolution around the array axis.



Let us now consider that all the elements of the linear array are equidistant and spaced by the distance d. For simplicity, we take the coordinates of the first element (n=0) of the array to be the origin. Then, in this scenario,

$$x_n' = nd, \qquad n = 0,1,...,N\text{-}1. \tag{7.13}$$

Array with equidistant elements

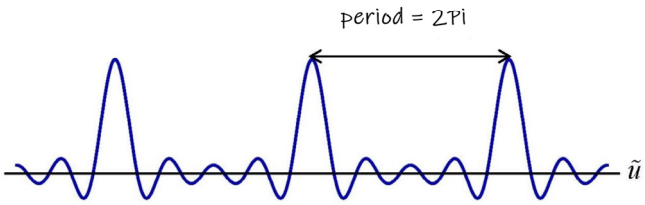


The array spatial factor can now be written as:

$$F = \sum_{n=0}^{N-1} \underline{I}_n e^{+j(k_0 d \cos\chi)n}, \qquad \text{(array with equidistant elements).} \tag{7.14}$$

The expression of F is alike to a truncated Fourier series in the variable $\tilde{u} = k_0 d \cos\chi$. The period of the function $F(\tilde{u})$ is 2π, i.e., $F(\tilde{u}) = F(\tilde{u} + 2\pi)$. In general, F is a complex valued function.

Sketch of the array spatial factor amplitude |F|

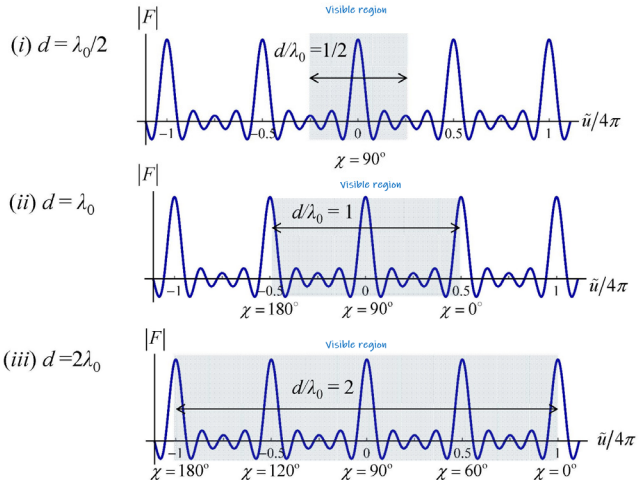


Evidently, only the interval $\left|\tilde{u}/\left(k_0 d\right)\right| = \left|\cos\chi\right| \le 1$ is relevant for the physical problem, as the range $\left|\cos\chi\right| > 1$ is not accessible for observation. Thus, the interval $-k_0 d \le \tilde{u} \le k_0 d$ is known as the "visible region":

$$-k_0 d \le \tilde{u} \le k_0 d\,, \qquad \text{(visible region)}. \tag{7.15}$$

A consequence of the periodicity of $F\left(\tilde{u}\right)$ is that the array spatial factor can be "replicated" multiple times in the visible region. This is illustrated in cases *ii*) and *iii*) of the figure below, for which the visible region encompasses 2 and 4 periods of the function F, respectively. Note that the period (2π) and profile of F are independent of λ_0 / d. The diameter of the visible region is $2k_0 d = 4\pi \times d/\lambda_0$. In the figure, the horizontal axis is $\tilde{u}/4\pi$ and thus one full period has width $1/2$.

Visible region of the array factor for different spacings between the elements



The appearance of "replicas" in the array spatial factor profile may occur when the diameter of the visible region ($4\pi d / \lambda_0$) exceeds the period (2π) of the function *F*, i.e., when the distance between the antennas exceeds half-wavelength $d \geq \lambda_0 / 2$. In these circumstances, the radiation pattern of the array spatial factor may be formed by multiple grating lobes with identical amplitude (see cases *ii*) and *iii*)). This feature is undesired as it will induce a multi-lobe structure in the overall radiation pattern of the system. Due to this reason the spacing between the array elements is usually chosen such that:

$$d \leq \lambda_0 / 2, \qquad \text{(ensures that } |F| \text{ does not have multiple lobes).} \tag{7.16}$$

The typical value for the distance between the array elements is $d = \lambda_0 / 2$, as it guarantees that the visible region consists of a full period of the array spatial factor. Very small values of *d* are not interesting because they lead to $|F| \approx const.$, i.e., only a small portion of the visible region is scanned and there is no array effect. Furthermore, very small values of *d* lead to a strong antenna coupling, and this may invalidate the hypothesis that the current distribution on the antennas is the same as when each element stands alone in free-space.

7.4 Progressive phaseshift array

A progressive phaseshift array is a linear array with equidistant elements such that the phases of the feeding currents vary in a linear progression: $\arg\left(\underline{I}_n\right) = \arg\left(\underline{I}_{n-1}\right) + \delta$. Here, δ is the current phaseshift for adjacent array elements. The feeding currents satisfy (for simplicity the phase of the current of the 1st array element is taken identical to zero):

$$\underline{I}_n = \left|\underline{I}_n\right| e^{jn\delta}, \qquad n = 0,1,..., N\text{-}1. \tag{7.17}$$

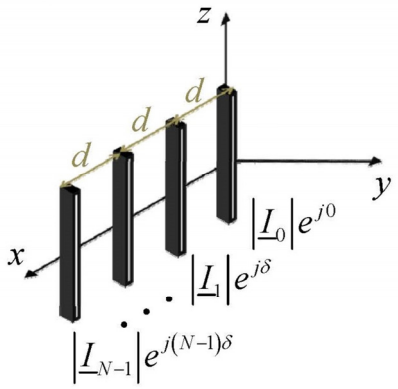
The corresponding array spatial factor is:

$$F = \sum_{n=0}^{N-1} \left|\underline{I}_n\right| e^{+j(k_0 d \cos\chi + \delta)n}. \tag{7.18}$$

The same result can be written in a more compact way as:

$$\boxed{F(u)=\sum_{n=0}^{N-1}\left|\underline{I}_n\right|e^{+jun}}, \qquad \boxed{u\equiv k_0 d\cos\chi+\delta}. \qquad \text{(progressive phaseshift array)} \tag{7.19}$$

Progressive phaseshift array



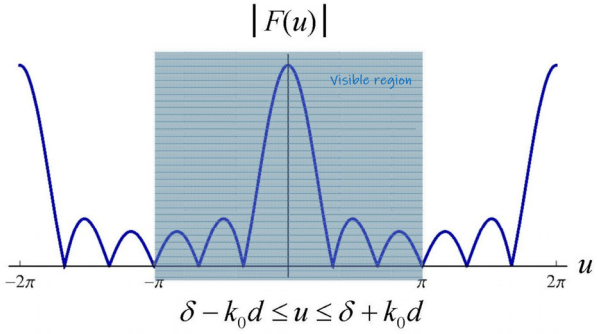
The array spatial factor can be regarded as a function of u. Furthermore, F is a periodic function u such that:

$$F(u)=F(u+2\pi). \tag{7.20}$$

Taking into account that $u=\tilde{u}+\delta$ (with $\tilde{u}$ defined as in the previous subsection) it is evident that the "visible region" is now determined by the interval:

$$\boxed{-k_0 d+\delta\le u\le k_0 d+\delta}, \qquad \text{(visible region)}. \tag{7.21}$$

Sketch of the array factor for a progressive phaseshift array



If the distance between the array elements is $d = \lambda_0 / 2$, the visible region is a full period of *F*. The condition $d \le \lambda_0 / 2$ prevents the emergence of grating lobes. These properties follow immediately from the analysis of Sect. 7.3.

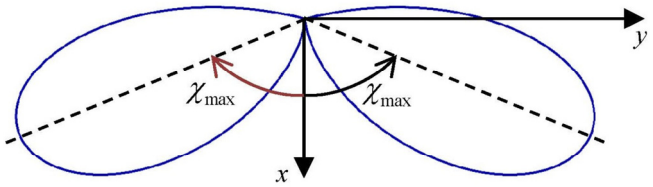
For a progressive phaseshift array, the direction of the maximum of $|F|$ can be found explicitly. In fact, from the "triangle inequality", $|a+b| \le |a| + |b|$, one obtains:

$$|F(u)| = \left| \sum_{n=0}^{N-1} |\underline{I}_n| e^{+jun} \right| \le \sum_{n=0}^{N-1} \left| |\underline{I}_n| e^{+jun} \right| = \sum_{n=0}^{N-1} |\underline{I}_n| = F(0). \tag{7.22}$$

This means that the maximum of $|F|$ occurs for $u = 0, \pm 2\pi, \pm 4\pi, \ldots$, where the periodicity of the function is taken into account:

$$\max |F(u)| = |F(0)|, \qquad \text{(reached for } u = 0, \pm 2\pi, \pm 4\pi, \ldots) \tag{7.23}$$

Maximum of the array factor



The maximum in the visible region is typically $u = 0$. The physical direction of observation ($\chi_{\max}$) associated with the maximum of the array spatial factor can be obtained by solving $0 = u = k_0 d \cos \chi_{\max} + \delta$. This yields the condition:

$$\boxed{\cos \chi_{\max} = \frac{-\delta}{k_0 d}}. \tag{7.24}$$

Thus, the direction of the maximum is controlled by the current phaseshift angle δ. The maximum can be sweeped full range $0 \le \chi_{\max} \le 180^\circ$, by varying δ in the range $-k_0 d \le \delta \le k_0 d$

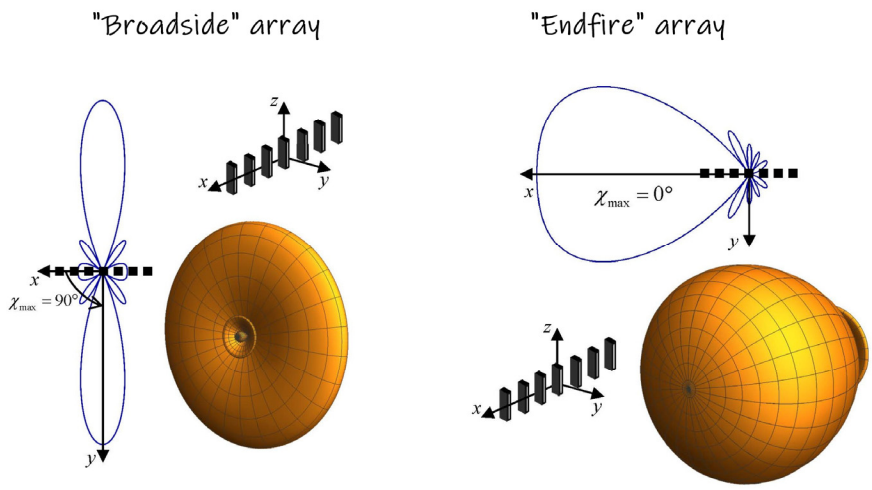
(here it is implicit that $k_0 d \le \pi$, i.e., $d \le \lambda_0 / 2$, so that the array spatial factor has a well defined maximum and there are no grating lobes).

When the direction of the array factor maximum is perpendicular to the array axis ($\chi_{\max} = 90^{\circ}$) the array is said to be "broadside". The current phaseshift is zero:

$$\boxed{\chi_{\max} = 90^{\circ} \Leftrightarrow \delta = 0}, \qquad \text{(broadside array).} \tag{7.25}$$

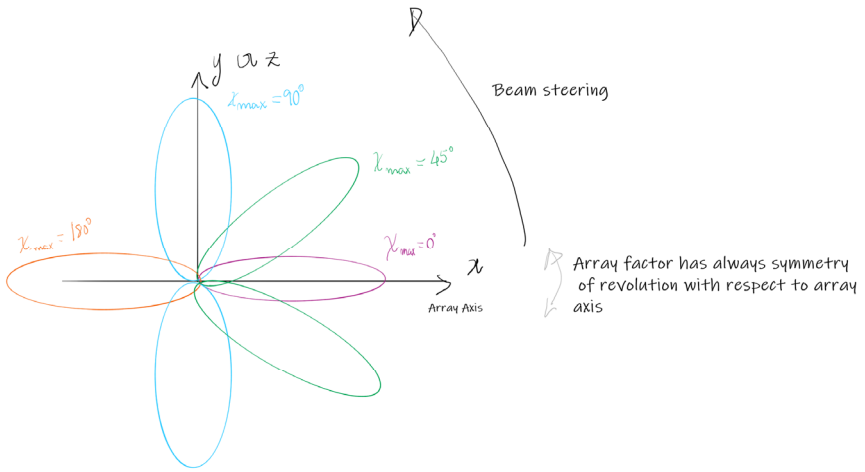
On the other hand, when the direction of the array factor maximum is along the array axis ($\chi_{\max} = 0^{\circ}$) the array is said to be "endfire". The corresponding current phaseshift is given by

$$\boxed{\chi_{\max} = 0^{\circ} \Leftrightarrow \delta = -k_0 d}, \qquad \text{(ordinary endfire array).} \tag{7.26}$$

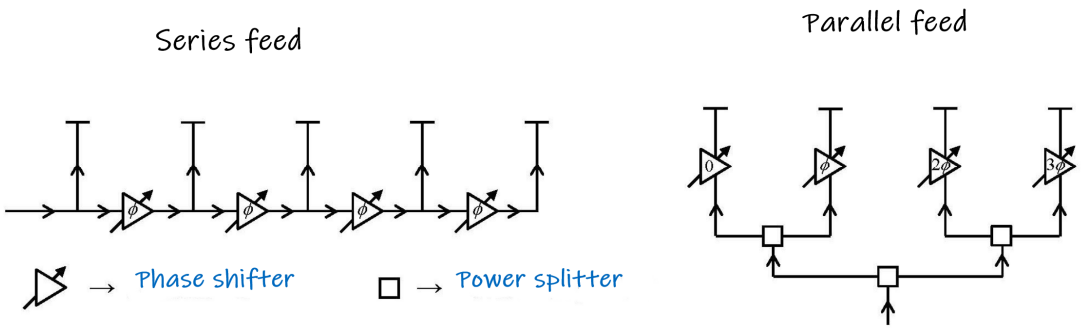


7.5 Electronic beam steering

The progressive phaseshift arrays create the opportunity to control *electronically* the direction of maximum radiation of an antenna system, without moving mechanically the antennas. This is particularly useful in RADAR systems where the antennas are supposed to sweep continuously the region of space under observation.



In practice, the required phaseshifts of the currents can be obtained either through a series or a parallel feed arrangement, as illustrated in the figure below. The circuitry of the series feed is simpler than the parallel feed, as it requires a single control signal to steer the beam (all the phaseshifts are identical). In contrast, the parallel feed requires *N-1* control signals. Nevertheless, the series arrangement is more sensitive to the insertion loss of the phase-shifters, as the signal that feeds the last antenna has to go through all the phase-shifters. In the parallel feed, all the signals go through a unique phase-shifter.



7.6 Uniform array

The uniform array is a particular case of a progressive phaseshift array in which the currents have identical amplitudes ($|\underline{I}_n| = |\underline{I}_0|$). The complex amplitudes of the currents satisfy

$$\underline{I}_n = |\underline{I}_0| e^{jn\delta}, \qquad n = 0,1,\ldots,\ N\text{-}1 \qquad \text{(uniform array).} \tag{7.27}$$

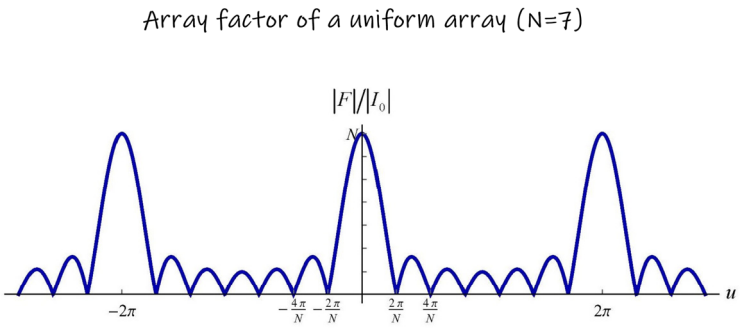
The array factor of the uniform array is a truncated geometrical series, and thereby can be evaluated explicitly:

$$F = \sum_{n=0}^{N-1} \left|\underline{I}_0\right| e^{+j\left(\underbrace{k_0 d\cos\chi+\delta}_{u}\right)n} = \\ = \left|\underline{I}_0\right| \sum_{n=0}^{N-1} e^{+jun} = \left|\underline{I}_0\right| \frac{1-e^{juN}}{1-e^{ju}} = \left|\underline{I}_0\right| e^{ju\frac{N-1}{2}} \frac{\sin Nu/2}{\sin u/2} \tag{7.28}$$

In particular, the absolute value of the array spatial factor satisfies

$$\boxed{\left|F\right| = \left|\underline{I}_0\right| \times \left|\frac{\sin Nu/2}{\sin u/2}\right|}, \qquad \text{(uniform array).} \tag{7.29}$$

The plot of this function is shown in the figure below for the case N=7.



As expected, due to the progressive phaseshift of the currents, the maxima of $\left|F\right|$ occur for $u = 0, \pm 2\pi, \pm 4\pi, \ldots$:

$$\max\left|F\right| = N, \qquad \text{(reached for } u = 0, \pm 2\pi, \pm 4\pi, \ldots\text{).} \tag{7.30}$$

The zeros of $\left|F\right|$ are zeros of $\sin Nu/2$, i.e., $Nu/2 = 0, \pm\pi, \pm 2\pi, \ldots$, so that $u = \cancel{0}, \pm\frac{1}{N}\times 2\pi, \pm\frac{2}{N}\times 2\pi, \ldots, \pm\frac{N-1}{N}2\pi, \cancel{\pm 2\pi}, \ldots$. The crossed elements must be discarded because they are also zeros of the denominator of $\left|F\right|$ (they correspond to the maxima). Hence, the nulls of the array spatial factor are:

$$\text{nulls } \left|F\right| = \pm\frac{1}{N}\times 2\pi, \pm\frac{2}{N}\times 2\pi, \ldots, \pm\frac{N-1}{N}2\pi, \ldots. \tag{7.31}$$

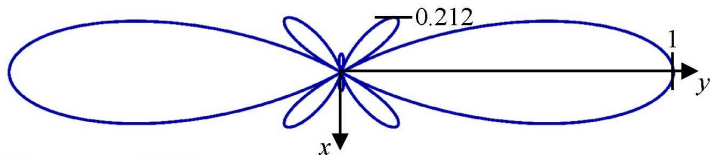
The secondary maxima of $|F|$ are associated with sidelobes. They occur approximately midway between the nulls. The sidelobe with largest amplitude is the one nearest to the maximum of $|F|$ and occurs roughly at $u \approx 3\pi / N$. The sidelobe level (SLL) characterizes the amplitude of the sidelobes relative to the main beam. It is defined as:

$$\text{SLL}=\frac{\text{maximum amplitude of the sidelobes}}{\text{amplitude of the main beam}}. \tag{7.32}$$

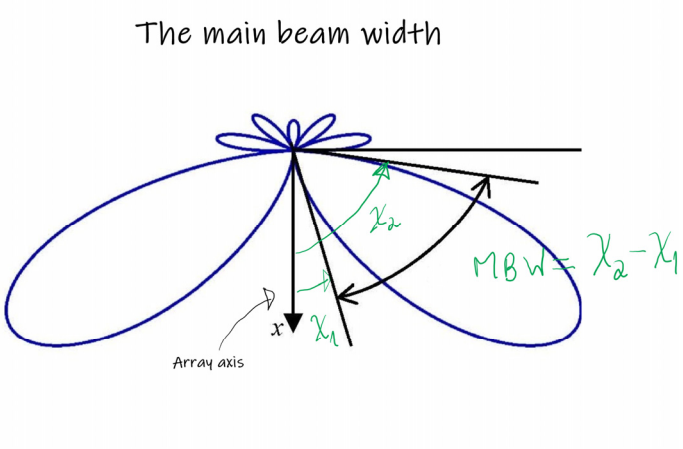
For a uniform array the sidelobe level is approximately:

$$\text{SLL}\Big|_{\substack{\text{uniform}\\ \text{array}}} \approx \frac{\left|F\left(3\pi/N\right)\right|}{\left|F\left(0\right)\right|} = \frac{1}{\left|N\sin\frac{3\pi}{2N}\right|} \underset{N>>1}{\approx} \frac{1}{\left|N\frac{3\pi}{2N}\right|} = \frac{2}{3\pi}. \tag{7.33}$$

Importantly, the sidelobe level of the uniform array cannot be made smaller than $\frac{2}{3\pi} = 0.212$ no matter how many elements are added to the array (in logarithmic units, $\text{SLL}\Big|_{\substack{\text{uniform}\\ \text{array}}} \approx 20\log_{20}\frac{2}{3\pi} = -13.5 \text{ dB}$).



In contrast, the directivity of the array can be increased as much as one may wish by increasing *N*. The directional properties of the array can be characterized in a simple way through the "main beam width" (MBW). The main beam width gives the angular width of the main beam determined by directions of the zeros.



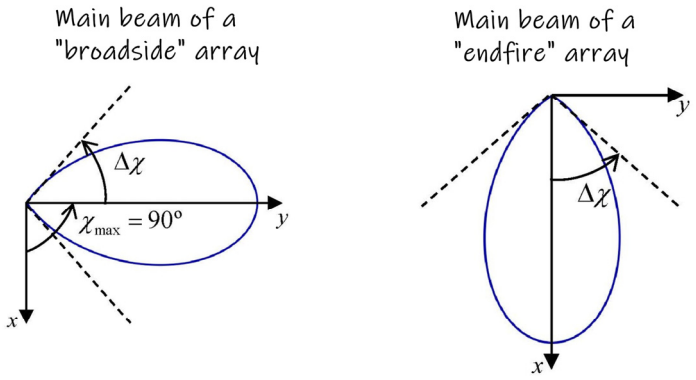
For a broadside array ($\chi_{\max} = 90^{\circ} \Leftrightarrow \delta = 0$), the MBW can be found noting that the two zeros that delimit the main beam are associated with $u = \pm 2\pi / N$. Since for a broadside array $u = k_0 d \cos\chi$, this implies that $\mp 2\pi / N = k_0 d \cos\left(90^{\circ} \pm \Delta\chi\right)$ where $\Delta\chi$ is the angle represented in the figure below (left). Since $\cos\left(90^{\circ} \pm \Delta\chi\right) = \mp \sin\left(\Delta\chi\right)$, it follows that:

$$\Delta\chi = \arcsin\frac{2\pi}{N \times k_0 d} = \arcsin\frac{\lambda_0}{N \times d}.$$

Noting that $\mathrm{MBW} = 2 \times \Delta\chi$, one concludes that:

$$\boxed{\mathrm{MBW}_{\mathrm{BS}} = 2\arcsin\frac{\lambda_0}{N \times d} \underset{N>>1}{\approx} \frac{2\lambda_0}{N \times d}}. \tag{7.34}$$

This formula confirms that by increasing the number of array elements, the main beam can be as narrow as one may wish.



A similar analysis shows for the ordinary endfire array $\Delta\chi = 2\arcsin\sqrt{\frac{\lambda_0}{2N\times d}}$. Thus, the corresponding main beam width is:

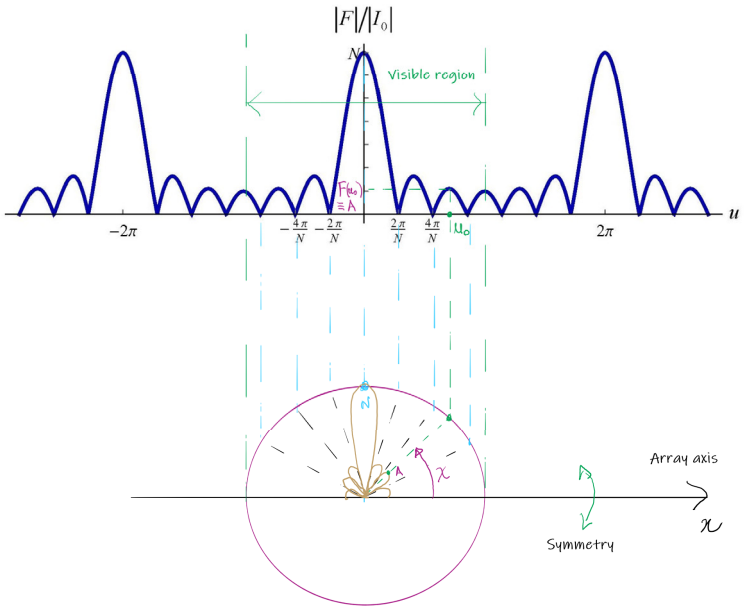
$$\boxed{\mathrm{MBW}_{\mathrm{OE}} = 4\arcsin\sqrt{\frac{\lambda_0}{2N\times d}} \underset{N>>1}{\approx} 2\sqrt{\frac{2\lambda_0}{N\times d}}}. \tag{7.35}$$

The MBW of the ordinary endfire array can also be arbitrarily small, but it approaches zero with $\sqrt{\frac{1}{N}}$, which is much slower than the $\frac{1}{N}$ scaling law of the broadside case. Due to this reason, broadside arrays are usually much more directive than endfire arrays with the same number of elements.

7.7 Graphical construction of the array radiation pattern

For any linear array with equidistant elements, the radiation pattern of $|F|$ can be found from the plot of $|F(u)|$ using the geometrical construction sketched in the figure below.

Graphical construction of the radiation pattern of the array



The construction is done as follows:

- Represent $|F(u)|$ as a function of u and identify the visible region. The visible region depends on the phaseshift δ and on the distance between the array elements. In the example of the figure, $\delta = 0^{\circ}$ (broadside array).
- Represent a circumference with diameter identical to the visible region on the array axis.
- The amplitude of $|F|$ along a generic direction of observation χ (represented in purple in the figure) is obtained by finding in the graphic the value of $|F|$ determined by the point u_0. The vertical projection of u_0 intersects the circumference along the direction χ.

7.8 The Dolph-Chebyshev array

From the practical point of view, the beam of a broadside array should be as narrow as possible, the gain a maximum, and the sidelobes, if any, should be at a low level. It is often a difficult matter to reconcile these demands. For example, the gain can be made a maximum by feeding all the antennas with currents with the same amplitude and phase (uniform array). Although it is true that this current distribution results in a narrow beam, it also results in high sidelobes which are only 13.5 dB down on the main beam.

Dolph[1] in 1946 proposed a solution for this problem based on the properties of Chebyshev polynomials. He introduced an *optimal* current distribution such that for a fixed number of array elements (i) if the sidelobe level (SLL) is specified the main beam width (MBW) is as narrow as possible. (ii) if the first null is specified, the sidelobe level is minimized. The design of the Dolph array is completely determined by either the SLL or, alternatively, by the MBW. In the Dolph array all the secondary lobes are at the same level.

[1] C. L. Dolph, "A Current Distribution for Broadside Arrays Which Optimizes the Relationship between Beam Width and Side-Lobe Level", Proceedings of the IRE, 34(6), 335–348, (1946).

Chebyshev polynomials

We start by reviewing some useful properties of Chebyshev polynomials. A Chebyshev polynomial of order *m* can be written as:

$$T_m(x) = \begin{cases} \cos(m \arccos x), & |x| \le 1 \\ \cosh(m \operatorname{arcosh} x), & x > 1 \end{cases}. \tag{7.36}$$

From the definition,

$$T_0(x) = 1, \qquad T_1(x) = x. \tag{7.37}$$

Denoting $w = \arccos x$, one sees that

$$T_{m+1} = \cos\big((m+1)w\big) = \cos(mw)\cos w - \sin(mw)\sin w$$

$$T_{m-1} = \cos\big((m-1)w\big) = \cos(mw)\cos w + \sin(mw)\sin w$$

Using $x = \cos w$, one obtains the recursion relation:

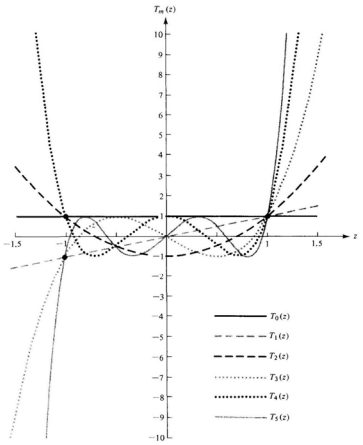
$$T_{m+1} + T_{m-1} = 2xT_m \tag{7.38}$$

which confirms that T_m are really polynomials. For example, from the recursion formula one finds that $T_2 + T_0 = 2xT_1$, so that $T_2 = 2x^2 - 1$. Similarly, from $T_3 + T_1 = 2xT_2$ one gets $T_3 = 4x^3 - 3x$, and so on. It is clear from the recursion formula and from (7.36) that the Chebyshev polynomials have the following properties:

- T_m is a polynomial with real-valued coefficients.
- T_m is an even function when *m* is even, and T_m is an odd function when *m* is odd.
- $|T_m(x)| \le 1$ for $|x| \le 1$ and $|T_m(x)| > 1$ for $|x| > 1$. Furthermore, $T_m(1) = 1$.
- All the zeros of T_m are in the interval $-1 < x < 1$.

A sketch of the first few Chebyshev polynomials is shown in the next figure.

Chebyshev polynomials



From $T_m(x)=\cos\left(m\underbrace{\arccos x}_{w}\right)$, the zeros $x_i=\cos w_i$ of the polynomial T_m are such that $w=w_i=\dfrac{\pi}{2m}(2i-1)$ with $i=0,\pm1,\pm2,\ldots$. Thus, the *m* distinct zeros are:

$$x_i=\cos\left(\frac{\pi}{2m}(2i-1)\right),\qquad i=1,2,\ldots,m,\quad (\text{zeros of } T_m,\ m\geq 1). \tag{7.39}$$

The extrema (local maxima or local minima) of T_m are coincident with the points inside the interval $|x|<1$ where $T_m(x)=\pm1$. The extreme points $\tilde{x}_i=\cos\tilde{w}_i$ are thus determined by $\tilde{w}_i=\dfrac{\pi}{m}i$ with *i* an integer, or equivalently by (the range of *i* is restricted to ensure that the condition $|x|<1$ is satisfied):

$$\tilde{x}_i=\cos\left(\frac{\pi}{m}i\right),\qquad i=1,2,\ldots,m-1,\qquad (\text{extrema of } T_m). \tag{7.40}$$

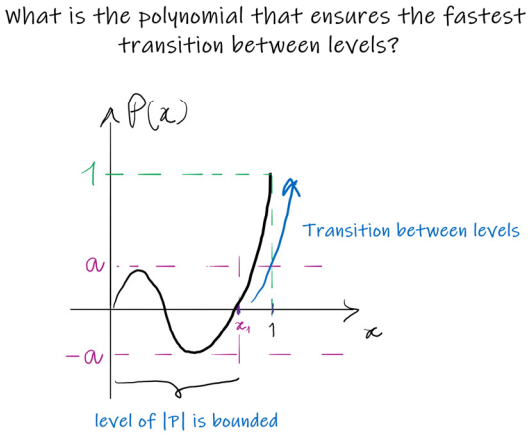
The Chebyshev polynomials are of particular relevance in engineering (array theory, filter theory, and others) and optimization problems due to the theorem enunciated next:

Theorem: Let $a<1$ be a fixed positive number, and consider the class of polynomials $P(x)$ of degree *N* with the following properties:

- $P(x)$ has real-valued coefficients and $P(x)$ is either an even or an odd function.

- All the zeros of $P(x)$ lie in the interval $|x|<1$.
- $P(1)=1$ and $|P(x)|\le a$ for $|x|\le|x_1|<1$ with x_1 the zero of P with the largest amplitude.

Then, among all the polynomials in this class, the one with the largest $|x_1|$ is a Chebyshev-type polynomial[1].



In simple terms, the theorem establishes that among all the polynomials of order *N* the one that can jump *faster* from a "zero" to some reference level "1" ensuring that in the rest of the domain the response amplitude is below some reference level "*a*" is a Chebyshev-type polynomial. Thus, Chebyshev polynomials are the ones that for a given complexity of a system (determined by the order of the polynomial) ensure faster transitions between states for a given level of "ripple" (in array theory, a given level of sidelobes).

[1] Specifically, the theorem establishes that the polynomial with the largest $|x_1|$ is $Q(x)=\frac{1}{a}T_N(xu_0)$ where $T_N(u_0)=a$ with $x_{1C}<u_0<1$ and $x_{1C}=\max|x_1|=\cos\left(\frac{\pi}{2N}\right)$. A proof of the theorem can be found in Dolph's article.

Dolph's optimal array

Dolph's array is a progressive phaseshift array with *N* elements described an array spatial factor $F(u)$ whose amplitude is a Chebyshev polynomial:

$$\boxed{|F| \sim |T_{N-1}(x)|}, \qquad \text{with } x = x_0 \cos\frac{u}{2} \qquad \text{(Dolph-Chebyshev array).} \qquad (7.41)$$

The parameter $x_0 > 1$ is the only degree of freedom in the array design; it is fixed either by the sidelobe level or, alternatively, by the main beam width. The parameter *u* is defined as in Eq. (7.19).

Before discussing the design of Dolph's array, let us verify that Eq. (7.41) really determines a progressive phaseshift array. To begin with, it is supposed that the array has an odd number of elements *N*. Then, $m = N-1$ is an even number and T_m is of the form $T_m = c_0 + c_1 x^2 + \ldots + c_{m/2} x^m = \sum_{n=0}^{m/2} c_n x^{2n}$ for some coefficients c_n. Noting that $\cos\frac{u}{2} = \frac{1}{2}\left(e^{ju/2} + e^{-ju/2}\right)$ it is possible to write:

$$\cos^2\left(\frac{u}{2}\right) = \frac{1}{4}\left(e^{ju} + 2 + e^{-ju}\right). \qquad (7.42)$$

This shows that for $x = x_0 \cos\frac{u}{2}$ one has

$$T_m = \sum_{n=0}^{m/2} c_n \left(x^2\right)^n = \sum_{n=0}^{m/2} \frac{c_n x_0^{2n}}{4^n}\left(e^{ju} + 2 + e^{-ju}\right)^n. \qquad (7.43)$$

Therefore, T_m is clearly a truncated Fourier series in the variable *u*, i.e., some some function of the form:

$$F = \sum a_n e^{jnu}.$$

Thereby, the Dolph's relation $|F| \sim |T_{N-1}(x)|$ really determines the array spatial factor of a linear array with equidistant elements (see Eq. (7.19)).

Let us illustrate the ideas with an example. Suppose that *N=3* so that $|F| \sim |T_2(x)|$. Using $T_2(x) = 2x^2 - 1$ and $x = x_0 \cos\frac{u}{2}$ one finds with the help of Eq. (7.42) that,

$$\begin{aligned} T_2 &= 2x_0^2 \frac{1}{4}\left(e^{ju} + 2 + e^{-ju}\right) - 1 \\ &= \frac{x_0^2}{2} e^{-ju} + \left(x_0^2 - 1\right) + \frac{x_0^2}{2} e^{ju} \end{aligned}$$

which is evidently a truncated Fourier series in *u*. From here and from Eq. (7.19), one finds that

$$F\big|_{\substack{\text{Dolph} \\ N=3}} = \sum_n \left|\underline{I}_n\right| e^{+jun} \sim \frac{x_0^2}{2} e^{-ju} + \left(x_0^2 - 1\right) + \frac{x_0^2}{2} e^{ju}. \tag{7.44}$$

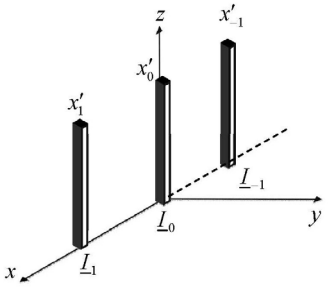
Note that the coefficients of the Fourier series determine the amplitudes of the feeding currents, apart from a scaling factor. They are all positive real numbers, as it should be, because $x_0 > 1$.

From the general theory of Sect. 7.3, the coefficient e^{jnu} is associated with an antenna positioned at $x_n' = nd$. If all the antennas are positioned along the *positive* *x*-axis only positive *n*'s should be allowed. It is always possible to enforce this, by factoring out exponential factors of the form e^{-ju}. For example, in the current example we could write:

$$F\big|_{\substack{\text{Dolph} \\ N=3}} = \sum_n \left|\underline{I}_n\right| e^{+jun} \sim \frac{x_0^2}{2} + \left(x_0^2 - 1\right) e^{+ju} + \frac{x_0^2}{2} e^{j2u}$$

However, enforcing positive *n*'s is not strictly required if one allows the array elements to be positioned on the negative *x*-axis, which is of course totally fine. For example, Eq. (7.44) may describe a situation where the antennas are placed at $x_{-1}' = -d$, $x_0' = 0$ and $x_1' = +d$, with the feeding currents satisfying the proportionality relations:

$$\left|\underline{I}_{-1}\right| : \left|\underline{I}_0\right| : \left|\underline{I}_1\right| = \frac{x_0^2}{2} : \left(x_0^2 - 1\right) : \frac{x_0^2}{2}$$



An interesting feature of the Dolph's array is that it is a *symmetric* array. Specifically, the current distribution is symmetric about the center of the array, such that elements equidistant from the center are fed with currents with the same amplitude. For instance, for the case *N=3* analyzed previously one has $\left|\underline{I}_{-1}\right| = \left|\underline{I}_{1}\right|$ [1].

Due to monotonic behaviour of the Chebyshev polynomials for $|x| > 1$, maximum of the array factor $|F| \sim \left|T_{N-1}\left(x_0 \cos\frac{u}{2}\right)\right|$ is clearly reached for $u = 0$, that is:

$$\max|F| \sim \left|T_{N-1}(x_0)\right|, \qquad (x_0 > 1). \tag{7.45}$$

This property is consistent with the fact that the array is a progressive phaseshift array.

The results discussed so far can be generalized in a straightforward way to the case wherein the number of array elements *N* is even. In that situation, T_m (with $m = N-1$) is an odd function, and thereby can be written as $T_m = x \times (\text{even function})$. Upon the replacement $x = x_0 \cos\frac{u}{2}$, the even function becomes a truncated Fourier series in the variable *u*. Noting that $xe^{+ju/2} = \frac{x_0}{2}\left(1 + e^{ju}\right)$, it is evident that $T_m e^{+ju/2}$ is also a truncated Fourier series, and hence can

[1] One can check that this a general property. For example, for *N* odd, $T_m\left(x_0 \cos\frac{u}{2}\right)$ is an even function, and thereby its Fourier series $\sum a_n e^{jnu}$ has coefficients given by $a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} T_m\left(x_0 \cos\frac{u}{2}\right)\cos(nu)$. From this formula it is clear that $a_n = a_{-n}$, i.e., that the array is symmetric.

model the array factor of a radiating system. Hence, for N even we take $F \sim T_m e^{+ju/2}$. For example, for N=4, one has

$$T_3 = 4x^3 - 3x = x \times \underbrace{\left(4x^2 - 3\right)}_{\text{even function}}$$
$$= x_0 \left(\frac{e^{ju/2} + e^{-ju/2}}{2} \right) \times \left(4x_0^2 \frac{1}{4} \left(e^{ju} + 2 + e^{-ju} \right) - 3 \right)$$

From here,

$$F\big|_{\substack{\text{Dolph} \\ N=4}} \sim T_3 e^{+ju/2} = x_0 \left(\frac{e^{ju} + 1}{2} \right) \times \left(4x_0^2 \frac{1}{4} \left(e^{ju} + 2 + e^{-ju} \right) - 3 \right)$$
$$= \frac{x_0}{2} \left(x_0^2 e^{-ju} + \left(3x_0^2 - 3\right) + \left(3x_0^2 - 3\right) e^{ju} + x_0^2 e^{j2u} \right)$$
$$\sim \left(x_0^2 e^{-ju} + \left(3x_0^2 - 3\right) + \left(3x_0^2 - 3\right) e^{ju} + x_0^2 e^{j2u} \right)$$

One can interpret this spatial factor as being associated with elements centered at $x_n' = nd$ with $n = -1, 0, 1, 2$ and with the feeding currents satisfying the proportionality relations:

$$\left|\underline{I}_{-1}\right| : \left|\underline{I}_0\right| : \left|\underline{I}_1\right| : \left|\underline{I}_2\right| = x_0^2 : \left(3x_0^2 - 3\right) : \left(3x_0^2 - 3\right) : x_0^2$$

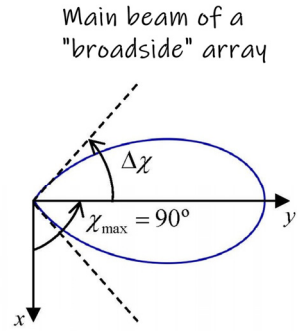
Note that similar to the first example, the feeding currents distribution is symmetric about the center of the array.

Design of a Dolph-Chebyshev broadside array

As previously discussed, for a Dolph-Chebyshev array the maximum of the array factor is $\left|F\right|_{\max} = \left|F\right|_{u=0} \sim \left|T_{N-1}\left(x_0\right)\right|$. For a broadside array ($\delta = 0$) the maximum is evidently in the visible zone. The amplitude of the sidelobes of the Dolph array is determined by the amplitude of the extrema of $\left|T_{N-1}\left(x\right)\right|$, which from the properties of the Chebyshev polynomials is "1". Thus, the sidelobe level is given by (here it is assumed that the visible zone is wide enough to contain that a least one sidelobe):

$$\boxed{\text{SLL}_{\text{Dolph}} = \frac{1}{\left|T_{N-1}\left(x_0\right)\right|}}. \tag{7.46}$$

The nulls of $|F|$ are determined by the nulls of the Chebyshev polynomial [Eq. (7.39)]. Since $x = x_0 \cos\frac{u}{2}$ and since for a broadside array $u = k_0 d \cos\chi$, one can write $x = x_0 \cos\left(\frac{k_0 d}{2}\cos(90^\circ + \Delta\chi)\right) = x_0 \cos\left(\frac{k_0 d}{2}\sin(\Delta\chi)\right)$ for directions $\chi = \chi_{\max} + \Delta\chi$ near the maximum $\chi_{\max} = 90^\circ$.



The null closest to the main beam is $x_1 = \cos\left(\frac{\pi}{2m}\right)$ with $m = N - 1$ [Eq. (7.39)]. Thus, the main beam width is determined by $\mathrm{MBW} = 2\Delta\chi$ with $x_0 \cos\left(\frac{k_0 d}{2}\sin(\Delta\chi)\right) = \cos\left(\frac{\pi}{2m}\right)$. Solving with respect to $\Delta\chi$ one finds that:

$$\boxed{\mathrm{MBW}_{\mathrm{Dolph}} = 2\arcsin\left(\frac{2}{k_0 d}\arccos\left(\frac{1}{x_0}\cos\left(\frac{\pi}{2(N-1)}\right)\right)\right)}. \tag{7.47}$$

The design of the Dolph array is based on Eqs. (7.46) and (7.47).

If the SLL is specified, one can find the required x_0 from Eqs. (7.36) and (7.46):

$$\boxed{x_0 = \cosh\left(\frac{1}{N-1}\operatorname{arccosh}^{-1}\frac{1}{\mathrm{SLL}}\right)}. \tag{7.48}$$

Alternatively, if the MBW is specified, one can find the required x_0 from Eq. (7.47):

$$\boxed{x_0 = \frac{\cos\left(\dfrac{\pi}{2(N-1)}\right)}{\cos\left(\dfrac{k_0 d}{2}\sin\left(\dfrac{\mathrm{MBW}}{2}\right)\right)}}. \tag{7.49}$$

Design example

As an illustration, consider the design of a broadside Dolph-Chebyshev array with $N = 5$ elements, $d = \lambda_0 / 2$ and $\mathrm{SLL} = 0.1$. From Eq. (7.48) one finds that:

$$x_0 = \cosh\left(\frac{1}{5-1}\operatorname{arccosh}^{-1}\frac{1}{0.1}\right) = 1.29$$

From the expression of T_4 one can show that the array factor is:

$$\begin{aligned} F\big|_{\substack{\mathrm{Dolph}\\N=5}} \sim T_4 &= \left(\frac{x_0^4}{2}e^{-2ju} + 2\left(x_0^4 - x_0^2\right)e^{-ju} + \left(3x_0^4 - 4x_0^2 + 1\right) + 2\left(x_0^4 - x_0^2\right)e^{ju} + \frac{x_0^4}{2}e^{2ju}\right) \\ &\sim \left(e^{-2ju} + \frac{4}{x_0^4}\left(x_0^4 - x_0^2\right)e^{-ju} + \frac{2}{x_0^4}\left(3x_0^4 - 4x_0^2 + 1\right) + \frac{4}{x_0^4}\left(x_0^4 - x_0^2\right)e^{ju} + e^{2ju}\right) \\ &\sim \left(1 + 1.61e^{ju} + 1.93e^{2ju} + 1.61e^{3ju} + e^{4ju}\right) \end{aligned}$$

We enforced that the first element of the array is centered at the origin. The array of 5 antennas must be fed with a current distribution obeying the proportionality relations $\left|\underline{I}_0\right| : \left|\underline{I}_1\right| : \left|\underline{I}_2\right| : \left|\underline{I}_3\right| : \left|\underline{I}_4\right| = 1 : 1.61 : 1.93 : 1.61 : 1$. The currents are in phase because the array is broadside.

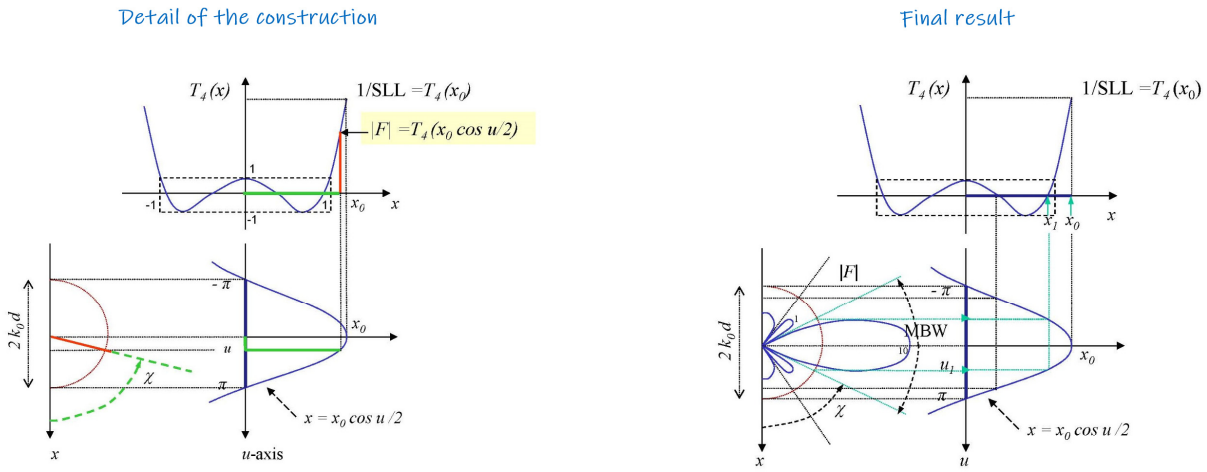
The main beam width is determined by Eq. (7.47):

$$\mathrm{MBW}_{\mathrm{Dolph}} = 2\arcsin\left(\frac{2}{\pi}\arccos\left(\frac{1}{1.29}\cos\left(\frac{\pi}{2\times 4}\right)\right)\right) = 59.1^{\mathrm{o}}$$

The radiation pattern of the array can be constructed graphically through a generalization of the procedure of Sect. 7.7. It is a two step procedure based on the plot of the Chebyshev polynomial and on the plot of the function $x = x_0 \cos\dfrac{u}{2}$. The amplitude of $\left|F\right|$ along a generic

direction χ is found starting from the intersection of χ with a circle with the same diameter as the visible zone ($2k_0 d$, which for $d = \lambda_0 / 2$ equals 2π). The intersection point suffers then a horizontal projection and a vertical projection, which yield the vertical line segment marked in red (left panel). The radial distance marked along the direction χ is identical to the length of the vertical line segment.

Graphical construction of the radiation pattern of the Dolph array

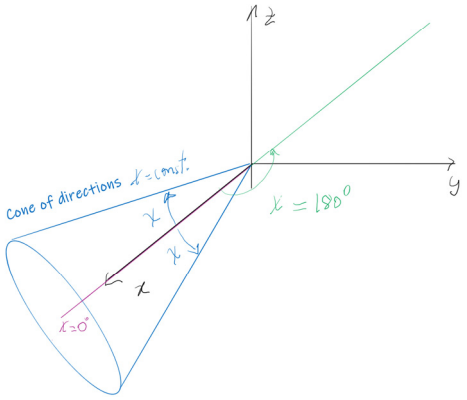


7.9 *Planar arrays*

The radiation pattern of a linear array has revolution symmetry about the array axis: $|F| = |F(\cos\chi)|$. The only direction of space that can be pinpointed by this type of radiation pattern is the array axis itself. Indeed, with the exceptions of $\chi = 0^\circ$ and $\chi = 180^\circ$, which correspond to the positive and negative *x*-axis, respectively, all the other angles $\chi = const.$ determine a conical surface.

In some applications, such as in beam steering, it is relevant to pinpoint an arbitrary direction of space. To break the revolution symmetry of $|F|$ and have additional control over the array spatial factor one needs to use a planar array.

The only directions that can be individually pinpointed by a linear array are the positive and negative array semi-axes

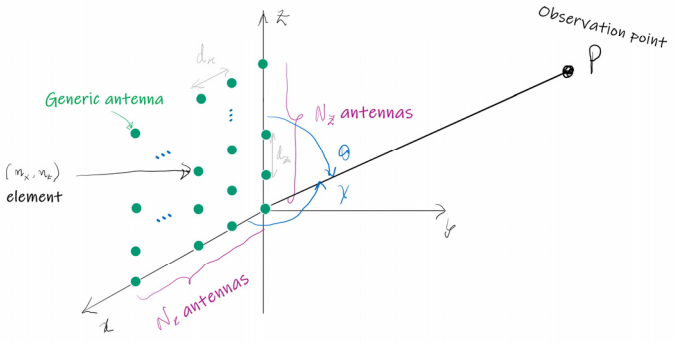


In a planar array, the array elements are positioned on a plane. The array plane is taken as the *xoz* plane. We focus on case wherein the array elements are equidistant along the two main (*x* and *z*) directions. The antennas are positioned at the points:

$$\mathbf{r}'_n = \mathbf{r}'_{(n_x,n_z)} = d_x n_x \hat{\mathbf{x}} + d_z n_z \hat{\mathbf{z}}, \qquad n_x = 0,...N_x - 1, \quad n_z = 0,...N_z - 1. \tag{7.50}$$

Here, d_x, d_z are the spacings along the *x* and *z* directions. The array is formed by $N_x \times N_z$ identical elements. Each array element is labelled by a pair of integers $n = (n_x, n_z)$. A generic direction of observation is specified uniquely by the angles χ, θ determined by the observation point and the two array axes (*x* and *z* axes): $\cos\chi = \hat{\mathbf{r}} \cdot \hat{\mathbf{x}}$ and $\cos\theta = \hat{\mathbf{r}} \cdot \hat{\mathbf{z}}$.

Planar array geometry



The array spatial factor of the planar array can be found from the general formula (7.7):

$$F(\theta,\chi)=\sum_{n=(n_x,n_z)} e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'_n}\underline{I}_{(n_x,n_z)}=\sum_{n=(n_x,n_z)}\underline{I}_{(n_x,n_z)}e^{+jn_x(k_0d_x\cos\chi)}e^{+jn_z(k_0d_z\cos\theta)}. \tag{7.51}$$

In the second identity, we used $\hat{\mathbf{r}}\cdot\mathbf{r}'_n=n_xd_x\cos\chi+n_zd_z\cos\theta$. Clearly, the planar array spatial factor is a function of two angles: $F=F(\theta,\chi)$.

A current distribution is *separable* if it is of the form:

$$\underline{I}_{(n_x,n_z)}=\frac{1}{\underline{I}_{(0,0)}}\underline{I}_{x,n_x}\times\underline{I}_{z,n_z}. \tag{7.52}$$

Here, $\underline{I}_{x,n_x}\equiv\underline{I}_{(n_x,0)}$ determines the profile of the currents in the *x*-axis, $\underline{I}_{z,n_z}\equiv\underline{I}_{(0,n_z)}$ determines the profile of the currents in the *z*-axis, and $\underline{I}_{(0,0)}$ is the feeding current of the element centered at the origin. Thus, a separable current distribution is completely determined by the feeding currents of the array elements placed along the *x* and *z* axis (there are N_x+N_z-1 independent currents). When the current distribution is separable, the array spatial factor can be written as a product of two array spatial factors:

$$F=\frac{1}{\underline{I}_{(0,0)}}\underbrace{\left(\sum_{n_x=0}^{N_x-1}\underline{I}_{x,n_x}e^{+jn_xk_0d_x\cos\chi}\right)}_{F_x(\chi)}\times\underbrace{\left(\sum_{n_z=0}^{N_z-1}\underline{I}_{z,n_z}e^{+jn_zk_0d_z\cos\theta}\right)}_{F_z(\theta)}. \tag{7.53}$$

This property can be regarded as consequence of the principle of multiplication of the pattern diagrams because a planar array with a separable current distribution may be regarded as an array of linear arrays! In particular, $F_x(\chi)$ ($F_z(\theta)$) is nothing but the array spatial factor of the linear array with equidistant elements formed by the antennas placed along the *x* (*z*) axis.

Let us now suppose that $\underline{I}_{x,n_x}=\left|\underline{I}_{(n_x,0)}\right|e^{j\delta_xn_x}$ and $\underline{I}_{z,n_z}=\left|\underline{I}_{(n_z,0)}\right|e^{j\delta_zn_z}$ so that the phases of the currents satisfy an arithmetic progression. Then, the array spatial factor becomes:

$$F=\frac{1}{\underline{I}_{(0,0)}}\underbrace{\left(\sum_{n_x=0}^{N_x-1}\left|\underline{I}_{x,n_x}\right|e^{+jn_xu_x}\right)}_{F_x(u_x)}\times\underbrace{\left(\sum_{n_z=0}^{N_z-1}\left|\underline{I}_{z,n_z}\right|e^{+jn_zu_z}\right)}_{F_z(u_z)}, \tag{7.54}$$

with $u_x = k_0 d_x \cos\chi + \delta_x$ and $u_z = k_0 d_z \cos\theta + \delta_z$. Thus, it is product of the array spatial factors of two progressive phaseshift linear arrays. The two spatial factors F_x, F_z are totally independent, and are controlled by the angles χ, θ, respectively.

For progressive phaseshift arrays the maximum occurs for the directions that satisfy $u = 0$. Thus, the maximum of $|F_x|$ occurs along the *cone* of directions that satisfies $u_x = 0$ ($\chi = \chi_{\max}$), and the maximum of $|F_z|$ occurs along the *cone* of directions that satisfies $u_z = 0$ ($\theta = \theta_{\max}$). These conditions yield (compare with Eq. (7.24)):

$$\boxed{\cos\chi_{\max} = \frac{-\delta_x}{k_0 d_x}}, \qquad \boxed{\cos\theta_{\max} = \frac{-\delta_z}{k_0 d_z}}. \tag{7.55}$$

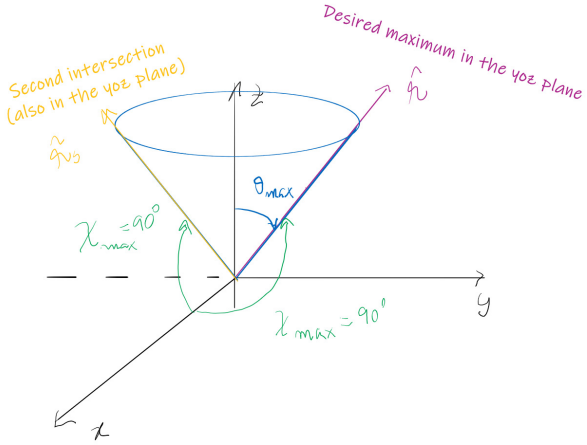
Provided the two cones determined by $\chi = \chi_{\max}$ and $\theta = \theta_{\max}$ have a non-trivial intersection[1], the maximum of $|F| = |F_x(u_x)||F_z(u_z)| / |\underline{I}_{(0,0)}|$ is reached when the two conditions are simultaneously satisfied,

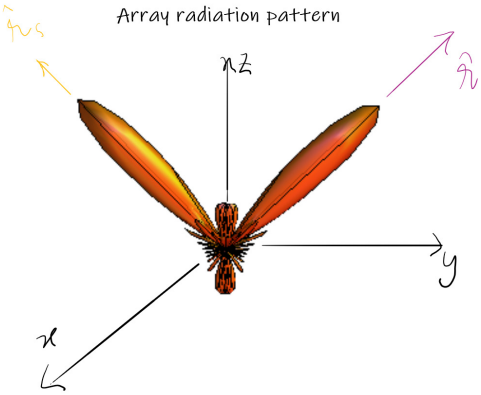
$$\max|F| \quad \Leftrightarrow \quad \chi = \chi_{\max} \text{ and } \theta = \theta_{\max}. \tag{7.56}$$

This discussion shows that for any given direction of observation $\hat{\mathbf{r}}$ it is possible to tailor the current phaseshifts δ_x and δ_z so that the array spatial factor pinpoints that direction.

It should be mentioned that even though the values of $\chi_{\max}$ and $\theta_{\max}$ are uniquely determined by $\hat{\mathbf{r}}$, the intersection of the cones $\chi = \chi_{\max}$ and $\theta = \theta_{\max}$ may not be unique. Indeed, two intersecting cones typically touch at two different semi-lines. Due to this reason, the spatial factor of a progressive phaseshift planar array typically has *two* maxima. The exceptions are the directions in the array plane, which can be pinpointed uniquely. The described limitation can only be overcome using space arrays.

[1]Even though any observation direction $\hat{\mathbf{r}}$ is associated with well defined values of θ, χ, not all the combinations of θ, χ lead to valid observation directions. Indeed, in order that $\cos\chi = \hat{\mathbf{r}} \cdot \hat{\mathbf{x}}$ and $\cos\theta = \hat{\mathbf{r}} \cdot \hat{\mathbf{z}}$ it is necessary that $\cos^2\theta + \cos^2\chi \le 1$.





To illustrate the ideas, let us consider a planar uniform array with a constant current amplitude distribution:

$$\underline{I}_{n_x,n_z} = \left|\underline{I}_{(0,0)}\right| e^{j\delta_x n_x} e^{j\delta_z n_z}, \qquad \text{(uniform planar array).} \tag{7.57}$$

The corresponding array spatial factor is the product of the array spatial factors of two uniform arrays:

$$\frac{|F|}{\left|\underline{I}_{(0,0)}\right|} = \left|\frac{\sin\left(N_x u_x/2\right)}{\sin\left(u_x/2\right)}\right| \left|\frac{\sin\left(N_z u_z/2\right)}{\sin\left(u_z/2\right)}\right|. \tag{7.58}$$

The figure above shows the array radiation pattern when the direction of maximal radiation is picked in the *yoz* plane and makes an angle of 45º with the *z*-axis. The number of antennas is 10×10 and the distance between antennas is $\lambda_0/2$. As seen, the direction of maximal radiation is not unique, as the array factor is formed by two main lobes. Nevertheless, this a considerable improvement as compared to a linear array, which has always a radiation pattern with revolution of symmetry about the array axis.

7.10 Power gain and gain in field intensity

The power gain of a generic antenna array is determined by

$$G_{\text{array}} = 4\pi \frac{U_{\text{array}}}{P_{\text{in,array}}}, \tag{7.59}$$

with U_{array} the radiation intensity of the array and $P_{\text{in,array}}$ the input power of the system. It can be related to the power gain of a generic array element (G_{single}) noting that:

$$\begin{aligned} G_{\text{array}} &= 4\pi \frac{U_{\text{single}}}{P_{\text{in,single}}} \frac{U_{\text{array}}}{U_{\text{single}}} \frac{P_{\text{in,single}}}{P_{\text{in,array}}} \\ &= G_{\text{single}} \frac{U_{\text{array}}}{U_{\text{single}}} \frac{P_{\text{in,single}}}{P_{\text{in,array}}} \\ &= G_{\text{single}} \frac{\left|\underline{\mathbf{E}}\right|^2_{\text{array}}}{\left|\underline{\mathbf{E}}\right|^2_{\text{single}}} \frac{P_{\text{in,single}}}{P_{\text{in,array}}} \end{aligned} \tag{7.60}$$

We introduce the "gain in field intensity" (G_f) as the parameter that gives the field radiated by the array divided by the field radiated by a single array element fed with the same input power as the array:

$$\boxed{\mathcal{G}_f = \text{gain in field intensity} = \frac{\left|\underline{E}\right|_{\text{array}}}{\left|\underline{E}\right|_{\substack{\text{single element fed} \\ \text{with same power}}}}}. \tag{7.61}$$

From this definition, it is clear that:

$$\boxed{G_{\text{array}} = G_{\text{single-element}} \mathcal{G}_f^2} \tag{7.62}$$

Let us derive an explicit formula for $\mathcal{G}_f$ in terms of the array spatial factor $\left|F\right|$. To begin with, we note that the power delivered to a single array element is $P_{\text{in,single}} = \frac{1}{2} R_a \left|\underline{I}(0)\right|^2$. Here, R_a is the input resistance of a single array element alone in free-space and $\underline{I}(0)$ is the feeding current. The condition $P_{\text{in,single}} = P_{\text{in,array}}$ requires that $\underline{I}(0)$ satisfies:

$$\left|\underline{I}(0)\right| = \sqrt{\frac{2P_{\text{in,array}}}{R_a}}. \tag{7.63}$$

Then, from the principle of multiplication of pattern diagrams [Eq. (7.6)], it follows that the gain in field intensity can be written as:

$$\mathcal{G}_f = \frac{|F|}{|\underline{I}(0)|} = \sqrt{\frac{R_a}{2P_{\text{in,array}}}}|F|. \tag{7.64}$$

which is the same as

$$\mathcal{G}_f = \sqrt{\frac{R_a}{2}\frac{|\underline{I}_0|^2}{P_{\text{in,array}}}}\frac{|F|}{|\underline{I}_0|}. \tag{7.65}$$

In the above $\underline{I}_0$ is the feeding current of the first (n=0) array element. Note that $P_{\text{in,array}}/|\underline{I}_0|^2, |F|/|\underline{I}_0|$ depend only on the *profile* of the current distribution.

The input power of the array can be written in terms of the impedance matrix $\mathbf{Z}$ determined by the self and mutual impedances. For a reciprocal system, one has $P_{\text{in,array}} = \frac{1}{2}\mathbf{I}^* \cdot \mathbf{R} \cdot \mathbf{I}$ with $\mathbf{R} = \text{Re}\{\mathbf{Z}\}$ the resistance matrix and $\mathbf{I} = [\underline{I}_0 \quad \underline{I}_1 \quad \ldots \quad \underline{I}_{N-1}]^T$ a column vector with the current distribution (see Sect. 4.5). Thus, the gain in field intensity can be found using:

$$\boxed{\mathcal{G}_f = \sqrt{R_a \frac{|\underline{I}_0|^2}{\mathbf{I}^* \cdot \mathbf{R} \cdot \mathbf{I}}}\frac{|F|}{|\underline{I}_0|}}. \tag{7.66}$$

Let us consider some particular cases. To begin, suppose that the coupling between the array elements is negligible. In such a case, the resistance matrix is approximately diagonal: $\mathbf{R} \approx R_a \mathbf{1}$. This leads to the simplified formula:

$$\mathcal{G}_f = \sqrt{\frac{|\underline{I}_0|^2}{\mathbf{I}^* \cdot \mathbf{I}}}\frac{|F|}{|\underline{I}_0|}, \qquad \text{(negligible coupling).} \tag{7.67}$$

Consider now a uniform array. Since current amplitudes are constants, $\mathbf{I}^* \cdot \mathbf{I} = N|\underline{I}_0|^2$, and thereby:

$$\mathcal{G}_f\Big|_{\substack{\text{uniform}\\ \text{array}}} = \frac{1}{\sqrt{N}}\frac{|F|}{|\underline{I}_0|} = \frac{1}{\sqrt{N}}\left|\frac{\sin\left(N\frac{u}{2}\right)}{\sin\left(\frac{u}{2}\right)}\right|, \quad \text{(negligible coupling).} \tag{7.68}$$

The maximum of the gain in field intensity occurs for $u = 0$ and is equal to $\mathcal{G}_{f,\max} = \sqrt{N}$. Thus, if the maximum of the array factor is aligned with the maximum of the radiation pattern of a generic element, it is possible to write the maximum power gain of the uniform array as:

$$\boxed{G_{\text{m,array}}\Big|_{\substack{\text{uniform}\\ \text{array}}} = G_{\text{m,generic-element}} \times N} \qquad \text{(negligible coupling + aligned maxima).} \qquad (7.69)$$

This proves that when the antenna coupling is negligible, the power gain is proportional to the number of array elements.

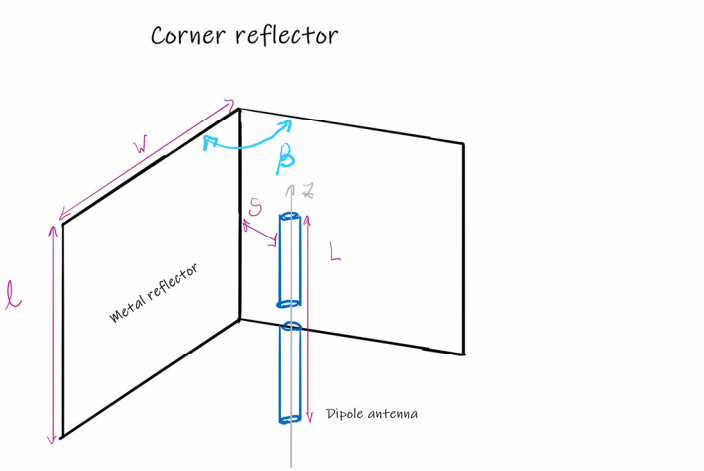
8. Wire-type antennas

8.1 Introduction

In this Chapter, we analyze several antennas of widespread use in communication systems, and that have in common the fact that they are formed by pieces of metallic wires (wire-type antennas). We will discuss the main principles of operation of the corner reflector, Yagi-Uda antenna, turnstile antenna, loop antenna and helical antenna.

8.2 Corner reflector

A simple way to increase the directivity of an antenna is to place it near a reflector that redirects the radiation to some desired angular region. This is the working principle of the corner reflector. It consists of two flat metal sheets intersecting at an angle β. The system is excited by a dipole antenna with height L placed at a distance S from the vertex of the reflector. The dipole is placed in the bisector plane of the reflector with the dipole axis aligned with the reflector edge.

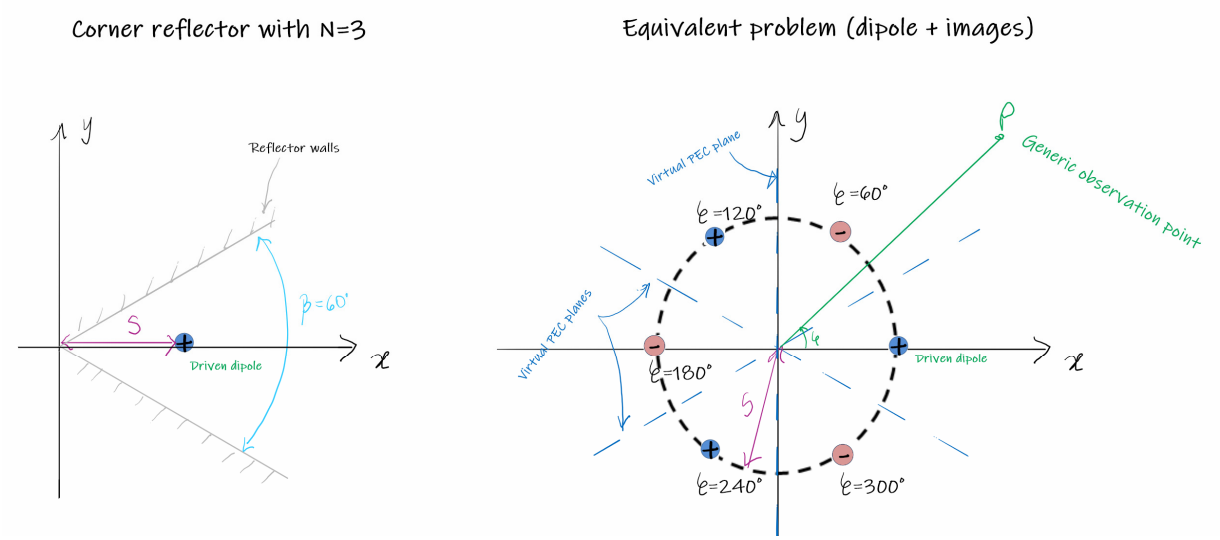


For analysis purposes, the metal sheets that form the reflector are assumed to be perfect conductors of infinite extent. This is usually a good approximation if $l \geq 1.2L$ and $w \geq 2S$. It is

also possible to consider designs where the metal sheets are replaced by a dense grid of metallic wires.

Illustration of corner reflector formed by an array of metal wires

It is difficult to determine the corner reflector radiation pattern for a generic β. Interestingly, when $\beta = 180^{\circ}/N$ with $N \geq 1$ one can obtain an analytical solution with the help of the image method. The idea is illustrated in the figure below for the case $N = 3$ ($\beta = 60^{\circ}$).



The fields radiated by the dipole in the vicinity of the corner reflector are identical to the fields radiated by a set of $2N$ dipoles uniformly distributed in a circle of radius S along the azimuthal direction ($\varphi_n = n\beta, n = 0,1,...,2N-1$), and fed with currents with alternating ($\pm$) sign. In fact, due to the image method, such a current distribution ensures that perfect electric conductor (PEC) planes can be placed along the directions $\varphi_m^{\mathrm{PEC}} = \beta/2 + m\beta$, $m = 0,1,...,$

without affecting the radiated fields. Note that the currents flowing in the dipoles are "horizontal" with respect to the virtual PEC planes.

The fields radiated by the dipole and its $2N-1$ images in the far-zone can be found using array theory:

$$\left|\underline{\mathbf{E}}\right|_{\text{dipole + images}} = \underbrace{\left|F\right|}_{\text{array spatial factor}} \times \left.\frac{\left|\underline{\mathbf{E}}\right|}{\left|\underline{I}(0)\right|}\right|_{\text{generic element}} . \tag{8.1}$$

The array spatial factor is computed using $F = \sum e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'_n}\underline{I}_n$. The antennas are centered at $\mathbf{r}'_n = S\left(\cos\varphi_n\hat{\mathbf{x}} + \sin\varphi_n\hat{\mathbf{y}}\right)$ with $\varphi_n = n\beta$. Taking into account, that the current phases change by π from element to element one obtains $F = \underline{I}_0\sum_{n=0}^{2N-1}(-1)^n e^{+jk_0S\hat{\mathbf{r}}\cdot(\cos n\beta\hat{\mathbf{x}}+\sin n\beta\hat{\mathbf{y}})}$. This formula can also be written as:

$$\boxed{F\big|_{\substack{\text{corner}\\\text{reflector}}} = \underline{I}_0\sum_{n=0}^{2N-1}(-1)^n e^{+jk_0S\cos(\varphi-n\beta)\sin\theta}}. \tag{8.2}$$

Hence, the field radiated by the dipole in the vicinity of the corner reflector is:

$$\left|\underline{\mathbf{E}}\right|_{\text{dipole +corner reflector}} = \left|F_{\substack{\text{corner}\\\text{reflector}}}\right| \times \frac{\left|\underline{\mathbf{E}}\right|_{\text{dipole}}}{\left|\underline{I}(0)\right|} . \tag{8.3}$$

The radiated field has the same polarization as a single dipole ($\underline{\mathbf{E}} = \underline{E}_\theta\hat{\boldsymbol{\theta}}$) because the array spatial factor cannot change the antenna polarization (because it is a scalar). When the dipole is has length $L = \lambda_0/2$, using the sinusoidal current approximation, one can write $\left|\underline{\mathbf{E}}\right|_{\text{dipole}} = \left|\underline{\mathbf{E}}\right|_{\lambda/2}$ with

$$\left|\underline{\mathbf{E}}\right|_{\text{dipole}} = \left|\underline{\mathbf{E}}\right|_{\lambda/2} = \left|\underline{I}(0)\right|\frac{60}{r}\left|\frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta}\right| . \tag{8.4}$$

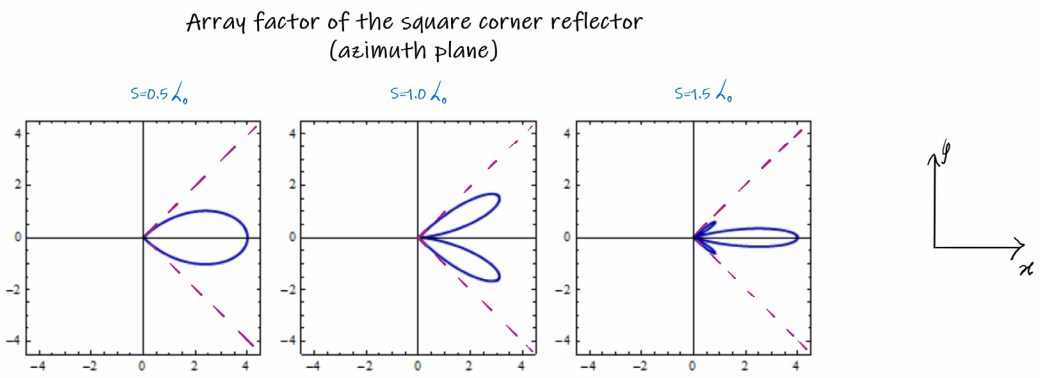
Let us now focus our attention on the *square corner reflector* $\beta = 90^\circ$ (N=2). For $\beta = 90^\circ$ the array spatial factor reduces to

$$F\Big|_{\substack{\text{square corner}\\ \text{reflector}}} = \underline{I}_0 \left(e^{+jk_0 S\cos\varphi\sin\theta} - e^{+jk_0 S\cos\left(\varphi-\frac{\pi}{2}\right)\sin\theta} + e^{+jk_0 S\cos(\varphi-\pi)\sin\theta} - e^{+jk_0 S\cos\left(\varphi-\frac{3\pi}{2}\right)\sin\theta} \right),$$
$$= \underline{I}_0 \left(e^{+jk_0 S\cos\varphi\sin\theta} - e^{+jk_0 S\sin\varphi\sin\theta} + e^{-jk_0 S\cos\varphi\sin\theta} - e^{-jk_0 S\sin\varphi\sin\theta} \right)$$

which is the same as

$$\boxed{\left.\frac{F}{\underline{I}_0}\right|_{\substack{\text{square corner}\\ \text{reflector}}} = 2\left(\cos\left(k_0 S\cos\varphi\sin\theta\right) - \cos\left(k_0 S\sin\varphi\sin\theta\right)\right).} \tag{8.5}$$

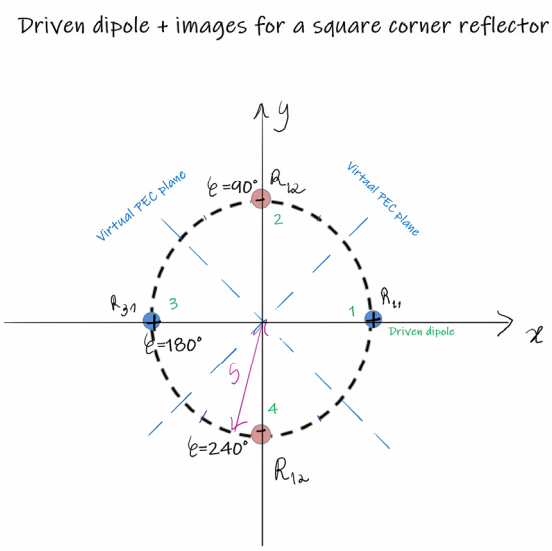
The figure below depicts the radiation pattern of the array spatial factor of the corner reflector for different values of the distance *S*. For $S = \lambda_0 / 2$, the maximum is $\left|F / \underline{I}_0\right|_{\max} = 4$, because the driven dipole radiation is redirected by the reflecting walls to a single quadrant of space.



The power gain of the corner reflector can be found from the gain in field intensity (see Sect. 7.10). For the square corner-reflector, the "gain in field intensity" is given by:

$$\mathcal{G}_f = \frac{\left|\underline{\mathbf{E}}\right|_{\substack{\text{corner}\\ \text{reflector}}}}{\left|\underline{\mathbf{E}}\right|_{\substack{\text{single dipole}\\ \text{fed with the same power}\\ \text{as the corner reflector}}}} = \sqrt{\frac{R_a}{2}\frac{\left|\underline{I}_0\right|^2}{P_{\text{in,corner-reflector}}}}\frac{\left|F_{\substack{\text{corner}\\ \text{reflector}}}\right|}{\left|\underline{I}_0\right|} \tag{8.6}$$
$$= 2\left|\left(\cos\left(k_0 S\cos\varphi\sin\theta\right) - \cos\left(k_0 S\sin\varphi\sin\theta\right)\right)\right|\sqrt{\frac{R_a}{R_{\text{in,corner}}}}$$

Here, $R_a = R_{\lambda/2} \approx 73.1\Omega$ is the input resistance of the dipole antenna alone in free-space, and $P_{\text{in,corner-reflector}} = \frac{1}{2} R_{\text{in,corner}} \left|\underline{I}_0\right|^2$ is the input power of the driven element of the corner-reflector. The input resistance $R_{\text{in,corner}}$ is not identical to $R_{\lambda/2}$ because of the near-field coupling with the metal walls, or equivalently because of the coupling with the image antennas.



In fact, the voltage at the input of the driven element can be written in terms of the self-impedance and mutual impedances as (for notational convenience here the antennas are labelled as in the figure above, so that the driven element is associated with the element *m*=1):

$$\underline{V}_1 = Z_{11}\underline{I}_1 + Z_{12}\underline{I}_2 + Z_{13}\underline{I}_3 + Z_{14}\underline{I}_4 . \tag{8.7}$$

Taking into account that $\underline{I}_1 = -\underline{I}_2 = \underline{I}_3 = -\underline{I}_4$ and that by symmetry $Z_{12} = Z_{14}$ (see the figure), one finds that:

$$\underline{V}_1 = \underbrace{\left(Z_{11} - 2Z_{12} + Z_{13}\right)}_{Z_{\text{in,corner}}} \underline{I}_1 . \tag{8.8}$$

Thus, the input impedance of the driven element ($Z_{\text{in,corner}}$) is determined by the impedance matrix of the dipole and associated images. In particular, the input resistance is $R_{\text{in,corner}} = R_{11} - 2R_{12} + R_{13} \approx 73.1 - 2R_{12} + R_{13}$, where the self-impedance was approximated by the

impedance of the antenna alone in free-space. This shows that the gain in field intensity is given by:

$$\mathcal{G}_f = 2\left|\left(\cos\left(k_0 S\cos\varphi\sin\theta\right)-\cos\left(k_0 S\sin\varphi\sin\theta\right)\right)\right|\sqrt{\frac{73.1}{73.1-2R_{12}+R_{13}}}\,. \tag{8.9}$$

For $S=\lambda_0/2$, the maximum of the gain in field intensity is ($\varphi=0^o,\theta=90^o$):

$$\mathcal{G}_{f,\max}=4\times\sqrt{\frac{73.1}{73.1-2R_{12}+R_{13}}}\,. \tag{8.10}$$

The values of the mutual resistances can be found with the perturbation method discussed in Chapter 4. The value of $\mathcal{G}_{f,\max}$ is on the order of 10dB for $S=\lambda_0/2$. Thus, the power gain of the corner reflector $G_{\text{corner-reflector}}$ is roughly 10dB larger than the power gain of a single dipole antenna $G_{\lambda/2}\approx 1.64$ [1].

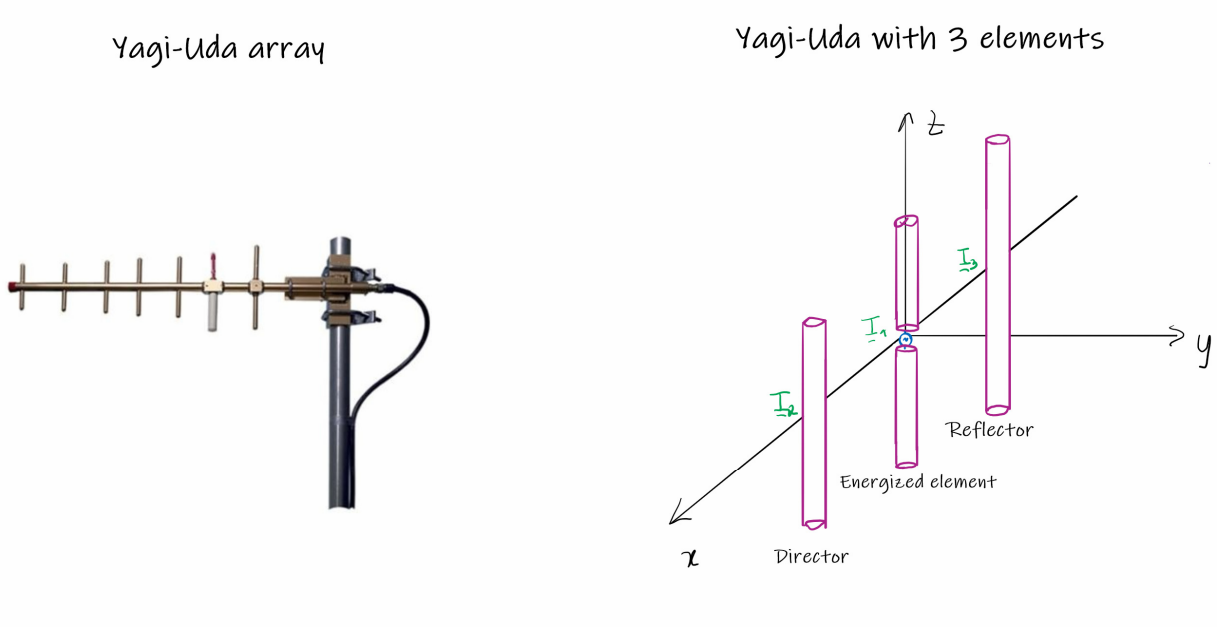
8.3 Yagi-Uda

The Yagi-Uda array is a classical antenna that is widely used as a home TV receiver and FM radio in the HF (3-30MHz), VHF (30-300MHz) and UHF (300MHz-3GHz) ranges. The antenna was originally invented by Shintaro Uda in 1926, and was further developed in a series of papers by his colleague Hidetsugu Yagi. Yagi's work popularized the antenna outside Japan.

The Yagi-Uda consists of a driven half-wavelength dipole coupled to an array of parasitic elements, which are not energized. The Yagi array is designed to operate as an endfire array. This is achieved by ensuring that the parasitic elements in the rear act as *reflectors*, whereas the parasitic elements in the front act as *directors*. Yagi referred to the row of directors as the "wave canal". The typical number of elements of the Yagi array is in the range of 3 to 7. The height of the array elements is on the order of $0.5\lambda_0$. To achieve the endfire operation, the directors must

[1] Recall that the power gain is related to the gain in field intensity as $G_{\text{array}}=G_{\text{single-element}}\mathcal{G}_f^2$. In the present problem, this formula implies that $G_{\text{corner-reflector}}=G_{\lambda/2}\times\mathcal{G}_f^2$.

be slightly shorter than the energized element, and the reflectors slightly longer. The typical spacing between the directors is on the order of $0.15\lambda_0 - 0.25\lambda_0$.



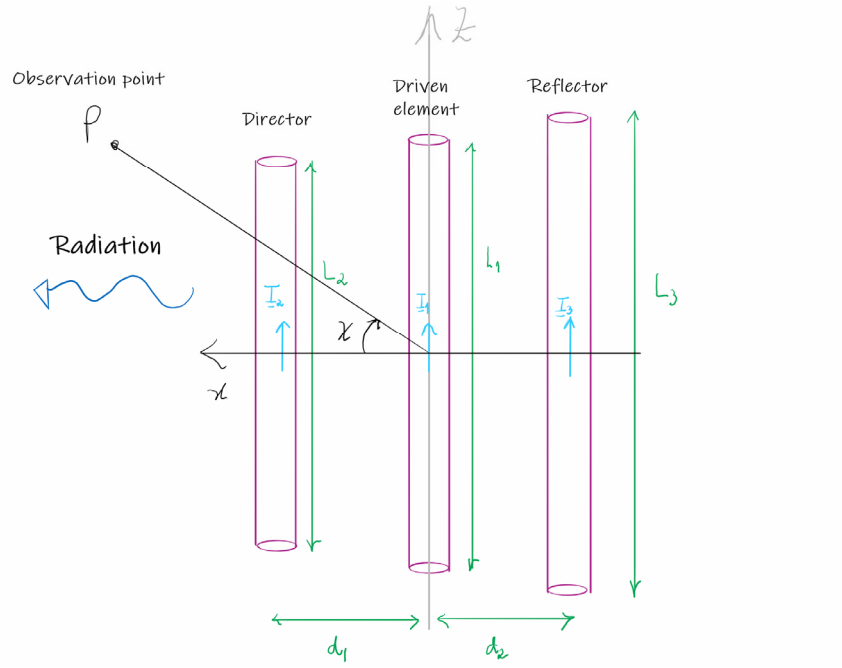
For simplicity, we will restrict our analysis to a Yagi antenna formed by a single director and single reflector, so that the number of array elements is 3. Since the elements of the array have roughly (but not exactly) the same size ($L_1 \approx L_2 \approx L_3$), one can use array theory to write:

$$\left|\underline{\mathbf{E}}\right|_{\text{Yagi}} \approx \left|F_{\text{Yagi}}\right| \times \frac{\left|\underline{\mathbf{E}}\right|_{\text{dipole}}}{\left|\underline{I}(0)\right|} \approx \left|F_{\text{Yagi}}\right| \frac{60}{r} \left| \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \right|. \tag{8.11}$$

In the second identity, we used the sinusoidal current approximation for a $0.5\lambda_0$ dipole [Eq. (8.4)]. The array spatial factor for the 3-element Yagi is given by:

$$F_{\text{Yagi}} = \underline{I}_1 + \underline{I}_2 e^{jk_0 d_1 \cos\chi} + \underline{I}_3 e^{-jk_0 d_2 \cos\chi}. \tag{8.12}$$

Here, d_1, d_2 are the distances between the driven element (centered at the origin, fed with the current $\underline{I}_1$), and the director and reflector, respectively. The currents at the center of the director ($\underline{I}_2$) and reflector ($\underline{I}_3$) are nontrivial due to the mutual coupling between the array elements. The angle χ is determined, as usual, by the array axis (x-axis) and the direction of observation.



The currents induced on the reflector and director can be found using the impedance matrix of the system, noting that the director and reflector are short circuited ($\underline{V}_2 = \underline{V}_3 = 0$):

$$\begin{aligned} \underline{V}_1 &= Z_{11}\underline{I}_1 + Z_{12}\underline{I}_2 + Z_{13}\underline{I}_3 \\ 0 &= Z_{12}\underline{I}_1 + Z_{22}\underline{I}_2 + Z_{23}\underline{I}_3 \\ 0 &= Z_{13}\underline{I}_1 + Z_{23}\underline{I}_2 + Z_{33}\underline{I}_3 \end{aligned} . \tag{8.13}$$

Using the last two equations, one can find the currents in the director and reflector as a function of the current in the driven element:

$$\underline{I}_2 = \underline{I}_1 \frac{Z_{13}Z_{23} - Z_{12}Z_{33}}{Z_{22}Z_{33} - Z_{23}^2}, \qquad \underline{I}_3 = \underline{I}_1 \frac{Z_{12}Z_{23} - Z_{13}Z_{22}}{Z_{22}Z_{33} - Z_{23}^2}. \tag{8.14}$$

To a first approximation, one can neglect the coupling between the elements of the array that are not neighbours (that are more distant), i.e. neglect the coupling between the director and the reflector: $Z_{23} \approx 0$. This approximation leads to:

$$\underline{I}_2 \approx \frac{-Z_{12}}{Z_{22}}\underline{I}_1, \qquad \underline{I}_3 \approx \frac{-Z_{13}}{Z_{33}}\underline{I}_1. \tag{8.15}$$

Thus, the array spatial factor of the Yagi antenna can be written as:

$$F_{\text{Yagi}} \approx \underline{I}_1 \left(1 - \frac{Z_{12}}{Z_{22}} e^{jk_0 d_1 \cos\chi} - \frac{Z_{13}}{Z_{33}} e^{-jk_0 d_2 \cos\chi} \right). \tag{8.16}$$

Furthermore, substituting Eq. (8.15) into Eq. (8.13), one sees that the input impedance of the Yagi antenna is:

$$Z_{\text{in,Yagi}} = Z_{11} - \frac{Z_{12}^2}{Z_{22}} - \frac{Z_{13}^2}{Z_{33}}. \tag{8.17}$$

A simple design criterion is to enforce that (*i*) the contributions the director and reflector to the direction $\chi = 0^{\text{o}}$ interfere constructively with that of the driven element. This ensures that the maximum of $\left|F_{\text{Yagi}}\right|$ occurs for $\chi = 0^{\text{o}}$. *(i)* the contributions the director and reflector to the direction $\chi = 180^{\text{o}}$ interfere destructively with that of the driven element. This ensures that the minimum of $\left|F_{\text{Yagi}}\right|$ occurs for $\chi = 180^{\text{o}}$.

The condition (i) requires that $\arg\left(\frac{Z_{12}}{Z_{22}} e^{jk_0 d_1 \cos 0^{\text{o}}} \right) = \pi$ and $\arg\left(\frac{Z_{13}}{Z_{33}} e^{-jk_0 d_2 \cos 0^{\text{o}}} \right) = \pi$, or equivalently:

$$\begin{aligned} \arg(Z_{12}) - \arg(Z_{22}) &= -k_0 d_1 + \pi \\ \arg(Z_{13}) - \arg(Z_{33}) &= k_0 d_2 - \pi \end{aligned} \quad \text{(contructive interference } \chi = 0^{\text{o}}\text{)}. \tag{8.18}$$

The condition (ii) requires that $\arg\left(\frac{Z_{12}}{Z_{22}} e^{jk_0 d_1 \cos 180^{\text{o}}} \right) = 0$ and $\arg\left(\frac{Z_{13}}{Z_{33}} e^{-jk_0 d_2 \cos 180^{\text{o}}} \right) = 0$, or equivalently:

$$\begin{aligned} \arg(Z_{12}) - \arg(Z_{22}) &= k_0 d_1 \\ \arg(Z_{13}) - \arg(Z_{33}) &= -k_0 d_2 \end{aligned} \quad \text{(destructive interference } \chi = 180^{\text{o}}\text{)}. \tag{8.19}$$

In order that the two conditions are compatible it is necessary that $k_0 d_i = -k_0 d_i + \pi$, i.e., the distances between elements should be on the order of $\lambda_0 / 4 = 0.25\lambda_0$ (in optimized designs they are in the range of $0.15\lambda_0 - 0.25\lambda_0$, as previously mentioned). Note that when $d_1 = d_2$ the Yagi array is formed by equidistant elements. A progressive phaseshift endfire array (with radiation

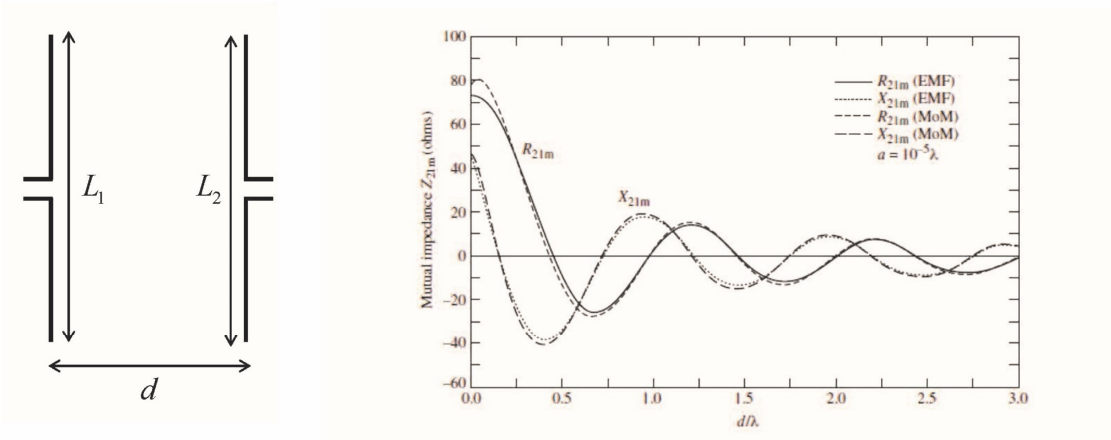
maximum along $\chi = 0^{\circ}$) requires that the phases of the currents vary in steps of $-k_0 d$. Such a condition is equivalent to Eq. (8.18).

It can be verified that the phase of the mutual impedance (e.g., Z_{12}) is little sensitive to variations in the dipole height, whereas the self-impedance (e.g., Z_{22}) varies appreciably with the dipole height. This means that:

$$\begin{aligned} &Z_{ii} \approx Z_{a,\text{dip}}(L_i), \qquad i = 1,2,3 \\ &Z_{12} \approx Z_m(d_1), \qquad Z_{13} \approx Z_m(d_2). \end{aligned} \tag{8.20}$$

In the above, $Z_{a,\text{dip}}$ is the impedance of a dipole antenna alone in free-space (which depends only on the antenna height) and $Z_m = R_m + jX_m$ is the mutual impedance of two side-by-side *half-wavelength* dipoles spaced by some distance *d*. The mutual impedance can be graphically obtained (see the figure below), or, alternatively, numerically calculated.

Mutual impedance of two side-by-side half-wavelength dipoles



With these approximations, Eq. (8.19) can be replaced by:

$$\boxed{\begin{aligned} &\arg\left(Z_m(d_1)\right) - \arg\left(Z_{a,\text{dip}}(L_2)\right) = k_0 d_1 \\ &\arg\left(Z_m(d_2)\right) - \arg\left(Z_{a,\text{dip}}(L_3)\right) = -k_0 d_2 \end{aligned}} \tag{8.21}$$

These equations must be solved with respect to to the heights of the director and reflector (L_2, L_3), which typically are few percent shorter and longer than the driver, respectively.

The gain in field intensity for the Yagi array is computed using the same procedure as for the corner reflector (see Eq. (8.6)):

$$\mathcal{G}_f = \frac{\left|\underline{\mathbf{E}}\right|_{\text{Yagi}}}{\left|\underline{\mathbf{E}}\right|_{\substack{\text{driver alone in free-space} \\ \text{fed with the same power} \\ \text{as the Yagi}}}} = \sqrt{\frac{R_a}{R_{\text{Yagi}}}} \frac{\left|F_{\text{Yagi}}\right|}{\left|\underline{I}_1\right|}. \tag{8.22}$$

It can be written explicitly as:

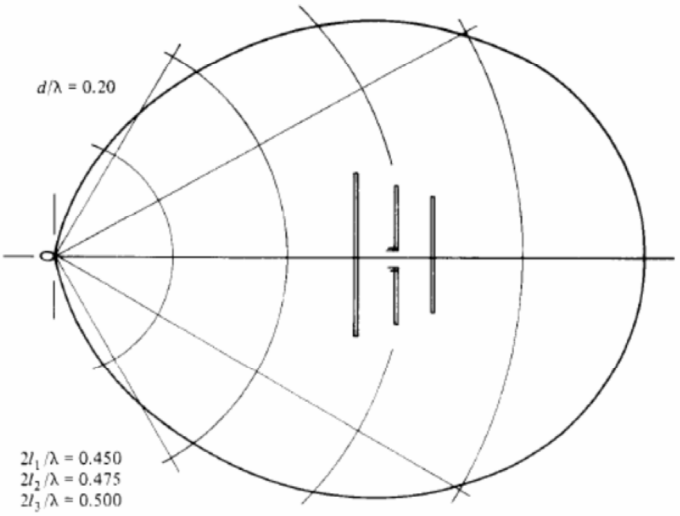
$$\mathcal{G}_f = \sqrt{\frac{73.1}{R_{\text{Yagi}}}} \left| 1 - \frac{Z_m(d_1)}{Z_{a,\text{dip}}(L_2)} e^{jk_0 d_1 \cos\chi} - \frac{Z_m(d_2)}{Z_{a,\text{dip}}(L_3)} e^{-jk_0 d_2 \cos\chi} \right|. \tag{8.23}$$

The input resistance of the Yagi antenna is

$$R_{\text{Yagi}} = \text{Re}\left\{ Z_{11} - \frac{Z_{12}^2}{Z_{22}} - \frac{Z_{13}^2}{Z_{33}} \right\} \approx \text{Re}\left\{ Z_{a,\text{dip}}(L_1) - \frac{Z_m(d_1)}{Z_{a,\text{dip}}(L_2)} - \frac{Z_m(d_2)}{Z_{a,\text{dip}}(L_3)} \right\}.$$

The typical value of $\mathcal{G}_f$ for a Yagi array with 3 elements is on the order of 7.5 dB, which corresponds to a power gain of 9.65dB. Larger gains can be obtained with arrays with a larger number of directors. The Yagi antenna is relatively narrowband and bandwidths on the order of 2% are common.

Radiation pattern of a N=3 Yagi array

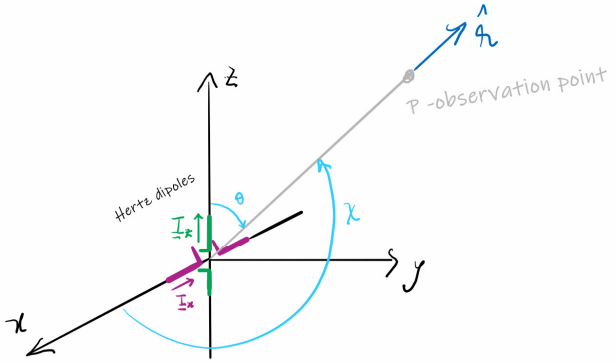


8.4 The turnstile antenna

The turnstile antenna or "cross-dipole" consists of two perpendicular oriented dipole antennas fed with currents in quadrature. In general, the polarization radiated by the cross-dipole is elliptical. For the direction perpendicular to the dipoles the polarization is circular which is useful, for example, for satellite communication through the ionosphere. Furthermore, when the turnstile antenna is oriented parallel to the ground its radiation pattern is omnidirectional with a horizontal polarization. The noise near the surface of the Earth is mainly vertically polarized, and thus horizontal polarizations are less sensitive to noise effects. Due to this reason the turnstile antenna is commonly used for TV broadcasting.

Turnstile (crossed-dipole) antenna

Turnstile formed by two Hertz dipoles



Turnstile formed by Hertz dipoles

To begin with, it is supposed that the two dipoles have infinitesimal length (Hertz dipoles). The dipoles are oriented along the x and z axis. The currents that feed the dipoles are in quadrature:

$$\underline{I}_z \equiv \underline{I}_0\,, \qquad \underline{I}_x = j\underline{I}_0\,. \tag{8.24}$$

The fields radiated by a cross-dipole formed by Hertzian dipoles can be found analytically using the superposition principle. We are mostly interested in the far-field region. From the general theory of Chapter 1, the far-field on any antenna system is determined by the (vectorial) effective length:

$$\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}} = \eta_0 jk_0 \underline{I}(0)\mathbf{h}_e(\theta,\varphi)\frac{e^{-jk_0 r}}{4\pi r}. \tag{8.25}$$

The effective length is such that

$$\underline{I}(0)\mathbf{h}_e(\hat{\mathbf{r}}) = \left[\left(\iiint dV' \underline{\mathbf{j}}(\mathbf{r}')e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'}\right)\times\hat{\mathbf{r}}\right]\times\hat{\mathbf{r}}. \tag{8.26}$$

For an antenna formed by Hertz dipoles the current distribution is given by:

$$\underline{\mathbf{j}}(\mathbf{r})\big|_{\text{Hertz}} = \underbrace{\underline{I}_x dl\delta(\mathbf{r})\hat{\mathbf{x}}}_{\text{contribution dipole along } x} + \underbrace{\underline{I}_z dl\delta(\mathbf{r})\hat{\mathbf{z}}}_{\text{contribution dipole along } z}. \tag{8.27}$$

For simplicity the lengths of both dipoles are assumed equal (dl). From here, it follows that the effective length of the Hertz turnstile antenna is such that:

$$\underline{I}(0)\mathbf{h}_e(\hat{\mathbf{r}}) = \underline{I}_x dl(\hat{\mathbf{x}}\times\hat{\mathbf{r}})\times\hat{\mathbf{r}} + \underline{I}_z dl(\hat{\mathbf{z}}\times\hat{\mathbf{r}})\times\hat{\mathbf{r}}. \tag{8.28}$$

Thereby, the far-field of the turnstile antenna is given by:

$$\underline{\mathbf{E}}(\mathbf{r})\big|_{\substack{\text{(Hertz)}\\ \text{turnstile}}} \approx \eta_0 jk_0\left(\underline{I}_x dl(\hat{\mathbf{x}}\times\hat{\mathbf{r}})\times\hat{\mathbf{r}} + \underline{I}_z dl(\hat{\mathbf{z}}\times\hat{\mathbf{r}})\times\hat{\mathbf{r}}\right)\frac{e^{-jk_0 r}}{4\pi r}. \tag{8.29}$$

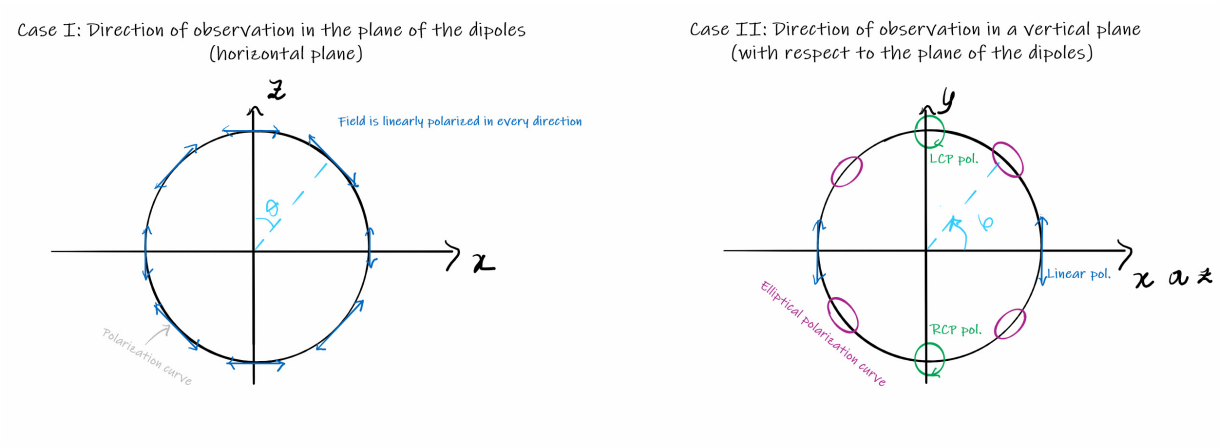
Straightforward calculations show that $(\hat{\mathbf{z}}\times\hat{\mathbf{r}})\times\hat{\mathbf{r}} = \sin\theta\hat{\boldsymbol{\theta}}$ and $(\hat{\mathbf{x}}\times\hat{\mathbf{r}})\times\hat{\mathbf{r}} = \sin\varphi\hat{\boldsymbol{\varphi}} - \cos\theta\cos\varphi\hat{\boldsymbol{\theta}}$. Using these formulas and Eq. (8.24), one can write:

$$\boxed{\underline{\mathbf{E}}(\mathbf{r})\big|_{\substack{\text{(Hertz)}\\ \text{turnstile}}} \approx \eta_0 jk_0 \underline{I}_0 dl\left[\left(\sin\theta - j\cos\theta\cos\varphi\right)\hat{\boldsymbol{\theta}} + j\sin\varphi\hat{\boldsymbol{\varphi}}\right]\frac{e^{-jk_0 r}}{4\pi r}}. \tag{8.30}$$

This shows that the field radiated by the antenna is proportional to $\underline{\mathbf{E}} \sim \left(\sin\theta - j\cos\theta\cos\varphi\right)\hat{\boldsymbol{\theta}} + j\sin\varphi\hat{\boldsymbol{\varphi}}$. Since the phase of the $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\varphi}}$ components is different, in general the antenna polarization is elliptical.

It is instructive to analyze two particular cases. In the first case, it is supposed that the direction of observation is in the plane of the dipoles (*xoz* plane; the direction of observation is horizontal with respect to the cross-dipole). This corresponds to $\varphi = 0^{\text{o}}$ or $\varphi = 180^{\text{o}}$. Clearly, in the *xoz* plane the radiated field is of the form $\underline{\mathbf{E}} \sim \left(\sin\theta \pm j\cos\theta\right)\hat{\boldsymbol{\theta}}$ which corresponds to a *linear* polarization (because $\underline{E}_\varphi = 0$).

In the second case, it is supposed that the direction of observation is in a vertical plane with respect to the plane of the dipoles, for example in the *xoy* plane. In the *xoy* plane $\theta = 90^{\circ}$ and the electric field is proportional to $\underline{\mathbf{E}} \sim \hat{\boldsymbol{\theta}} + j \sin\varphi \hat{\boldsymbol{\varphi}}$. Since the $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\varphi}}$ components of the field are in quadrature, the polarization is elliptical. There are two exceptions (i) $\varphi = 0^{\circ}, 180^{\circ}$, i.e., the *x*-axis, which as already seen in the first case corresponds to a linear polarization, and (ii) $\varphi = 90^{\circ}, 270^{\circ}$, i.e., the *y*-axis, for which the amplitude of the $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\varphi}}$ components of the field is identical. This means that the antenna polarization along the *y* axis (perpendicular to the dipoles) is circular. It is simple to check that the radiated wave is polarized to the left (right) along the positive (negative) *y*-axis.



In summary, the turnstile antenna emits an elliptically polarized wave in all directions of space, with the exception of the *xoy* plane wherein the polarization is linear, and the *y*-axis for which the polarization is circular.

The radiation intensity of the the antenna can be found in the usual way through (see also Eq. 1.24 in Chapter 1):

$$U = S \times r^2 \Big|_{\text{far-field}} = \frac{|\underline{\mathbf{E}}|^2}{2\eta_0} \times r^2 \Bigg|_{\text{far-field}} = \frac{\eta_0}{32\pi^2} k_0^2 \left|\underline{I}(0)\right|^2 |\mathbf{h}_e|^2 . \tag{8.31}$$

There is now an interesting point. From Eq. (8.28), the real-part of the effective length of the antenna is determined by the dipole oriented along *z*, whereas the imaginary part of the effective length is determined by the dipole oriented along *x*. Specifically, one can write:

$$\frac{1}{\underline{I}_0}\underline{I}(0)\mathbf{h}_e(\hat{\mathbf{r}}) = \underbrace{dl(\hat{\mathbf{z}}\times\hat{\mathbf{r}})\times\hat{\mathbf{r}}}_{\text{real-vector}=\mathbf{h}_{e,\text{dip z}}} + j\underbrace{dl(\hat{\mathbf{x}}\times\hat{\mathbf{r}})\times\hat{\mathbf{r}}}_{\text{real-vector}=\mathbf{h}_{e,\text{dip x}}}. \tag{8.32}$$

In the above, $\mathbf{h}_{e,\text{dip z}}, \mathbf{h}_{e,\text{dip x}}$ are the standard effective lengths of Hertz dipoles oriented along *z* and *x*, respectively. Taking into account that for any complex vector $\mathbf{A}$ one has $|\mathbf{A}|^2 = |\mathbf{A}_R|^2 + |\mathbf{A}_I|^2$ with $\mathbf{A}_R = \text{Re}\{\mathbf{A}\}$ and $\mathbf{A}_I = \text{Im}\{\mathbf{A}\}$, it follows that:

$$|\underline{I}(0)\mathbf{h}_e(\hat{\mathbf{r}})|^2 = |\underline{I}_0|^2\left(|\mathbf{h}_{e,\text{dip z}}|^2 + |\mathbf{h}_{e,\text{dip x}}|^2\right). \tag{8.33}$$

Substituting this result into Eq. (8.34), it is found that:

$$U = \frac{\eta_0}{32\pi^2}k_0^2|\underline{I}_0|^2\left(|\mathbf{h}_{e,\text{dip z}}|^2 + |\mathbf{h}_{e,\text{dip x}}|^2\right) = U_{\text{dip,z}} + U_{\text{dip,x}}, \tag{8.34}$$

where $U_{\text{dip,z}}, U_{\text{dip,x}}$ are the radiation intensities for the individual dipoles alone in the free-space. In particular, integrating over all solid angles ($P_{\text{rad}} = \iint U d\Omega$) one finds that the radiated power is:

$$P_{\text{rad}}\big|_{\text{crossed-dipole}} = P_{\text{rad}}\big|_{\text{dip-z}} + P_{\text{rad}}\big|_{\text{dip-x}}, \tag{8.35}$$

This means that the energies radiated by two dipoles fed by currents in *quadrature* combine additively. This result is rather general and can be extended to any two antennas fed by currents in quadrature, provided the effective length of the antennas is a *real-valued function*. In such conditions, the fields radiated by the antennas *do not interfere* (there are no minima or maxima due to destructive or constructive interference!).

The radiation intensity for a Hertz dipole oriented along *z* and fed by a current $\underline{I}_0$ was studied in Chapter 1. It is given by

$$U_{\text{dip-z}} = \frac{\eta_0}{32\pi^2}\left(k_0 dl\right)^2 \left|\underline{I}_0\right|^2 \sin^2\theta = \frac{\eta_0}{32\pi^2}\left(k_0 dl\right)^2 \left|\underline{I}_0\right|^2 \left(1-\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{z}}\right)^2\right).$$

The rightmost identity can be immediately generalized to a dipole with an arbitrary orientation. For example, if the dipole is oriented along x the radiation intensity becomes:

$$U_{\text{dip-x}} = \frac{\eta_0}{32\pi^2}\left(k_0 dl\right)^2 \left|\underline{I}_0\right|^2 \left(1-\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}\right)^2\right).$$

The previous formulas show that the radiation intensity of the turnstile antenna can be written as:

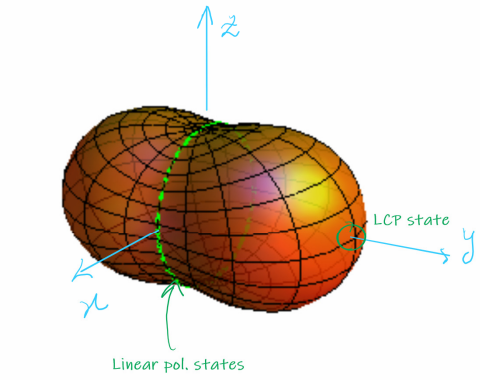
$$\begin{aligned} U &= \frac{\eta_0}{32\pi^2}\left(k_0 dl\right)^2 \left|\underline{I}_0\right|^2 \left[2-\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}\right)^2-\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{z}}\right)^2\right] \\ &= \frac{\eta_0}{32\pi^2}\left(k_0 dl\right)^2 \left|\underline{I}_0\right|^2 \left(\sin^2\chi+\sin^2\theta\right) \end{aligned}. \qquad (8.36)$$

As usual, the angle χ is the angle between the x-axis and the observation point ($\cos\chi = \hat{\mathbf{r}}\cdot\hat{\mathbf{x}}$). Alternatively, from $\hat{\mathbf{r}}\cdot\hat{\mathbf{x}} = \sin\theta\cos\varphi$ one can also write:

$$\boxed{U = \frac{\eta_0}{32\pi^2}\left(k_0 dl\right)^2 \left|\underline{I}_0\right|^2 \left(1+\sin^2\theta\sin^2\varphi\right)}. \qquad (8.37)$$

Clearly the direction of maximum radiation is for $\varphi = 0, \theta = 0, 180^{\circ}$, which corresponds to the y axis. As could be expected, the radiation pattern has rotation symmetry around this axis and is shaped like a "potato". It has no nulls.

Radiation pattern of the turnstile antenna formed by Hertz dipoles



From the definition, the directive gain of the antenna is $g = 4\pi U / P_{\text{rad}}$. The directivity is the maximum of the directive gain:

$$D = \max \frac{4\pi U}{P_{\text{rad}}} = \max \frac{4\pi\left(U_{\text{dip-z}} + U_{\text{dip-x}}\right)}{P_{\text{rad}}\big|_{\text{dip-z}} + P_{\text{rad}}\big|_{\text{dip-x}}} = \frac{4\pi 2 \times \max U_{\text{dip-z}}}{2 \times P_{\text{rad}}\big|_{\text{dip-z}}} = D_{\text{dip-z}}. \tag{8.38}$$

In the third identity, we used the fact that $U_{\text{dip-z}}, U_{\text{dip-x}}$ are identical along the *y*-axis (direction of maximum gain) and that $P_{\text{rad}}\big|_{\text{dip-z}} = P_{\text{rad}}\big|_{\text{dip-x}}$. Therefore, the directivity of the cross-dipole is exactly the same as the directivity of a simple Hertz dipole:

$$D\big|_{\substack{\text{(Hertz)} \\ \text{turnstile}}} = 1.5. \tag{8.39}$$

From the expression of $P_{\text{rad}}\big|_{\text{dip-z}}$ for a simple Hertz dipole ($P_{\text{rad}}\big|_{\text{dip-z}} = \frac{\eta_0}{12\pi}\left(k_0 dl\right)^2 \left|\underline{I}_0\right|^2$), the directive gain of the turnstile antenna can be explicitly evaluated:

$$\boxed{g\big|_{\substack{\text{(Hertz)} \\ \text{turnstile}}} = \frac{3}{4}\left(1 + \sin^2\theta \sin^2\varphi\right)}. \tag{8.40}$$

Turnstile formed by half-wavelength dipoles

In practice, the turnstile antenna is made of two half-wavelength dipoles. As in the case of infinitesimal dipoles, the effective length of the antenna is the vector sum of the effective lengths of the two elements:

$$\underline{I}(0)\mathbf{h}_e\left(\hat{\mathbf{r}}\right) = \underline{I}_x \mathbf{h}_{\text{x-dip}} + \underline{I}_z \mathbf{h}_{\text{z-dip}}. \tag{8.41}$$

The effective length of a dipole antenna oriented along *z* can be written as (sinusoidal current approximation)

$$\begin{aligned} \mathbf{h}_{\text{z-dip}} &= h_{e,z}\left(\theta\right)\hat{\boldsymbol{\theta}} = \frac{h_{e,z}\left(\theta\right)}{\sin\theta}\left(\hat{\mathbf{z}} \times \hat{\mathbf{r}}\right) \times \hat{\mathbf{r}} \\ &= \frac{2}{k_0} \frac{\cos\left(\frac{\pi}{2}\hat{\mathbf{r}} \cdot \hat{\mathbf{z}}\right)}{1 - \left(\hat{\mathbf{r}} \cdot \hat{\mathbf{z}}\right)^2}\left(\hat{\mathbf{z}} \times \hat{\mathbf{r}}\right) \times \hat{\mathbf{r}} \end{aligned} \tag{8.42}$$

with $h_{e,z}(\theta)=\dfrac{2}{k_0}\dfrac{\cos\left(\dfrac{\pi}{2}\cos\theta\right)}{\sin\theta}$. In the second and third identities, we used $\cos\theta=\hat{\mathbf{r}}\cdot\hat{\mathbf{z}}$ and $\left(\hat{\mathbf{z}}\times\hat{\mathbf{r}}\right)\times\hat{\mathbf{r}}=\sin\theta\hat{\boldsymbol{\theta}}$. The previous formula can be readily generalized to a dipole oriented along an arbitrary direction $\hat{\mathbf{l}}$, simply by replacing $\hat{\mathbf{z}}\to\hat{\mathbf{l}}$. In particular, the effective length of a dipole oriented along *x* is:

$$\mathbf{h}_{\text{x-dip}}=\frac{2}{k_0}\frac{\cos\left(\dfrac{\pi}{2}\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}\right)}{1-\left(\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}\right)^2}\left(\hat{\mathbf{x}}\times\hat{\mathbf{r}}\right)\times\hat{\mathbf{r}}=\frac{2}{k_0}\frac{\cos\left(\dfrac{\pi}{2}\cos\chi\right)}{\sin^2\chi}\left(\hat{\mathbf{x}}\times\hat{\mathbf{r}}\right)\times\hat{\mathbf{r}},\tag{8.43}$$

where $\cos\chi=\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}=\sin\theta\cos\varphi$. Combining Eqs. (8.41)-(8.43) and substituting the result in Eq. (8.25) with $\underline{I}_z\equiv\underline{I}_0$ and $\underline{I}_x=j\underline{I}_0$ it is found that the far-field of the antenna is given by:

$$\underline{\mathbf{E}}\big|_{\text{far-field}}\approx\eta_0 j\,k_0\underline{I}_0\frac{e^{-jk_0r}}{4\pi r}\left[\frac{h_{e,z}(\theta)}{\sin\theta}\left(\hat{\mathbf{z}}\times\hat{\mathbf{r}}\right)\times\hat{\mathbf{r}}+j\frac{h_{e,x}(\chi)}{\sin\chi}\left(\hat{\mathbf{x}}\times\hat{\mathbf{r}}\right)\times\hat{\mathbf{r}}\right],\tag{8.44}$$

where $h_{e,x}(\chi)=\dfrac{2}{k_0}\dfrac{\cos\left(\dfrac{\pi}{2}\cos\chi\right)}{\sin\chi}$. The fields radiated by the two dipoles are in quadrature. Thus, similar to the case of the infinitesimal dipoles discussed previously, the radiated powers combine additively. In fact, one can check that:

$$\left|\underline{\mathbf{E}}\right|^2\approx\left(\eta_0\,k_0\left|\underline{I}_0\right|\frac{1}{4\pi r}\right)^2\left[\left|h_{e,z}(\theta)\right|^2+\left|h_{e,x}(\chi)\right|^2\right].\tag{8.45}$$

This implies that, analogous to the case of infinitesimal dipoles (see Eq. (8.34)) one can write $U=U_{\text{dip,z}}+U_{\text{dip,x}}$, with $U_{\text{dip,z}}$ and $U_{\text{dip,x}}$ the radiation intensities of the dipoles alone in free-space. Furthermore, since the powers radiated by the dipoles are identical, the directive gain of the turnstile antenna is given by:

$$g\big|_{\text{turnstile}}=\frac{1}{2}\left(g_z(\theta)+g_x(\chi)\right),\tag{8.46}$$

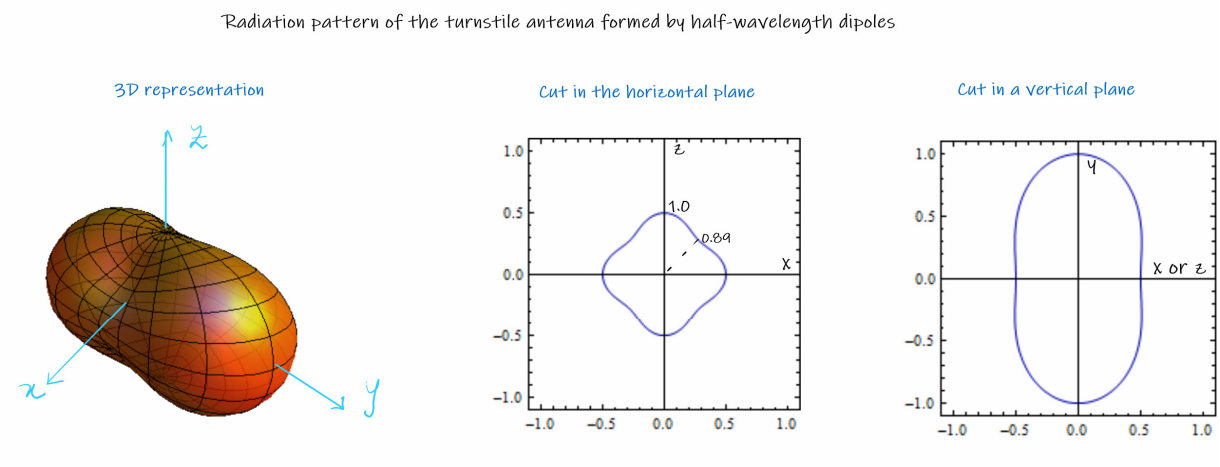
where $g_z(\theta)$ ($g_x(\chi)$) is the directive gain for a dipole oriented along *z* (*x*):

$$g_z(\theta) \approx 1.64 \left| \frac{\cos\left(\frac{\pi}{2}\cos\theta\right)}{\sin\theta} \right|^2, \qquad g_x(\chi) \approx 1.64 \left| \frac{\cos\left(\frac{\pi}{2}\cos\chi\right)}{\sin\chi} \right|^2. \tag{8.47}$$

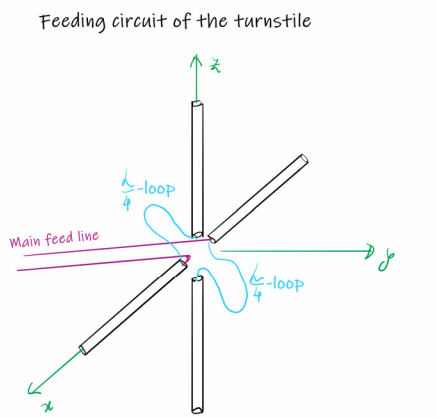
The maximum of the directive gain occurs along the y-axis ($\theta = \chi = 90^\circ$). The directivity of the turnstile is the same as the directivity of a half-wavelength dipole:

$$D\big|_{\text{turnstile}} = 1.64. \tag{8.48}$$

The radiation pattern of the turnstile is sketched in the figure below and has features analogous to the turnstile formed by infinitesimal dipoles. For half-wavelength dipoles the radiation pattern in the horizontal plane is not a circle, and so the radiation pattern is only approximately omnidirectional. The antenna polarization is also analogous to the case of infinitesimal dipoles.



In practice, the cross-dipoles are fed with a single transmission line, using the configuration shown below. The signal that feeds the dipole oriented along x travels through a $\lambda/4$ loop before feeding the dipole oriented along z. This ensures that the two currents are in quadrature.

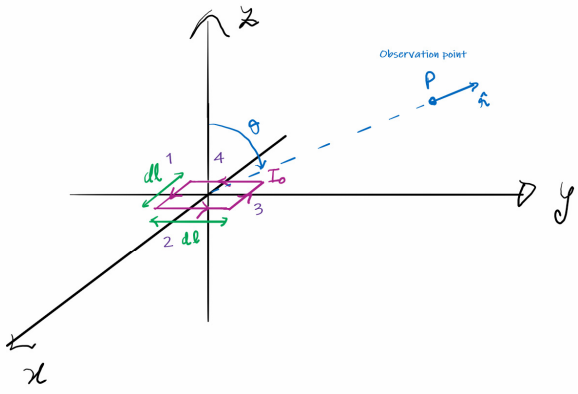


8.5 Loop antenna

Antennas with the shape of a wire loop are widely used, particularly as receivers of AM radio. The loop can take many forms. For simplicity here we consider a square loop in the *xoy* plane. Furthermore, we restrict our analysis to small loops with a diameter much less than the wavelength.

Loop antenna

Square loop



A small loop antenna can be regarded as a set of 4 segments of constant current, i.e., as a set four Hertz dipoles. Recalling that the current distribution associated with a Hertz dipole centered at the point $\mathbf{r}_0 = (x_0, y_0, z_0)$, with length dl, and oriented along the direction $\hat{\mathbf{u}}$ is of the form $\mathbf{j}_{\text{Hertz}} = \underline{I}\, dl\,\delta(x-x_0)\delta(y-y_0)\delta(z-z_0)\hat{\mathbf{u}}$, one sees that the current that describes the square loop can be taken equal to:

$$\mathbf{j}_{\text{loop}} = \underbrace{\underline{I}_0\, dl\delta(x)\delta\left(y+\frac{dl}{2}\right)\delta(z)\hat{\mathbf{x}}}_{\text{arm 1}} + \underbrace{\underline{I}_0\, dl\delta\left(x-\frac{dl}{2}\right)\delta(y)\delta(z)\hat{\mathbf{y}}}_{\text{arm 2}} + \underbrace{-\underline{I}_0\, dl\delta(x)\delta\left(y-\frac{dl}{2}\right)\delta(z)\hat{\mathbf{x}}}_{\text{arm 3}} - \underbrace{\underline{I}_0\, dl\delta\left(x+\frac{dl}{2}\right)\delta(y)\delta(z)\hat{\mathbf{y}}}_{\text{arm 4}} \tag{8.49}$$

The different arms of the square loop are numbered as indicated in the figure. Since the length of the arms of the dipoles is small, we can expand the delta functions in a Taylor series. For example, one can write $\delta\left(x+\frac{dl}{2}\right) \approx \delta(x)+\frac{dl}{2}\delta'(x)$. This procedure yields:

$$\mathbf{j}_{\text{loop}} = \underbrace{\underline{I}_0 (dl)^2 \delta(x)\delta'(y)\delta(z)\hat{\mathbf{x}}}_{\text{arms 1+3}} - \underbrace{\underline{I}_0 (dl)^2 \delta'(x)\delta(y)\delta(z)\hat{\mathbf{y}}}_{\text{arms 2+4}}. \tag{8.50}$$

The previous formula can be written in a compact form as:

$$\begin{aligned}\mathbf{j}_{\text{loop}} &= \nabla\times\left\{m\hat{\mathbf{z}}\delta(x)\delta(y)\delta(z)\right\} \\ &= \nabla\times\left\{\mathbf{m}\delta(\mathbf{r})\right\}\end{aligned} \tag{8.51}$$

where $\mathbf{m} = m\hat{\mathbf{z}}$ is the *magnetic dipole moment* of the loop with $m = \underline{I}_0 A$ and $A = (dl)^2$ the area of the loop. In fact, it can be shown that the result $\mathbf{j}_{\text{loop}} = \nabla\times\left\{\mathbf{m}\delta(\mathbf{r})\right\}$ with $\mathbf{m} = \underline{I}_0 A\hat{\mathbf{z}}$ is rather general, and independent of the shape of the loop.

The fields radiated by the small loop ($\underline{\mathbf{E}}', \underline{\mathbf{H}}'$) can be found by solving the Maxwell's equations:

$$\begin{aligned}\nabla\times\underline{\mathbf{E}}' &= -j\omega\mu_0\underline{\mathbf{H}}' \\ \nabla\times\underline{\mathbf{H}}' &= \nabla\times\left\{m\hat{\mathbf{z}}\delta(\mathbf{r})\right\} + j\omega\varepsilon_0\underline{\mathbf{E}}'\end{aligned} \tag{8.52}$$

It is convenient to introduce $\underline{\mathbf{H}} = \underline{\mathbf{H}}' - \mathbf{m}\delta(\mathbf{r})$ and $\underline{\mathbf{E}} = \underline{\mathbf{E}}'$, which are coincident with the radiated (primed) fields everywhere in space, except at the position of the loop. The unprimed fields satisfy:

$$\begin{aligned}\nabla\times\underline{\mathbf{E}} &= -j\omega\mu_0 m\hat{\mathbf{z}}\delta(\mathbf{r}) - j\omega\mu_0\underline{\mathbf{H}} \\ \nabla\times\underline{\mathbf{H}} &= +j\omega\varepsilon_0\underline{\mathbf{E}}\end{aligned} \qquad \text{(small loop).} \tag{8.53}$$

Remarkably, the structure of the above equations is rather similar to the structure of the equations satisfied by an electric Hertz dipole:

$$\begin{aligned}\nabla\times\underline{\mathbf{H}}_{\text{dip}} &= j\omega p_e\hat{\mathbf{z}}\delta(\mathbf{r})+j\omega\varepsilon_0\underline{\mathbf{E}}_{\text{dip}}\\ \nabla\times\underline{\mathbf{E}}_{\text{dip}} &= -j\omega\mu_0\underline{\mathbf{H}}_{\text{dip}}\end{aligned}\qquad\text{(electric Hertz dipole).}\tag{8.54}$$

In the above, p_e is the electric dipole moment of the Hertz dipole ($j\omega p_e=\underline{I}(0)dl$). The interesting thing is that by interchanging the roles of the electric and magnetic fields in equations (8.53) and (8.54), they transform into one another. Specifically, one can identify the following *duality* mapping of the fields, sources, and physical constants:

$$\begin{array}{ccc}\underline{\mathbf{E}} & \leftrightarrow & -\underline{\mathbf{H}}_{\text{dip}}\\ \underline{\mathbf{H}} & \leftrightarrow & \underline{\mathbf{E}}_{\text{dip}}\\ \mu_0 m & \leftrightarrow & p_e\\ \varepsilon_0 & \leftrightarrow & \mu_0\\ \underbrace{\mu_0}_{\substack{\text{Solution}\\ \text{Eq. 8.53}}} & \leftrightarrow & \underbrace{\varepsilon_0}_{\substack{\text{Solution}\\ \text{Eq. 8.54}}}\end{array}\tag{8.55}$$

The duality link between the fields radiated by the electric and magnetic dipoles implies that the solution of equation (8.53) can be formally obtained from the fields of the electric Hertz dipole (which were derived in the beginning of the course) using the duality transformation.

For simplicity, we will illustrate the concept for the far-field zone. The far-field of a conventional electric Hertz dipole is:

$$\underline{\mathbf{E}}_{\text{dip}}=\underline{E}_{\theta,\text{dip}}\hat{\boldsymbol{\theta}},\qquad\qquad \underline{\mathbf{H}}_{\text{dip}}=\underline{H}_{\varphi,\text{dip}}\hat{\boldsymbol{\varphi}}\tag{8.56a}$$

$$\underline{E}_{\theta,\text{dip}}=\eta_0\underline{H}_{\varphi,\text{dip}}=\eta_0 jk_0\left(j\omega p_e\right)\sin\theta\frac{e^{-jk_0r}}{4\pi r}\tag{8.56b}$$

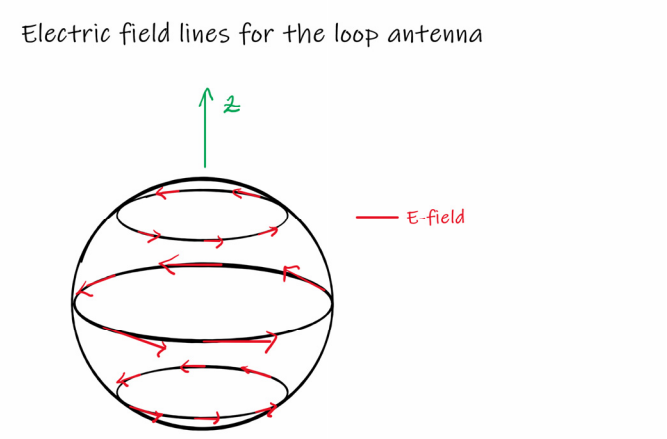
From here, using the duality transformation $\underline{\mathbf{H}}=\underline{\mathbf{E}}_{\text{dip}}\Big|_{\substack{p_e\to\mu_0 m\\ \eta_0\to 1/\eta_0}}$ and $\underline{\mathbf{E}}=-\underline{\mathbf{H}}_{\text{dip}}\Big|_{\substack{p_e\to\mu_0 m\\ \eta_0\to 1/\eta_0}}$, one finds that the far-field of the loop antenna (magnetic dipole) is given by:

$$\underline{\mathbf{E}}=\underline{E}_{\varphi}\hat{\boldsymbol{\varphi}},\qquad\qquad \underline{\mathbf{H}}=\underline{H}_{\theta}\hat{\boldsymbol{\theta}}.\tag{8.57a}$$

$$\underline{H}_\theta = \frac{-\underline{E}_\varphi}{\eta_0} = \frac{1}{\eta_0} jk_0 \left(j\omega\mu_0 m \right) \sin\theta \frac{e^{-jk_0 r}}{4\pi r}. \tag{8.57b}$$

Using $m = \underline{I}_0 A$, the same result can be written as:

$$\boxed{\underline{E}_\varphi = -\eta_0 \underline{H}_\theta = \eta_0 k_0^2 \underline{I}_0 A \sin\theta \frac{e^{-jk_0 r}}{4\pi r}} \qquad \text{(far-field loop antenna).} \tag{8.58}$$



Because of the duality link, the field lines of the loop antenna are identical to the field lines of the Hertz dipole, with the electric and magnetic fields interchanged. The shape of the radiation pattern of the two radiators is also identical. In particular, similar to the Hertz dipole, the directive gain of the loop antenna is:

$$\boxed{g\big|_{\substack{\text{loop}\\\text{antenna}}} = \frac{3}{2}\sin^2\theta} \tag{8.59}$$

The directivity of a small loop is $D = \max g = 1.5$.

Comparing the far-field formula of the small loop [Eq. (8.58)] with the generic expression

$$\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}} = \eta_0 jk_0 \underline{I}(0) \mathbf{h}_e(\theta,\varphi) \frac{e^{-jk_0 r}}{4\pi r}$$

one sees that the effective length of the loop antenna is given by:

$$\mathbf{h}_e\big|_{\text{loop}} = -j\,k_0 A \sin\theta\, \hat{\boldsymbol{\varphi}}\,. \tag{8.60}$$

The radiation resistance of the small loop can be found in a standard way and satisfies:

$$R_{\text{rad}}\big|_{\text{loop}} = \frac{\eta_0}{6\pi}\left(k_0^2 A\right)^2 . \tag{8.61}$$

So far, it was assumed that the antenna is formed by a single loop of current, which is a rather inefficient antenna with a small radiation resistance. The magnetic dipole moment of the antenna can be enhanced using additional (N) loops current. It can also be increased by winding the wire around a ferromagnetic core. In this case, the magnetic dipole moment is enhanced by a factor (μ_{ef} / μ_0) proportional to the permeability of the core μ. The permeability enhancement factor is:

$$\frac{\mu_{\text{ef}}}{\mu_0} = \frac{\mu}{\mu_0}\frac{1}{1 + D\left(\dfrac{\mu}{\mu_0} - 1\right)} \tag{8.62}$$

where D is a demagnetization factor that depends on the ratio between the length and the diameter of the ferromagnetic rod. The demagnetization factor approaches zero when the length of the ferromagnetic rod is much larger than its diameter, so that the leakage of magnetic field lines is negligible.

The effective length of a loop antenna with N loop turns wound around a ferromagnetic rod is enhanced by the same factor $\left(N\,\mu_{\text{ef}} / \mu_0\right)$ as the magnetic dipole moment:

$$\boxed{\mathbf{h}_e\big|_{\substack{\text{loop}\\ N\text{ turns}}} = -jN\frac{\mu_{\text{ef}}}{\mu_0}k_0 A \sin\theta\,\hat{\boldsymbol{\varphi}}}. \tag{8.63}$$

On the other hand, the radiation resistance is boosted by a factor of $\left(N\,\mu_{\text{ef}} / \mu_0\right)^2$:

$$\boxed{R_{\text{rad}}\big|_{\substack{\text{loop}\\ N\text{ turns}}} = \frac{\eta_0}{6\pi}\left(k_0^2 A \times N \times \frac{\mu_{\text{ef}}}{\mu_0}\right)^2}. \tag{8.64}$$

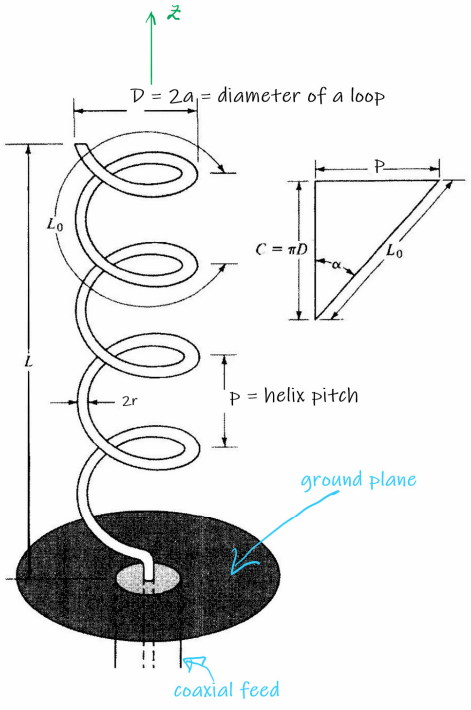
8.6 The helical antenna

The helical antenna consists of a conductor wound in the shape of a screw thread (helix) and a flat metal sheet serving as the ground plane. The antenna is typically fed using a coaxial line,

with the inner conductor connected to the helix and the outer conductor connected to the metal plate. The ground plane should be at least half-wavelength in diameter. The helical antenna supports two modes of operation: (i) the *normal mode* and (ii) the *axial mode*. In the normal mode the maximum of radiation occurs for directions perpendicular to the helix axis (*xoy* plane). The dimensions of the antenna are electrically small, which results in a low efficiency and small bandwidth. In the axial mode the maximum of radiation is along the helix axis (*z*-direction) and the antenna polarization is approximately circular. In most designs the helical antenna is used in the "axial mode", which is characterized by a high gain and large bandwidth (about 50%). Due to the circular polarization, the helix antenna is widely used for space communications applications. For example, helical antennas were placed on the Moon by the astronauts of the Apollo mission to transmit telemetry data back to Earth.

Array of 4 helical antennas

Geometry of a helical antenna



The geometry of the helical antenna is depicted in the figure above (right). The diameter of each loop is $D = 2a$ with a the loop radius. The helix pitch is the distance between successive loops

and is denoted by p (in the literature the helix pitch is often denoted by S). The helix antenna is formed by N loops (i.e., N wire turns). A generic helix with axis along z can be parameterized as:

$$\mathbf{r}(s) = a\cos\left(\frac{2\pi}{L_0}s\right)\hat{\mathbf{x}} \pm a\sin\left(\frac{2\pi}{L_0}s\right)\hat{\mathbf{y}} + p\frac{s}{L_0}\hat{\mathbf{z}}, \tag{8.65}$$

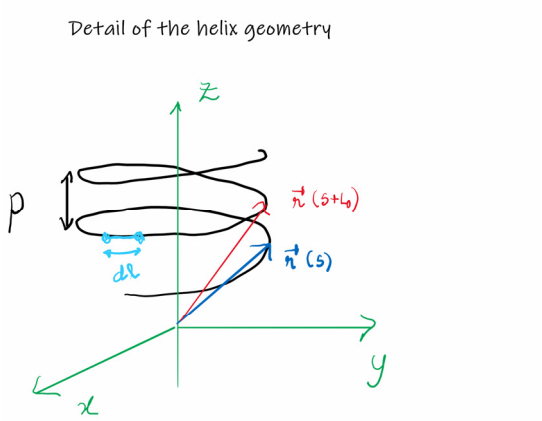
where we introduce the parameter L_0 given by:

$$L_0 = \sqrt{(2\pi a)^2 + p^2} = \sqrt{(\pi D)^2 + p^2}\,. \tag{8.66}$$

This sign $\pm$ in the Eq. (8.65) depends if the helix is right-handed (+) or left-handed (-) with respect to the +z-axis. The parameter s identifies a generic point on the helix. The distance between two nearby points $\mathbf{r}(s)$ and $\mathbf{r}(s+ds)$ is $dl = |\mathbf{r}(s+ds) - \mathbf{r}(s)| \approx |\mathbf{r}'(s)|ds$, or equivalently the infinitesimal arc length is:

$$dl = \sqrt{\left(\frac{2\pi a}{L_0}\right)^2 + \left(\frac{p}{L_0}\right)^2}\,ds = ds \tag{8.67}$$

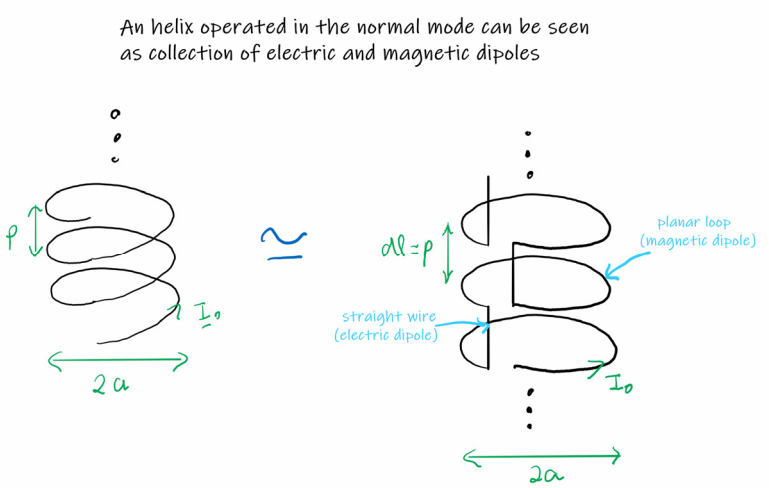
This means that the parameter s measures the arc length along the helix.



Noting that $\mathbf{r}(s+L_0) = \mathbf{r}(s) + p\hat{\mathbf{z}}$, we see that the points corresponding to s and $s+L_0$ are separated by a full loop. The length of one helix turn can thus be found by integrating the infinitesimal arc length from s to $s+L_0$: $\int_{\text{one turn}} dl = \int_{s}^{s+L_0} ds = L_0$. This means that L_0 defined as in Eq. (8.66) gives the length of one turn of the helix.

Theory of the normal mode

The helix is operated in the "normal mode" when its total length $N \times L_0$ is much smaller than the wavelength. In particular, the pitch and the helix diameter satisfy $p << \lambda_0$, $D << \lambda_0$. Due to this reason a single helix turn can be regarded as the superposition of an electric dipole and a magnetic dipole.



The equivalent electric dipole is an Hertz dipole oriented along *z*, has length $dl = p$, and is fed by the current $\underline{I}_0$. The effective length of the electric Hertz dipole is: $\mathbf{h}_e\big|_{\text{dip}} = dl \sin\theta \hat{\boldsymbol{\theta}} = p \sin\theta \hat{\boldsymbol{\theta}}$ The equivalent magnetic dipole is also oriented along *z* because the equivalent current loop is parallel to the *xoy* plane. The corresponding magnetic dipole moment is $m = \pm \underline{I}_0 A = \pm \underline{I}_0 \pi a^2$. Note that the leading sign of the magnetic dipole moment depends on the orientation of the helix. The effective length of the loop is [see Eq. (8.60)] $\mathbf{h}_e\big|_{\text{loop}} = -j k_0 \dfrac{m}{\underline{I}_0} \sin\theta \hat{\boldsymbol{\varphi}} = -(\pm) j k_0 \pi a^2 \sin\theta \hat{\boldsymbol{\varphi}}$.

The effective length of a single turn of the helix is the vector sum of the effective lengths of the equivalent electric and magnetic dipoles (in this analysis we neglect the effect of the ground plane, as its diameter is typically small as compared to the wavelength):

$$\mathbf{h}_e\big|_{\substack{\text{single}\\\text{turn helix}}} = \sin\theta\left(p\,\hat{\boldsymbol{\theta}} - (\pm)\,j\,k_0\pi a^2\,\hat{\boldsymbol{\varphi}}\right). \tag{8.68}$$

The effective length of the entire helix is obtained from $\mathbf{h}_e\big|_{\substack{\text{single}\\\text{turn helix}}}$ by multiplication of the number of turns (this possible because total length of the helix is much smaller than the wavelength, and hence, effectively, all the equivalent dipoles are centered on the same point):

$$\boxed{\mathbf{h}_e\big|_{\text{helix}} = N\sin\theta\left(\,p\,\hat{\boldsymbol{\theta}} - (\pm)\,j\,k_0\pi a^2\,\hat{\boldsymbol{\varphi}}\right).} \tag{8.69}$$

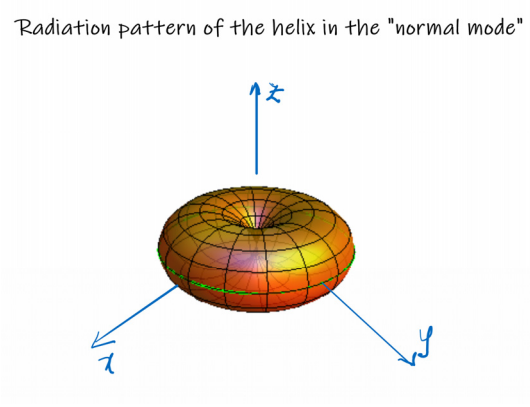
The electric far-field of the helix is given by (see Eq. (8.25)):

$$\boxed{\underline{\mathbf{E}}\big|_{\text{helix}} = \eta_0 j k_0 \underline{I}_0 N\sin\theta\left(p\,\hat{\boldsymbol{\theta}} - j(\pm)k_0\pi a^2\,\hat{\boldsymbol{\varphi}}\right)\frac{e^{-jk_0r}}{4\pi r}.} \tag{8.70}$$

It is useful to note that the electric and magnetic dipoles radiate fields in quadrature and perpendicular polarizations. Due to this reason, the powers radiated by the electric and magnetic dipoles combine additively, analogous to the turnstile antenna. In particular, it is evident the radiation pattern of the helix antenna operated in normal mode is the same as for an electric Hertz dipole with the directive gain given by:

$$\boxed{g\big|_{\substack{\text{helix}\\\text{normal mode}}} = \frac{3}{2}\sin^2\theta.} \tag{8.71}$$

Radiation pattern of the helix in the "normal mode"



The power radiated by the helix can be found by integrating the radiation intensity $U = \dfrac{\eta_0}{8}\left|\underline{I}_0\right|^2\dfrac{\left|\mathbf{h}_e\right|^2}{\lambda_0^2}$ over all solid angles:

$$\begin{aligned} P_{\text{rad}} &= \iint d\Omega \ \frac{\eta_0}{8}\left|\underline{I}_0\right|^2 \frac{1}{\lambda_0^2} N^2 \left(p^2 + \left(k_0 \pi a^2\right)^2 \right) \sin^2\theta \\ &= \frac{\eta_0}{8}\left|\underline{I}_0\right|^2 \frac{1}{\lambda_0^2} N^2 \left(p^2 + \left(k_0 \pi a^2\right)^2 \right) \times 2\pi \times \frac{4}{3} \end{aligned} \tag{8.72}$$

From this result, one sees that the radiation resistance of the helix is:

$$R_{\text{rad}}\big|_{\text{helix}} = \frac{2\pi\eta_0}{3} N^2 \frac{1}{\lambda_0^2} \left(p^2 + \left(k_0 \pi a^2\right)^2 \right). \tag{8.73}$$

From Eq. (8.70), it follows that the polarization of the helix is typically elliptical with the main axes of the polarization ellipse along the directions $\hat{\boldsymbol{\theta}}, \hat{\boldsymbol{\varphi}}$. The axial ratio (AR) of the polarization ellipse is independent of the observation direction. The polarization of the antenna can be made circular by enforcing that:

$$p = k_0 \pi a^2 \qquad \text{(condition for circular polarization).} \tag{8.74}$$

The sense of rotation of the emitted field is determined by the orientation of the helix. For the + sign (right-handed helix) the antenna polarization is RCP, and for the "–" sign (left-handed helix) the antenna polarization is LCP. In practice, the bandwidth of the circular polarization regime is very narrow in normal mode, and hence this regime is rarely exploited in practice.

Theory of the axial mode

For the axial mode, the perimeter of each helix loop ($C = 2\pi a$) is on the order of the wavelength:

$$\frac{3}{4}\lambda_0 < C < \frac{4}{3}\lambda_0 \qquad \text{(axial mode operation).} \tag{8.75}$$

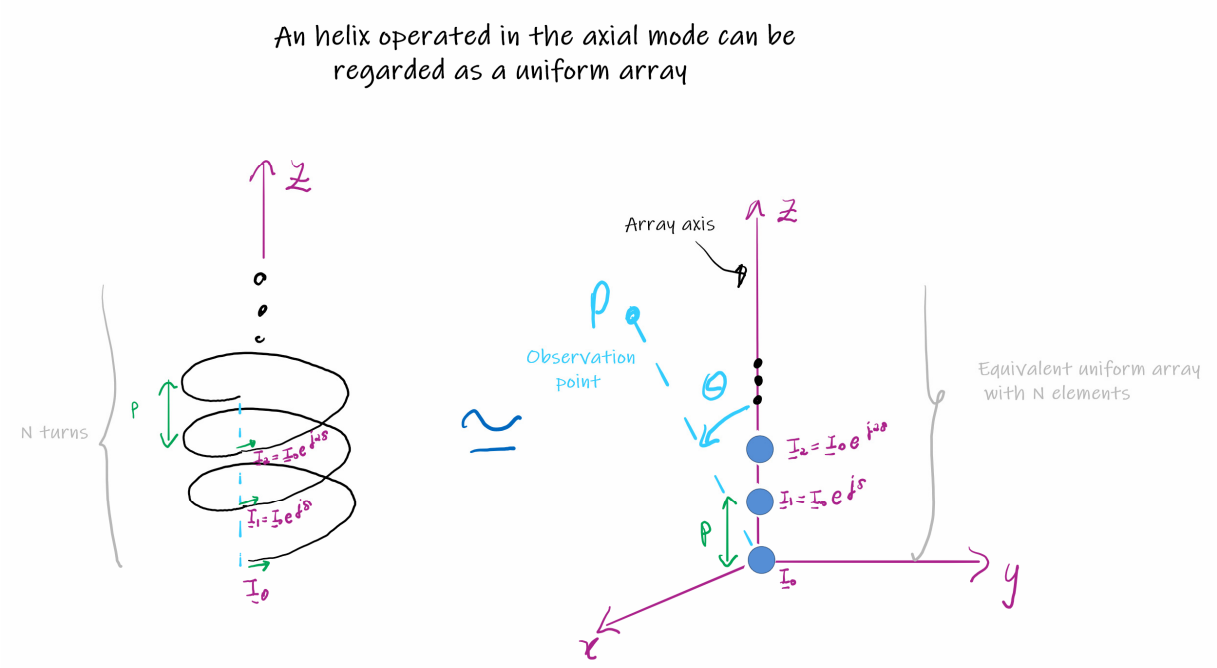
The fields radiated by the helix can be determined based on the semi-empirical observation that for a helix with $N >> 1$ turns, the current along the helix is approximately a travelling wave:

$$\underline{I}(s) = \underline{I}_0 e^{-jk_h s}. \tag{8.76}$$

The propagation constant of the current along the wire is k_h. Evidently, near the end of the helix the travelling wave hypothesis is inaccurate because the current must vanish. We will neglect

such an effect. The travelling wave hypothesis implies that all the turns of the helix are fed with currents of identical amplitude with the currents phase following an arithmetic progression. The feeding currents are

$$\underline{I}_n \equiv \underline{I}\left(s = nL_0\right) = \underline{I}_0 e^{jn\delta}, \quad \text{with } \delta = -k_h L_0, \qquad n = 0,1,\ldots N-1 \tag{8.77}$$



Thus, the helical antenna can be regarded as a uniform (progressive phaseshift) array formed by N identical elements. Each element of the array is a helix turn. The phaseshift between adjacent elements is $\delta = -k_h L_0$ (note that one loop turn corresponds to the shift $s \to s + L_0$).

From the principle of the multiplication of pattern diagrams:

$$\boxed{\left|\underline{\mathbf{E}}\right|_{\text{helix}} = \underbrace{\left|F\right|}_{\text{array spatial factor}} \times \frac{\left|\underline{\mathbf{E}}\right|_{\substack{\text{single turn}\\ \text{of the helix}}}}{\left|\underline{I}_0\right|}}. \tag{8.78}$$

The array spatial factor is the same as for a uniform array:

$$\left|F\right|_{\text{helix}} = \left|\underline{I}_0\right| \left| \frac{\sin\left(\frac{Nu}{2}\right)}{\sin\left(\frac{u}{2}\right)} \right|, \qquad u = k_0 p \cos\theta + \delta. \tag{8.79}$$

Note that the angle between the observation point and the array axis is θ because the array elements are placed along the *z*-axis. The distance between the array elements is $d = p$.

The value of the phaseshift $\delta = -k_h L_0$ depends on the (unknown) propagation constant k_h of the travelling wave of current. The value of k_h can be estimated taking into account the experimental fact that the maximum of radiation is near $\theta = 0^o$ (endfire array). This indicates that $u_{\theta=0^o} = 2\pi m$ for some *m* integer, or equivalently that $k_0 p - k_h L_0 = 2\pi m$. Since for typical helices the length of one turn is much larger than the pitch, $L_0 >> p$, the natural choice for *m* is $m = -1$. This yields the estimate:

$$k_h \approx k_0 \frac{p}{L_0} + \frac{2\pi}{L_0} \qquad \text{(ordinary endfire array approx.).} \qquad (8.80)$$

An alternative formula is obtained by supposing that the helix behaves as an "endfire with increased directivity", rather than as an ordinary endfire array. The endfire array with increased directivity was introduced by Hansen and Woodyard and corresponds to a standard uniform array with the array maximum determined by the condition $u_{\theta=\theta_{\max}} = -\frac{\pi}{N} + 2\pi m$. This second theory yields a different estimate for the helix propagation constant (picking again $m = -1$):

$$k_h \approx k_0 \frac{p}{L_0} + \frac{2\pi}{L_0}\left(1 + \frac{1}{2N}\right) \qquad \text{(Hansen-Woodyard array approx.).} \qquad (8.81)$$

This approximation is found to agree better with the experimental results. The above formula can also be written as:

$$\frac{k_0}{k_h} \approx \frac{\frac{L_0}{\lambda_0}}{\frac{p}{\lambda_0} + \left(1 + \frac{1}{2N}\right)} \qquad \text{(for} \quad 0.75\lambda_0 < C < 1.1\lambda_0\text{).} \qquad (8.82)$$

Within the vality of this formula, the array spatial factor of the helix is given by:

$$\boxed{\left|F\right|_{\text{helix}} = \left|\underline{I}_0\right| \left| \frac{\sin\left(\frac{Nu}{2}\right)}{\sin\left(\frac{u}{2}\right)} \right|, \qquad u = k_0 p\left(\cos\theta - 1\right) - \frac{\pi}{N} - 2\pi.} \tag{8.83}$$

The factor of -2π in the definition of u does not affect the shape of the array spatial factor.

Now that we characterized the equivalent array spatial factor for the helix, we turn our attention to the field radiated by a single helix turn. This field is completely determined by the effective length [Eq. (8.26)]. For a wire antenna it can be written as:

$$\mathbf{h}_e\left(\hat{\mathbf{r}}\right) = \frac{1}{\underline{I}_0}\left[\left(\int\limits_{\substack{\text{single helix}\\ \text{turn}}} dl\,\hat{\mathbf{t}}\left(s\right)\underline{I}\left(s\right)e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}\left(s\right)}\right)\times\hat{\mathbf{r}}\right]\times\hat{\mathbf{r}}\,. \tag{8.84}$$

In the above, $\mathbf{r}\left(s\right)$ determines a generic point along the helix and is given by Eq. (8.65). The vector $\hat{\mathbf{t}}\left(s\right)$ is a unit vector tangent to the helix and gives the direction of the current flow. Since the parameter s can be identified with the length along the curve, one has $\hat{\mathbf{t}}\left(s\right) = \mathbf{r}'\left(s\right)$, or equivalently:

$$\hat{\mathbf{t}}\left(s\right) = \frac{2\pi}{L_0}\left(-a\sin\left(\frac{2\pi}{L_0}s\right)\hat{\mathbf{x}} \pm a\cos\left(\frac{2\pi}{L_0}s\right)\hat{\mathbf{y}} + \frac{p}{2\pi}\hat{\mathbf{z}}\right). \tag{8.85}$$

Using $\underline{I}\left(s\right) = \underline{I}_0 e^{-jk_h s}$, one finds that:

$$\int\limits_{\substack{\text{single helix}\\ \text{turn}}} dl\,\hat{\mathbf{t}}\left(s\right)\frac{\underline{I}\left(s\right)}{\underline{I}_0}e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}\left(s\right)} = \int\limits_0^{L_0} ds e^{-jk_h s}e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}\left(s\right)}\frac{2\pi}{L_0}\left(-a\sin\left(\frac{2\pi}{L_0}s\right)\hat{\mathbf{x}} \pm a\cos\left(\frac{2\pi}{L_0}s\right)\hat{\mathbf{y}} + \frac{p}{2\pi}\hat{\mathbf{z}}\right). \tag{8.86}$$

The exact analytical evaluation of the integral is cumbersome, and will not be discussed here. To simplify the problem and gain some physical insight, it is enough to use the rough assumption $e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}\left(s\right)} \approx 1$. The phase of the exponent is $\left|\phi\right| \le \max k_0\left|\mathbf{r}\left(s\right)\right| \approx k_0\sqrt{a^2+p^2}$ (the parameter s varies along a single helix turn). An helix operated in the axial mode has typical electrical length such

that $k_h L_0 \approx 2\pi$, which is a few times larger than $k_0\sqrt{a^2+p^2}$ when $p << L_0$[1]. Thus, $|\phi|$ is typically a few times smaller than 2π, which justifies the considered approximation.

Using $e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}(s)} \approx 1$ the integration in Eq. (8.86) can be carried out analytically. To further simplify the problem we restrict our attention to the resonant case $k_h L_0 \approx 2\pi$. In this situation, we have:

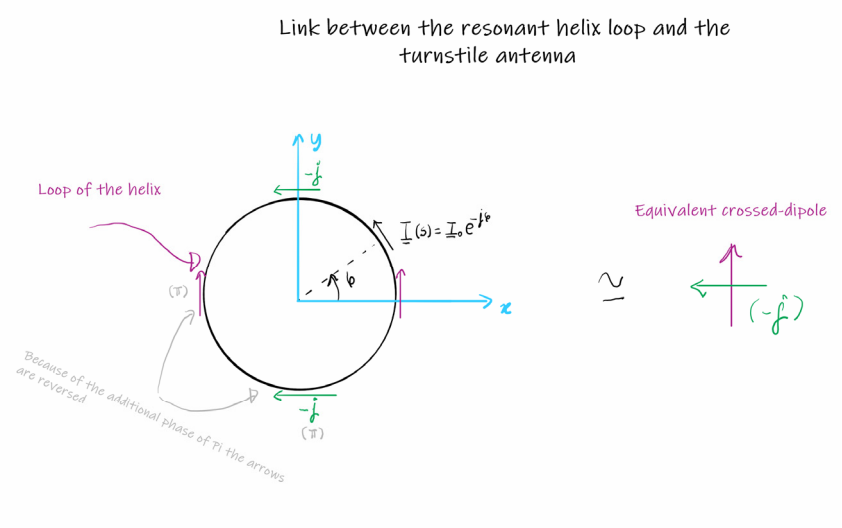
$$\int\limits_{\substack{\text{single helix}\\ \text{turn}}} dl\,\hat{\mathbf{t}}(s)\frac{\underline{I}(s)}{\underline{I}_0}e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}(s)} \approx \int\limits_0^{L_0} ds\, e^{-j\frac{2\pi}{L_0}s}\,\frac{2\pi}{L_0}\left(-a\sin\left(\frac{2\pi}{L_0}s\right)\hat{\mathbf{x}} \pm a\cos\left(\frac{2\pi}{L_0}s\right)\hat{\mathbf{y}} + \frac{p}{2\pi}\hat{\mathbf{z}}\right) \qquad (8.87)$$
$$= j\pi a\left(\hat{\mathbf{x}} - (\pm)\,j\hat{\mathbf{y}}\right)$$

Therefore, the effective length of the single helix turn is:

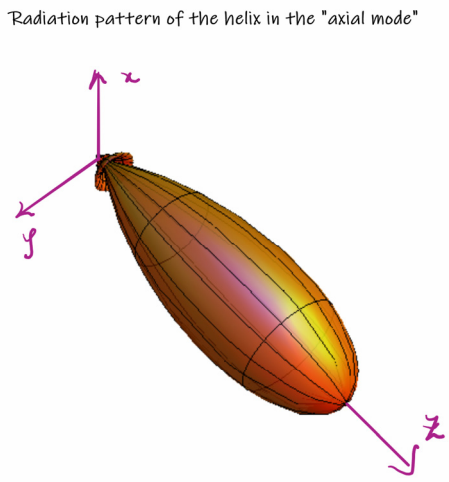
$$\boxed{\mathbf{h}_e\big|_{\substack{\text{single}\\ \text{helix turn}}} \approx j\pi a\left[\left(\left(\hat{\mathbf{x}} - (\pm)\,j\hat{\mathbf{y}}\right)\times\hat{\mathbf{r}}\right]\times\hat{\mathbf{r}}\right.}, \quad \text{(at resonance } k_h L_0 \approx 2\pi\text{)}. \qquad (8.88)$$

Comparing this result with Eq. (8.28), one sees that a single helix loop behaves as turnstile antenna formed by Hertz dipoles oriented along the *x* and *y* directions! The physical justification is that when $p << L_0$ an helix turn is roughly a loop with perimeter L_0. When $k_h L_0 \approx 2\pi$ the current gains a 90º phase delay when the azimuthal angle varies by 90º. Thus, the currents in the horizontal parts of the loop are in quadrature with the respect to the currents in the vertical parts of the loop, analogous to a turnstile antenna.

[1] The value of $k_0\sqrt{a^2+p^2}$ is roughly $k_0\sqrt{a^2+p^2} \approx 1.0$ when $k_h L_0 \approx 2\pi$ and $p << L_0$.



Evidently, within the considered assumptions, the radiation pattern of a single helix loop is the same as for a turnstile antenna. The antenna polarization along the $+z$-axis (maximum of radiation) is circular. The sense of rotation is RCP (LCP) when the helix is right (left) handed with respect to the $+z$ axis. Below, we show the radiation pattern for a resonant helix formed by 10 turns. The radiation pattern was evaluated using Eqs. (8.83) and (8.88) with $k_0 p = \pi / 2$.



To conclude this section, we quote some useful empirical formulas that are useful to design helix antennas operating in the axial mode:

$$Z_{in} \approx \frac{140C}{\lambda_0}\ [\Omega], \qquad \text{(input impedance, } C = 2\pi a\text{).} \tag{8.89a}$$

$$G \approx K_G \frac{N \times p \times C^2}{\lambda_0^3}, \quad K_G = 15, \quad \text{(antenna gain).} \tag{8.89b}$$

$$\text{HPBW} \approx \frac{K_B}{C}\sqrt{\frac{\lambda_0^3}{N \times p}}\ [\text{deg}], \quad K_B = 52, \quad \text{(half-power beamwidth).} \tag{8.89c}$$

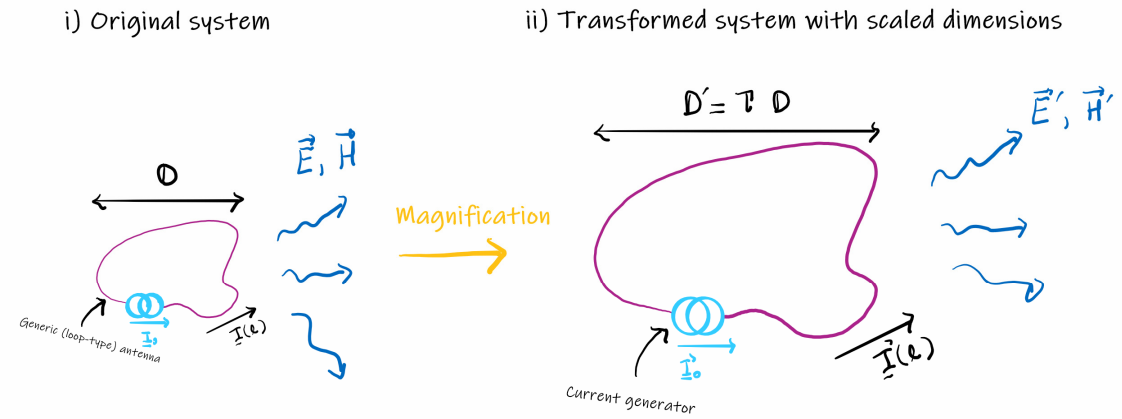
$$AR = \frac{2N+1}{2N}, \quad \text{(axial ratio).} \tag{8.89d}$$

9. Frequency independent antennas

9.1 Transformation of scale

The bandwidth of an antenna is typically limited by some "resonant condition". For example, for the dipole antenna the metal rod is resonant when its length equals half-wavelength. When the frequency is detuned away from the resonance, the emitted power drops sharply due to the change in the input impedance of the antenna. Typical bandwidths for a dipole antenna are on the order of 10%. As discussed in the previous chapter, larger bandwidths (more than 50%) can be achieved with a helical antenna operated in the axial mode. This due to the fact that the axial mode is associated with a travelling wave that propagates along the helix. The travelling wave is less sensitive to a frequency detuning than the standing wave existing on a dipole antenna.

Is it possible to design a radiating system whose radiating properties are frequency independent? To analyze this problem, let us study how the response of an electromagnetic system changes when its elements are scaled up or scaled down. For example, the figure below illustrates a loop-type antenna with arbitrary shape that is magnified (scaled up) by some factor τ.



The fields radiated by the original system are described by the Maxwell's equations:

$$\begin{aligned} \nabla \times \underline{\mathbf{E}} &= -j\omega\mu_0 \underline{\mathbf{H}} \\ \nabla \times \underline{\mathbf{H}} &= \mathbf{j}_e + j\omega\varepsilon_0 \varepsilon(\mathbf{r}) \underline{\mathbf{E}} \end{aligned} \tag{9.1}$$

Here, $\mathbf{j}_e$ describes the currents impressed by the external excitation, for example by a current generator. For simplicity the materials are assumed non-magnetic. The material structures that form the antenna are described by the space-dependent permittivity $\varepsilon(\mathbf{r})$.

Let us consider a scale transformation of the system such that:

$$\mathbf{r} \equiv (x, y, z) \xrightarrow{\tau} \mathbf{r}' \equiv (x', y', z') = (\tau x, \tau y, \tau z) \tag{9.2}$$

The parameter $\tau > 0$ determines the change of scale. When $\tau > 1$ all the dimensions of the system are scaled up (magnified) by a factor of τ, and when $\tau < 1$ all the dimensions are scaled down (demagnified) by τ. Let us introduce the primed fields,

$$\underline{\mathbf{E}}'(\mathbf{r}') \equiv \underline{\mathbf{E}}(\mathbf{r}), \qquad \underline{\mathbf{H}}'(\mathbf{r}') \equiv \underline{\mathbf{H}}(\mathbf{r}). \tag{9.3}$$

Using $\underline{\mathbf{E}}'(\mathbf{r}') = \underline{\mathbf{E}}(\mathbf{r}'/\tau)$, it is clear that $\dfrac{\partial \underline{\mathbf{E}}'}{\partial x'} = \dfrac{1}{\tau}\dfrac{\partial \underline{\mathbf{E}}}{\partial x}$, etc. This implies that $\nabla' \times \underline{\mathbf{E}}' = \dfrac{1}{\tau}(\nabla \times \underline{\mathbf{E}})$ and $\nabla' \times \underline{\mathbf{H}}' = \dfrac{1}{\tau}(\nabla \times \underline{\mathbf{H}})$, with the right-hand sides of the equations evaluated at the point $\mathbf{r} = \mathbf{r}'/\tau$. Using this result in Eq.(9.1) one sees that:

$$\begin{aligned} \tau \nabla' \times \underline{\mathbf{E}}' &= -j\omega\mu_0 \underline{\mathbf{H}}' \\ \tau \nabla' \times \underline{\mathbf{H}}' &= \mathbf{j}_e(\mathbf{r}) + j\omega\varepsilon_0 \varepsilon(\mathbf{r}) \underline{\mathbf{E}}' \end{aligned}$$

which is equivalent to

$$\begin{aligned} \nabla' \times \underline{\mathbf{E}}' &= -j\omega'\mu_0 \underline{\mathbf{H}}' \\ \nabla' \times \underline{\mathbf{H}}' &= \mathbf{j}'_e + j\omega'\varepsilon_0 \varepsilon' \underline{\mathbf{E}}' \end{aligned}$$

where we defined:

$$\begin{aligned}\omega' &= \omega/\tau \\ \varepsilon'(\mathbf{r}') &= \varepsilon(\mathbf{r}) \\ \mathbf{j}_e'(\mathbf{r}') &= \frac{1}{\tau}\mathbf{j}_e(\mathbf{r})\end{aligned} \tag{9.4}$$

The transformed (primed) fields satisfy the Maxwell's equations in the system described by the permittivity $\varepsilon'(\mathbf{r}') = \varepsilon(\mathbf{r})$ with $\mathbf{r} = \mathbf{r}'/\tau$ subject to the transformed external excitation $\mathbf{j}_e'$. In other words, the transformed fields satisfy the Maxwell's equations in the "τ-scaled system". The scaled system is identical to the original system except that the dimensions of all the materials bodies are multiplied by the scaling factor τ. If the external excitation is a lumped current generator (Hertz dipole current distribution) then the transformed current is also a lumped current generator with a scaled amplitude.

Importantly, the operating frequency is transformed as $\omega \to \omega' = \omega/\tau$. Hence, when a system is *scaled up* (magnified) by a factor τ the corresponding operating frequency is *scaled down* by the same factor. This means that if the dimensions of any antenna are scaled up then all the resonance frequencies are scaled down in the same proportion. For example, if the length of a dipole antenna is increased, the corresponding resonance frequency decreases. In particular, we conclude that the resonance frequencies of a radiating system depends on the scale, i.e., depends on the size of the material structures that form the antenna.

The previous result may suggest that it is impossible to design a radiating system whose response is frequency independent. In abstract, there is however a "way out" of the problem: the radiating systems that are *invariant under a change of scale.*

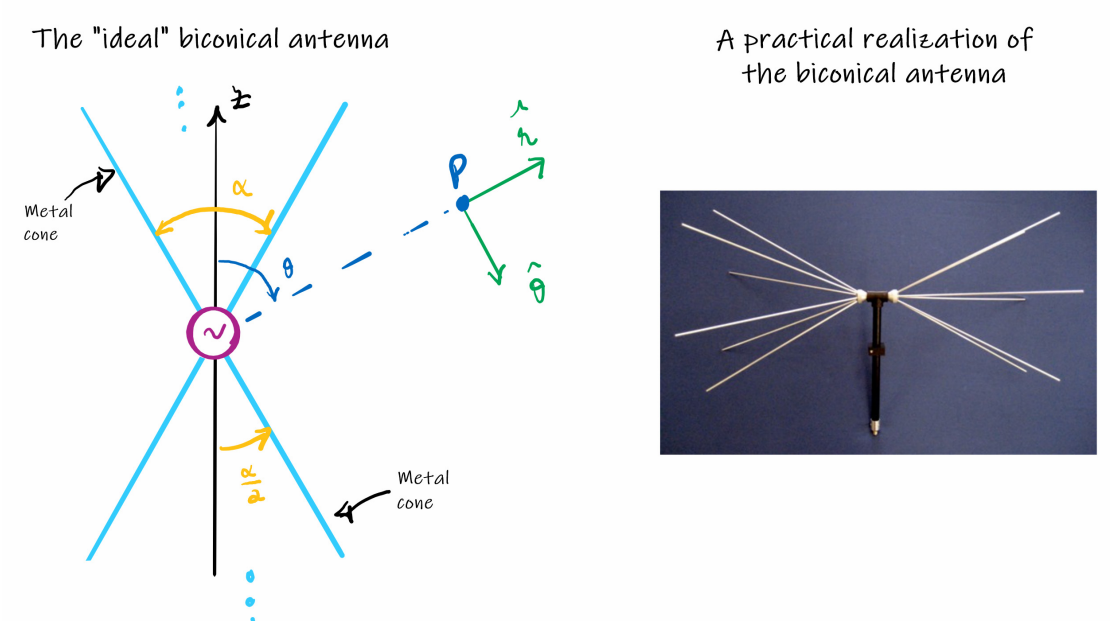
A system that is invariant under a change of scale is a system that remains invariant upon an arbitrary magnification or demagnification. Formally, the material structures that comprise the system are subject to the constraint $\varepsilon(\mathbf{r}) = \varepsilon(\mathbf{r}/\tau)$ with τ *arbitrary*. Thus, the permittivity cannot depend on the distance to the origin, it can only depend on the direction $\hat{\mathbf{r}}$. This implies

that a frequency independent antenna is defined by *angles* rather than by *lengths*. We will see an example in the next subsection.

For systems invariant under a change of scale the unprimed ($\underline{\mathbf{E}}$) and primed ($\underline{\mathbf{E}}'$) fields are defined over the *same* material structure. Since $\underline{\mathbf{E}}' = \underline{\mathbf{E}}$ and $\omega' = \omega / \tau$ the response of such systems does not depend on frequency because the field distribution for the frequency ω is the same as the frequency response for any other frequency ω' [1]. Thus, the bandwidth of such idealized systems is infinitely large!

9.2 Biconical antenna

To illustrate the ideas and give an example of a frequency independent antenna we consider the biconical antenna. The antenna is formed by two identical metallic cones (ideally perfect electric conductors) that are fed by a point (infinitesimal) generator placed at the origin. The antenna is clearly invariant under a change of scale. Note that the antenna can be regarded as a dipole antenna with infinite length and with a tapered "wire" diameter (increasing with the distance to the feeding point).



To find the fields radiated by the antenna we look for a solution of the type:

$$\underline{\mathbf{E}} = \underline{E}_\theta \hat{\boldsymbol{\theta}} \ , \qquad \underline{\mathbf{H}} = \underline{H}_\varphi \hat{\boldsymbol{\varphi}} \ . \tag{9.5}$$

We adopt a spherical coordinate system. Since the antenna has symmetry of revolution around the *z*-axis, the fields must be independent of the azimuth angle φ. Taking this into account, one finds that:

$$\nabla \times \underline{\mathbf{E}} = \frac{1}{r^2 \sin\theta} \begin{vmatrix} \hat{\mathbf{r}} & r\hat{\boldsymbol{\theta}} & r\sin\theta\hat{\boldsymbol{\varphi}} \\ \partial_r & \partial_\theta & \partial_\varphi \\ \underline{E}_r & r\underline{E}_\theta & r\sin\theta\underline{E}_\varphi \end{vmatrix} = \frac{1}{r^2 \sin\theta} \begin{vmatrix} \hat{\mathbf{r}} & r\hat{\boldsymbol{\theta}} & r\sin\theta\hat{\boldsymbol{\varphi}} \\ \partial_r & \partial_\theta & 0 \\ 0 & r\underline{E}_\theta & 0 \end{vmatrix} = \frac{1}{r}\partial_r\left(r\underline{E}_\theta\right)\hat{\boldsymbol{\varphi}} ,$$

$$\nabla \times \underline{\mathbf{H}} = \ldots = \frac{1}{r^2 \sin\theta} \begin{vmatrix} \hat{\mathbf{r}} & r\hat{\boldsymbol{\theta}} & r\sin\theta\hat{\boldsymbol{\varphi}} \\ \partial_r & \partial_\theta & 0 \\ 0 & 0 & r\sin\theta\underline{H}_\varphi \end{vmatrix} = \frac{1}{r\sin\theta}\left[\partial_\theta\left(\sin\theta\underline{H}_\varphi\right)\hat{\mathbf{r}} - \partial_r\left(r\sin\theta\underline{H}_\varphi\right)\hat{\boldsymbol{\theta}}\right]$$

Substituting these formulas into the Maxwell's equations one finds that in the air region:

$$\begin{aligned} \frac{1}{r}\partial_r\left(r\underline{E}_\theta\right)\hat{\boldsymbol{\varphi}} &= -j\omega\mu_0\underline{H}_\varphi\hat{\boldsymbol{\varphi}} \\ \frac{1}{r\sin\theta}\left[\partial_\theta\left(\sin\theta\underline{H}_\varphi\right)\hat{\mathbf{r}} - \partial_r\left(r\sin\theta\underline{H}_\varphi\right)\hat{\boldsymbol{\theta}}\right] &= j\omega\varepsilon_0\underline{E}_\theta\hat{\boldsymbol{\theta}} \end{aligned} \ . \tag{9.6}$$

This is equivalent to:

$$\begin{aligned} \frac{\partial}{\partial r}\left(r\underline{E}_\theta\right) &= -j\omega\mu_0\left(r\underline{H}_\varphi\right) \\ \frac{\partial}{\partial r}\left(r\underline{H}_\varphi\right) &= -j\omega\varepsilon_0\left(r\underline{E}_\theta\right) \end{aligned} , \qquad \text{and} \quad \partial_\theta\left(\sin\theta\underline{H}_\varphi\right) = 0 . \tag{9.7}$$

The equation system on the left is formally equivalent to "telegrapher" equations with $r\underline{E}_\theta \to \underline{V}$, $r\underline{H}_\varphi \to \underline{I}$, $\mu_0 \to L$ and $\varepsilon_0 \to C$:

$$\begin{aligned} \frac{\partial \underline{V}}{\partial x} &= -j\omega L\underline{I} \\ \frac{\partial \underline{I}}{\partial x} &= -j\omega C\underline{V} \end{aligned} \ . \tag{9.8}$$

[1] This result holds true only if the permittivity is frequency independent. Dispersive systems remain frequency dependent, even if the radiating structure is defined only by angles.

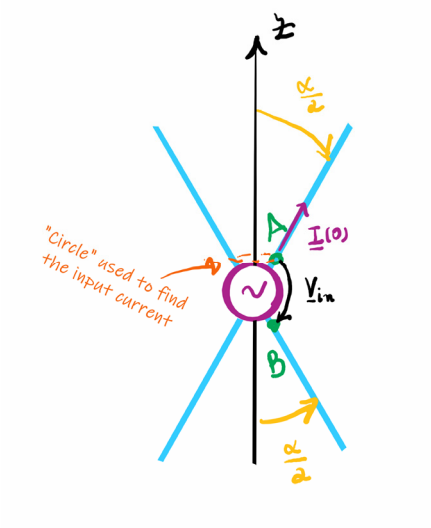
The solution of the telegrapher equations is of the form $\underline{V} = \underline{V}_0^+ e^{-j\beta x} + \underline{V}_0^- e^{+j\beta x}$ and $\underline{I} = \frac{\underline{V}_0^+}{Z_c} e^{-j\beta x} - \frac{\underline{V}_0^-}{Z_c} e^{+j\beta x}$. For a radiating system the emitted fields must be outgoing waves and thereby one can exclude the terms associated with the propagation factor $e^{+j\beta x}$. Hence, by analogy, we see that the fields emitted by the biconical antenna are of the form:

$$r\underline{E}_\theta = \alpha_0^+ e^{-jk_0 r} \ , \qquad r\underline{H}_\varphi = \frac{\alpha_0^+}{\eta_0} e^{-jk_0 r} \ . \tag{9.9}$$

It was taken into account that the propagating constant and impedance of the equivalent line are $\beta = \sqrt{LC}\omega = \sqrt{\varepsilon_0 \mu_0}\,\omega = k_0$ and $Z_c = \sqrt{L/C} = \sqrt{\mu_0/\varepsilon_0} = \eta_0$. The parameter α_0^+ is independent of the radial distance but can depend on the elevation angle: $\alpha_0^+ = \alpha_0^+(\theta)$. To find this dependence, we use $\partial_\theta\left(\sin\theta \underline{H}_\varphi\right) = 0$. This equation implies that the magnetic field is of the form $\underline{H}_\varphi = \frac{f(r)}{\sin\theta}$ where f is some function of the radial distance. This shows that $\alpha_0^+(\theta) = \frac{\underline{V}_0}{\sin\theta}$ with $\underline{V}_0$ some constant independent of the system coordinates. In summary, the fields emitted by the biconical antenna are:

$$\underline{E}_\theta = \frac{\underline{V}_0}{r\sin\theta} e^{-jk_0 r} \ , \qquad \underline{H}_\varphi = \frac{\underline{V}_0}{\eta_0 r\sin\theta} e^{-jk_0 r} \ . \tag{9.10}$$

We do not need to worry with the boundary conditions on the metal cones because the electric field is already perpendicular to the conical surfaces. The derived result is *exact*.



Let us now find the input impedance of the antenna. To do this, we need to find the voltage and the current near the point generator. The input voltage is given by:

$$\underline{V}_{in} = \underline{V}_{AB} = \int_{A}^{B} \underline{\mathbf{E}} \cdot \mathbf{dl} = \int_{\alpha/2}^{\pi-\alpha/2} \underline{E}_{\theta}\big|_{r=0^{+}} (r d\theta). \tag{9.11}$$

The integration can be carried out explicitly noting that $\underline{V}_{in} = \underline{V}_0 \int_{\alpha/2}^{\pi-\alpha/2} \frac{1}{\sin\theta} d\theta = \underline{V}_0 \left(\ln \tan \frac{\theta}{2} \right) \Big|_{\alpha/2}^{\pi-\alpha/2}$. This yields after a little simplication:

$$\underline{V}_{in} = 2\underline{V}_0 \ln \cot \frac{\alpha}{4}. \tag{9.12}$$

On the other hand, the input current $\underline{I}(0)$ can be found from the radial surface current flowing on the conical surface. For a perfect electric conductor the surface current density is $\mathbf{K}_s = \hat{\mathbf{n}} \times \underline{\mathbf{H}}$ with $\hat{\mathbf{n}}$ the outward unit vector normal to the metal surface. In the present problem, $\hat{\mathbf{n}} = \hat{\boldsymbol{\theta}}$ for the top metallic conical surface. Thus, the surface current density is $\mathbf{K}_s = \hat{\boldsymbol{\theta}} \times \underline{H}_{\varphi} \hat{\boldsymbol{\varphi}} = \underline{H}_{\varphi} \hat{\mathbf{r}}$, which as expected, is a radial vector field. The input current is given by the flux of $\mathbf{K}_s$ through a cross-section of the cone (a circumference in "orange" in the figure) near the generator:

$$\underline{I}(0) = \int_{\text{circle with } \theta=\frac{\alpha}{2}} \mathbf{K}_s \cdot \hat{\mathbf{r}} dl = \int_{\text{circle with } \theta=\frac{\alpha}{2}} \mathbf{K}_s \cdot \hat{\mathbf{r}} \underbrace{r \sin\theta d\varphi}_{dl}. \tag{9.13}$$

Using $\mathbf{K}_s \cdot \hat{\mathbf{r}} = \underline{H}_\varphi$ one gets:

$$\underline{I}(0) = 2\pi \left(\underline{H}_\varphi r \sin\theta \right)_{r=0^+} = \frac{2\pi \underline{V}_0}{\eta_0}. \tag{9.14}$$

Thus, the input impedance of the antenna is:

$$\boxed{Z_{in}\big|_{\text{biconical}} = \frac{\underline{V}_{in}}{\underline{I}(0)} = \frac{\eta_0}{\pi} \ln\cot\frac{\alpha}{4}}. \tag{9.15}$$

Remarkably, the input impedance is purely resistive and independent of frequency. This is a consequence of the antenna being invariant under a change of scale!

Next, we obtain the radiation pattern of the antenna. From the definition, the radiation intensity can be written as:

$$U = S_r r^2\big|_{\text{far-field}} = \frac{\left|\underline{\mathbf{E}}\right|^2}{2\eta_0} r^2 = \frac{\left|\underline{V}_0\right|^2}{2\eta_0 \sin^2\theta}. \tag{9.16}$$

The radiated power can be found by direct integration $P_{rad} = \iint U d\Omega$ or, more simply, using $P_{rad} = \frac{1}{2}\mathrm{Re}\left\{ Z_{in} \left|\underline{I}(0)\right|^2 \right\}$ (because the system is lossless). This gives:

$$P_{rad} = \frac{\eta_0}{2\pi} \ln\cot\frac{\alpha}{4} \left|\underline{I}(0)\right|^2 = \left|\underline{V}_0\right|^2 \frac{2\pi}{\eta_0} \ln\cot\frac{\alpha}{4}. \tag{9.17}$$

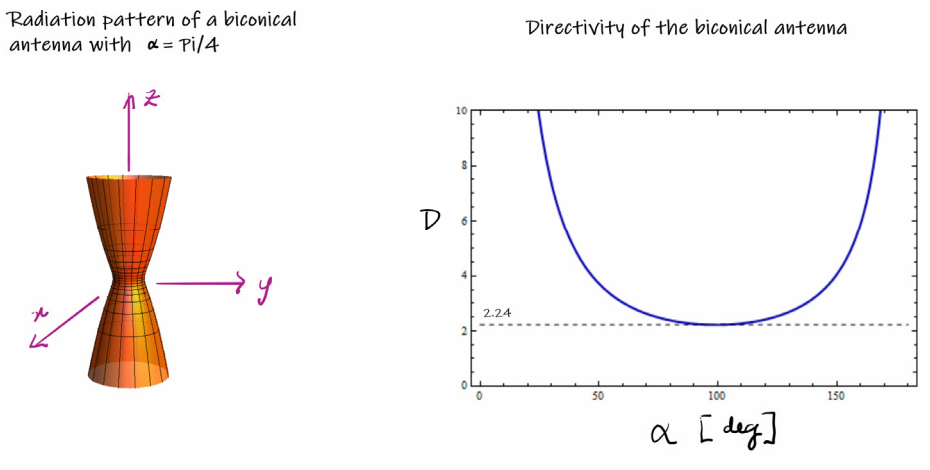
Thus, the directive gain of the antenna is:

$$\boxed{g\big|_{\text{biconical}} = \frac{4\pi U}{P_{rad}} = \frac{1}{\sin^2\theta} \frac{1}{\ln\cot\frac{\alpha}{4}}}. \tag{9.18}$$

Note that the directive gain and the radiation pattern are independent of frequency. The maximum of the directive gain occurs for $\theta = \alpha/2$, i.e. for the directions that define the conical surfaces. The directivity of the antenna is:

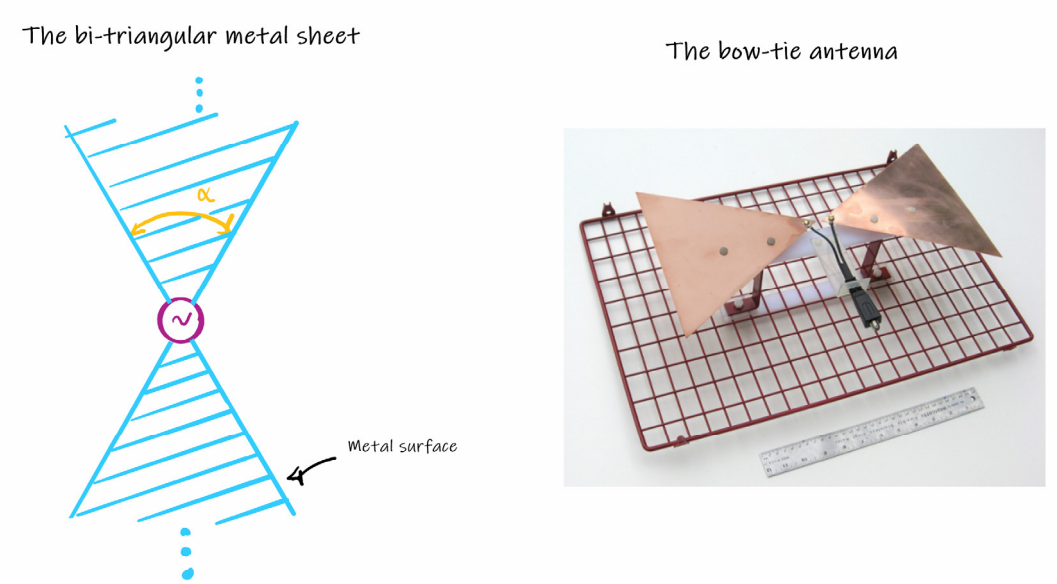
$$\boxed{D\big|_{\text{biconical}} = \frac{1}{\sin^2\frac{\alpha}{2}} \frac{1}{\ln\cot\frac{\alpha}{4}}}. \tag{9.19}$$

The directivity diverges ($D \to \infty$) when $\alpha \to 0$ (conical surface is "needle"-like) or when $\alpha \to \pi$ (the two conical surfaces almost touch).



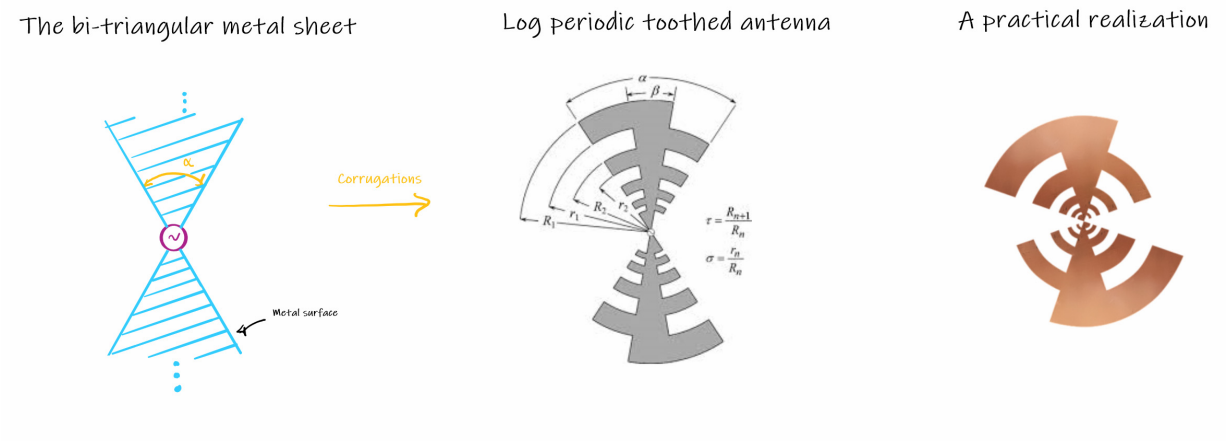
Evidently, a practical version of a biconical antenna has finite dimensions, and thereby a finite bandwidth. The bandwidth ($f_L < f < f_U$) of a realistic biconical antenna can be as large as $f_U / f_L = 8$. However, because a solid biconical antenna is so massive it is impractical to use at most frequencies. A planar version of the antenna is the bi-triangular metal sheet. For an unbounded structure, this antenna is also frequency independent provided the thickness of the metal sheet is negligible.

The "bow-tie" antenna is a truncated version of the bi-triangular metal sheet. The bow-tie antenna has a linear polarization with two broad main beams in the plane perpendicular to the bow-tie. Unfortunately, as compared to the biconical antenna, the bow-tie is relatively narrowband due to the truncation effects. The truncation can affect significantly the performance of the system when the current at the truncating points is significant.



9.3 Log-periodic antennas

As discussed in the end of previous subsection, truncating a frequency-independent antenna can have disastrous consequences on its performance. This can be avoided with a design that enforces that the current decays with distance to the input terminals, so that the antenna, even though finite sized, behaves as "effectively infinite". One way of making the currents decay rapidly away from the feed is to introduce "discontinuities", for example, for the bowtie antenna, introducing "teeth" in the fins. However, the corrugations deteriorate the "self-scaling" properties of the antenna.



Interestingly, it is possible to design radiating systems that are not defined by "angles" and nonetheless still have *some* "frequency independent" features. These systems are known are "log-periodic antennas". The concept is as follows.

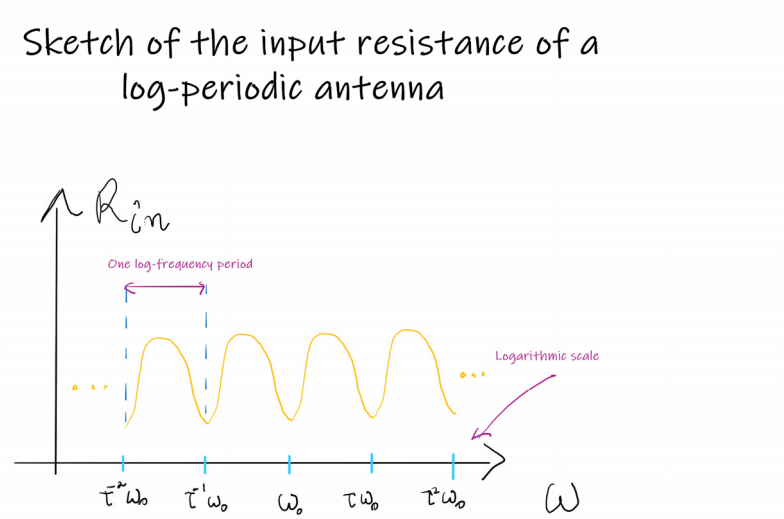
It was demonstrated in Sect. 9.1, that if a system is invariant under a scale transformation $(x,y,z)\xrightarrow{\tau}(\tau x,\tau y,\tau z)$ then the solution of Maxwell's equations at frequency $\omega'=\omega/\tau$ is fully determined by the solution of Maxwell's equations at frequency ω. If one wishes to have a frequency independent response, the system needs to be invariant under *any* scale transformation (invariant for any τ), and in that case, as discussed in Sect. 9.1, the structure can only be defined by "angles". However, one can impose a less severe constraint on the system, and just enforce that it is invariant under some *fixed* scale transformation, i.e., the radiating system is scale invariant for some fixed τ. In such a case, the antenna response certainly varies with frequency, but, from the previous discussion, its response is necessarily repeated for frequencies related by the scaling factor τ:

$$...., \ \tau^{-2}\omega_0, \ \tau^{-1}\omega_0, \ \omega_0, \ \tau\omega_0, \ \tau^2\omega_0, \ ....$$

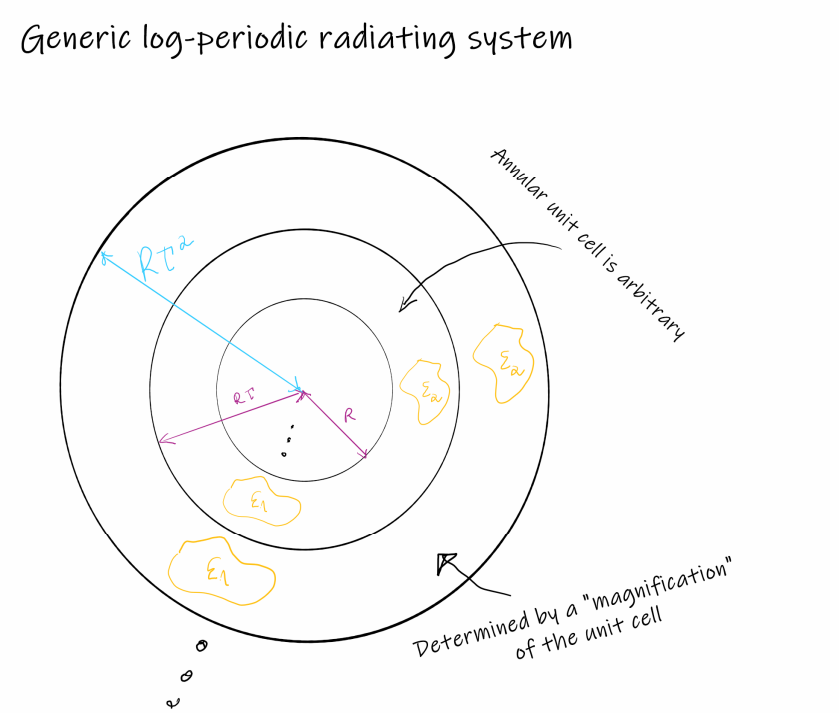
For example, the input impedance or the radiation pattern of a τ-scale invariant radiating system are repeated for the sequence of frequencies $...., \ \tau^{-2}\omega_0, \ \tau^{-1}\omega_0, \ \omega_0, \ \tau\omega_0, \ \tau^2\omega_0, ...$ Thus, the system response is a periodic function of the *logarithm of the frequency*

$$...., \ -2\log\tau+\log\omega_0, \ -\log\tau+\log\omega_0, \ \log\omega_0, \ \log\tau+\log\omega_0, \ +2\log\tau+\log\omega_0, \ ....$$

Accordingly, this class of self-scalable antennas is known as "log-periodic". Due to the fact that the radiation properties are repeated with a logarithmic-periodicity the log-periodic antennas are typically broadband when the magnification factor τ is near 1.



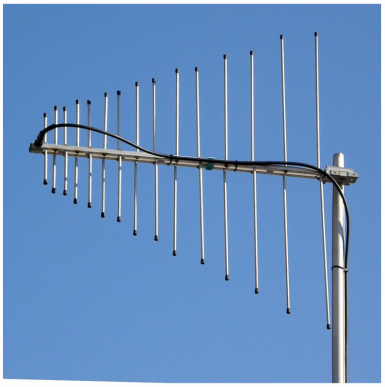
A log-periodic radiating system has material parameters constrained by $\varepsilon(\mathbf{r}) = \varepsilon(\mathbf{r}/\tau)$ for a fixed τ. Geometrically, the constraint means that the system does not change under a magnification or demagnification by the fixed scale factor τ. The geometry of the system is completely determined by any annular region of the form $R < r < \tau R$.



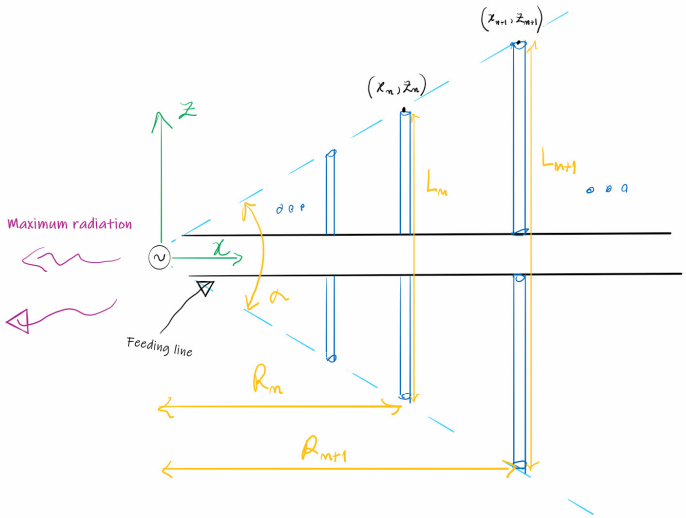
The log periodic toothed antenna (see the figure in page 164) is a modified bi-triangular metal sheet where the corrugations are used to make the antenna "effectively infinite", as previously discussed. The dimensions of the antenna are chosen to satisfy the log-periodic constraint. For example, the radial distances of two consecutive "teeth" must satisfy $\tau = R_{n+1} / R_n$. If the structure is ideally extended to infinity (in the outward direction) and to zero (in the inward direction) its response will be repeated for frequencies of the form $\tau^n \omega_0$ with $n = 0, \pm 1, \pm 2, \ldots$. It is found that the current extends up to the region where the length of the antenna is roughly one quarter of wavelength. The truncation does not affect the antenna response provided the current profile of the infinite structure (about one quarter of wavelength) fits into the truncated structure. Thus, the frequency limits of operation are set by the frequencies for which the largest and smallest teeth are about one quarter of wavelength. The radiation pattern is bi-directional with maxima in the directions perpendicular to the plane of the antenna.

The log-periodic dipole array is another example of a log-periodic antenna of widespread use in terrestrial TV reception. It is formed by an array of dipole antennas whose lengths and positions are scaled "log-periodically".

The log-periodic dipole array



Geometry of the antenna



It is supposed that the antennas lie in the *y*=0 plane and that the dipole axes are along *z*. The coordinates of the top edge of the *n*-th dipole are denoted by $(x_n, 0, z_n)$ $(n = 0, \pm 1, \pm 2, ...)$. The coordinates of different elements must be related as $(x_{n+1}, z_{n+1}) = \tau (x_n, z_n)$. This implies that (see the figure)

$$\frac{R_{n+1}}{R_n} = \tau = \frac{L_{n+1}}{L_n}. \tag{9.20}$$

The radial distances and the lengths are related to the aperture angle α as

$$\tan\frac{\alpha}{2} = \frac{L_n}{2R_n}. \tag{9.21}$$

Typically the array is fed by a single transmission line. The number of elements of the array controls the bandwidth (not the gain). All the antenna elements are directly fed by the transmission line. The active region of the antenna is determined by the few dipoles near the one with half-wavelength. Due to this reason the antenna behaves as "effectively infinite". The operating band is determined by the length of the shorter (maximum operating frequency) and longer (smallest operating frequency) dipoles. The maximum of radiation is in the direction of the apex similar to the Yagi-Uda antenna. The longer dipoles behave as reflectors and the shorter dipoles as directors.

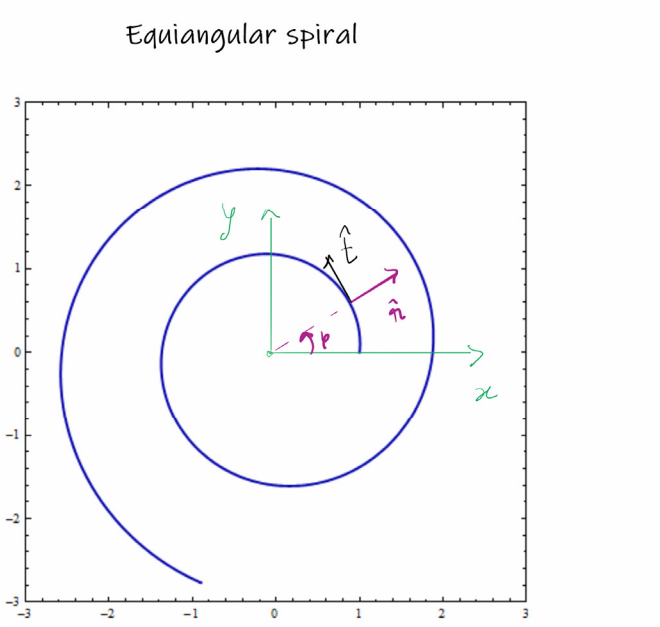
9.4 Spiral antennas

There are other radiating structures which are not frequency independent but that which behave in a similar way. An example is the spiral antenna.

Consider a curve in the *xoy* plane defined by

$$r = r_0 e^{a(\varphi - \varphi_0)}, \tag{9.22}$$

with r_0, a, φ_0 some constants. The curve is known as an equiangular spiral. The name of the curve comes from the fact that the angle between the tangent vector ($\hat{\mathbf{t}}$) and the radial vector ($\hat{\mathbf{r}}$) is the same for all points of the curve.

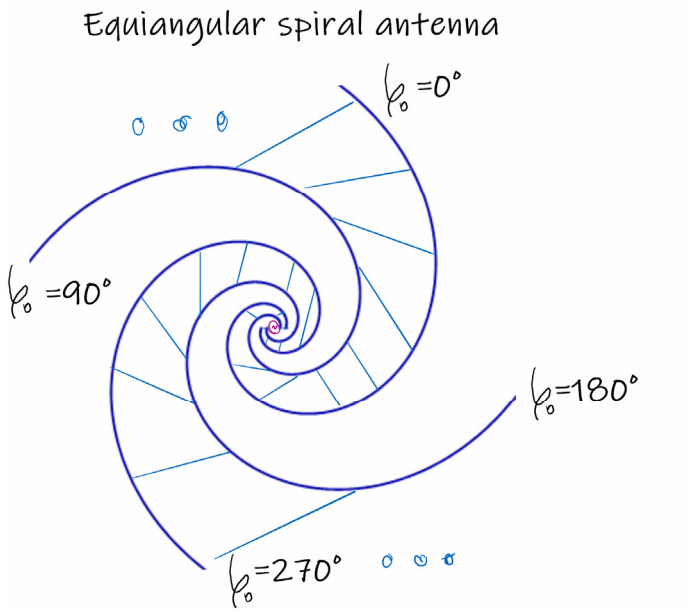


The equiangular spiral has a truly remarkable property: it stays invariant under *any* magnification (or demagnification) apart from a *rotation* about the *z*-axis. The proof is simple. Consider an arbitrary scaling factor $\tau = e^{\alpha}$ (here $\alpha \equiv \ln \tau$) . Under a scale transformation the spiral is transformed as $\mathbf{r} \to \mathbf{r}\tau$, or equivalently $r \to \tau r = e^{\alpha} r_0 e^{a(\varphi - \varphi_0)}$. Thus, the transformed curve is described by

$$r' = r_0 e^{a\left(\varphi - \varphi_0 + \frac{\alpha}{a}\right)} = r_0 e^{a(\varphi - \varphi_0')}, \qquad \text{with } \varphi_0' = \varphi_0 - \frac{\alpha}{a} = \varphi_0 - \frac{\ln \tau}{a}. \tag{9.23}$$

This shows that the transformed spiral is obtained from the original one through a *rotation* by the angle $-\dfrac{\ln \tau}{a}$ around the *z*-axis. If the rotation angle is multiple of 2π the spiral transforms into itself. Thereby, the spiral is log-periodic for the scaling factor τ that satisfies $\dfrac{\ln \tau}{a} = 2\pi$.

Consider now a planar antenna (e.g., a printed metallic antenna with the metal sheet having negligible thickness) whose shape is defined by equiangular spirals. Specifically, the equiangular spiral antenna is formed by two metallic arms connected by a point generator at the origin, with each arm defined by two spirals. The entire antenna is generated by 4 spirals with the same r_0, a but with $\varphi_0 = \underbrace{0, \pi/2}_{\text{1st arm}}, \underbrace{\pi, 3\pi/2}_{\text{2nd arm}}$.



The antenna is rigourously self-scaling apart from a rotation about the *z*-axis. This implies that the fields radiated by the antenna are also independent of frequency, apart, possibly, from a rotation about the *z*-axis (and apart from some irrelevant phase or amplitude factors). For example, changing the frequency can only *rotate* the radiation pattern of the antenna, but not change its shape or its directionality! In particular, the directivity of the antenna is rigorously independent of frequency. The input impedance is also frequency independent.

Evidently, in a realistic design the antenna needs to be truncated both in the outward ($r = R_{\max}$) and inward directions ($r = R_{\min}$). The radiation from the spiral comes mainly from the region with radius $r = \lambda_0/2$. Consequently, if the structure is truncated at $r = \lambda_0$, the truncated structure, although finite, is effectively infinite. The bandwidth of the structure is roughly

determined by $R_{\min} < \lambda_0 < R_{\max}$. Thus, the upper and lower frequency limits satisfy $f_U / f_L = R_{\max} / R_{\min}$.

The spiral antenna has a bi-directional radiation pattern with the maximum in the broadside direction (z) relative to the plane of the spiral. The radiated fields follow approximately a $\cos\theta$ law ($|\underline{\mathbf{E}}| \sim \cos\theta$), and thereby the half-power beamwidth is 90º. The antenna polarization is circular.

10. Aperture antennas

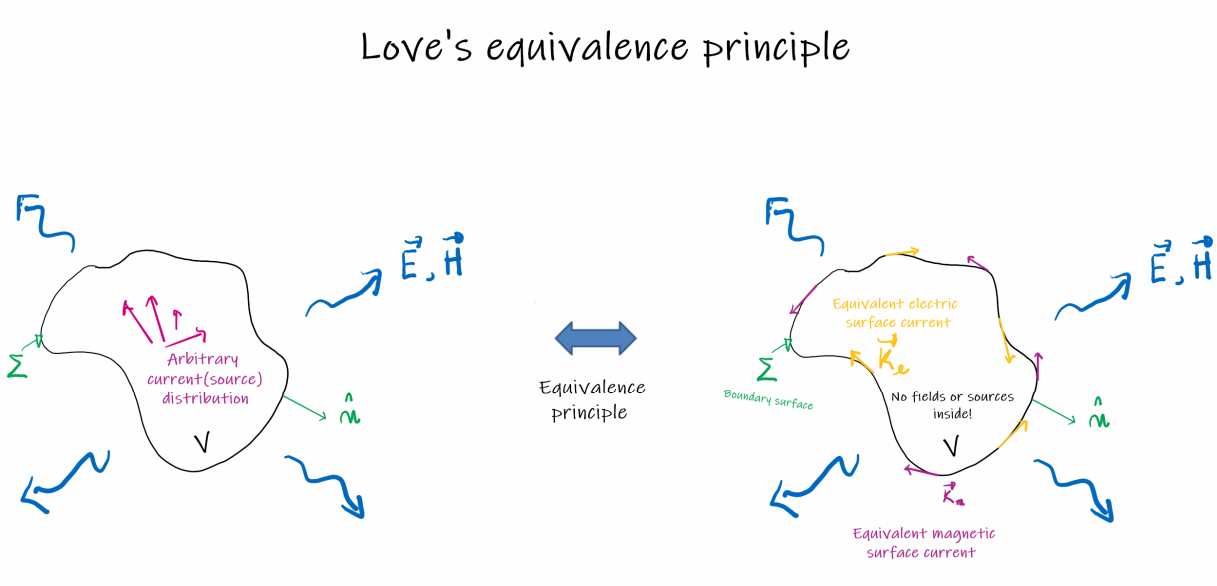
10.1 Introduction

For all the antennas studied so far in this course, the current distribution was assumed to be known with some degree of accuracy. For example, for the dipole antenna, the Yagi-Uda, the corner-reflector, the cross-dipole or the loop antenna the current distributions are approximately sinusoidal standing waves; for the helical antenna the current is a travelling wave. The radiation fields are determined by the antenna effective length and hence by the integration of the current distribution. This family of antennas is known as “wire-type antennas”.

There is another class of antennas for which the current is difficult to “guess” or calculate but for which the fields are known with reasonable accuracy on some surface enclosing the antenna. For such antennas, it is simpler to find the antenna far-field from the electromagnetic fields defined on the aforementioned surface, rather than from the current distribution. These antennas are referred to as aperture antennas. The most prominent of these are the horn antenna, the slot antenna and the parabolic reflector.

10.2 The equivalence principle

The equivalence principle (formulated by A. Love in 1901) is an important result of electromagnetic theory that establishes that the fields radiated by an arbitrary source distribution fully contained inside some volume V are univocally determined by the (tangential) electromagnetic fields on the surface Σ that encloses V.



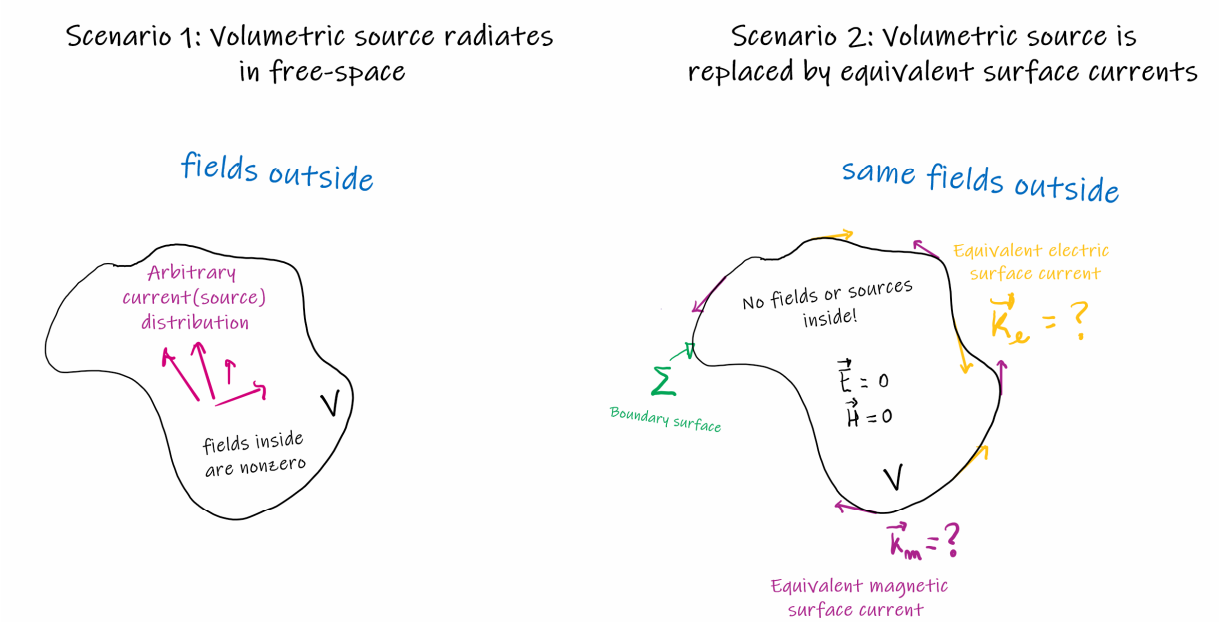
Futhermore, the theorem establishes that the field in the outside region is radiated by equivalent surface currents $\underline{\mathbf{K}}_e$ and $\underline{\mathbf{K}}_m$ defined on Σ, with no currents inside V. The equivalent currents are determined by the electromagnetic fields on the surface Σ

$$\boxed{\underline{\mathbf{K}}_e = \hat{\mathbf{n}} \times \underline{\mathbf{H}}\big|_\Sigma}, \qquad \boxed{\underline{\mathbf{K}}_m = -\hat{\mathbf{n}} \times \underline{\mathbf{E}}\big|_\Sigma}. \tag{10.1}$$

Here, $\hat{\mathbf{n}}$ is the surface unit normal vector oriented in the outward direction. Note that the equivalent currents only depend on the components of the electromagnetic fields tangential to the surface.

Demonstrating the equivalence principle is equivalent to solving the following "puzzle": Given some arbitrary (volumetric) source distribution contained in some volume V what are the equivalent surface currents defined on the boundary surface Σ that will create the same field as the original source distribution in the outside region and create vanishing fields inside V?

Even though the problem may look rather nontrivial, its solution requires little more than basic electromagnetic theory.



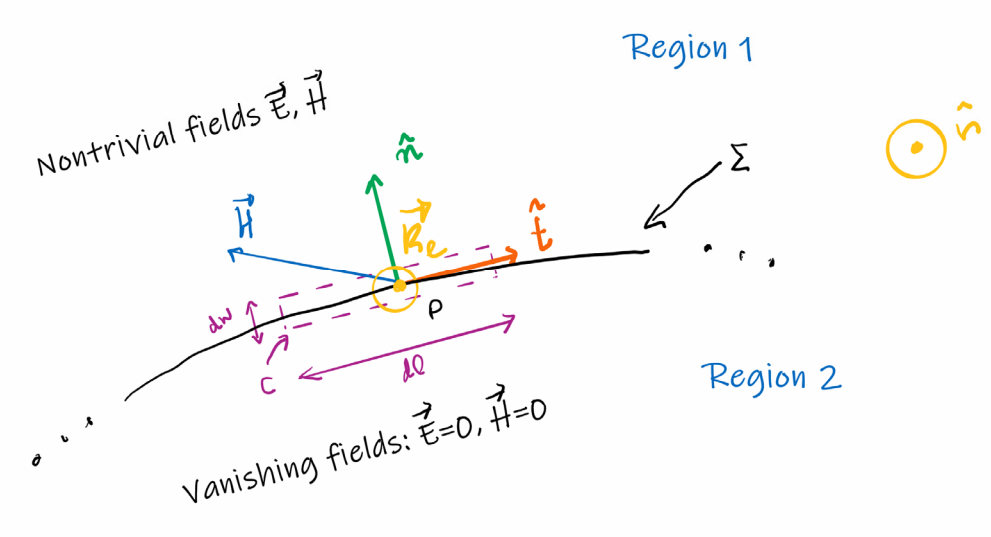
To solve the "puzzle" let us consider the frequency domain Maxwell's equations in free-space with electric and magnetic current sources:

$$\begin{aligned} \nabla \times \underline{\mathbf{E}} &= -\mathbf{j}_m - j\omega\mu_0 \underline{\mathbf{H}} \\ \nabla \times \underline{\mathbf{H}} &= \mathbf{j}_e + j\omega\varepsilon_0 \underline{\mathbf{E}} \end{aligned} \tag{10.2}$$

Here, $\mathbf{j}_m$ is a magnetic current density that describes the transport of hypothetical magnetic charges (magnetic monopoles). Even though there is no evidence in nature that magnetic monopoles do exist, for theoretical developments it is useful to admit that the current $\mathbf{j}_m$ may be nonzero. In fact, the equivalence principle is formulated in terms of equivalent *electric* and *magnetic* currents!

We want to find the surface electric and magnetic currents $\underline{\mathbf{K}}_e, \underline{\mathbf{K}}_m$ that guarantee that the radiated fields outside the volume V are unperturbed, whereas the fields inside the volume V vanish. These equivalent currents can be obtained with the same approach that is used to derive boundary conditions at material interfaces in basic electromagnetic courses. Specifically, consider the situation depicted in the next figure. The point "P" is a generic point on the surface

Σ. The vector $\hat{\mathbf{t}}$ is some arbitrary vector tangent to Σ. Consider the contour C contained in the plane generated by $\hat{\mathbf{t}}$ and $\hat{\mathbf{n}}$.



Applying Ampere's law to the contour, one finds that:

$$\oint_C \mathbf{H}\cdot\mathbf{dl} = \int ds\,\hat{\mathbf{v}}\cdot\left(\mathbf{j}_e + j\omega\varepsilon_0\underline{\mathbf{E}}\right) \approx \int ds\,\hat{\mathbf{v}}\cdot\mathbf{j}_e = \int dl\,\hat{\mathbf{v}}\cdot\underline{\mathbf{K}}_e \tag{10.3}$$

The surface integral is over the area enclosed by C and $\hat{\mathbf{v}} = \hat{\mathbf{t}}\times\hat{\mathbf{n}}$ is the vector perpendicular to the plane of the figure. In the second identity it is implicit that the contour C has infinitesimal length. In this case, the displacement current $j\omega\varepsilon_0\underline{\mathbf{E}}$ is negligible as compared to $\mathbf{j}_e$, because the electric field does not have singularities on the surface Σ. In the third identity, we used the fact that the current distribution is localized (and singular) on the surface and hence can be modelled by a surface current density ($\underline{\mathbf{K}}_e$). Supposing now that the width dw of the contour is much smaller than the length dl, one finds that:

$$\left(-\underline{\mathbf{H}}_1 + \underline{\mathbf{H}}_2\right)\cdot\hat{\mathbf{t}}\,dl = dl\;\hat{\mathbf{v}}\cdot\underline{\mathbf{K}}_e\,. \tag{10.4}$$

The electric field in the interior of volume V (region 2) must vanish: $\mathbf{H}_2 = 0$. Noting that $\hat{\mathbf{v}}\cdot\underline{\mathbf{K}}_e = \left(\hat{\mathbf{t}}\times\hat{\mathbf{n}}\right)\cdot\underline{\mathbf{K}}_e = \hat{\mathbf{t}}\cdot\left(\hat{\mathbf{n}}\times\underline{\mathbf{K}}_e\right)$, it follows that $\hat{\mathbf{t}}\cdot\left(\hat{\mathbf{n}}\times\underline{\mathbf{K}}_e + \underline{\mathbf{H}}_1\right) = 0$. But since $\hat{\mathbf{t}}$ is an arbitrary vector tangent to the surface, this means that $\hat{\mathbf{n}}\times\underline{\mathbf{K}}_e + \underline{\mathbf{H}}_{1,\tan} = 0$ where $\underline{\mathbf{H}}_{1,\tan}$ is the tangential

component of the magnetic field. Applying the $\hat{\mathbf{n}}\times$ to both sides of the equation and using $\hat{\mathbf{n}}\times\left(\hat{\mathbf{n}}\times\underline{\mathbf{K}}_e\right)=-\underline{\mathbf{K}}_e$, one finally finds that $\underline{\mathbf{K}}_e=\hat{\mathbf{n}}\times\underline{\mathbf{H}}_{1,\tan}=\hat{\mathbf{n}}\times\underline{\mathbf{H}}\big|_\Sigma$.

By manipulating the other Maxwell equation (Faraday's law) in the exact same manner one finds that $\underline{\mathbf{K}}_m=-\hat{\mathbf{n}}\times\underline{\mathbf{E}}\big|_\Sigma$.

In summary, it was proved that the equivalent electric surface current is necessarily $\underline{\mathbf{K}}_e=\hat{\mathbf{n}}\times\underline{\mathbf{H}}\big|_\Sigma$ and that the equivalent magnetic surface current is necessarily of the form $\underline{\mathbf{K}}_m=-\hat{\mathbf{n}}\times\underline{\mathbf{E}}\big|_\Sigma$ [Eq. (10.1)]. This demonstrates the equivalence principle[1].

10.3 Fields radiated by the equivalent currents

Let us now consider the problem of finding the fields radiated by the equivalent currents. For simplicity, we restrict our analysis to the far-field region, but all the ideas can be extended to the near-field region as well.

Because of the superposition principle, the fields radiated by $\underline{\mathbf{K}}_e$ and $\underline{\mathbf{K}}_m$ can be calculated separately. The fields radiated by the electric current $\underline{\mathbf{K}}_e$ can be easily found using the general formalism studied in Chapter 1.

$$\underline{\mathbf{E}}(\mathbf{r})\big|_{\text{far-field}}=\eta_0 jk_0\underline{I}(0)\mathbf{h}_e(\theta,\varphi)\frac{e^{-jk_0r}}{4\pi r}$$

with the effective length defined as $\underline{I}(0)\mathbf{h}_e(\hat{\mathbf{r}})=\left[\left(\iiint dV'\underline{\mathbf{j}}(\mathbf{r}')e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'}\right)\times\hat{\mathbf{r}}\right]\times\hat{\mathbf{r}}$. Evidently, for a surface current distribution the volume integral in $\mathbf{h}_e(\hat{\mathbf{r}})$ becomes a surface integral. Thereby, the far-field radiated by the surface electric current $\underline{\mathbf{K}}_e$ is:

$$\underline{\mathbf{E}}(\mathbf{r})\big|_{\underline{\mathbf{K}}_e}=\eta_0 jk_0\frac{e^{-jk_0r}}{4\pi r}\left[\left(\iint_\Sigma ds'\underline{\mathbf{K}}_e(\mathbf{r}')e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'}\right)\times\hat{\mathbf{r}}\right]\times\hat{\mathbf{r}}\,. \qquad (10.5a)$$

[1] Strictly speaking, it was only proved that if the problem has a solution then it is necessarily as in Eq. (10.1). A more sophisticated analysis shows that these equivalent currents really ensure that the fields vanish inside the volume V and do radiate the same fields as the original field distribution.

The corresponding magnetic field is given by $\underline{\mathbf{H}} = \frac{1}{\eta_0}\hat{\mathbf{r}} \times \underline{\mathbf{E}}$, which is the same as:

$$\underline{\mathbf{H}}(\mathbf{r})\big|_{\underline{\mathbf{K}}_e} = jk_0 \frac{e^{-jk_0 r}}{4\pi r}\left(\iint_{\Sigma} ds' \, \underline{\mathbf{K}}_e(\mathbf{r}') e^{+jk_0 \hat{\mathbf{r}}\cdot\mathbf{r}'} \right) \times \hat{\mathbf{r}} \,. \tag{10.5b}$$

To find the fields radiated by the surface magnetic current $\underline{\mathbf{K}}_m$, we use the duality symmetry of the Maxwell's equations already discussed in Chapter 8.5. Specifically, from any given solution of Eq. (10.2) (primed fields), one can construct another solution (unprimed fields) using the duality transformation:

$$\begin{array}{ccc} \underline{\mathbf{E}} & \leftrightarrow & -\underline{\mathbf{H}}' \\ \underline{\mathbf{H}} & \leftrightarrow & \underline{\mathbf{E}}' \\ \mathbf{j}_e & \leftrightarrow & -\mathbf{j}'_m \\ \mathbf{j}_m & \leftrightarrow & \mathbf{j}'_e \\ \varepsilon_0 & \leftrightarrow & \mu_0 \\ \mu_0 & \leftrightarrow & \varepsilon_0 \end{array} \tag{10.6}$$

The duality transformation interchanges the electric and magnetic fields and the electric and magnetic currents. This implies that the duality mapping transforms the fields radiated by an electric current, into the fields radiated by a magnetic current, and vice-versa. Thus, one can find the fields emitted by $\underline{\mathbf{K}}_m$ from the general formula [Eq. (10.5)] that gives the fields radiated by the electric current $\underline{\mathbf{K}}_e$. The relevant duality transformation is $\underline{\mathbf{E}}(\mathbf{r})\big|_{\underline{\mathbf{K}}_m} = -\underline{\mathbf{H}}'(\mathbf{r})\big|_{\substack{\underline{\mathbf{K}}'_e \to \underline{\mathbf{K}}_m \\ \varepsilon_0 \to \mu_0 \\ \mu_0 \to \varepsilon_0}}$. It leads to:

$$\underline{\mathbf{E}}(\mathbf{r})\big|_{\underline{\mathbf{K}}_m} = -jk_0 \frac{e^{-jk_0 r}}{4\pi r}\left(\iint_{\Sigma} ds' \, \underline{\mathbf{K}}_m(\mathbf{r}') e^{+jk_0 \hat{\mathbf{r}}\cdot\mathbf{r}'} \right) \times \hat{\mathbf{r}} \,. \tag{10.7}$$

Combining this result with Eq. (10.5a), one sees that the far-field due to the combined equivalent currents is given by:

$$\underline{\mathbf{E}}(\mathbf{r}) = jk_0 \frac{e^{-jk_0 r}}{4\pi r}\left(\iint_{\Sigma} ds' \left(-\underline{\mathbf{K}}_m(\mathbf{r}') + \eta_0 \underline{\mathbf{K}}_e(\mathbf{r}') \times \hat{\mathbf{r}} \right) e^{+jk_0 \hat{\mathbf{r}}\cdot\mathbf{r}'} \right) \times \hat{\mathbf{r}} \,. \tag{10.8}$$

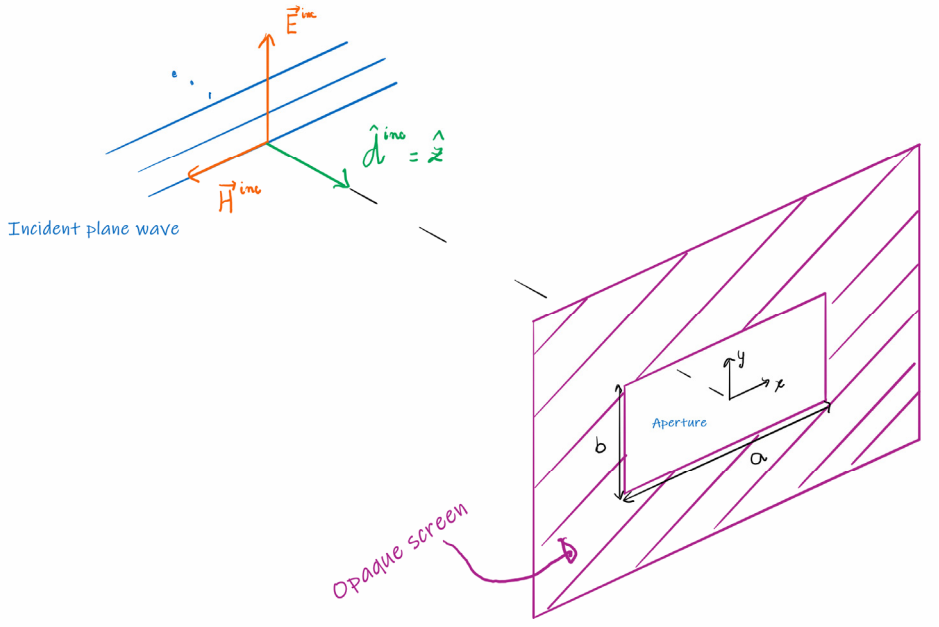
Using Eq. (10.1), the electric far-field can finally be written in terms of the aperture fields:

$$\boxed{\underline{\mathbf{E}}(\mathbf{r}) = jk_0 \frac{e^{-jk_0 r}}{4\pi r}\left(\iint_{\Sigma} ds'\left(\hat{\mathbf{n}}\times\underline{\mathbf{E}}_{\Sigma} + \eta_0\left(\hat{\mathbf{n}}\times\underline{\mathbf{H}}_{\Sigma}\right)\times\hat{\mathbf{r}}\right)e^{+jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'}\right)\times\hat{\mathbf{r}}}. \qquad (10.9)$$

The fields $\underline{\mathbf{E}}_{\Sigma}, \underline{\mathbf{H}}_{\Sigma}$ are evaluated on the surface Σ.

10.4 Uniform rectangular aperture

To illustrate the developed formalism, next we calculate the fields radiated by a rectangular aperture in an opaque screen. The opaque screen lies in the z=0 plane. One can imagine that the fields on the aperture are created by an incident plane wave propagating along the $+z$ direction, as illustrated in the figure.



If the reflections on the screen are ignored, the fields on the rectangular aperture are (the aperture is denoted by S):

$$\underline{\mathbf{E}}_S = E_0\hat{\mathbf{y}}, \qquad \underline{\mathbf{H}}_S = -\frac{E_0}{\eta_0}\hat{\mathbf{x}}, \qquad (|x| \le a/2, \quad |y| \le b/2). \qquad (10.10)$$

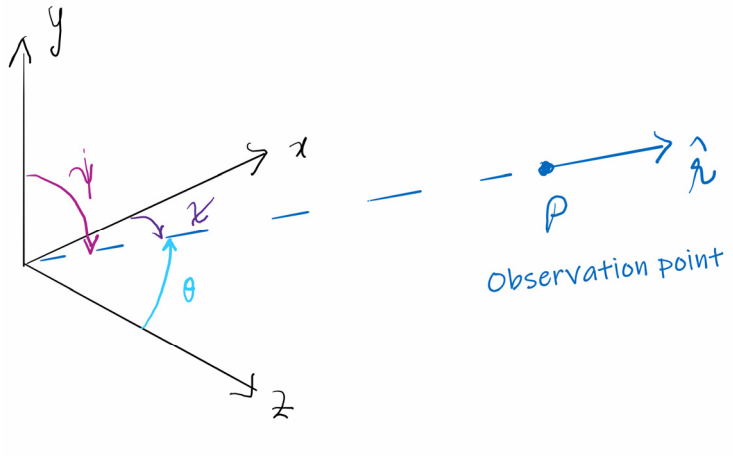
The fields on the $z = 0^+$ side of the opaque screen are assumed negligible. The fields radiated by the aperture can be found by substituting $\underline{\mathbf{E}}_S, \underline{\mathbf{H}}_S$ in Eq. (10.9) with S in the place of Σ. Using $\hat{\mathbf{n}} = +\hat{\mathbf{z}}$, a little analysis shows that the electric field is:

$$\underline{\mathbf{E}}(\mathbf{r}) = jk_0 \frac{e^{-jk_0 r}}{4\pi r} E_0 \left(-\hat{\mathbf{x}} - \hat{\mathbf{y}} \times \hat{\mathbf{r}}\right) \times \hat{\mathbf{r}} \left(\iint_S dx'dy'\, e^{+jk_0 \hat{\mathbf{r}} \cdot \mathbf{r}'} \right). \tag{10.11}$$

Writing $k_0 \hat{\mathbf{r}} \cdot \mathbf{r}' = x'k_x + y'k_y$ with $k_x = k_0 \hat{\mathbf{r}} \cdot \hat{\mathbf{x}} = k_0 \cos\chi$ and $k_y = k_0 \hat{\mathbf{r}} \cdot \hat{\mathbf{y}} = k_0 \cos\psi$, the integral can be evaluated explicitly:

$$\begin{aligned} \iint_S dx'dy'\, e^{+jk_0 \hat{\mathbf{r}} \cdot \mathbf{r}'} &= \int_{-a/2}^{a/2} dx' \int_{-b/2}^{b/2} dy'\, e^{+j\left(x'k_x + y'k_y\right)} \\ &= a \times b\, \mathrm{sinc}\left(\frac{k_x a}{2}\right) \times \mathrm{sinc}\left(\frac{k_y b}{2}\right) \end{aligned} \tag{10.12}$$

In the above, $\mathrm{sinc}(u) = \sin u / u$ and χ, ψ are the angles between the direction of observation $\hat{\mathbf{r}}$ and the x and y directions, respectively.



The field radiated by the aperture satisfies:

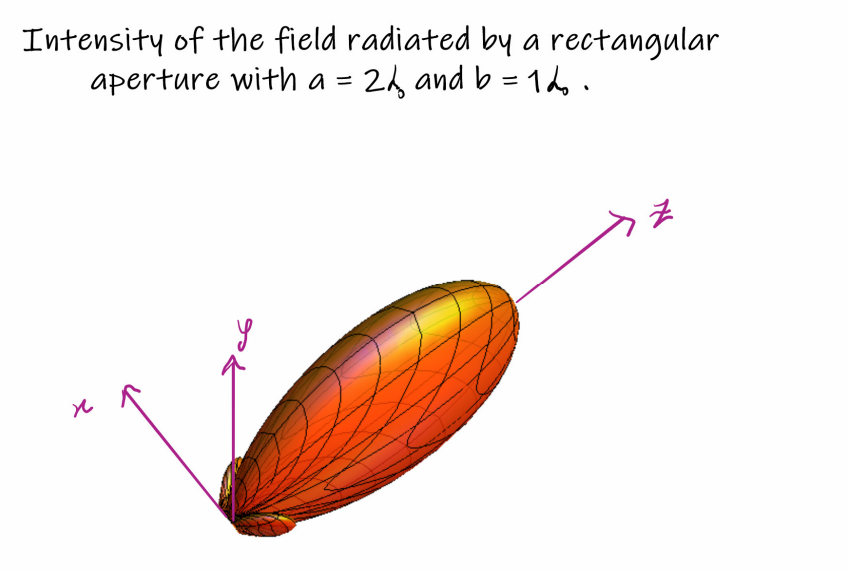
$$\underline{\mathbf{E}}\big|_{\text{far-field}} = -jk_0 E_0 \times a \times b\, \mathrm{sinc}\left(\frac{1}{2} k_0 a \cos\chi\right) \times \mathrm{sinc}\left(\frac{1}{2} k_0 b \cos\psi\right) \left[\hat{\mathbf{x}} \times \hat{\mathbf{r}} + \left(\hat{\mathbf{y}} \times \hat{\mathbf{r}}\right) \times \hat{\mathbf{r}}\right] \frac{e^{-jk_0 r}}{4\pi r}. \tag{10.13}$$

The emitted field can be written in spherical coordinates by projecting the unit vectors $\hat{\mathbf{x}}, \hat{\mathbf{y}}$ in the spherical coordinate system: $\hat{\mathbf{x}} = \alpha_1 \hat{\mathbf{r}} + \alpha_2 \hat{\boldsymbol{\theta}} + \alpha_3 \hat{\boldsymbol{\varphi}}$ and $\hat{\mathbf{y}} = \beta_1 \hat{\mathbf{r}} + \beta_2 \hat{\boldsymbol{\theta}} + \beta_3 \hat{\boldsymbol{\varphi}}$. The coefficients of the expansions are $\alpha_1 = \hat{\mathbf{x}} \cdot \hat{\mathbf{r}} = \sin\theta \cos\varphi$, etc. From here:

$$\begin{aligned} \hat{\mathbf{x}} \times \hat{\mathbf{r}} + \left(\hat{\mathbf{y}} \times \hat{\mathbf{r}}\right) \times \hat{\mathbf{r}} &= \left(\alpha_3 - \beta_2\right) \hat{\boldsymbol{\theta}} - \left(\alpha_2 + \beta_3\right) \hat{\boldsymbol{\varphi}} \\ &= -\left(1 + \cos\theta\right)\left(\sin\varphi \hat{\boldsymbol{\theta}} + \cos\varphi \hat{\boldsymbol{\varphi}}\right) \end{aligned}. \tag{10.14}$$

Since the $\hat{\boldsymbol{\theta}}$ and $\hat{\boldsymbol{\varphi}}$ components of the field are in phase, the radiated field is linearly polarized in every direction of space. The amplitude of the emitted field is:

$$|\underline{\mathbf{E}}| = k_0 |E_0| \times a \times b \left|\operatorname{sinc}\left(\frac{1}{2} k_0 a \cos\chi\right)\right| \times \left|\operatorname{sinc}\left(\frac{1}{2} k_0 b \cos\psi\right)\right| (1+\cos\theta) \frac{1}{4\pi r}. \qquad (10.15)$$



The radiated field is maximum along the z-axis: $\theta = 0^\circ$, $\chi = \psi = 90^\circ$. The maximum is:

$$|\underline{\mathbf{E}}|_{\substack{\max \\ (\theta=0^\circ)}} = k_0 |E_0| \times a \times b \frac{1}{2\pi r}. \qquad (10.16)$$

The electric field emitted to the z direction is directed along y. The half-power beamwidths in the H-plane (xoz plane) and in the E-plane (yoz plane) are approximately determined by the "sinc" functions. Using $\operatorname{sinc}(1.39) \approx 1/\sqrt{2}$ one can show that:

$$(\mathrm{HPBW})_{yoz} \approx 2 \times \arcsin \frac{1.39\lambda_0}{\pi b} \approx 0.886 \frac{\lambda_0}{b} \quad [\mathrm{rad}]. \qquad (10.17a)$$

$$(\mathrm{HPBW})_{xoz} \approx 2 \times \arcsin \frac{1.39\lambda_0}{\pi a} \approx 0.886 \frac{\lambda_0}{a} \quad [\mathrm{rad}]. \qquad (10.17b)$$

The rightmost identities assume that the aperture dimensions are large as compared to the wavelength. Clearly, the larger is the aperture the more directive is the system.

The directivity of the radiating system can be found from:

$$D = \frac{4\pi U_{\max}}{P_{rad}}\,. \tag{10.18}$$

Here, $U = \left|\underline{\mathbf{E}}\right|^2 r^2 / \left(2\eta_0\right)$ is the radiation intensity, whose maximum value is given by $U_{\max} = \frac{1}{2\eta_0} k_0^2 \left|E_0\right|^2 \left(\frac{ab}{2\pi}\right)^2$. The simplest way to determine the radiated power is by integrating the Poynting vector flux through the aperture:

$$P_{rad} = \iint\limits_{\text{aperture}} \mathbf{S}\cdot\hat{\mathbf{n}}\, ds = \frac{\left|E_0\right|^2}{2\eta_0} \times a \times b\,. \tag{10.19}$$

Combining the previous formulas, one finds an explicit formula for the directivity:

$$\boxed{D = k_0^2 \frac{ab}{\pi} = 4\pi \frac{ab}{\lambda_0^2}}. \tag{10.20}$$

Neglecting the antenna loss ($e = 1$), the maximum effective area ($A_{ef} = \frac{\lambda_0^2}{4\pi} G = \frac{\lambda_0^2}{4\pi} e \times g$) can be expressed through the directivity as

$$A_{ef,\max} = \frac{\lambda_0^2}{4\pi} D = a \times b = A_{\text{ap}}\,. \tag{10.21}$$

As seen, the effective area of the radiating system is exactly coincident with the physical area of the aperture (A_{ap}).

For a generic aperture antenna, one defines the aperture efficiency as:

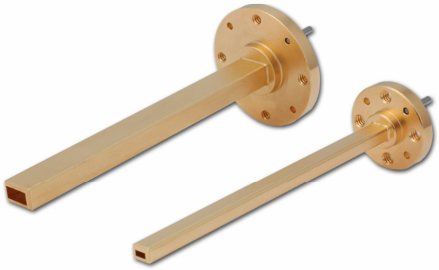
$$\mathcal{E}_{\text{ap}} = \frac{A_{ef,\max}}{A_{\text{ap}}}\,. \tag{10.22}$$

Typical values for the aperture efficiency range from 30% to 90% (e.g., for a horn antenna the aperture efficiency is on the order of 50%).

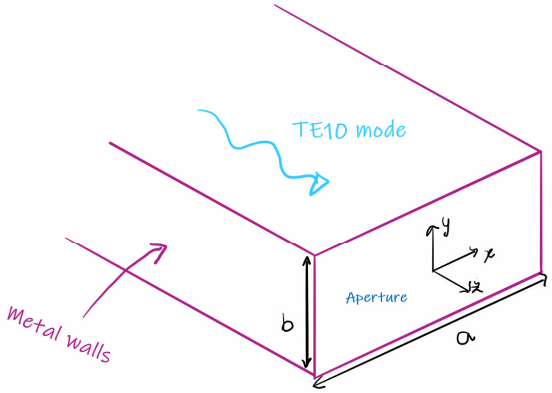
10.5 Open-ended rectangular waveguide

One can readily generalize the analysis of the previous section to the case of an open-ended metallic waveguide with rectangular cross-section.

Open-ended rectangular waveguides



Sketch of the geometry



For this system, the fields inside the guide can be assumed a superposition of two counter-propagating TE10 modes. Thus, the electric field in the guide (region $z<0$) can be taken equal to:

$$\underline{\mathbf{E}} = \cos\left(\frac{\pi x}{a}\right)\hat{\mathbf{y}}\left(\underbrace{E_0^+ e^{-j\beta z}}_{\text{incident wave}} + \underbrace{E_0^- e^{+j\beta z}}_{\text{reflected wave}}\right), \qquad \beta = \sqrt{\left(\frac{\omega}{c}\right)^2 - \left(\frac{\pi}{a}\right)^2}. \tag{10.23}$$

Note that the electric field profile is determined by a "cosine" rather than by a "sine" because the origin of the coordinate system is *centered* with respect to the guide cross-section. The propagation constant β is the calculated for the fundamental guide mode (TE10). It is assumed that the guide is filled with air.

The fields in the aperture ($z=0$) are given by (compare with Eq. (10.10)):

$$\underline{\mathbf{E}}_S \approx E_0 \cos\left(\frac{\pi x}{a}\right)\hat{\mathbf{y}}, \qquad \underline{\mathbf{H}}_S \approx -\frac{1}{Z_1} E_0 \cos\left(\frac{\pi x}{a}\right)\hat{\mathbf{x}}, \qquad (|x| \le a/2, \quad |y| \le b/2) \tag{10.24}$$

In the above, $E_0 = E_0^+ + E_0^-$ is the total electric field on the aperture and $H_0 = E_0 / Z_1$ is the total magnetic field. Here, Z_1 is the impedance calculated at the output plane, which depends on the

wave impedance of the TE10 mode $\eta_{\mathrm{TE10}} = \omega\mu_0 / \beta$ and on the reflection coefficient $\Gamma = E_0^- / E_0^+$ in the usual way $Z_1 = \eta_{\mathrm{TE10}} \dfrac{1+\Gamma}{1-\Gamma}$.

The field radiated by the open-ended guide can be found by substituting the aperture fields in Eq. (10.9). The calculations are rather similar to ones in the previous section. Now, we find that (compare with Eq. (10.11)):

$$\underline{\mathbf{E}} = -jk_0 \frac{e^{-jk_0 r}}{4\pi r} E_0 \left[\left(\hat{\mathbf{x}} + \frac{\eta_0}{Z_1} \hat{\mathbf{y}} \times \hat{\mathbf{r}} \right) \times \hat{\mathbf{r}} \right] \left(\iint_S dx'dy' \cos\left(\frac{\pi x}{a} \right) e^{+jk_0 \hat{\mathbf{r}} \cdot \mathbf{r}'} \right). \tag{10.25}$$

Using the integration formula

$$\begin{aligned} \int_{-a/2}^{a/2} dx' \cos\left(\frac{\pi x}{a} \right) e^{+jk_x x'} &= \frac{1}{2} a \left[\operatorname{sinc}\left(\left(k_x + \frac{\pi}{a} \right) \frac{a}{2} \right) + \operatorname{sinc}\left(\left(k_x - \frac{\pi}{a} \right) \frac{a}{2} \right) \right] \\ &= \frac{1}{2} a \cos\left(\frac{k_x a}{2} \right) \left[\frac{1}{\frac{k_x a}{2} + \frac{\pi}{2}} - \frac{1}{\frac{k_x a}{2} - \frac{\pi}{2}} \right] \\ &= \frac{2a}{\pi} \cos\left(\frac{k_x a}{2} \right) \frac{1}{1 - \left(\frac{k_x a}{\pi} \right)^2} \end{aligned} \tag{10.26}$$

and $k_0 \hat{\mathbf{r}} \cdot \mathbf{r}' = x'k_x + y'k_y = x'k_0 \cos\chi + y'k_0 \cos\psi$, one can show that the radiated field is:

$$\underline{\mathbf{E}} = -jk_0 E_0 a \times b \left[\left(\hat{\mathbf{x}} + \frac{\eta_0}{Z_1} \hat{\mathbf{y}} \times \hat{\mathbf{r}} \right) \times \hat{\mathbf{r}} \right] \left[\frac{\frac{2}{\pi} \cos\left(\frac{k_0}{2} a \cos\chi \right)}{1 - \left(\frac{k_0 a \cos\chi}{\pi} \right)^2} \operatorname{sinc}\left(\frac{k_0 b}{2} \cos\psi \right) \right] \frac{e^{-jk_0 r}}{4\pi r}. \tag{10.27}$$

The radiation diagram of the open-ended guide is qualitatively analogous to that of uniform rectangular aperture. For example, the half-power beamwidth in the E-plane is exactly the same for the two geometries. The directivity of the radiating system can be found as in the previous section. A detailed analysis shows that:

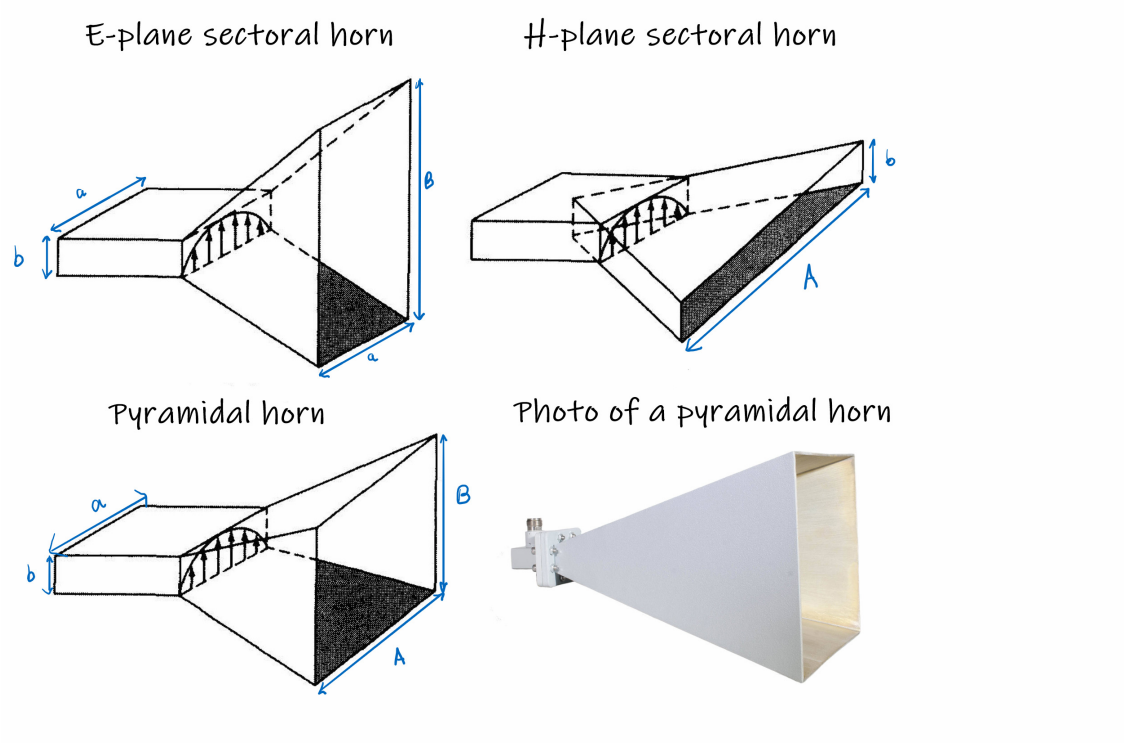
$$\boxed{D = \frac{8}{\pi} \frac{Z_1}{\eta_0} \frac{a \times b}{\lambda_0^2} \left(1 + \frac{\eta_0}{Z_1} \right)^2}. \tag{10.28}$$

When the guide is matched to free-space ($Z_1 = \eta_0$), the directivity satisfies $D = \frac{32}{\pi}\frac{a \times b}{\lambda_0^2}$ and is slightly lower (by a factor of $2\left(\frac{2}{\pi}\right)^2 = 0.81$) than the directivity of a uniform rectangular aperture. Thus, in theory it is possible to have a very directional emission. However, this radiating system has a number of problems: i) the abrupt transition from the rectangular waveguide to the free-space region typically leads to strong reflections and to a poorly matched system. ii) to obtain a large directivity one needs an electrically large aperture (one of the waveguide dimensions must be much larger than the wavelength). A large aperture implies the excitation of higher order modes in the waveguide. In other words, it is impossible to ensure the single mode operation of the rectangular guide. In the next section, we will see how to overcome these difficulties using a horn antenna.

10.6 Horn antenna

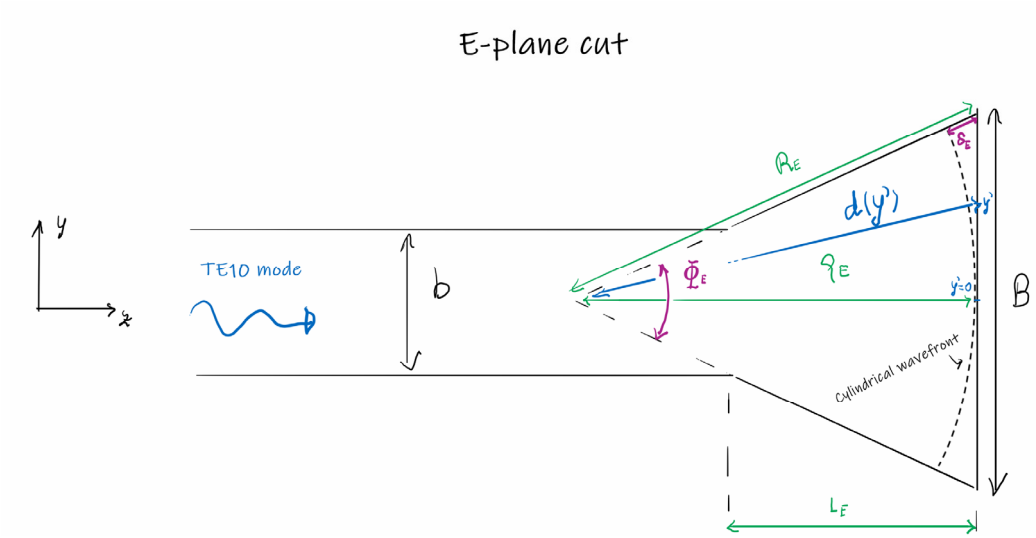
The matching of the open-ended guide with the free-space region can be much improved by flaring (tapering) the walls of the guide, so that they gradually become wider. This allows for a gradual transition from the guided wave to the free-space wave, with the result that the reflections are minimized. Furthermore, for a gradual transition the higher order modes generated at the "throat" of the structure cannot reach the waveguide region, as these modes are in cut-off in the waveguide region. The flaring of the walls results in a larger aperture, and thereby in a higher directivity and narrower beam as compared to an open-ended guide. A rectangular guide with flared (tapered) walls is known as a "horn antenna".

There are different types of horn antennas. When only the E-plane or the H-plane walls are flared the antenna is known as an E-sectoral or H-sectoral horn, respectively. When both walls are flared the antenna is a pyramidal horn antenna.



Horn antennas are directive radiators primarily used in the microwave range (~ 1GHz). They have a relatively high gain (10-100) and wide bandwidth ($f_U : f_L \sim 10:1$) and are simple to construct. The beamwidth in the E and H planes is controlled by the dimensions B and A, respectively. For sectoral horns, the beamwidth can be controlled in a single plane; the beamwidth in the other plane it is the same as for an open-ended guide. Thus, sectoral horns are characterized by fan-shaped beams. For pyramidal horns, the beamwidth can be controlled independently in the two principal planes. For simplicity, in the following we restrict our analysis to E-plane sectoral horns.

The fields on the aperture of the E-sectoral horn can be found under the approximation that the wave in the throat of the horn is a cylindrical wave. The center of curvature of the wave is at a distance ρ_E from the aperture, as illustrated in the figure below. The curvature center depends on the tapering angle Φ_E of the E-plane walls.



One can obtain the following geometrical relations from the figure:

$$\tan\frac{\Phi_E}{2}=\frac{(B-b)/2}{L_E}=\frac{B/2}{\rho_E}. \tag{10.29}$$

The fields at the aperture can be obtained from the fields in the waveguide, introducing a phase correction due to the cylindrical wavefront (there is also an amplitude correction due to the divergence of the cylindrical wave, but it is irrelevant for the characterization of the directional properties of the antenna). Supposing that the antenna is well-matched to free-space (there is no reflected wave in the antenna throat) one can write (compare with Eq. (10.24)):

$$\underline{\mathbf{E}}_S \approx E_0\cos\left(\frac{\pi x}{a}\right)e^{-j\phi}\hat{\mathbf{y}}, \qquad \underline{\mathbf{H}}_S \approx -\frac{1}{\eta_0}E_0\cos\left(\frac{\pi x}{a}\right)e^{-j\phi}\hat{\mathbf{x}}, \qquad (|x|\le a/2, \quad |y|\le B/2) \tag{10.30}$$

The impedance is taken equal to η_0 (rather than the wave impedance of the guide) because the antenna is matched to free-space. Here, $\phi=k_0 d$ is the phase acquired by the wave when it travels a distance d from the curvature center to the aperture. The phase ϕ is not a constant because $d=d(y')$ (see the figure). One can write $d=\sqrt{\rho_E^2+y'^2}\approx\rho_E+\frac{1}{2}\frac{y'^2}{\rho_E}$ when $B/2<<\rho_E$. With such an approximation the propagation phase becomes $\phi=\phi_0+k_0\frac{1}{2}\frac{y'^2}{\rho_E}$ with

$\phi_0 = k_0\rho_E$ and irrelevant phase term that will be dropped. Thus, the curvature of the cylindrical wavefront introduces a quadratic phase distortion ($\delta\phi = k_0 \frac{1}{2}\frac{y'^2}{\rho_E}$) on the aperture fields:

$$\underline{\mathbf{E}}_S \approx E_0 \cos\left(\frac{\pi x}{a}\right) e^{-jk_0\frac{1}{2}\frac{y'^2}{\rho_E}}\hat{\mathbf{y}}, \qquad \underline{\mathbf{H}}_S \approx -\frac{1}{\eta_0} E_0 \cos\left(\frac{\pi x}{a}\right) e^{-jk_0\frac{1}{2}\frac{y'^2}{\rho_E}}\hat{\mathbf{x}} \tag{10.31}$$

By substituting these formulas into Eq. (10.9), one can find the fields radiated by the horn antenna. It is simple to check that one obtains the same result as in Eq. (10.24) with the replacements: $Z_1 \to \eta_0$, $b \to B$, and $\text{sinc}\left(\frac{k_0 b}{2}\cos\psi\right) \to I_D$, so that:

$$\underline{\mathbf{E}} = -jk_0 E_0 a \times B\left[\left(\hat{\mathbf{x}} + \hat{\mathbf{y}} \times \hat{\mathbf{r}}\right) \times \hat{\mathbf{r}}\right] \left[\frac{\frac{2}{\pi}\cos\left(\frac{k_0}{2} a \cos\chi\right)}{1 - \left(\frac{k_0 a \cos\chi}{\pi}\right)^2} I_D\right] \frac{e^{-jk_0 r}}{4\pi r}. \tag{10.32}$$

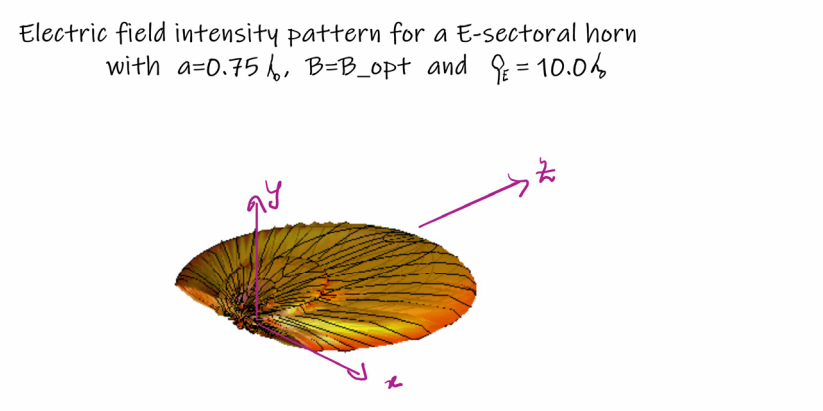
where I_D is the diffraction integral:

$$I_D = \frac{1}{B}\int_{-B/2}^{B/2} dy' e^{-jk_0\frac{1}{2}\frac{y'^2}{\rho_E}} e^{+jk_y y'}, \qquad k_y = k_0 \cos\psi\,. \tag{10.33}$$

The diffraction integral can be expressed in terms of special functions known as the “Fresnel integrals”.

A larger *B* implies a larger aperture, which in general favours a more directional beam. However, a larger *B* also implies a larger quadratic phase error $\delta\phi$. The phase error describes the departure from the uniform phase (plane wave) aperture. Large quadratic phase errors cause the divergence of the radiated wave. Due to this reason there is an *optimal value* for *B*. A detailed analysis shows that the optimal *B* that maximizes the directivity is such that:

$$\boxed{B_{\text{opt}} = \sqrt{2\lambda_0\rho_E}}. \tag{10.34}$$



It can be shown that the directivity of the E-plane sectoral horn is given by:

$$\boxed{D_E = \frac{4\pi}{\lambda_0^2} \times a \times B \times \mathcal{E}_{\text{ap}}^E}. \tag{10.35}$$

where $\mathcal{E}_{\text{ap}}^E$ is the aperture efficiency. It can be estimated using the formula:

$$\mathcal{E}_{\text{ap}}^E = \frac{8}{\pi^2}\mathcal{E}_{\text{ph}}^E, \qquad \mathcal{E}_{\text{ph}}^E = 1.0033 - 0.119s - 2.752s^2, \qquad 0 < s < 0.262, \tag{10.36}$$

where the parameter s is defined by:

$$\boxed{s = \frac{B^2}{8\lambda_0 \rho_E}}. \tag{10.37}$$

For a optimal E-sectoral horn antenna, $s = s_{\text{opt}} = 1/4$ and $\mathcal{E}_{\text{ap,opt}}^E = 0.649$.

The half-power beamwidth of the antenna in the E-plane can be found using the formula:

$$\boxed{\text{HPBW}_{\text{E-plane}} = 54^\circ \times \frac{\lambda_0}{B}}. \tag{10.38}$$