TECHNICAL UNIVERSITY GHEORGHE ASACHI OF IAȘI

FACULTY OF ELECTRONICS, TELECOMMUNICATIONS, AND INFORMATION TECHNOLOGY

Prof. Ion BOGDAN

Lecture notes on

ANTENNAS AND PROPAGATION

2017

C O N T E N T

C h a p t e r I

FUNDAMENTALS OF ELECTROMAGNETISM THEORY

1.1 - A Short History of Antennas' Evolution

The IEEE standard no. 145-1983 (Standard Definitions of Terms for Antennas) defines antenna as "*a means of transmitting or receiving of radio waves*", that is antenna is a block of a radio communication equipment enabling its electromagnetic power exchange with the environment (free air, usually, but it could be also outer space vacuum – for satellite antennas, plasma – for special applications in physics, organic tissue – for biological applications etc.). An antenna may be seen as a matching element between the environment and the receiver or transmitter that converses an electromagnetic power into an electrical one (when receiving) or an electrical power into an electromagnetic one (when transmitting).

The foundation of antennas' theory was built by James Clark Maxwell which unified the previous separated theories of electricity and magnetism and derived the equations that govern radio wave propagation. Based on physical considerations, J. C. Maxwell proved mathematically in his seminal work *A Treatise on Electricity and Magnetism*, published in 1873, that the electromagnetic power propagates by means of waves and that travelling velocity in vacuum (and free air) equals the velocity of light.

Heinrich Rudolph Hertz proved experimentally that electromagnetic waves really exist by sensing with a small loop electric sparks between the branches of a dipole. Guglielmo Marconi proved experimentally in 1901 that the electromagnetic waves could carry an information at large distance. The experiment, known as *the first transmission by means of electromagnetic waves* in communication history, was realized between Poldhu (UK) and Newfoundland (Canada). 50 vertical wires electrically connected at both of their ends (like a cage) materialized the transmitting antenna, while a 200 meter length vertical wire elevated by a kite represented the receiving one.

This wire shape of antennas was largely used until 1940, although the operation frequencies became higher and higher reaching the 1 GHz limit. But, once the klystron and the magnetron were invented, frequencies above 1 GHz came into use, smaller size of antennas became appropriate and more complex shapes (wave guide antenna, horn antenna, reflector antenna etc.) were used. Powerful computers built after World War II allowed the design of even more complex antennas, adapted to sophisticated applications. Also, new, more efficient, analytical and numerical techniques were developed, allowing for a detailed analysis and a precise design of antennas: method of moments (MoM), finite difference method (FDM), finite element method (FEM), optical theory of electromagnetic wave propagation or geometrical optics (GO), geometrical theory of diffraction (GTD), uniform theory of diffraction (UTD), finite difference time domain theory (FDTD) etc.

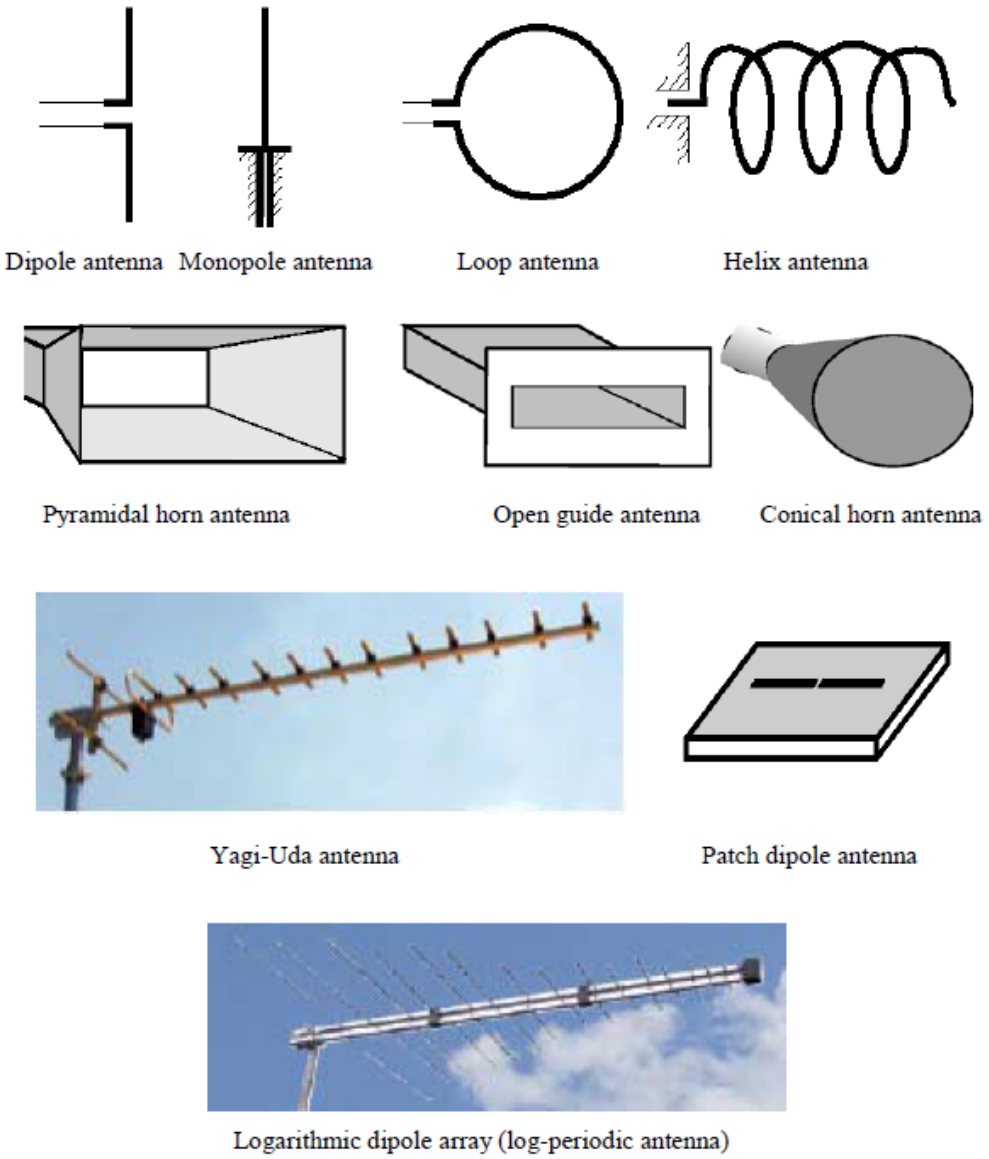


Figure no. 1.1 *– Typical antennas*

Types of Antennas

Most of the antennas built before World War II were cylindrical conductors with a very small cross sectional size (thin wire, dipole etc.) or regular geometrical shapes (helix, rhomb etc.). Fast radar technique advancements during the war and its sophisticated applications asked for new and more efficient antennas to be developed: open guide antenna, slot antenna, lens antenna, horn antenna, reflector antenna etc. In 1950 the so called frequency independent antennas were developed (equiangular antenna, spiral antenna, log-periodic antenna). They had a ratio of the maximum to the minimum operating frequencies of about 40:1, much more than the maximum 2:1 value obtained until then. Patch antennas were implemented beginning with 1970. They are lightweight, flexible, and cheap, and they are mainly used in spatial applications. Technological advancements allow nowadays the embedding of patch antennas into the integrated circuits they have to work with (monolithic antennas).

High design requirements, especially very small main lobe beam width, could not be met by using one single antenna, because large size structures are needed. Grouping more small size antennas (antenna arrays) is a better solution, because the associated mechanical problems like weight, fixture, and wind resistance are easier to solve. A new difficulty

appears: building an efficient feeding network for the antennas in the array. Technological advances and complex software design tools allowed for implementation of very efficient feeding solutions. Antenna arrays allow for the introduction of new control features like independently setting the directions for the main lobe and for the nulls the radiation pattern, respectively, or electrically modifying the main lobe direction (very useful for radar applications). Antenna arrays associated with powerful signal processors implement usually the modern concept of smart antenna that permanently modifies its radiation pattern in order to match the electromagnetic environment variable properties and to meet a preset objective function. Smart antennas are used by the most sophisticated modern communication systems: mobile cellular systems of the third (UMTS) and the fourth (LTE) generations, software defined radio, cognitive radio, spectrum sensing etc.

1.2 - Maxwell's Equations

Maxwell's equations describe the evolution in time and space of the characteristic variables of an electromagnetic wave. They constitute a complete system; this means that any other relation in the electromagnetism's theory can be derived starting from this group of four equations. In their differential form (there is also an integral form), Maxwell's equations are as follows:

$$\begin{aligned} \nabla \times \boldsymbol{E} &= -\frac{\partial \boldsymbol{B}}{\partial t} \\ \nabla \times \boldsymbol{H} &= \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J} \\ \nabla \cdot \boldsymbol{D} &= \boldsymbol{\rho} \\ \nabla \cdot \boldsymbol{B} &= \boldsymbol{0} \end{aligned} \qquad (1.1)$$

where: $\boldsymbol{E}$ – electric field intensity (in *Volts / meter*),
$\boldsymbol{H}$ – magnetic field intensity (in *Amperes / meter* ≡ *Henry*),
$\boldsymbol{D}$ – electric displacement field (in *Coulombs / meter*^{2}),
$\boldsymbol{B}$ – magnetic field (in *Amperes / meter*^{3} ≡ *Tesla*),
$\boldsymbol{J}$ – current density (in *Amperes / meter*^{2}),
ρ – electric charge density (in *Coulombs / meter*^{3}).

All of the above variables are functions of point ($\boldsymbol{r}$ – position vector) and time (t). They are also related through the so called *laws of material* that describe the influence of the propagation medium upon the electromagnetic waves:

$$\begin{aligned} \boldsymbol{D} &= \varepsilon \boldsymbol{E} + \boldsymbol{P} \\ \boldsymbol{B} &= \mu(\boldsymbol{H} + \boldsymbol{M}) \end{aligned} \qquad (1.1)$$

where ε and μ are the medium's electric permittivity and the magnetic permeability, respectively, $\boldsymbol{P}$ – its electrical Polarization, and $\boldsymbol{M}$ – its Magnetization. Usually ε and μ are complex variables dependent on point ($\boldsymbol{r}$), while $\boldsymbol{P}$ and $\boldsymbol{M}$ are vectorial functions of point and time ($\boldsymbol{r}$, t).

In the above relations and in the followings we use the standard notations, i.e.:

- vectorial variables and functions – **bold** letters;
- scalar variables and functions – *italic* letters;
- $\nabla \times \boldsymbol{a}$ – curl of a vector $\boldsymbol{a}$;
- $\nabla \cdot \boldsymbol{a}$ – divergence of a vector $\boldsymbol{a}$;
- ∇f – gradient of a scalar function f;
- ∇ – operator "nabla"; it has a double meaning as a *vector* and as a *derivative*

operator of its argument along with spatial coordinates; for instance, in Cartesian coordinates its expression is

$$\nabla \stackrel{\text{def}}{=} \frac{\partial}{\partial x}\hat{\boldsymbol{x}} + \frac{\partial}{\partial y}\hat{\boldsymbol{y}} + \frac{\partial}{\partial z}\hat{\boldsymbol{z}}$$

- $\hat{\boldsymbol{a}}$ – versor (unitary norm vector);
- $\boldsymbol{a} \cdot \boldsymbol{b}$ – scalar product of vectors **a** and **b**;
- $\boldsymbol{a} \times \boldsymbol{b}$ – vectorial product of vectors **a** and **b**;

Knowing that:

$$\nabla \cdot (\nabla \times \boldsymbol{a}) \equiv \boldsymbol{0}, \ \ \forall \boldsymbol{a}\,, \tag{1.3}$$

For mediums with non-time variable electrical parameters and without polarization one gets from the 2nd and the 3rd equation in (1.1) that:

$$0 \equiv \nabla \cdot (\nabla \times \boldsymbol{H}) = \nabla \cdot \left(\frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J}\right) = \frac{\partial (\nabla \cdot \boldsymbol{D})}{\partial t} + \nabla \cdot \boldsymbol{J} = \frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{J} \tag{1.4}$$

or:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot \boldsymbol{J} \equiv 0 \tag{1.5}$$

which is the *continuity equation* or the *conservation law of the electrical charge*. A conservation law is a fundamental one in any domain of science. The above result showing that the conservation law of the electrical charge is a direct consequence of the Maxwell's equations is a strong argument for the statement that the Maxwell's equations are a complete system.

Maxwell's equations are asymmetrical due to the absence of physical magnetic source. They become symmetrical only in mediums where electric charge and electric currents are absent. But radiating and receiving radio waves are only possible if electric charge and electric currents are present in antennas. On the other hand, symmetry is very useful in finding solutions for complex differential equations. This is why Maxwell's equations are artificially symmetrized by postulating the existence of a non-zero magnetic charge with the spatial density ρ_m and, consequently, of a non-zero magnetic current with surface density $\boldsymbol{J_m}$.

The symmetrized Maxwell's equations are:

$$\begin{aligned} \nabla \times \boldsymbol{E} &= -\frac{\partial \boldsymbol{B}}{\partial t} - \boldsymbol{J_m} \\ \nabla \times \boldsymbol{H} &= \frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J} \\ \nabla \cdot \boldsymbol{D} &= \rho \\ \nabla \cdot \boldsymbol{B} &= \rho_m \end{aligned} \tag{1.6}$$

In mediums without magnetization and with magnetic variables independent of time the first equation and the 4th one in the group (1.6) yield:

$$\frac{\partial \rho_m}{\partial t} + \nabla \cdot \boldsymbol{J_m} \equiv 0 \tag{1.7}$$

which is the *conservation law* of the magnetic charge.

When the variables in the Maxwell's equations have a harmonic evolution in time one can use their representation in the Fourier transform domain, which is the following:

$$\begin{aligned} \nabla \times \boldsymbol{E} &= -j\omega \boldsymbol{B} - \boldsymbol{J_m} \\ \nabla \times \boldsymbol{H} &= j\omega \boldsymbol{D} + \boldsymbol{J} \end{aligned} \tag{1.8}$$

$$\nabla \cdot \boldsymbol{D} = \rho$$

$$\nabla \cdot \boldsymbol{B} = \rho_m$$

where $j = \sqrt{-1}$ is the complex operator and $\omega = 2\pi f$ is the angular frequency of the electromagnetic wave.

1.3 - Energy and Power

The magnitude and the direction of the power transmitted by an electromagnetic wave are given by the Poynting vector, that is:

$$\boldsymbol{S} = \boldsymbol{E} \times \boldsymbol{H}^* \tag{1.9}$$

where $(^*)$ means complex conjugation.

In order to evaluate the energy entering in a volume *V* inside a closed surface *Σ*, one has to integrate the Poynting vector along this surface:

$$\tfrac{1}{2}\int_{\Sigma} \boldsymbol{S} \cdot d\boldsymbol{\sigma} = \tfrac{1}{2}\int_{\Sigma} (\boldsymbol{E} \times \boldsymbol{H}^*) \cdot d\boldsymbol{\sigma} \tag{1.10}$$

where vector $d\boldsymbol{\sigma}$ is the area infinitesimal element having the positive orientation towards the interior of the volume *V*. The coefficient ½ is used because relation (1.10) equates rms values of power, while $\boldsymbol{S}$, $\boldsymbol{E}$, and $\boldsymbol{H}$ are complex amplitudes of the respective variables and they are supposed to be harmonic.

Taking into account the divergence theorem:

$$\int_V (\nabla \cdot \boldsymbol{a})\, dV = -\int_{\Sigma} \boldsymbol{a} \cdot d\boldsymbol{\sigma}, \tag{1.11}$$

the mathematical identity:

$$\nabla \cdot (\boldsymbol{a} \times \boldsymbol{b}) \equiv \boldsymbol{b} \cdot (\nabla \times \boldsymbol{a}) - \boldsymbol{a} \cdot (\nabla \times \boldsymbol{b}), \tag{1.12}$$

and the Maxwell's equations, the left hand side of the equation (1.10) becomes:

$$\begin{aligned} \tfrac{1}{2}\int_{\Sigma} \boldsymbol{S} \cdot d\boldsymbol{\sigma} &= -\tfrac{1}{2}\int_V (\nabla \cdot \boldsymbol{S}) dV = -\tfrac{1}{2}\int_V [\nabla \cdot (\boldsymbol{E} \times \boldsymbol{H}^*)] dV = \\ &= -\tfrac{1}{2}\int_V [\boldsymbol{H}^* \cdot (\nabla \times \boldsymbol{E}) - \boldsymbol{E} \cdot (\nabla \times \boldsymbol{H}^*)] dV = \\ &= \tfrac{1}{2}\int_V j\omega(\boldsymbol{E} \cdot \varepsilon^* \boldsymbol{E}^* + \boldsymbol{H}^* \cdot \mu \boldsymbol{H}) dV + \tfrac{1}{2}\int_V \boldsymbol{E} \cdot \boldsymbol{J}^* dV\,. \end{aligned} \tag{1.13}$$

That is:

$$\tfrac{1}{2}\int_{\Sigma} \boldsymbol{S} \cdot d\boldsymbol{\sigma} = 2j\omega \int_V \tfrac{1}{4}(\boldsymbol{E} \cdot \varepsilon^* \boldsymbol{E}^* + \boldsymbol{H}^* \cdot \mu \boldsymbol{H}) dV + \tfrac{1}{2}\int_V \boldsymbol{E} \cdot \boldsymbol{J}^* dV \tag{1.14}$$

The above equation containing complex variables is equivalent to a set of two equations containing real variables:

$$\begin{aligned} Re\left(\tfrac{1}{2}\int_{\Sigma} \boldsymbol{S} \cdot d\boldsymbol{\sigma}\right) &= Im\left[2\omega \int_V \tfrac{1}{4}(\boldsymbol{E} \cdot \varepsilon^* \boldsymbol{E}^* + \boldsymbol{H}^* \cdot \mu \boldsymbol{H}) dV\right] + \\ &+ Re\left(\tfrac{1}{2}\int_V \boldsymbol{E} \cdot \boldsymbol{J}^* dV\right) \end{aligned} \tag{1.15}$$

meaning that *the active power entering in a volume V equals the sum between the power dissipated in the conductors in V and the power needed to polarize and to magnetize the medium in V*

and

$$Im\left(\tfrac{1}{2}\int_{\Sigma} \boldsymbol{S} \cdot d\boldsymbol{\sigma}\right) = Re\left[2\omega \int_V \tfrac{1}{4}(\boldsymbol{E} \cdot \varepsilon^* \boldsymbol{E}^* + \boldsymbol{H}^* \cdot \mu \boldsymbol{H}) dV\right] \tag{1.16}$$

meaning that *the flux of the reactive power of the Poynting vector through the surface Σ equals the sum of powers stored in the electric field and in the magnetic field in V times 2ω.*

1.4 – Wave Equation

Among the variables that describe an electromagnetic filed, $\boldsymbol{E}$ and $\boldsymbol{H}$ are considered as primary variables, while $\boldsymbol{B}$ and $\boldsymbol{D}$ are as secondary ones. This is why, finding a solution for the electromagnetic wave radiated by an antenna means to find out the formulas for the variables $\boldsymbol{E}$ and $\boldsymbol{H}$. These formulas are solutions of a specific differential equation denoted as the *wave equation*.

Applying the curl operator to the first Maxwell's equation one gets:

$$\nabla \times (\nabla \times \boldsymbol{E}) = \nabla \times \left(-\frac{\partial \boldsymbol{B}}{\partial t}\right) \tag{1.17}$$

In mediums with no polarization and no magnetization the right hand side of the above equation becomes as follows:

$$\nabla \times \left(-\frac{\partial \boldsymbol{B}}{\partial t}\right) = -\frac{\partial}{\partial t}(\nabla \times \boldsymbol{B}) = -\mu \frac{\partial}{\partial t}(\nabla \times \boldsymbol{H}) =$$

$$= -\mu \frac{\partial}{\partial t}\left(\frac{\partial \boldsymbol{D}}{\partial t} + \boldsymbol{J}\right) = -\mu\varepsilon \frac{\partial^2 \boldsymbol{E}}{\partial t^2} - \mu \frac{\partial \boldsymbol{J}}{\partial t} \tag{1.18}$$

Taking into account the mathematical identity:

$$\nabla \times (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{a}(\nabla \cdot \boldsymbol{b}) - \boldsymbol{b}(\nabla \cdot \boldsymbol{a}) + (\boldsymbol{b} \cdot \nabla)\boldsymbol{a} - (\boldsymbol{a} \cdot \nabla)\boldsymbol{b} \tag{1.19}$$

the left hand side of the equation (1.17) becomes:

$$\nabla \times (\nabla \times \boldsymbol{E}) = \nabla(\nabla \cdot \boldsymbol{E}) - \boldsymbol{E}(\nabla \cdot \nabla) + (\boldsymbol{E} \cdot \nabla)\nabla - (\nabla \cdot \nabla)\boldsymbol{E} \tag{1.20}$$

The 2nd and the 3rd terms in the above relation are not meaningful mathematically, because the vector ∇, which is also a derivative operator, has no argument, and have to be eliminated. The scalar product $\nabla \cdot \nabla$ is the Laplace operator ∇^2 and $\nabla \cdot \boldsymbol{E} = \frac{1}{\varepsilon}(\nabla \cdot \boldsymbol{D}) = \frac{\rho}{\varepsilon}$. So that, the above relation becomes:

$$\nabla \times (\nabla \times \boldsymbol{E}) = \frac{\nabla \rho}{\varepsilon} - \nabla^2 \boldsymbol{E} \tag{1.21}$$

Finally, from (1.17), (1.18), and (1.21) one gets:

$$\nabla^2 \boldsymbol{E} - \mu\varepsilon \frac{\partial^2 \boldsymbol{E}}{\partial t^2} = \mu \frac{\partial \boldsymbol{J}}{\partial t} + \frac{\nabla \rho}{\varepsilon} \tag{1.22}$$

Similarly, one obtains:

$$\nabla^2 \boldsymbol{H} - \mu\varepsilon \frac{\partial^2 \boldsymbol{H}}{\partial t^2} = -\nabla \times \boldsymbol{J} \tag{1.23}$$

Mathematically, a differential equation of the above type is known as a *Helmholtz equation* and it represents the *wave equation* in the electro-magnetic theory.

When the variables $\boldsymbol{E}$ and $\boldsymbol{H}$ are harmonic the wave equations could be written in the Fourier transform domain and they become:

$$\nabla^2 \boldsymbol{E} + k^2 \boldsymbol{E} = j\omega\mu \boldsymbol{J} + \frac{\nabla \rho}{\varepsilon} \tag{1.24}$$

$$\nabla^2 \boldsymbol{H} + k^2 \boldsymbol{H} = -\nabla \times \boldsymbol{J} \tag{1.25}$$

where $k \stackrel{\text{def}}{=} \omega\sqrt{\mu\varepsilon}$ and is denoted as the wave number or the propagation constant.

For mediums with no charge and no current the solutions of the homogeneous wave equations are:

$$\boldsymbol{E} = \boldsymbol{E}_0 e^{\pm j\boldsymbol{k}\cdot\boldsymbol{r}} \qquad \boldsymbol{H} = \boldsymbol{H}_0 e^{\pm j\boldsymbol{k}\cdot\boldsymbol{r}} \tag{1.26}$$

where $\boldsymbol{E}_0$ and $\boldsymbol{H}_0$ are integration constants, $\boldsymbol{k}$ is a vector with modulus *k*, denoted as *propagation vector*, and $\boldsymbol{r}$ is the position vector.

The solutions having the plus sign at the exponent are not physically acceptable as they do not fulfill the restrictions imposed by the *Sommerfeld conditions*. These conditions impose to a solution ψ of a wave equation the following two requirements:

$$|r\psi| < K$$
$$lim_{r\to\infty}\left[r\left(\frac{\partial\psi}{\partial r}+jK\psi\right)\right]=0 \tag{1.27}$$

where *K* is an arbitrary constant.

The first relation requires that the modulus of the function ψ at least as fast as distance *r* from the radiation source powered at -1. The second relation requires that function ψ travels away from the radiation source.

In conclusion, the only physically acceptable solutions for the wave equation are:

$$\boldsymbol{E}=\boldsymbol{E}_0 e^{-j\boldsymbol{k}\cdot\boldsymbol{r}} \qquad \boldsymbol{H}=\boldsymbol{H}_0 e^{-j\boldsymbol{k}\cdot\boldsymbol{r}} \tag{1.28}$$

Note that these solutions have only one independent variable (*r*) from the ones in the spherical coordinate system (r,θ,ϕ) and thus:

$$\nabla e^{-j\boldsymbol{k}\cdot\boldsymbol{r}} = -j\boldsymbol{k}e^{-j\boldsymbol{k}\cdot\boldsymbol{r}} \tag{1.29}$$

From the last two equations in the Maxwell's equations' group one gets:

$$\boldsymbol{k}\cdot\boldsymbol{E}_0=0 \qquad \boldsymbol{k}\cdot\boldsymbol{H}_0=0 \tag{1.30}$$

Hence:

$$\boldsymbol{k}\times(\boldsymbol{k}\times\boldsymbol{E}_0)=(\boldsymbol{k}\cdot\boldsymbol{E}_0)\boldsymbol{k}-k^2\boldsymbol{E}_0=-k^2\boldsymbol{E}_0 \tag{1.31}$$

On the other hand, from (1.28) and the Maxwell's equations one gets:

$$\boldsymbol{k}\times\boldsymbol{E}_0=\omega\mu\boldsymbol{H}_0 \qquad \boldsymbol{k}\times\boldsymbol{H}_0=-\omega\varepsilon\boldsymbol{E}_0 \tag{1.32}$$

and, hence:

$$\boldsymbol{k}\times(\boldsymbol{k}\times\boldsymbol{E}_0)=\boldsymbol{k}\times\omega\mu\boldsymbol{H}_0=\omega\mu(\boldsymbol{k}\times\boldsymbol{H}_0)=-\omega^2\mu\varepsilon\boldsymbol{E}_0 \tag{1.33}$$

Thus, from (1.31) and (1.33) one gets:

$$-k^2\boldsymbol{E}_0=-\omega^2\mu\varepsilon\boldsymbol{E}_0 \tag{1.34}$$

or:

$$k^2=\omega^2\mu\varepsilon \tag{1.35}$$

The above relation reveals the reason why *k* is denoted as *propagation constant*.

Knowing that $v=1/\sqrt{\varepsilon\mu}$ is the traveling speed of a wave in a medium having material parameters ε and μ one could obtain that:

$$k=\omega\sqrt{\varepsilon\mu}=\omega/v=(2\pi f)/v=(2\pi)/(Tv)=(2\pi)/\lambda \tag{1.36}$$

where λ is the wavelength of a wave with frequency *f*.

The above relation reveals the reason why *k* is denoted as *wave number*.

From (1.32):

$$\boldsymbol{H}_0=\frac{\boldsymbol{k}\times\boldsymbol{E}_0}{\omega\mu}=\frac{\omega\sqrt{\varepsilon\mu}}{\omega\mu}\widehat{\boldsymbol{k}}\times\boldsymbol{E}_0=\sqrt{\frac{\varepsilon}{\mu}}\widehat{\boldsymbol{k}}\times\boldsymbol{E}_0=\frac{1}{\eta}\widehat{\boldsymbol{k}}\times\boldsymbol{E}_0 \tag{1.37}$$

The parameter $\eta=\sqrt{\mu/\varepsilon}$ is denoted as the *intrinsic impedance* of the medium.

For the vacuum (and, by extension, also for the free air) the material parameters and the propagation constant receive a subscript "0". So, we have:

$$\begin{aligned}&\varepsilon_0 \approx 8.8541878 \cdot 10^{-12}\ Farads/meter\\&\mu_0 = 4\pi \cdot 10^{-7}\ Henry/meter\\&\eta_0 = 120\pi \approx 377\ ohmi\\&k_0 = 2\pi/\lambda_0.\end{aligned} \tag{1.38}$$

1.5 –Vector Potential

Solving the wave equation in order to obtain a formula for the wave radiated by a source implies the knowledge of the distribution of the spatial electric charge ρ_s and/or that of the electric current $\boldsymbol{J}$. Even so, the available methods to find a solution are quite complex and very difficult to apply in some cases. A simplification could appear if auxiliary variables are used. One of these variables is the *electric vector potential* $\boldsymbol{A}$.

Its usefulness is based on the physical fact that a magnetic charge does not exist and thus the right hand side of the last equation in the Maxwell's equations group is always zero: $\nabla \cdot \boldsymbol{B} \equiv 0$. Taking into account the mathematical identity:

$$\nabla \cdot (\nabla \times \boldsymbol{a}) \equiv 0, \qquad \forall \boldsymbol{a} \tag{1.39}$$

one can conclude that there exists a vector $\boldsymbol{A}$ such that:

$$\nabla \times \boldsymbol{A} = \boldsymbol{B} \tag{1.40}$$

The vector $\boldsymbol{A}$ defined through the above relation is denoted as *electric vector potential* or, shortly, *vector potential*.

Based on this definition, for harmonic variables, the first of Maxwell' equations can be rewritten as:

$$\nabla \times \boldsymbol{E} = -j\omega(\nabla \times \boldsymbol{A}) \tag{1.41}$$

and, after integration:

$$\boldsymbol{E} = -j\omega\boldsymbol{A} - \nabla\Phi \tag{1.42}$$

In the above relation Φ is an arbitrary scalar function, and $\nabla\Phi$ is the integration constant. It is written like this because:

$$\nabla \times (\nabla\Phi) \equiv 0, \quad \forall\Phi \tag{1.43}$$

and the scalar function Φ can be selected to fulfill other restrictions than the ones imposed by the integration of equation (1.41).

Using the definition of $\boldsymbol{A}$ and considering mediums without polarization and magnetization, the second of Maxwell's equations becomes:

$$\begin{aligned}\nabla \times (\nabla \times \boldsymbol{A}) &= j\omega\mu\varepsilon(-j\omega\boldsymbol{A} - \nabla\Phi) + \mu\boldsymbol{J} =\\&= k^2\boldsymbol{A} - j\omega\mu\varepsilon\nabla\Phi + \mu\boldsymbol{J}\end{aligned} \tag{1.44}$$

We showed previously that (relation 1.20):

$$\nabla \times (\nabla \times \boldsymbol{A}) = \nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2\boldsymbol{A} \tag{1.45}$$

so, that:

$$\nabla(\nabla \cdot \boldsymbol{A}) - \nabla^2\boldsymbol{A} = k^2\boldsymbol{A} - j\omega\mu\varepsilon\nabla\Phi + \mu\boldsymbol{J} \tag{1.46}$$

or:

$$\nabla^2\boldsymbol{A} + k^2\boldsymbol{A} = \nabla(\nabla \cdot \boldsymbol{A}) + j\omega\mu\varepsilon\nabla\Phi - \mu\boldsymbol{J} \tag{1.47}$$

Because the operator ∇ is linear the above equation could be rewritten as:

$$\nabla^2 \boldsymbol{A} + k^2 \boldsymbol{A} = \nabla(\nabla \cdot \boldsymbol{A} + j\omega\mu\varepsilon\Phi) - \mu \boldsymbol{J} \tag{1.48}$$

and it becomes a Helmoltz type equation if we require that:

$$\nabla \cdot \boldsymbol{A} + j\omega\mu\varepsilon\Phi = 0 \tag{1.49}$$

The scalar function Φ is chosen to meet the above condition and it is denoted as *electric scalar potential* or, shortly, *scalar potential*. The relation (1.49) is known as the ***Lorentz condition***.

Summarizing, the vector potential is the solution of the Helmholtz type equation:

$$\nabla^2 \boldsymbol{A} + k^2 \boldsymbol{A} = -\mu \boldsymbol{J} \tag{1.50}$$

Based on the expression (1.42) of the electric field, in mediums without polarization, the third of the Maxwell's equations becomes:

$$\rho = \nabla \cdot \boldsymbol{D} = \varepsilon(\nabla \cdot \boldsymbol{E}) = -j\omega\varepsilon(\nabla \cdot \boldsymbol{A}) - \varepsilon(\nabla^2 \Phi) \tag{1.51}$$

Using the Lorentz condition the above equation becomes:

$$\rho = -j\omega\varepsilon(-j\omega\mu\varepsilon\Phi) - \varepsilon(\nabla^2 \Phi) \tag{1.52}$$

or:

$$\nabla^2 \Phi + k^2 \Phi = \frac{\rho}{\varepsilon} \tag{1.53}$$

Note that the scalar potential is also a solution of a Helmholtz type equation.

In summary, if a solution for the vector potential is found, then the primary variables of the radiated wave are determined by algebraic computation:

$$\begin{aligned} \boldsymbol{E} &= -j\omega \boldsymbol{A} - \nabla(\nabla \cdot \boldsymbol{A})/(j\omega\mu\varepsilon) \\ \boldsymbol{H} &= \tfrac{1}{\mu}(\nabla \times \boldsymbol{A}) \end{aligned} \tag{1.54}$$

while the scalar potential is computed from the Lorentz condition:

$$\Phi = -(\nabla \cdot \boldsymbol{A})/(j\omega\mu\varepsilon) \tag{1.55}$$

For mediums with no polarization $\nabla \cdot \boldsymbol{D} = 0$ and an auxiliary magnetic vector potential $\boldsymbol{A}_m$ could be defined, such that:

$$\boldsymbol{D} = -\nabla \times \boldsymbol{A}_m \tag{1.56}$$

One can prove similarly that magnetic vector potential is a solution of the Helmholtz type equation:

$$\nabla^2 \boldsymbol{A}_m + k^2 \boldsymbol{A}_m = -\varepsilon \boldsymbol{J}_m \tag{1.57}$$

if a Lorentz type condition is fulfilled:

$$\nabla \cdot \boldsymbol{A}_m + j\omega\mu\varepsilon\Phi_m = 0 \tag{1.58}$$

If a solution for the magnetic vector potential $\boldsymbol{A}_m$ is found, then the primary variables of the radiated wave are determined by algebraic computation:

$$\begin{aligned} \boldsymbol{E} &= -\tfrac{1}{\varepsilon}(\nabla \times \boldsymbol{A}_m) \\ \boldsymbol{H} &= -j\omega \boldsymbol{A}_m + \nabla(\nabla \cdot \boldsymbol{A}_m)/(j\omega\mu\varepsilon) \end{aligned} \tag{1.59}$$

while the magnetic scalar potential is computed from the Lorentz condition:

$$\Phi_m = -(\nabla \cdot \boldsymbol{A}_m)/(j\omega\mu\varepsilon) \tag{1.60}$$

1.6 – Boundary Conditions

When traversing the boundary between mediums with different electric and magnetic properties the components tangential to the boundary of the primary variables (***E***, ***H***) and the components normal to the boundary of the secondary variables (***D***, ***B***) remain constant (see figure no.1.2). This property is expressed analytically as follows:

$$\hat{\boldsymbol{n}} \times \boldsymbol{E}_1 = \hat{\boldsymbol{n}} \times \boldsymbol{E}_2 \qquad \hat{\boldsymbol{n}} \cdot \boldsymbol{D}_1 = \hat{\boldsymbol{n}} \cdot \boldsymbol{D}_2$$
$$\hat{\boldsymbol{n}} \times \boldsymbol{H}_1 = \hat{\boldsymbol{n}} \times \boldsymbol{H}_2 \qquad \hat{\boldsymbol{n}} \cdot \boldsymbol{B}_1 = \hat{\boldsymbol{n}} \cdot \boldsymbol{B}_2 \tag{1.61}$$

But if there are electric charges and currents or magnetic charges and currents on the boundary, then the above components remain constant no longer; their magnitudes are modified in accordance with the electric charge density (ρ_s), the electric current density ($\boldsymbol{J}_s$), the magnetic charge density (ρ_{ms}), and the magnetic current density ($\boldsymbol{J}_{ms}$), respectively. The relations expressing these changes are as follows:

$$\hat{\boldsymbol{n}} \times (\boldsymbol{E}_1 - \boldsymbol{E}_2) = -\boldsymbol{J}_{ms} \qquad \hat{\boldsymbol{n}} \cdot (\boldsymbol{D}_1 - \boldsymbol{D}_2) = \rho_s$$
$$\hat{\boldsymbol{n}} \times (\boldsymbol{H}_1 - \boldsymbol{H}_2) = \boldsymbol{J}_s \qquad \hat{\boldsymbol{n}} \cdot (\boldsymbol{B}_1 - \boldsymbol{B}_2) = \rho_{ms} \tag{1.62}$$

When one of the mediums is a perfect conductor (the second one, for instance), the electromagnetic field inside it vanishes and the above relations become:

$$\hat{\boldsymbol{n}} \times \boldsymbol{E}_1 = -\boldsymbol{J}_{ms} \qquad \hat{\boldsymbol{n}} \cdot \boldsymbol{D}_1 = \rho_s$$
$$\hat{\boldsymbol{n}} \times \boldsymbol{H}_1 = \boldsymbol{J}_s \qquad \hat{\boldsymbol{n}} \cdot \boldsymbol{B}_1 = \rho_{ms} \tag{1.63}$$

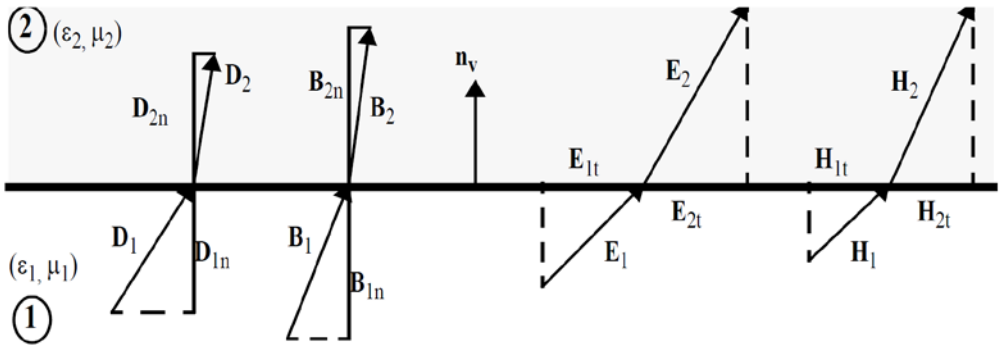


Figure no. 1.2 – *Boundary conditions*

1.7 – General Principle of Reciprocity

Antennas are reciprocal devices, this meaning that they have identical radiation diagrams in both of their modes of operation: transmitting and receiving. This statement is proved by using the general principle of reciprocity that we present in the followings.

Let *V* be a volume bounded by a surface *Σ*. Let ε and μ be the electric and magnetic parameters of the anisotropic medium in *V*. The electromagnetic field radiated by the electric and magnetic current distributions $\boldsymbol{J}_1$ and $\boldsymbol{J}_{m1}$ in *V* has to be in accordance with the Maxwell's equations, that is:

$$\nabla \times \boldsymbol{E}_1 = -j\omega\mu \boldsymbol{H}_1 - \boldsymbol{J}_{m1}$$
$$\nabla \times \boldsymbol{H}_1 = j\omega\varepsilon \boldsymbol{E}_1 + \boldsymbol{J}_1 \tag{1.64}$$

Replacing the medium in *V* with the one having the parameters ε_t and μ_t ("*t*" means transposition) and placing a new distribution of electric and magnetic currents $\boldsymbol{J}_2$ and $\boldsymbol{J}_{m2}$

one gets a new radiated electromagnetic field, also in accordance with the Maxwell's equations:

$$\nabla \times \boldsymbol{E}_2 = -j\omega\mu_t \boldsymbol{H}_2 - \boldsymbol{J}_{m2}$$
$$\nabla \times \boldsymbol{H}_2 = j\omega\varepsilon_t \boldsymbol{E}_2 + \boldsymbol{J}_2 \tag{1.65}$$

We apply the divergence theorem to the vector:

$$\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1 \tag{1.66}$$

and we obtain:

$$\int_V \nabla \cdot (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) dV =$$
$$= \int_\Sigma (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) \cdot d\boldsymbol{\sigma} \tag{1.67}$$

But:

$$\begin{aligned} \nabla \cdot (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) &= \nabla \cdot (\boldsymbol{E}_1 \times \boldsymbol{H}_2) - \nabla \cdot (\boldsymbol{E}_2 \times \boldsymbol{H}_1) = \\ = \mathbf{H}_2 \cdot (\nabla \times \boldsymbol{E}_1) - \mathbf{E}_1 \cdot (\nabla \times \boldsymbol{H}_2) - \mathbf{H}_1 \cdot (\nabla \times \boldsymbol{E}_2) + \mathbf{E}_2 \cdot (\nabla \times \boldsymbol{H}_1) = \\ = \mathbf{H}_2 \cdot (-j\omega\mu \boldsymbol{H}_1 - \boldsymbol{J}_{m1}) - \mathbf{E}_1 \cdot (j\omega\varepsilon_t \boldsymbol{E}_2 + \boldsymbol{J}_2 \nabla \times \boldsymbol{H}_2) - \\ -\mathbf{H}_1 \cdot (-j\omega\mu_t \boldsymbol{H}_2 - \boldsymbol{J}_{m2}) + \mathbf{E}_2 \cdot (j\omega\varepsilon \boldsymbol{E}_1 + \boldsymbol{J}_1) = \\ = \boldsymbol{E}_2 \cdot \boldsymbol{J}_1 - \boldsymbol{E}_1 \cdot \boldsymbol{J}_2 + \boldsymbol{H}_1 \cdot \boldsymbol{J}_{m2} - \boldsymbol{H}_2 \cdot \boldsymbol{J}_{m1} \end{aligned} \tag{1.68}$$

In developing the above relations we took into account that:

$$\boldsymbol{H}_2 \cdot (\mu \boldsymbol{H}_1) = \boldsymbol{H}_1 \cdot (\mu_t \boldsymbol{H}_2)$$
$$\boldsymbol{E}_2 \cdot (\varepsilon \boldsymbol{E}_1) = \boldsymbol{E}_1 \cdot (\varepsilon_t \boldsymbol{E}_2) \tag{1.69}$$

Based on (1.68) the equation (1.67) becomes:

$$\int_\Sigma (\boldsymbol{E}_1 \times \boldsymbol{H}_2 - \boldsymbol{E}_2 \times \boldsymbol{H}_1) \cdot d\boldsymbol{\sigma} =$$
$$= \int_V (\boldsymbol{E}_2 \cdot \boldsymbol{J}_1 - \boldsymbol{E}_1 \cdot \boldsymbol{J}_2 + \boldsymbol{H}_1 \cdot \boldsymbol{J}_{m2} - \boldsymbol{H}_2 \cdot \boldsymbol{J}_{m1})\, dV \tag{1.70}$$

The above relation represents the *general principle of reciprocity* or the *Harrington-Villeneuve principle.*

When there is no radiation source (charges or currents) inside the volume *V*, then the argument of the volume integral in equation (1.70) is zero, an so is the integral itself. The equation (1.70) becomes:

$$\int_\Sigma (\boldsymbol{E}_1 \times \boldsymbol{H}_2) \cdot d\boldsymbol{\sigma} = \int_\Sigma (\boldsymbol{E}_2 \times \boldsymbol{H}_1) \cdot d\boldsymbol{\sigma} \tag{1.71}$$

and this is the *Lorentz principle* of the reciprocity.

When there is no radiation source outside the volume *V*, then the surface integral in the equation (1.70) vanishes and the equation becomes:

$$\int_V (\boldsymbol{E}_2 \cdot \boldsymbol{J}_1 - \boldsymbol{E}_1 \cdot \boldsymbol{J}_2)\, dV = \int_V (\boldsymbol{H}_2 \cdot \boldsymbol{J}_{m1} - \boldsymbol{H}_1 \cdot \boldsymbol{J}_{m2})\, dV \tag{1.72}$$

The equation (1.72) represents the *Rayleigh-Carson principle* of the reciprocity.

1.8 – Duality Theorem

Obviously, solutions of identical type o differential equations associated to different variables have similar appearance. Those variables are denoted as *dual variables*. The practical consequence of duality property is that knowing the solution for a variable, the

solution for the dual one could be obtained simply by systematic replacing the corresponding parameters.

There are multiple dual variables in the electromagnetic theory. For instance, let consider the symmetrized Maxwell's equations with harmonic variables for mediums with no polarization and no magnetization:

$$\nabla \times \boldsymbol{E} = -j\omega\mu\boldsymbol{H} - \boldsymbol{J}_m$$
$$\nabla \times \boldsymbol{H} = j\omega\varepsilon\boldsymbol{E} + \boldsymbol{J} \qquad (1.73)$$

and rearrange them as follows:

$$\nabla \times (-\boldsymbol{E}) = j\omega\mu\boldsymbol{H} + \boldsymbol{J}_m$$
$$\nabla \times \boldsymbol{H} = j\omega\varepsilon\boldsymbol{E} + \boldsymbol{J} \qquad (1.74)$$

Note that the second equation becomes identical to the first one after making the following replacements:

$$\boldsymbol{H} \to -\boldsymbol{E} \quad \boldsymbol{E} \to \boldsymbol{H} \quad \varepsilon \to \mu \quad \boldsymbol{J} \to \boldsymbol{J}_m \qquad (1.75)$$

Making the same replacements in the solution obtained for the vector **H** in the second equation, one gets the solution for the vector **E** in the first one. Those four pairs of variables are dual with each other.

By analyzing the expressions of other differential equations and their solutions one concludes that knowing the expressions for the radiated field ($\boldsymbol{E}$, $\boldsymbol{H}$, $\boldsymbol{A}$) of a distribution of electric charges (ρ) and currents ($\boldsymbol{J}$) in an medium with parameters ε, μ, and η, one could write directly the expressions for the radiated field ($\boldsymbol{E}$, $\boldsymbol{H}$, $\boldsymbol{A}_m$) of a distribution of magnetic charges (ρ_m) and currents ($\boldsymbol{J}_m$) in the same medium simply by making the following replacements:

$$\boldsymbol{H} \to -\boldsymbol{E} \quad \boldsymbol{E} \to \boldsymbol{H} \quad \boldsymbol{J} \to \boldsymbol{J}_m \quad \boldsymbol{A} \to \boldsymbol{A}_m$$
$$\rho \to \rho_m \quad \varepsilon \to \mu \quad \mu \to \varepsilon \quad \eta \to 1/\eta \qquad (1.76)$$

Also, knowing the expressions for the radiated field ($\boldsymbol{E}$, $\boldsymbol{H}$, $\boldsymbol{A}_m$) of a distribution of magnetic charges (ρ_m) and currents ($\boldsymbol{J}_m$) in an medium with parameters ε, μ, and η, one could write directly the expressions for the radiated field ($\boldsymbol{E}$, $\boldsymbol{H}$, $\boldsymbol{A}$) of a distribution of electric charges (ρ) and currents ($\boldsymbol{J}$) in the same medium simply by making the following replacements:

$$\boldsymbol{E} \to -\boldsymbol{H} \quad \boldsymbol{H} \to \boldsymbol{E} \quad \boldsymbol{J}_m \to \boldsymbol{J} \quad \boldsymbol{A}_m \to \boldsymbol{A}$$
$$\rho_m \to \rho \quad \mu \to \varepsilon \quad \varepsilon \to \mu \quad 1/\eta \to \eta \qquad (1.77)$$

1.9 – Image Theorem

When a current source placed in point S close to an electric perfect conducting plane surface, the electromagnetic field in a point P above this surface is the vector sum of the field directly radiated by the current source and the one reflected by the plane surface. Note that (figure no. 1.3) the directions of the reflected field for different positions of point P converge in point S' under the plane surface that is symmetrical to the point S. The reflected field appears to arrive from a fictitious current source placed in this point that is denoted as *image source*.

From the point of view of the field in the upper semi space one could eliminate the plane surface and place an adequate current source in point S' that would radiate a field identical to the reflected one. The electromagnetic field in the lower semi space will remain zero. Using reflection laws and supposing the electromagnetic filed is a plane wave one

could prove that the fictitious image source for an electric current source has the same magnitude as the real source and:

- the same orientation if the source is normal to the reflecting surface;
- the opposite orientation if the source is parallel with the reflecting surface.

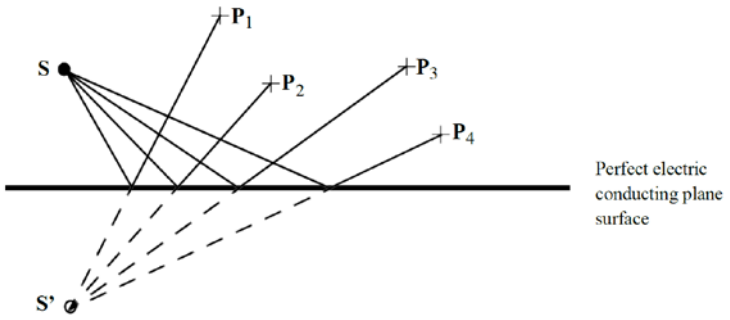


Figure no. 1.3 – *A current source radiating close to a perfect electric conducting plane surface*

For a magnetic current source radiating close to a perfect electric conducting plane surface the image source has the same magnitude as the real source and:

- the opposite orientation if the source is normal to the reflecting surface;
- the same orientation if the source is parallel with the reflecting surface.

For an electric current source radiating close to a perfect magnetic conducting plane surface the image source has the same magnitude as the real source and:

- the opposite orientation if the source is normal to the reflecting surface;
- the same orientation if the source is parallel with the reflecting surface.

For a magnetic current source radiating close to a perfect magnetic conducting plane surface the image source has the same magnitude as the real source and:

- the same orientation if the source is normal to the reflecting surface;
- the opposite orientation if the source is parallel with the reflecting surface.

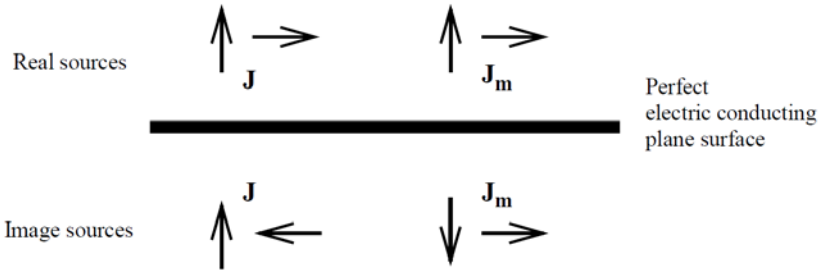


Figure no. 1.4 – *Image source orientation for a perfect electric conducting plane surface*

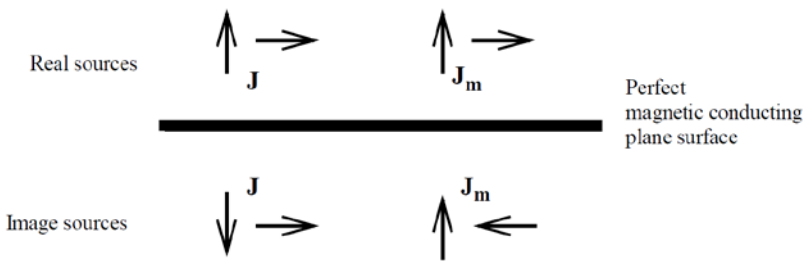


Figure no. 1.5 - *Image source orientation for a perfect magnetic conducting plane surface*

Summarizing, the fictitious image source is placed symmetrical from the reflecting plane surface and has the same magnitude as the real source. It keeps the real source orientation when:

- they are normal to the reflecting surface and the radiating source and the reflecting surface have the same type of conduction (electric or magnetic);
- they are parallel with the reflecting surface and the radiating source and the reflecting surface have the different types of conduction (one – electric and the other one – magnetic).

The image source orientation is opposite to the one of the real source when:

- they are parallel with the reflecting surface and the radiating source and the reflecting surface have the same type of conduction (electric or magnetic);
- they are normal to the reflecting surface and the radiating source and the reflecting surface have the different types of conduction (one – electric and the other one – magnetic).

These conclusions are illustrated in Figures no. 1.4 and 1.5.

C h a p t e r I I

RADIATION OF SIMPLE SOURCES

2.1 – Radiation of Electric Dipole

An *electric dipole* is an infinitesimal small linear unit current source of radiation. Analytically, it is described by $\hat{\boldsymbol{a}}\delta(\boldsymbol{r}-\boldsymbol{r}')$, where $\hat{\boldsymbol{a}}$ is the versor along the direction of current, $\delta(\cdot)$ is the spatial Dirac function, $\boldsymbol{r}$ is the position vector of an arbitrary point, and $\boldsymbol{r}'$ is the position vector of the current source. The spatial Dirac function placed in the origin of a coordinate system has the following properties:

$$\begin{cases}\delta(\boldsymbol{r}) = 0, \quad \forall \boldsymbol{r} \neq \boldsymbol{0} \\ \int_V \delta(\boldsymbol{r})dV = 1\end{cases} \tag{2.1}$$

where *V* is *any* volume containing the origin.

In order to obtain simpler forms for the equations we develop in the followings, we consider that the dipole is placed in the origin of Cartesian system of coordinates ($\boldsymbol{r}' = 0$) and it is oriented along the *Oz* axis ($\hat{\boldsymbol{a}} \equiv \hat{\boldsymbol{z}}$), that is the electric dipole we study is analytically described by $\hat{\boldsymbol{z}}\delta(\boldsymbol{r})$. We consider also that current has a sinusoidal time variation.

The vector potential $\boldsymbol{A}$ has the same direction as the current source, so we could write that $\boldsymbol{A} = A_z\hat{\boldsymbol{z}}$.

Based on the above arguments we could write that, in the domain of the Fourier transform, the wave equation (1.48) for the vector potential $\boldsymbol{A}$ is:

$$\nabla^2(A_z\hat{\boldsymbol{z}}) + k_0^2 A_z\hat{\boldsymbol{z}} = -\mu_0\delta(\boldsymbol{r})\hat{\boldsymbol{z}}. \tag{2.2}$$

This vectorial differential equation is equivalent to the following scalar one:

$$\nabla^2 A_z + k_0^2 A_z = -\mu_0\delta(\boldsymbol{r}) \tag{2.3}$$

As the radiation source (electric dipole) has a spherical symmetry (according to its definition), we expect that its radiation field has the same symmetry and, as a consequence, its expression to have one single independent variable (*r*) from the three ones (r, θ, ϕ) of a spherical coordination system. In a spherical coordination system the equation (2.3), taking into account only the coefficient of $\hat{\boldsymbol{r}}$ in the Laplacian, becomes:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial A_z}{\partial r}\right) + k_0^2 A_z = -\mu_0\delta(\boldsymbol{r}) \tag{2.4}$$

The solutions of the *homogeneous* differential equation associated to the equation (2.4) are the 0^{th} order spherical Bessel functions:

$$j_0(k_0 r) = \frac{sin(k_0 r)}{k_0 r} \qquad n_0(k_0 r) = -\frac{cos(k_0 r)}{k_0 r} \tag{2.5}$$

As the above solutions do not comply with the Sommerfeld conditions, we have to build other functions, based on these ones, that are in accordance with the above mentioned conditions. From the two possible functions:

$$h_0^1(k_0 r) \stackrel{\text{def}}{=} j_0(k_0 r) + j n_0(k_0 r) = -j \frac{e^{j k_0 r}}{k_0 r} \tag{2.6}$$

$$h_0^2(k_0 r) \stackrel{\text{def}}{=} j_0(k_0 r) - j n_0(k_0 r) = j \frac{e^{-j k_0 r}}{k_0 r} \tag{2.7}$$

(which are denoted as spherical Hankel functions) only the second one fulfills the Sommerfeld conditions. Consequently, the wave equation solution should appear as:

$$A_z = C j \frac{e^{-j k_0 r}}{k_0 r} \tag{2.8}$$

where *C* is an integration constant and it is given by the function in the right hand side of the wave equation (2.4). Because it includes the spatial Dirac function is we have to use its properties in order to obtain the value of *C*.

We consider an arbitrary volume *V* as a sphere with radius *r* centered in the origin of the coordinate system, we integrate de equation (2.4) on it and then we compute the limit of the integration result for $r \to 0$, because the second property of the spatial Dirac function requires that the result of the integration should equal 1 for *any* volume V containing de origin of the coordinate system.

By integrating the right hand side of the equation (2.4) we obtain:

$$\int_V [-\mu_0 \delta(\boldsymbol{r})]\, dV = -\mu_0 \int_V \delta(\boldsymbol{r})\, dV = -\mu_0 \tag{2.9}$$

As the integration of the left hand side of the equation (2.4) is regarding, we note from the expression (2.8) that $A_z \sim \frac{1}{r}$, while $dV \sim r^3$; so, the result of the integration of $k_0^2 A_z$ term is proportional to r^3 and its limit for $r \to 0$ is zero. Thus, the significant part of the integration of the left hand side of the equation (2.4) is:

$$\int_V \nabla^2 A_z dV = \int_V \nabla \cdot (\nabla A_z) dV = \int_\Sigma (\nabla A_z) \cdot d\sigma \tag{2.10}$$

In the above development we used the property that the Laplacian of a scalar function is the divergence of its gradient and then we applied the divergence theorem. Note that the surface Σ is the surface of the sphere with the radius *r*, while $d\boldsymbol{\sigma}$ is the vectorial area infinitesimal element, normal to the sphere surface and directed towards the exterior of the sphere.

Let's consider as vectorial area infinitesimal element $d\boldsymbol{\sigma}$ a vector having the modulus equal to the area determined on the sphere by an elementary solid angle $d\Omega$ with its vertex in center of the sphere. Thus:

$$d\boldsymbol{\sigma} = r^2 d\Omega \hat{\boldsymbol{r}} \tag{2.11}$$

Based on the expression (2.8) of the vector potential one gets:

$$\nabla A_z = \frac{\partial A_z}{\partial r} \hat{\boldsymbol{r}} = \frac{jC}{k_0} \left(-\frac{1}{r^2} - \frac{j k_0}{r} \right) e^{-j k_0 r} \hat{\boldsymbol{r}} \tag{2.12}$$

Thus:

$$\int_\Sigma (\nabla A_z) \cdot d\sigma = \int_{4\pi} \left[\frac{jC}{k_0} \left(-\frac{1}{r^2} - \frac{j k_0}{r} \right) e^{-j k_0 r} \hat{\boldsymbol{r}} \right] \cdot (r^2 d\Omega \hat{\boldsymbol{r}}) =$$

$$= -\frac{4\pi j C}{k_0} (1 + j k_0 r) e^{-j k_0 r} \tag{2.13}$$

In the above relation the symbol $\int_{4\pi}$ means an integration on the entire space (notation is based on the fact that the total solid angle around a point is equal to 4π).

Based on the equations (2.4), (2.8), and (2.13) the following result is derived:

$$lim_{r\to 0} \int_V (\nabla^2 A_z + k_0^2 A_z)\, dV = lim_{r\to 0} \left[-\frac{4\pi jC}{k_0}(1 + jk_0 r)e^{-jk_0 r} \right] = -\frac{4\pi jC}{k_0} \quad (2.14)$$

Based on (2.4), (2.9), and (2.14) we could write that:

$$-\frac{4\pi jC}{k_0} = -\mu_0 \quad \Rightarrow \quad C = \frac{\mu_0 k_0}{4\pi j} \quad (2.15)$$

Thus, the expression (2.8) of the vector potential becomes:

$$A_z = \frac{\mu_0}{4\pi}\frac{e^{-jk_0 r}}{k_0 r} \quad (2.16)$$

and:

$$\boldsymbol{A} = A_z \hat{\boldsymbol{z}} = \frac{\mu_0}{4\pi}\frac{e^{-jk_0 r}}{k_0 r}\hat{\boldsymbol{z}} \quad (2.17)$$

This expression can be used to directly write the solution of the wave equation when the radiating electric dipole is not placed in the origin of the coordinate system, but in a point having the position vector $\boldsymbol{r'}$, and it is oriented in the direction of an arbitrary versor $\hat{\boldsymbol{a}}$:

$$\boldsymbol{A} = \frac{\mu_0}{4\pi}\frac{e^{-jk_0|\boldsymbol{r}-\boldsymbol{r'}|}}{k_0|\boldsymbol{r}-\boldsymbol{r'}|}\hat{\boldsymbol{a}} \quad (2.18)$$

Based on the expression (2.17) of the vector potential, the primary variables $\boldsymbol{E}$ and $\boldsymbol{H}$ of the electromagnetic field radiated by the electric dipole can be derived using the relations (1.57). After all the calculus is performed one obtains:

$$\boldsymbol{E} = -\frac{j\eta_0}{2\pi k_0}\left(\frac{jk_0}{r^2} + \frac{1}{r^3}\right)e^{-jk_0 r}\, cos\,\theta\, \hat{\boldsymbol{r}} + \frac{j\eta_0}{4\pi k_0}\left(\frac{k_0^2}{r} - \frac{jk_0}{r^2} - \frac{1}{r^3}\right)e^{-jk_0 r}\, sin\,\theta\, \hat{\boldsymbol{\theta}} \quad (2.19)$$

$$\boldsymbol{H} = \frac{1}{4\pi}\left(\frac{jk_0}{r} + \frac{1}{r^2}\right)e^{-jk_0 r}\, sin\,\theta\, \hat{\boldsymbol{\phi}} \quad (2.20)$$

The above formulas yield the value of the field in any point in space, but they are used especially for points very close to the radiation source (*near region*). For points far away from the radiation source (*Fraunhoffer region*, *far region* or *radiation region*) the terms in $1/r^2$ and $1/r^3$ are negligible small as compared to those in $1/r$ and the formulas are simpler:

$$\boldsymbol{E} = \frac{jk_0\eta_0}{4\pi}\frac{e^{-jk_0 r}}{r}\, sin\,\theta\, \hat{\boldsymbol{\theta}} \quad (2.21)$$

$$\boldsymbol{H} = \frac{jk_0}{4\pi}\frac{e^{-jk_0 r}}{r}\, sin\,\theta\, \hat{\boldsymbol{\phi}} \quad (2.22)$$

The radiation region is the main area for any antenna, as its basic function consists in transmitting the electromagnetic power as far as possible; the terms in $1/r$ fulfill this function and they are denoted as *radiation terms*.

We note from the last two expressions that:

$$\boldsymbol{H} = \frac{1}{\eta_0}\hat{\boldsymbol{r}} \times \boldsymbol{E} \quad (2.23)$$

This type of a relationship defines a transversal electromagnetic wave (TEM), that is a wave with its two components $\boldsymbol{E}$ and $\boldsymbol{H}$ perpendicular to each other and both of them perpendicular on the direction of propagation (which is the one of $\hat{\mathbf{r}}$).

We derive in the followings some formulas we need to exemplify the parameters we define in the next paragraph.

The *power density* which represents the radiated power per unit area (*W/m*^{2}) is the real part of the Poynting vector. For an electric dipole:

$$\boldsymbol{P_\Sigma} = \frac{1}{2} Re(\boldsymbol{E} \times \boldsymbol{H}^*) = \frac{1}{2} Re\left(\frac{jk_0\eta_0}{4\pi}\frac{e^{-jk_0 r}}{r} sin\,\theta\,\hat{\boldsymbol{\theta}} \times \frac{-jk_0}{4\pi}\frac{e^{jk_0 r}}{r} sin\,\theta\,\hat{\boldsymbol{\phi}}\right) =$$

$$= \frac{1}{2}\frac{\eta_0 k_0^2}{(4\pi r)^2}(sin\,\theta)^2\hat{\boldsymbol{r}} \qquad (2.24)$$

Obviously, the modulus of the power density is:

$$|\boldsymbol{P_\Sigma}| \triangleq P_\Sigma = \frac{1}{2}\frac{\eta_0 k_0^2}{(4\pi r)^2}(sin\,\theta)^2 \quad [W/m^2] \qquad (2.25)$$

Generally, for a TEM:

$$P_\Sigma = \frac{1}{2} Re(\boldsymbol{E} \times \boldsymbol{H}^*) = \frac{1}{2} Re\left[\boldsymbol{E} \times \left(\frac{1}{\eta_0}\hat{\boldsymbol{r}} \times \boldsymbol{E}^*\right)\right] = \frac{|\boldsymbol{E}|^2}{2\eta_0}\hat{\boldsymbol{r}} = \frac{E^2}{2\eta_0}\hat{\boldsymbol{r}} \qquad (2.26)$$

and thus:

$$P_\Sigma = \frac{E^2}{2\eta_0} \qquad (2.27)$$

The *radiation intensity* is the radiated power per solid angle unity (*W/ste*) and thus:

$$P_\Omega = r^2 P_\Sigma \qquad [W/ste] \qquad (2.28)$$

For electric dipole:

$$P_\Omega = r^2\frac{1}{2}\frac{\eta_0 k_0^2}{(4\pi r)^2}(sin\,\theta)^2 = \frac{\eta_0 k_0^2}{32\pi^2}(sin\,\theta)^2 \qquad (2.29)$$

The *total radiated power* is computed by integration of the power density on the surface of a sphere centered in the origin (where the radiation source is) and with arbitrary radius *r*, but sufficiently large for the surface to be located in the radiation region of the antenna. Thus:

$$P_{rad} = \int_\Sigma \boldsymbol{P_\Sigma} \cdot \boldsymbol{d\sigma} = \int_0^\pi \int_0^{2\pi}\left[\frac{1}{2}\frac{\eta_0 k_0^2}{(4\pi r)^2}(sin\,\theta)^2\hat{\boldsymbol{r}}\right] \cdot (r^2\,sin\,\theta\,d\theta\,d\phi\,\hat{\boldsymbol{r}}) =$$

$$= \frac{\boldsymbol{\eta_0 k_0^2}}{\boldsymbol{32\pi^2}}\left[\int_0^\pi (sin\,\theta)^3\,d\theta\right]\left(\int_0^{2\pi} d\phi\right) = \frac{\boldsymbol{\eta_0 k_0^2}}{\boldsymbol{12\pi}} \quad [W] \qquad (2.30)$$

In the above development the area infinitesimal element $d\boldsymbol{\sigma}$ is replaced by the Lagrangean of the transformation from the Cartesian coordinates to the spherical ones.

The total radiated power can be computed using the formula for the radiation intensity (2.28) and integrating it for all possible values of Ω (that is, 4π), knowing that $d\Omega = \sin\theta\,d\theta\,d\phi$:

$$\int_{4\pi} P_\Omega d\Omega = \int_0^\pi \int_0^{2\pi}\left[\frac{\eta_0 k_0^2}{32\pi^2}(sin\,\theta)^2\right](sin\,\theta\,d\theta\,d\phi) = \frac{\eta_0 k_0^2}{12\pi}\ [W] \qquad (2.31)$$

2.2– Basic antenna parameters

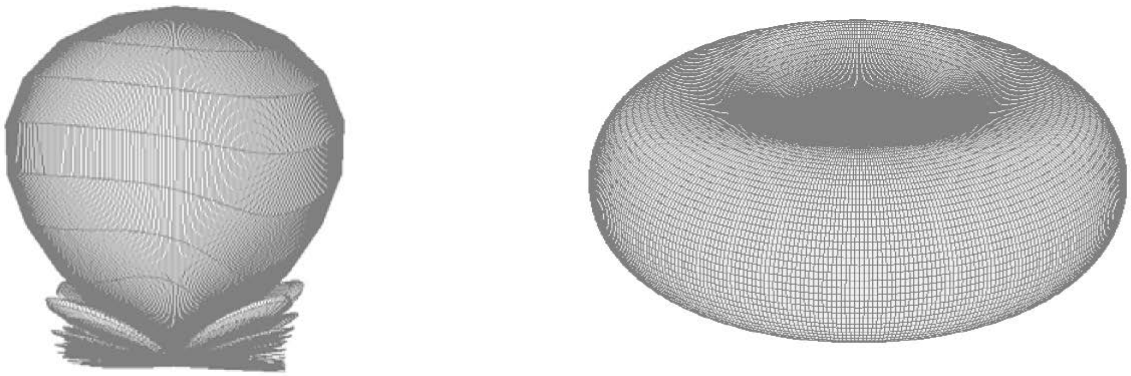
Radiation pattern

Real antennas do not radiate identical power in different spatial directions. As a consequence, the power density and the radiation intensity vary with the spatial angle (θ, ϕ).

We could draw vectors in any direction in space starting in the point where the antenna is located with modulus proportional to the power density (or the radiation intensity) of the radiated field. The envelope of the vertices of all the vectors represents the antenna *radiation pattern*. Obviously, a radiation pattern is a three dimensional closed surface. Usually, in order to fairly compare radiation patterns for different antennas, the relative values are used for representation, that is the ratio of the actual value of the considered variable to the maximum one from the set of all values.

The directions with zero values of the radiation intensity are denoted as *nulls* or *zeros* of the radiation pattern. The region of the radiation pattern between two neighbor zeros is denotes as a *lobe*. The maximum value of the radiation intensity within a lobe is denoted as the *lobe level* and the spatial angle (θ, ϕ) in which it is obtained is denoted as *lobe direction*. The plane angle between the directions in which the radiation intensity is 3 dB smaller than the lobe level is denoted as *lobe beamwidth* (the considered plane should include lobe direction). The above mentioned value of 3 dB is the default one, but other application dependent values could be used (10 dB, 80 dB etc.), too; for instance, in radar applications the beamwidth is considered as the angle between the adjacent zeros of the lobe, because this is the minimum angle separation of two targets, in order to be displayed as distinct objects (radar resolution).

When using relative values for representation, the radiation pattern has one or more lobes with level 1 or 0 dB, denoted as *main lobes*, and other lobes with levels smaller than 1 (negative values in dB), denoted as *secondary*, *auxiliary* or *side lobes*. If a secondary lobe is precisely in the opposite direction of the main lobe, than it is denoted as the *back lobe* and the inverse of its relative level is denoted as the *front/back ratio* of the radiation pattern.



directive antenna omnidirectional antenna

Figure no. 2.1 – *Examples of radiation patterns for directive antennas and omnidirectional antennas*

A 2D representation of a 3D surface (like radiation pattern) could hide significant regions of the surface. This is why plane cross sections of it are used frequently. The sectional plane could be the plane of vector $\boldsymbol{E}$, the plane of vector $\boldsymbol{H}$, the vertical plane, the horizontal plane or any other plane of interest for an application. Any of these plane representations are denoted also as the antenna radiation pattern, but one has to specify the plane used as a cross sectional plane of the 3D radiation pattern. The above mentioned parameters (null, lobe, beamwidth, front/back ratio etc.) could be used also to characterize a 2D radiation pattern.

For some applications it is more useful to monitor directly the value of the electric component $\boldsymbol{E}$ of the radiated electromagnetic wave. In these cases the radiation pattern is built based on the modulus of $\boldsymbol{E}$, but all the above parameters are identically defined. Note,

however, that decreasing the value of the power density or that of the radiation intensity by 3 dB means dividing its value by 2, while decreasing the value of $|\boldsymbol{E}|$ by 3 dB means dividing its value by $\sqrt{2}$.

An *isotropic antenna* is an ideal antenna that distributes uniformly its radiated power along all of the directions in space and, consequently, has a sphere as radiation pattern. The radius of the sphere is equal to 1 when the representation is done based on relative values. Although the isotropic radiator is just a concept and not a real antenna, directive properties of the real antennas are evaluated by comparison with those of the isotropic antenna.

A *directive antenna* is, practically, any real antenna, as they radiate more power in some directions than in the others. But this concept is usually associated with those antennas that radiate much more power in one direction than in the rest of the space. Figure no. 2.1 illustrates the radiation pattern of a directive antenna.

An *omnidirectional antenna* is a particular directive antenna that radiates uniformly in all directions in a plane, but radiates much less in directions outside that plane. Its radiation pattern has a main lobe as a torus and, possibly, some auxiliary lobes outside the torus. Figure no. 2.1 illustrates typical radiation pattern for omnidirectional antennas.

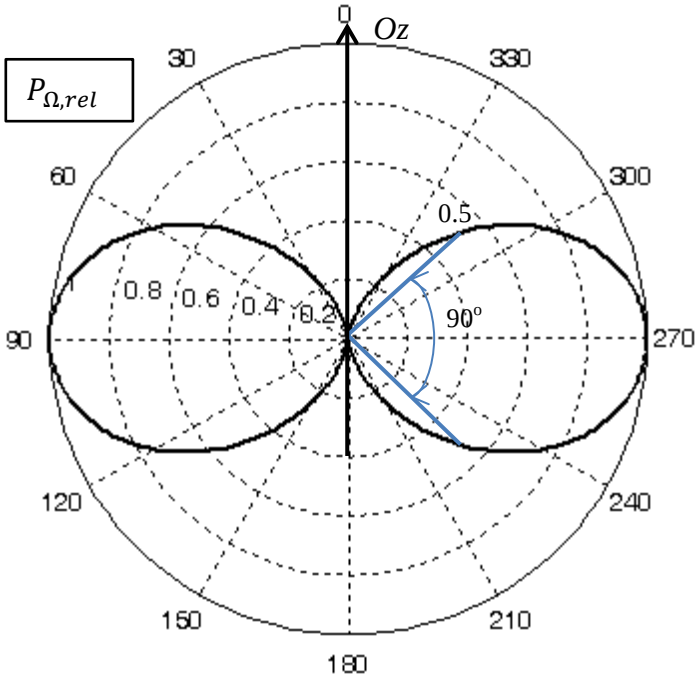


Figure no. 2.2 – *Radiation pattern of the electric dipole in vertical plane containing the Oz axis*

Based on the above relation one can build the radiation pattern of an electric dipole which is torus with *Oz* as its axis of symmetry (identical to the one presented in Figure no. 2.1 as an example of omnidirectional radiation pattern). The radiation pattern in any vertical plane containing the axis *Oz* has two opposite main lobes directed horizontally, with beamwidth equal to $\pi/2$, and two nulls along the *Oz* axis (Figure no. 2.2). In the horizontal plane the radiation pattern is circle with radius equal to 1.

For the electric dipole the maximum radiation intensity is obtained in the direction $\theta = \pi/2$, irrespective of the value of ϕ, (relation 2.29) with the value:

$$P_{\Omega,max} = P_{\Omega}(\theta = \pi/2) = \frac{\eta_0 k_0^2}{32\pi^2}\left(sin\frac{\pi}{2}\right)^2 = \frac{\eta_0 k_0^2}{32\pi^2} \qquad (2.32)$$

and, thus, the relative value of the radiation intensity is:

$$P_{\Omega,rel} = \left[\frac{\eta_0 k_0^2}{32\pi^2}(sin\theta)^2\right] / \left(\frac{\eta_0 k_0^2}{32\pi^2}\right) = (sin\theta)^2 \qquad (2.33)$$

We could use relation (2.21) to build the radiation pattern of the electric dipole based on the modulus of the electric component of its electromagnetic radiated field. We note that the maximum value of *E* is obtained in the direction $\theta = \pi/2$, irrespective of the value of ϕ, and is:

$$E_{max} = |\boldsymbol{E}|_{\theta=\pi/2} = \frac{k_0\eta_0}{4\pi}\frac{1}{r}sin\frac{\pi}{2} = \frac{k_0\eta_0}{4\pi r} \qquad (2.34)$$

Thus:

$$|\boldsymbol{E}_{rel}| = |\boldsymbol{E}|/E_{max} = \left|\frac{jk_0\eta_0}{4\pi}\frac{e^{-jk_0 r}}{r}sin\,\theta\right| / \left(\frac{k_0\eta_0}{4\pi r}\right) = |sin\,\theta| \qquad (2.35)$$

In a Cartesian coordinate system having the electric dipole in its center and directed along the *Oz* axis, the above relation represent a torus with *Oz* axis as axis of symmetry. In vertical planes containing the *Oz* axis the radiation pattern has two opposite identical horizontal main lobes with level 1 and beamwidth of $\pi/2$ (Figure no. 2.3). Note that the lobes have circular shape. The radiation pattern in the horizontal plane is a circle of radius 1.

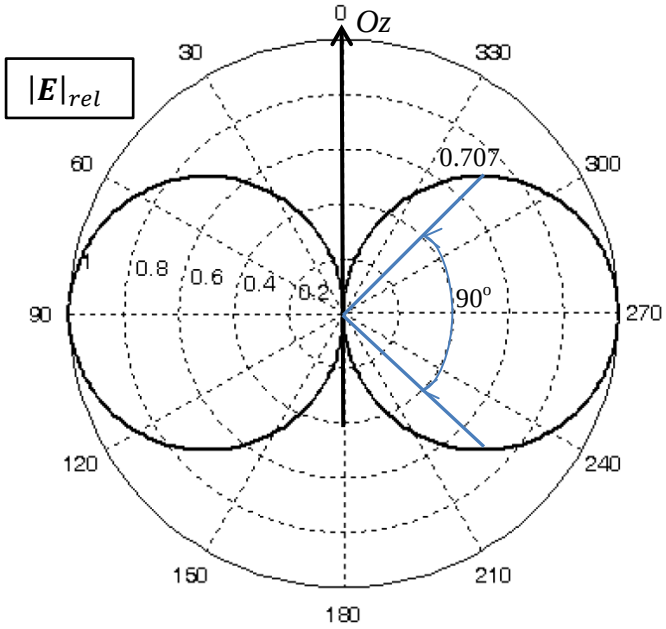


Figure no. 2.3 – *Radiation pattern of electric dipole in a vertical plane containing the Oz axis*

Directivity

This parameter expresses the capability of a given antenna to focus its radiated power on some directions in space. It is defined based on a reference antenna. As reference antenna is used the half-wave dipole, the isotropic radiator or the horn antenna (the latter, especially in the microwave domain).

When using the isotropic radiator as reference the directivity of an antenna in the spatial direction (θ, ϕ) is the ratio of its radiation intensity in the considered direction and the radiation intensity of an isotropic antenna with identical total radiated power. So:

$$D(\theta,\phi) \stackrel{\text{def}}{=} \frac{P_\Omega(\theta,\phi)}{P_{\Omega,isotrop}} = \frac{P_\Omega(\theta,\phi)}{P_{rad}/(4\pi)} = 4\pi\frac{P_\Omega(\theta,\phi)}{P_{rad}} \qquad (2.36)$$

In the above formula, 4π is the total solid angle around a point and P_{rad} is the total radiated power by both the isotropic antenna and the considered antenna.

The directivity is a non-dimensional variable and it could be expressed also in *dB*:

$$D(\theta,\phi) = 10\,log\left[4\pi\frac{P_\Omega(\theta,\phi)}{P_{rad}}\right] \quad [dB] \tag{2.37}$$

The reference antenna used for de computation of directivity is highlighted by adding a suffix to the measuring units: *dBd* – for half-wave dipole and *dBi* – for isotropic radiator.

Although the directivity can be computed in *any* direction, usually when characterizing an antenna by its directivity, the *maximum* value is used and this corresponds to the direction of the main lobe.

The parameter:

$$\Omega_A \overset{\text{def}}{=} \frac{\int_{4\pi} P_\Omega d\Omega}{P_{\Omega,max}} = \int_{4\pi} P_{\Omega,rel} d\Omega \tag{2.38}$$

represents a solid angle where an antenna would radiate the same total power using a constant radiation intensity of $P_{\Omega,max}$ and is denoted as the *equivalent solid angle* of the antenna. One can prove that:

$$D_{max} = \frac{4\pi}{\Omega_A} \tag{2.39}$$

So, the equivalent solid angle of antenna could be used instead of the maximum directivity in order to characterize its directive properties.

For an electric dipole, from (2.29), (2.31), and (2.36):

$$D(\theta,\phi) = 4\pi\left[\frac{\eta_0 k_0^2}{32\pi^2}(sin\,\theta)^2\right]\Big/\left(\frac{\eta_0 k_0^2}{12\pi}\right) = 1.5(sin\,\theta)^2 \tag{2.40}$$

yielding a maximum value of 1.5 or $10\log 1.5 \approx 1.76$ *dBi* that is obtained in the direction $\theta = \pi/2$, irrespective of the value of ϕ .

Approximate formulas of directivity

Computing directivity using its definition (2.36) is not an easy task for most of the practical antenna, because the radiation intensity formula could be very complex or not known with acceptable precision. For these situations, approximate formulas were derived.

For instance, when the radiation pattern contains a single narrow main lobe and some relative small auxiliary lobes, the equivalent solid angle of the antenna is approximated by the product of the beamwidth of the main lobe in two reciprocal perpendicular planes:

$$\Omega_A \approx \theta_1 \cdot \theta_2 \tag{2.41}$$

and, then, the maximum directivity is computed using the formula (2.39):

$$D_{max} \approx \frac{4\pi}{\theta_1 \cdot \theta_2} \tag{2.42}$$

When the beamwidth is measured in degrees:

$$D_{max} \approx \frac{41{,}253}{\theta_1 \cdot \theta_2} \tag{2.42}$$

For a planar array of antennas:

$$D_{max} \approx \frac{32{,}400}{\theta_1 \cdot \theta_2} \tag{2.43}$$

When the level of the secondary lobes is not sufficiently small the values yielded by the above formulas a quite optimistic. If there are two main lobes, the real directivity is only half from the value obtained by using these formulas.

Another formula of approximation uses the main lobe 3 *dB* beamwidth in two specific planes: the plane of vector $\boldsymbol{E}$ (θ_E) and the plane of vector $\boldsymbol{H}$ (θ_H):

$$D_{max} \approx \frac{22{,}181}{\theta_E^2+\theta_H^2} \tag{2.44}$$

when the beamwidth is measured in radians, and:

$$D_{max} \approx \frac{72{,}185}{\theta_E^2+\theta_H^2} \tag{2.45}$$

when the beamwidth is measured in degrees.

In order to estimate de errors introduced by the above approximate formulas in computing the maximum directivity of an antenna, let's consider an antenna having a closed form for its directivity for which we could compute exactly its maximum value and compare it with the approximate value yielded by the each of the formula. An antenna radiating an electromagnetic field with the radiation intensity:

$$P_{\Omega,rel} = \begin{cases} (cos\,\theta)^n, & 0 \le \theta \le \frac{\pi}{2}, \quad 0 \le \phi \le 2\pi \\ 0, & rest\ of\ the\ space \end{cases} \tag{2.46}$$

has a radiation pattern consisting in a single lobe symmetric around the positive direction of axis *Oz*. Its maximum directivity is in the direction of *Oz* $(\theta = 0)$ and it increases when the exponent *n* increases. Table 2.1 presents comparatively the exact value of the maximum directivity computed analytically based on the definition (the true value) and the approximate values yielded by the above formulas.

Table no. 2.1 – *Errors of the approximate formulas for directivity*

N	True value	Approximate value of directivity			
		Rel. (2.42)	Error (%)	Rel. (2.44)	Error (%)
1	4	2.86	*–28.5*	2.53	*–36.75*
2	6	5.09	*–15.27*	4.49	*–25.17*
3	8	7.35	*–8.12*	6.48	*–19*
4	10	9.61	*–3.9*	8.48	*–15.2*
5	12	11.87	*–1.08*	10.47	*–12.75*
6	14	14.13	*0.93*	12.46	*–11*
7	16	16.39	*2.48*	14.47	*–9.56*
8	18	18.66	*3.68*	16.47	*–8.5*
9	20	20.93	*4.64*	18.47	*–7.95*
10	22	23.19	*5.41*	20.47	*–6.96*
11.28	24.56	26.08	*6.24*	23.02	*–6.24*
15	32	34.52	*7.88*	30.46	*–4.81*
20	42	45.89	*9.26*	40.46	*–3.67*

One can see that the error introduced by the formula (2.44) is always negative, that is the approximate value of the directivity is always less than the true value. The error absolute value is smaller for greater values of the exponent *n*, that is for antennas with greater directivity. The error introduced by the formula (2.42) is negative (the approximate value is smaller than the true one) for small values of the exponent *n*, that is for antennas with small directivity. For big values of the exponent *n*, the error is positive (the approximate value is bigger than the true one). The error is about 0 for $n \approx 5.5$. For $n \approx 11.28$ the two approximate formulas yield equal values for error, but with opposite signs.

An approximate value for the maximum directivity for an omnidirectional antenna having some small side secondary lobes could be computed with the following formula:

$$D_{max} \approx \frac{101}{\theta_1 - 0.027\theta_1^2} \tag{2.47}$$

For omnidirectional radiation patterns without secondary lobes a more appropriate approximation for its maximum directivity is given by:

$$D_{max} \approx -172.4 + 191\sqrt{0.118 + 1/\theta_1} \tag{2.48}$$

In the above formulas θ_1 is the 3 dB beamwidth of the main lobe.

For most of the real antennas the radiation pattern is not known in a closed analytical form or it could be extremely complex; so, the total radiated power cannot be computed based, nor can its directivity based on the definition (2.36). Approximate values could be obtained by numerical integration.

Knowing that:

$$P_{rad} = \int_0^{\pi} \int_0^{2\pi} P_{\Omega}(\theta, \phi) \sin\theta \, d\theta \, d\phi \tag{2.49}$$

we could approximate the integral by a sum of a quite large, but finite, number of terms. We divide the interval $[0, \pi]$ of θ in M equal subintervals and the interval $[0, 2\pi]$ of ϕ in N equal subintervals. Thus:

$$P_{rad} \approx \sum_{i=1}^{M} \sum_{j=1}^{N} P_{\Omega}(\theta_i, \phi_j) \sin\theta_i \, \Delta\theta_i \, \Delta\phi_j \tag{2.50}$$

where θ_i and ϕ_j are arbitrarily chosen values in the subintervals i and j, respectively, $\Delta\theta_i = \pi/M$, and $\Delta\phi_j = 2\pi/N$. It results that:

$$P_{rad} \approx \frac{2\pi^2}{MN} \sum_{i=1}^{M} \sum_{j=1}^{N} P_{\Omega}(\theta_i, \phi_j) \sin\theta_i \tag{2.51}$$

Gain

The antenna gain is a parameter that describes its directive properties, but, unlike the directivity, the gain takes into consideration the antenna efficiency, that is the fraction of the input power that is radiated into space. This ratio between the total radiated power and the input power represents the antenna efficiency. The standard defines the antenna *absolute gain* in a spatial direction (θ, ϕ) as the ratio between the radiation intensity in the direction (θ, ϕ) and the radiation intensity that would have been obtained if antenna radiates isotropically all the input power. Analytically:

$$G(\theta, \phi) \overset{\text{def}}{=} \frac{P_{\Omega}(\theta,\phi)}{P_{\Omega,isotropic}} = \frac{P_{\Omega}(\theta,\phi)}{P_{in}/4\pi} = 4\pi \frac{P_{\Omega}(\theta,\phi)}{P_{in}} \tag{2.52}$$

This definition of antenna gain is similar to the definition (2.36) of antenna directivity, but the total radiated power P_{rad} is replaced with the input power P_{in}. This difference, apparently small, has a great impact on the use of these parameters. As the P_{rad} could only be computed, but not practically measured, the directivity could not be practically measured. On the other hand, P_{in} could be measured and the antenna gain is a parameter that can be practically measured.

In practice, it is mainly used the *relative gain* of an antenna, which is the ratio between the antenna absolute gain and the absolute gain of a reference antenna, both of them radiating the same input power. As a reference antenna one of the following is used: isotropic radiator, half wave dipole or horn antenna. When using the isotropic radiator as a reference antenna the relative gain is equal to the absolute one.

As mentioned above, the ratio between the total radiation power and the input power represents the efficiency of the antenna: $\varepsilon = P_{rad}/P_{in}$. As a consequence:

$$G(\theta,\phi) \stackrel{\text{def}}{=} 4\pi \frac{P_\Omega(\theta,\phi)}{P_{in}} = 4\pi \frac{P_\Omega(\theta,\phi)}{P_{rad}/\varepsilon} = \varepsilon 4\pi \frac{P_\Omega(\theta,\phi)}{P_{rad}} = \varepsilon D(\theta,\phi) \qquad (2.53)$$

As the efficiency of any practical antenna is always less than 1, the antenna gain is always less than the antenna directivity. An ideal antenna radiates all of its input power, its efficiency is 1 and, consequently, the gain and the directivity are equal.

For most of the practical antennas the maximum gain is approximately given by:

$$G_{max} \approx \frac{30{,}000}{\theta_E \theta_H} \qquad (2.54)$$

where θ_E and θ_H are the 3 dB main lobe beamwidth (in degrees) in the plane of vector $\boldsymbol{E}$ and $\boldsymbol{H}$, respectively.

Usually, the gain values are expressed in dB:

$$G(dB) = 10\, log\, G \qquad (2.55)$$

Polarization

The polarization of an antenna is the same with the polarization of the electromagnetic field it radiates. The latter one is the curve described by the vertex of the electric component $\boldsymbol{E}$ when its origin is fixed in an observation point. For a small observation time the modulus of $\boldsymbol{E}$ is constant and, as a consequence, the curve described by its vertex is situated in a plane. Generally, this curve is an ellipse and the field is said to have an *elliptical polarization*. The vector vertex could move on the ellipse in any of the two possible senses, so we have a *right* or *clockwise elliptical polarization* and a *left* or *anti clockwise* one. When referring to the associated antenna polarization we consider the sense from the point of view of an observer situated in the point of observation and looking towards the antenna.

The ellipse could be a circle in some particular cases and the polarization is denoted as *circular* with one of the senses presented above.

The ellipse could be a line in some particular cases and the polarization is denoted as a *linear* one. When the line is vertically or horizontally oriented we have a *vertical polarization* or a *horizontal polarization*, respectively. Note that we cannot define a *sense* for a linear polarization.

At large distance from antenna, that is in its radiation region, the radiated electromagnetic field is a transversal wave, that is the two components $\boldsymbol{E}$ and $\boldsymbol{H}$ are reciprocally perpendicular and both of them are perpendicular on the direction of propagation. In an isotropic medium (like the free air, for instance), this direction is the one of the versor $\hat{\boldsymbol{r}}$. As a consequence, the vector $\boldsymbol{E}$ has only two spatial components:

$$\boldsymbol{E} = E_\theta \hat{\boldsymbol{\theta}} + E_\phi \hat{\boldsymbol{\phi}} \qquad (2.56)$$

If a phase difference φ exists between the two components, then their evolution in time could be described as follows:

$$E_\theta = E_1 \cos \omega t \qquad E_\phi = E_2 \cos(\omega t + \varphi) \qquad (2.57)$$

where E_1 and E_2 are the amplitudes of the components and ω is the angular frequency of the radiated field.

These relations could be rewritten as:

$$\frac{E_\theta}{E_1} = \cos \omega t \qquad \frac{E_\phi}{E_2} = \cos \omega t \cos \varphi - \sin \omega t \sin \varphi \qquad (2.58)$$

After some simple mathematical manipulations, one gets:

$$\left(\frac{E_\theta}{E_1}\right)^2 + \left(\frac{E_\phi}{E_2}\right)^2 - 2\frac{E_\theta}{E_1}\frac{E_\phi}{E_2}\cos\varphi = (\sin\varphi)^2 \qquad (2.59)$$

which is the equation of an ellipse in a spherical coordinate system (r, θ, ϕ).

This result proves our previous statement that, in general, a transversal electromagnetic wave has an elliptical polarization. The ellipse of polarization becomes a line (linear polarization) for a phase difference $\varphi = 0$ or $\varphi = \pi$ and it becomes a circle (circular polarization) for a phase difference $\varphi = \pm\pi/2$ and $E_1 = E_2$.

The ellipse of polarization is completely characterized by its axial ratio:

$$AR = 20\,log\frac{E_{max}}{E_{min}} \quad [\text{dB}] \qquad (2.60)$$

where $E_{max} = \max\ (E_1, E_2)$ and $E_{min} = \min\ (E_1, E_2)$

This parameter is 0 dB for circular polarization and it has an infinite value for linear polarization.

Because the components E_θ and E_ϕ of a vector $\boldsymbol{E}$ are independent of each other, one can use them to transmit different information. The receiver should be able to detect simultaneously both of the components by using independent demodulators. As both of the components are present at the input of each of the demodulators, they have to discriminate between the component it intends to detect (denoted as *co-polar component*) ant the other one which it has to reject (denoted as *cross-polar component*).

The relative *level of the cross-polar component* is measured in dB:

$$XP = 20\,log\frac{E_{crosspolar}}{E_{copolar}} \quad [\text{dB}] \qquad (2.61)$$

and the inverse ratio is denoted as the *cross-polar rejection ratio*:

$$XPD = 20\,log\frac{E_{copolar}}{E_{crosspolar}} \quad [\text{dB}] \qquad (2.62)$$

Input Impedance

An antenna is connected with the transmitter or receiver by means of a waveguide or a transmission line. The efficiency of power transferring between the antenna and the waveguide / transmission line is given by the impedances of the two connected elements. The input impedance of an antenna is defined by standard as the *impedance at the antenna terminals* or the *ratio between the voltage and current at a pair of terminals* or the *ratio between the electric component and the magnetic component of the electromagnetic field at an appropriate chosen point*.

The input impedance Z_{in} is a complex variable and it has two parts:

-*the (imaginary) reactive part* (X_{in}) – due to the field induced in the objects in the vicinity of the antenna; the energy stored in these objects is periodically exchanged with the antenna.

-*the (real) resistive part* (R_{in}) – due to power loss in antenna conductors, in antenna dielectric, and as radiation power. The latter part is the "useful" loss and the corresponding fraction in the input resistive part is denoted as the *radiation resistance* (R_{rad}). The radiation resistance is defined as the value of the resistance of physical resistor that would dissipate a power equal to the total power radiated by the antenna when the current flowing through it equals the antenna input current. Obviously, this fraction of the input resistance should be maximized and, for that, the other types of power loss should be minimized.

If R_L is the fraction of the input resistance that corresponds to all of other types of power loss besides the radiation, then we could write that $R_{in} = R_{rad} + R_L$ and the ratio:

$$\varepsilon = \frac{R_{rad}}{R_{rad}+R_L} \tag{2.63}$$

is a measure of the *radiation efficiency* of the antenna as it shows how much of the input power is spread in space as an electromagnetic wave.

Let's evaluate the radiation resistance of *short dipole*; this real antenna is a direct implementation of the previously studied ideal concept of *electric dipole*. The short dipole has a very small, but not infinitesimally small, length *l* and a sinusoidal current with constant amplitude *I*, not equal to 1, flows through it. As a result, the total power radiated by the short dipole could be computed by using the formula (2.30) developed for the electric dipole and taking into account the differences presented above:

$$P_{rad} = \frac{\eta_0 k_0^2}{12\pi} I^2 l^2 \tag{2.64}$$

So, its radiation resistance is:

$$R_{rad} \overset{\text{def}}{=} P_{rad}/\left(\tfrac{1}{2}I^2\right) = \frac{\eta_0 k_0^2}{6\pi} l^2 = 80\pi^2 \left(\frac{l}{\lambda_0}\right)^2 \tag{2.65}$$

In the above development we replaced $\kappa_0 = 2\pi/\lambda_0$ and $\eta_0 = 120\pi$ ohms.

For $l = 0.01\lambda_0$ (*l* is "very small", as we previously stated) we obtain an approximate value of 0.08 ohms for the radiation resistance, which is extremely small: the radiation efficiency of the short dipole is very small. This conclusion remains true for all the antennas: the radiation efficiency is small if the physical dimensions of antenna are small as compared to wavelength of the radiated wave. *In order to obtain acceptable values for the radiation efficiency an antenna should have dimensions comparable to the radiated field wavelength.*

Frequency Bandwidth

According to the standard the frequency bandwidth of an antenna is the frequency interval where the antenna performance associated to a chosen parameter remains inside specified limits. The antenna frequency bandwidth means also the frequency domain around a central frequency (usually, the resonant frequency) where the main antenna parameters (radiation pattern, gain, input impedance, main lobe direction or beamwidth, secondary lobes' level, radiation efficiency – all of them or a restricted group) have non-significant variations. Based on the frequency bandwidth, antennas fall into three categories: r*esonant antennas*, *large bandwidth antennas* and *frequency independent antennas*. The bandwidth of a resonant antenna is a very small fraction of its central frequency. The ratio between the maximum frequency and the minimum one is about 10 for large bandwidth antennas and is more than 100 for frequency independent antennas. The above mentioned parameters that could be taken into account in estimating an antenna bandwidth have not similar changes with frequency, so the bandwidth of a particular antenna depends on the reference parameter(s). The reference parameter is application specific, but the most used are the radiation pattern, the gain, and the input impedance.

Receiving Cross Section

The receiving cross section of an antenna in a given direction is the ratio between the available power at its output port when the antenna is in its receiving mode of operation and power density of the incident plane wave form the considered direction, the antenna polarization matching the one of the incident wave. If no direction is specified, then the direction of the maximum power radiation is considered by default.

When the antenna has a physical radiation surface, as is the case for horn antenna or reflector antenna, then the ratio between the receiving cross section and its physical radiation surface is denoted as the *aperture efficiency* and it has a maximum value of 1.

Effective Height

The effective height of a linear polarized antenna when it receives a plane electromagnetic wave is defined as the ratio between the open circuit voltage at the output port and the rms value of the electric component of the electromagnetic wave on the direction of polarization. Also, the effective height is defined as the length of a thin linear conductor, perpendicular on the electromagnetic wave direction of arrival and parallel with the direction of antenna polarization, that radiates an electromagnetic wave identical to the one radiated by the antenna when a current equal to the antenna input current flows through it.

According to the definition, when $\boldsymbol{E_{in}}$ is electric component of the incident electromagnetic wave and $\boldsymbol{h_{ef}}$ is the effective height of the antenna, then the open circuit voltage at the antenna output port is $V_0 = \boldsymbol{h_{ef}} \cdot \boldsymbol{E_{in}}$.

If $\boldsymbol{E_a} = E_\theta \widehat{\boldsymbol{\theta}} + E_\phi \widehat{\boldsymbol{\phi}}$ is the field created by an antenna in its radiation region, then one could prove that it is related to the antenna effective height by the following formula:

$$\boldsymbol{E_a} = -j\eta_0 \frac{k_0 I_{in}}{4\pi r} \boldsymbol{h_{ef}} e^{-jk_0 r} \tag{2.66}$$

where I_{in} is antenna input current and the other variables have the usual meaning.

2.3– Radiation of Magnetic Dipole

A magnetic dipole is a small circular loop with a total area *S* carrying a constant current with amplitude *I*. We could associate to it a *magnetic moment*, which is a vector perpendicular to the loop, with the origin in the loop center and defined as $\boldsymbol{M} = \boldsymbol{SI}$.

By solving the wave equation for this vector as the radiation source one obtains the following formulas for the electromagnetic field created in the radiation region:

$$\boldsymbol{E} = \frac{\eta_0 k_0^2 SI}{4\pi} \frac{e^{-jk_0 r}}{r} sin\,\theta\, \widehat{\boldsymbol{\phi}} \qquad \boldsymbol{H} = -\frac{k_0^2 SI}{4\pi} \frac{e^{-jk_0 r}}{r} sin\,\theta\, \widehat{\boldsymbol{\theta}} \tag{2.67}$$

The above formulas reveal that radiated field of a magnetic dipole is a transversal electromagnetic wave (that is, they are in accordance with the relation 2.23).

One could notice that these relations are dual to the relations (2.21) and (2.22) and, based on this, we conclude that the radiation pattern of a magnetic dipole is identical to the one of an electric dipole perpendicular to the loop and with its current oriented along the loop magnetic moment.

The total power radiated by a magnetic dipole is:

$$P_{rad} = Re\left(\frac{1}{2}\int_\Sigma \boldsymbol{E} \times \boldsymbol{H}^* \cdot d\boldsymbol{\sigma}\right) = \frac{\eta_0^2 k_0^2 S^2 I^2}{12\pi} \tag{2.68}$$

and its radiation resistance is:

$$R_{rad} \overset{\text{def}}{=} \frac{P_{rad}}{\frac{1}{2}I^2} = 320\pi^4 \left(\frac{S}{\lambda_0^2}\right)^2 \tag{2.69}$$

2.4– Radiation of Arbitrary Current Distribution

The field radiated by an arbitrary current distribution could be computed as a weighted sum of fields radiated by an equivalent unitary sources of current (electric dipoles), based on the hypothesis that the radiation of an electric dipole is not perturbed by

the neighbor electric dipoles and so, we could use the results previously obtained for the free air radiation of an electric dipole.

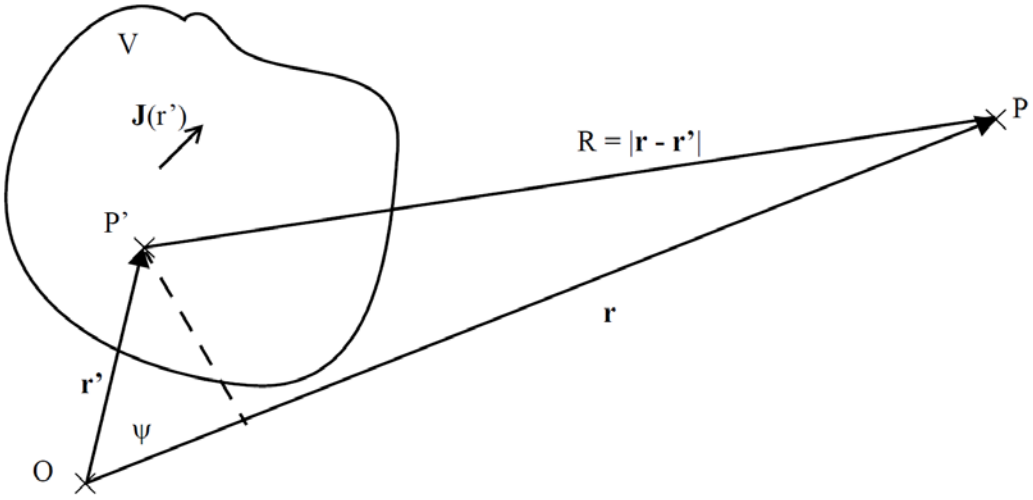


Figure no. 2.4 – *Geometry of arbitrary current distribution*

Let be a volume V bounded by a surface Σ and a coordinate system having the origin O outside the volume V (figure no. 2.4). The position vector of an arbitrary point outside the volume V is denoted by $\boldsymbol{r}$, while the position vector of an arbitrary point inside the volume V is denoted by $\boldsymbol{r'}$.The current distribution inside de volume V is analytically described by a point dependent vectorial function $\boldsymbol{J}(\boldsymbol{r'})$.

Based on the above hypotheses and using the formula (2.18) for the vector potential created by electric dipole we could write directly the formula for the vector potential created in a point P***(r)*** outside the volume V by the arbitrary distribution current in volume V as:

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\int_V \boldsymbol{J}(\boldsymbol{r'})\frac{e^{-jk_0R}}{R}dV \qquad (2.70)$$

where we denoted by $R = |\boldsymbol{r} - \boldsymbol{r'}|$ the distance of an arbitrary point inside the volume V to the point P (see also figure no. 2.5).

The corresponding formulas for the components $\boldsymbol{H}$ and $\boldsymbol{E}$ of the radiated electromagnetic field are obtained by means of the relations (1.54) and (1.8), respectively:

$$\boldsymbol{H} = \frac{1}{\mu_0}\nabla\times\boldsymbol{A} = \frac{1}{4\pi}\nabla\times\left[\int_V \boldsymbol{J}(\boldsymbol{r'})\frac{e^{-jk_0R}}{R}dV\right] =$$

$$= \frac{1}{4\pi}\int_V \left[\nabla\times\boldsymbol{J}(\boldsymbol{r'})\frac{e^{-jk_0R}}{R}\right]dV =$$

$$= \frac{1}{4\pi}\int_V \left[\frac{e^{-jk_0R}}{R}\nabla\times\boldsymbol{J}(\boldsymbol{r'}) - \boldsymbol{J}(\boldsymbol{r'})\times\nabla\frac{e^{-jk_0R}}{R}\right]dV =$$

$$= -\frac{1}{4\pi}\int_V \left[\boldsymbol{J}(\boldsymbol{r'})\times\nabla\frac{e^{-jk_0R}}{R}\right]dV \qquad (2.71)$$

$$\boldsymbol{E} = \frac{1}{j\omega\varepsilon_0}\nabla\times\boldsymbol{H} = -\frac{1}{4\pi j\omega\varepsilon_0}\nabla\times\int_V \left[\boldsymbol{J}(\boldsymbol{r'})\times\nabla\frac{e^{-jk_0R}}{R}\right]dV =$$

$$= -\frac{1}{4\pi j\omega\varepsilon_0}\int_V \left[\nabla\cdot\left(\nabla\frac{e^{-jk_0R}}{R}\right)\boldsymbol{J}(\boldsymbol{r'}) - [\boldsymbol{J}(\boldsymbol{r'})\cdot\nabla]\left(\nabla\frac{e^{-jk_0R}}{R}\right)\right]dV \qquad (2.72)$$

In developing the above relations we used the usual formulas from the vectorial analysis and some simplifications due to the following observations:

- the derivation operator ∇ and the integration one $\int$ are independent of each other as they are associated to two independent coordination systems: the derivation operator is related to arbitrary points *outside* the volume *V*, while the integration one is related to arbitrary points *inside* the volume *V*;
- one consequence of the above observation is that the order of application of these operators could be interchanged;
- another consequence is that $\nabla \times \boldsymbol{J}(\boldsymbol{r}') \equiv 0$ and $\nabla \cdot \boldsymbol{J}(\boldsymbol{r}') \equiv 0$;
- there is no current outside the volume *V* and thus $\boldsymbol{J} \equiv 0$ in relation (1.8);
- all variables have harmonic time variations and, thus, the involved equations could be solved in the Fourier transform domain.

The formula for the electric component $\boldsymbol{E}$ can be further simplified by noting that the divergence of the gradient of a scalar function is the Laplacian of that function $[\nabla \cdot (\nabla\varphi) = \nabla^2\varphi]$ and that the Green function is a solution of the homogeneous wave equation. Thus:

$$\nabla \cdot \left(\nabla \frac{e^{-jk_0R}}{R}\right) = \nabla^2 \left(\frac{e^{-jk_0R}}{R}\right) = -k_0^2 \frac{e^{-jk_0R}}{R} \tag{2.73}$$

and

$$\boldsymbol{E} = \frac{1}{4\pi j\omega\varepsilon_0} \int_V \left[k_0^2 \frac{e^{-jk_0R}}{R} \boldsymbol{J}(\boldsymbol{r}') + [\boldsymbol{J}(\boldsymbol{r}') \cdot \nabla] \left(\nabla \frac{e^{-jk_0R}}{R}\right)\right] dV \tag{2.74}$$

The above formulas for the components $\boldsymbol{E}$ and $\boldsymbol{H}$ cannot be further simplified and they are true for every point *outside* the volume V that contains the radiation sources.

Simpler formulas can be obtained for points at large distance from the volume *V* (the *radiation region*), where the terms in $1/r$ are much greater than the terms in $1/r^2, 1/r^3, \ldots$. When evaluating the modulus of the Green function in this region we approximate *R* as $R = |\boldsymbol{r} - \boldsymbol{r}'| \approx r$. When evaluating the phase of the Green function we approximate *R* as its projection on $\boldsymbol{r}$ (see figure no. 2.4):

$$R \approx r - r' \cos\psi = r - \frac{r\, r' \cos\psi}{r} = r - \frac{\boldsymbol{r} \cdot \boldsymbol{r}'}{r} = r - \hat{\boldsymbol{r}} \cdot \boldsymbol{r}' \tag{2.75}$$

It results that:

$$\frac{e^{-jk_0R}}{R} \approx \frac{e^{-jk_0(r - \hat{\boldsymbol{r}} \cdot \boldsymbol{r}')}}{r} = e^{jk_0\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} \frac{e^{-jk_0r}}{r} \tag{2.76}$$

$$\nabla \frac{e^{-jk_0R}}{R} \approx \nabla \left(e^{jk_0\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} \frac{e^{-jk_0r}}{r}\right) = e^{jk_0\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} \nabla \frac{e^{-jk_0r}}{r} \approx -jk_0 e^{jk_0\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} \frac{e^{-jk_0r}}{r} \hat{\boldsymbol{r}} \tag{2.77}$$

In developing the last of the above formulas we used the previous observation that the variable $\mathbf{r}'$ is independent of the derivation operator ∇ and that the versor $\hat{\mathbf{r}}$ is a constant vector; as a result the entire exponential $e^{jk_0\hat{\mathbf{r}} \cdot \mathbf{r}'}$ is a constant and its derivative by ∇ is zero. Also, the term in $1/r^2$ is neglected in the expression of gradient of $\frac{e^{-jk_0r}}{r}$.

Taking into account these approximations we obtain the following simpler formulas for the electromagnetic field created in the radiation region by arbitrary distribution of currents:

$$\boldsymbol{H} \approx \frac{jk_0}{4\pi} \frac{e^{-jk_0r}}{r} \int_V [\boldsymbol{J}(\boldsymbol{r}') \times \hat{\boldsymbol{r}}] e^{jk_0\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} dV \tag{2.78}$$

$$\boldsymbol{E} \approx -\frac{1}{4\pi j\omega\varepsilon_0} \int_V \left\{k_0^2 e^{jk_0\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} \frac{e^{-jk_0r}}{r} \boldsymbol{J}(\boldsymbol{r}') - [\boldsymbol{J}(\boldsymbol{r}') \cdot \nabla] \left(-jk_0 e^{jk_0\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} \frac{e^{-jk_0r}}{r} \hat{\boldsymbol{r}}\right)\right\} dV =$$

$$= -\frac{1}{4\pi j\omega\varepsilon_0} \int_V \left[k_0^2 \frac{e^{-jk_0r}}{r} \boldsymbol{J}(\boldsymbol{r}') - J_r \frac{\partial}{\partial r} \left(-jk_0 \frac{e^{-jk_0r}}{r} \hat{\boldsymbol{r}}\right)\right] e^{jk_0\hat{\boldsymbol{r}} \cdot \boldsymbol{r}'} dV \approx$$

$$\approx -\frac{jk_0\eta_0}{4\pi}\frac{e^{-jk_0r}}{r}\int_V \left(J_\theta\hat{\boldsymbol{\theta}} + J_\phi\hat{\boldsymbol{\phi}}\right)e^{jk_0\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'}dV \qquad (2.79)$$

These relations for $\boldsymbol{E}$ and $\boldsymbol{H}$ are in accordance with the condition (2.23) for an electromagnetic field to be a transversal wave and, thus, we conclude that *any radiation source creates in its radiation region a transversal wave*.

This conclusion allows one studying the radiation of a given antenna to derive a formula for one of the variable and, afterwards, to obtain the formula for the other one simply by using the property (2.23).

Note, again, that these approximations and this property are true only in the radiation region of an antenna, that is, at large distance from the antenna.

2.5– Radiation of Thin Wire Antenna

We showed previously that the electric dipole has acceptable radiation efficiency when its physical dimension is comparable with the wavelength of the radiated wave. This is true for any other antenna. At small frequency this could mean dimensions of tens or even hundreds of meters (for instance, the free air wavelength of a 1 MHz wave is 300 meters). The simplest shape of an efficient antenna at low frequency is a cylindrical conductor having the length comparable with the wavelength and much greater than the diameter of the conductor cross section. This type of antenna is the physical implementation of the theoretical concept of *thin wire antenna*.

Let's consider a thin wire antenna oriented along the *Oz* axis of a Cartesian coordinate system, with its median point positioned in the origin of the coordinate system and connected to the feeding transmission line in this point (figure no. 2.5). Because the diameter of the antenna cross section is much smaller than its length we could consider that the current flowing through antenna is concentrated in the center of its cross section. Due to the open circuit termination of the antenna at both of its ends, the current wave through the antenna has a total reflection at the antenna ends and a standing wave regime is installed in the antenna. Analytically, this standing wave is described by the following relation:

$$I(z) = I_0\, sin[k_0(l - |z|)], \quad -l \le z \le l \qquad (2.80)$$

The formulas for the electromagnetic field created by the thin wire antenna are obtained from the ones derived for an arbitrary current distribution taking into account the following peculiarities:

- the integration is not performed on a volume, but along the *Oz* axis between $-l$ and l, due to the peculiar shape of the antenna;
- the vector $\boldsymbol{J}(\boldsymbol{r}')$, which originally is a current density on the unit area, is replaced by the vector $I(z)\hat{\boldsymbol{z}}$, which is a current;
- the component J_θ of the current density along the versor $\hat{\boldsymbol{\theta}}$ is replaced by the similar component of the current $I(z)$ and its expression is:

$$J_\theta \to I(z)\hat{\boldsymbol{z}}\cdot\hat{\boldsymbol{\theta}} = I_0\, sin[k_0(l - |z|)]\,(-sin\,\theta) \qquad (2.81)$$

- the component J_ϕ of the current density along the versor $\hat{\boldsymbol{\phi}}$ is replaced by the similar component of the current $I(z)$, which is zero because the versor $\hat{\boldsymbol{\phi}}$ lies in the plane *xOy* and so, it is perpendicular on $\hat{\mathbf{z}}$: $J_\phi \to I(z)\hat{\mathbf{z}}\cdot\hat{\boldsymbol{\phi}} \equiv 0$;
- the vector $\boldsymbol{r}'$ is replaced by the vector $z\hat{\boldsymbol{z}}$ because all the points of the thin wire antenna lie along the *Oz* axis. Thus:

$$\hat{\boldsymbol{r}}\cdot\boldsymbol{r}' \to \hat{\boldsymbol{r}}\cdot(z\hat{\boldsymbol{z}}) = z\,cos\,\theta \qquad (2.82)$$

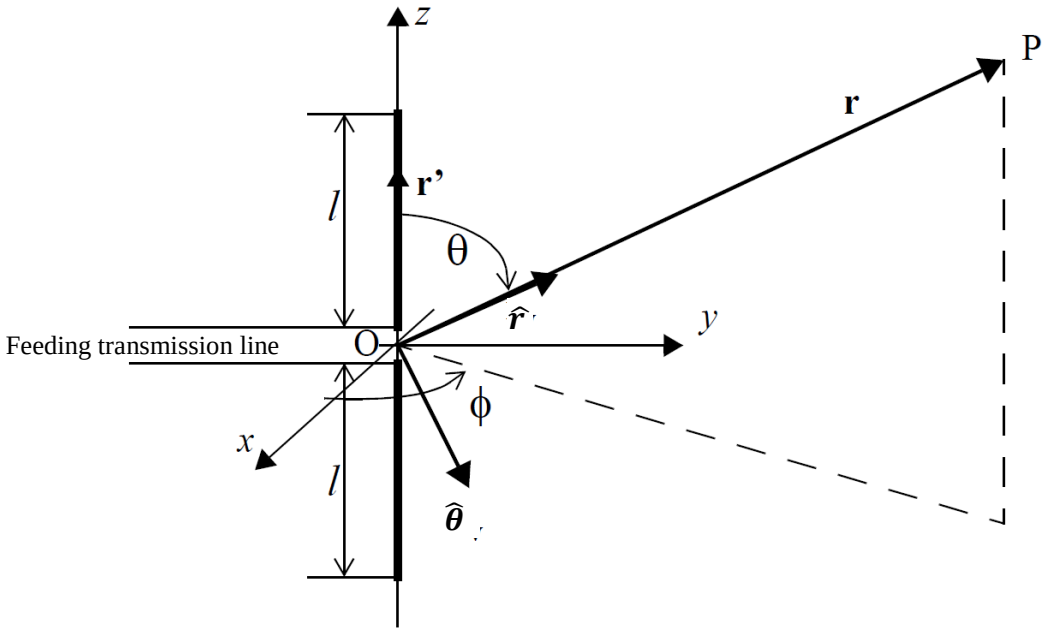


Figure no. 2.5 – *Geometry of thin wire antenna*

Based on the above considerations we obtain:

$$\boldsymbol{E} \approx -\frac{jk_0\eta_0}{4\pi}\frac{e^{-jk_0r}}{r}\int_{-l}^{l} I_0\, sin[k_0(l-|z|)]\,(-\sin\theta)\hat{\boldsymbol{\theta}}e^{jk_0z\cos\theta}dz =$$

$$= \frac{jk_0\eta_0}{4\pi}\frac{e^{-jk_0r}}{r} I_0 \sin\theta\,\hat{\boldsymbol{\theta}}\int_{-l}^{l} sin[k_0(l-|z|)]\, e^{jk_0z\cos\theta}dz =$$

$$= \frac{j\eta_0 I_0}{2\pi}\frac{e^{-jk_0r}}{r}\frac{cos(k_0l\cos\theta)-cos(k_0l)}{\sin\theta}\hat{\boldsymbol{\theta}} \qquad (2.83)$$

As the radiated field in the radiation region of any antenna is a transversal wave, we could write directly the formula for the magnetic field component:

$$\boldsymbol{H} = \frac{1}{\eta_0}\hat{\boldsymbol{r}}\times\boldsymbol{E} = \frac{jI_0}{2\pi}\frac{e^{-jk_0r}}{r}\frac{cos(k_0l\cos\theta)-cos(k_0l)}{\sin\theta}\hat{\boldsymbol{\phi}} \qquad (2.83)$$

The formulas for the radiated field in the radiation region do not include the spherical variable ϕ and, as a consequence, the radiation pattern of a thin wire antenna has symmetry of revolution around the *Oz* axis of the coordinate system, that is around the antenna.

The product $k_0 l$ can be written as:

$$k_0 l = \frac{2\pi}{\lambda_0} l = \frac{2l}{\lambda_0}\pi = n\pi \qquad (2.84)$$

where $n \triangleq \frac{2l}{\lambda_0}$ is the ratio between the physical length $2l$ of the antenna and the free air wavelength of the radiated field and is denoted as the *electrical length* of the antenna.

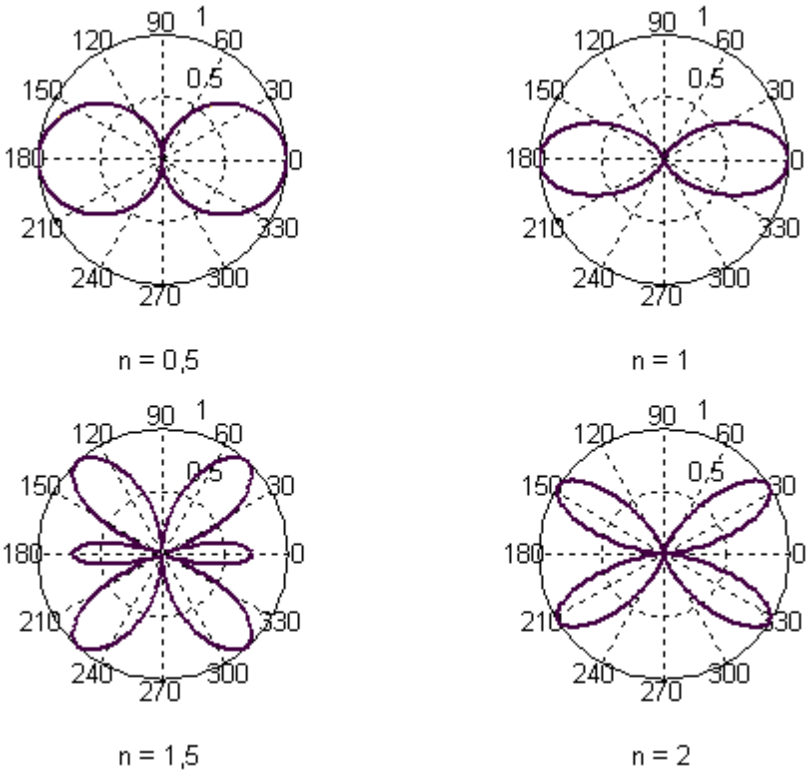


Figure no. 2.6 – *Radiation patterns of standing wave thin wire antenna*

Figure no. 2.6 presents the radiation pattern of a thin wire antenna for different values of its electrical length. Note that for values less than one of the electrical length the radiation pattern has a single main lobe which is perpendicular to the antenna; for values greater than one, secondary lobes and multiple main lobes appear; the latter ones are no longer perpendicular to the antenna.

Formulas for the *input resistance* of a thin wire antenna are difficult to obtain as it implies using of complex variational techniques. The formulas yield theoretical infinite values of the input resistance for physical lengths equaling multiples of λ_0, when the standing wave current through the antenna has a null precisely in the feeding point (middle point of the antenna). The input resistance of practical thin wire antennas is not infinite, but could reach unusual great values (1000 ohms or more).

For $n = 1/2$ (half wave dipole antenna) the input resistance is:

$$R_{in} = \frac{\eta_0}{4\pi}[ln(2\pi\gamma) - Ci(2\pi)] \approx 73\ ohmi \qquad (2.86)$$

where $\gamma = 0.5772\ldots$ is the Euler constant, while $Ci(x) \stackrel{\text{def}}{=} -\int_x^\infty \frac{\cos x}{x} dx$ is denoted as *integral cosine* function.

Traveling Wave Thin Wire Antenna

When the thin wire antenna is terminated on a matched resistance, the current wave through suffers no reflection at the antenna end and a traveling wave regime is established in the antenna. If we admit that the ground plane at which the matching resistance is connected does not influence the antenna radiation and that the ohmic loss and the radiation loss along the wire are negligible small, then the current wave through the antenna has constant amplitude along the antenna and its phase speed equals the one in free space. Thus:

$$I(z) = I_0 e^{-jk_0 z} \qquad (2.87)$$

and the radiated field in the radiation region is:

$$\boldsymbol{E} \approx -\frac{jk_0\eta_0}{4\pi}\frac{e^{-jk_0r}}{r}\int_{-l}^{l} I_0 e^{-jk_0z}(-\sin\theta)\widehat{\boldsymbol{\theta}} e^{jk_0z\cos\theta}dz =$$

$$= \frac{jk_0\eta_0}{4\pi}\frac{e^{-jk_0r}}{r} I_0 \sin\theta\,\widehat{\boldsymbol{\theta}}\int_{-l}^{l} e^{-jk_0z(1-\cos\theta)}dz =$$

$$= \frac{jk_0I_0}{2\pi}\frac{e^{-jk_0r}}{r}\sin\theta\,\frac{\sin[k_0l(1-\cos\theta)]}{1-\cos\theta}\widehat{\boldsymbol{\theta}} \qquad (2.88)$$

$$\boldsymbol{H} = \frac{1}{\eta_0}\hat{\boldsymbol{r}}\times\boldsymbol{E} = \frac{jI_0}{2\pi}\frac{e^{-jk_0r}}{r}\sin\theta\,\frac{\sin[k_0l(1-\cos\theta)]}{1-\cos\theta}\widehat{\boldsymbol{\phi}} \qquad (2.89)$$

The radiation pattern of the traveling wave thin wire antenna has symmetry of revolution around the *Oz* axis, that is around the antenna, and its shape depends on the electrical length $n = 2l/\lambda_0$ of the antenna.

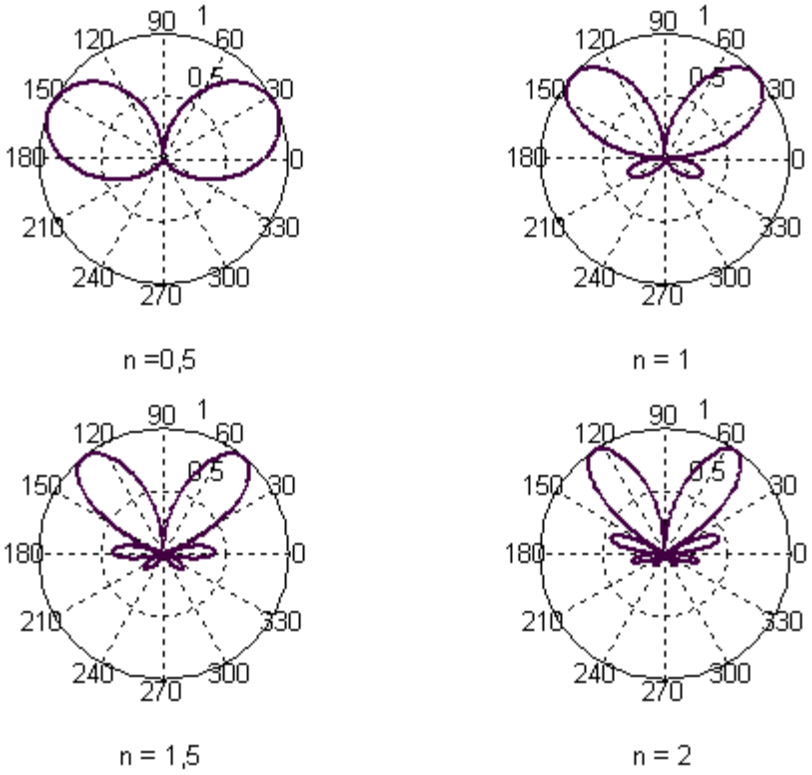


Figure no. 2.7 – *Radiation patterns of traveling wave thin wire antenna*

Figure no. 2.7 presents the radiation pattern of a traveling wave thin wire antenna for different values of is electrical length. As compared to the corresponding radiation patterns of the standing wave thin wire antenna (figure no. 2.6) these ones have a greater number of lobes, but a single main lobe. All the lobes are tilted in the sense the current wave through antenna flows, the main lobe is the closest one to the antenna and the level of the secondary lobes decreases as long as their angular distance from the main lobe increases.

The input impedance of a traveling wave thin wire antenna lies in the range 200 – 300 ohms and is almost resistive because the characteristic resistance of any traveling wave is identical to the characteristic resistance of the transmission line it flows through.

2.6– Practical Types of Thin Wire Antennas

The practical antennas that materialize the theoretical concept of thin wire antenna could be grouped into two main categories: *dipole antennas* and *long-wire antennas*. They differentiate through the electrical length: the first category includes antennas with small electrical length, while the latter – those antennas with great electrical length. The limit value that separates these two categories does not meet a consensual value among researchers, but this does not induce a great amount of ambiguity as the vast majority of dipole antennas have electrical lengths of 0.5 to 1, while most of long the wire antennas have electrical lengths greater than 3.

Dipole antennas

Cylindrical dipole

The cylindrical dipole is a direct implementation of the concept of thin wire antenna. Its properties differ slightly from the ones of the ideal theoretical reference due to the fact that its length *is not much larger* than its cross section diameter, as the theoretical analysis assumes. The main differences are the following:

- the radiated field in the null directions of the radiation pattern is not rigorously zero, but quite a small value as compared to the ones in the neighboring directions;
- the actual shape of the radiation pattern depends also on the absolute value of the cross section diameter, not only on the electrical length;
- the input resistance approaches the theoretical value only if the dipole is at large distance from the ground plane. Otherwise, it is strongly influenced by the conditions at the feeding point and by the dimensions and the conducting properties of the ground plane. For a free space dipole, the an approximate value of the input resistance could be computed with the following formula:

$$R_{in} \approx \begin{cases} 20G^2, & 0 < n < 1/4 \\ 24.7G^{2.5}, & 1/4 \le n < 1/2 \\ 11.4G^{4.17}, & 1/2 \le n < 0.6366 \end{cases} \tag{2.90}$$

where $\triangleq \pi n$.

The cylindrical with an electrical length equal to $n = 1/2$ has a physical length of $\lambda_0/2$ and is denoted as *half wave dipole*. It is one of the most used antennas due to its simple physical structure and to its parameters that could be adapted to requirements of many applications. For this particular length, the radiated field in the radiation region is:

$$\boldsymbol{E} \approx j60I_0 \frac{e^{-jk_0 r}}{r} \frac{cos\left(\frac{\pi}{2} cos\,\theta\right)}{sin\,\theta} \widehat{\boldsymbol{\theta}}$$

$$\boldsymbol{H} \approx \frac{j}{2\pi} I_0 \frac{e^{-jk_0 r}}{r} \frac{cos\left(\frac{\pi}{2} cos\,\theta\right)}{sin\,\theta} \widehat{\boldsymbol{\phi}} \tag{2.91}$$

The electrical field component has a maximum value in the direction of $\theta = \pi/2$, irrespective of the value of ϕ :

$$|\boldsymbol{E}|_{max} \triangleq E_{max} = \frac{60I_0}{r} \tag{2.92}$$

and a normalized expression:

$$E_{rel} \overset{\text{def}}{=} \frac{|\boldsymbol{E}|}{E_{max}} = \frac{cos\left(\frac{\pi}{2} cos\,\theta\right)}{sin\,\theta} \tag{2.93}$$

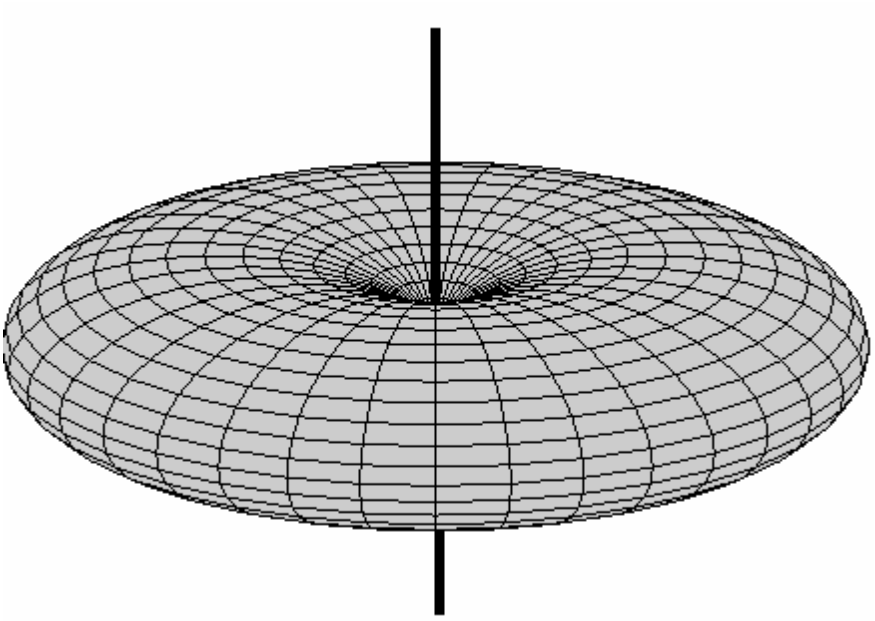


Figure no. 2.8 – 3D r*adiation pattern of a half wave cylindrical dipole*

As a consequence the radiation pattern of the half wave dipole in terms of the normalized radiated field is a torus with the axis *Oz* as its axis of symmetry (figure no. 2.8), that is the radiation pattern is omnidirectional. In any plane containing the axis *Oz* the radiation pattern contains two opposite lobes with level one, that is both of them are main lobes (figure no. 2.9). The beamwidth of the lobes equals 78 degrees.

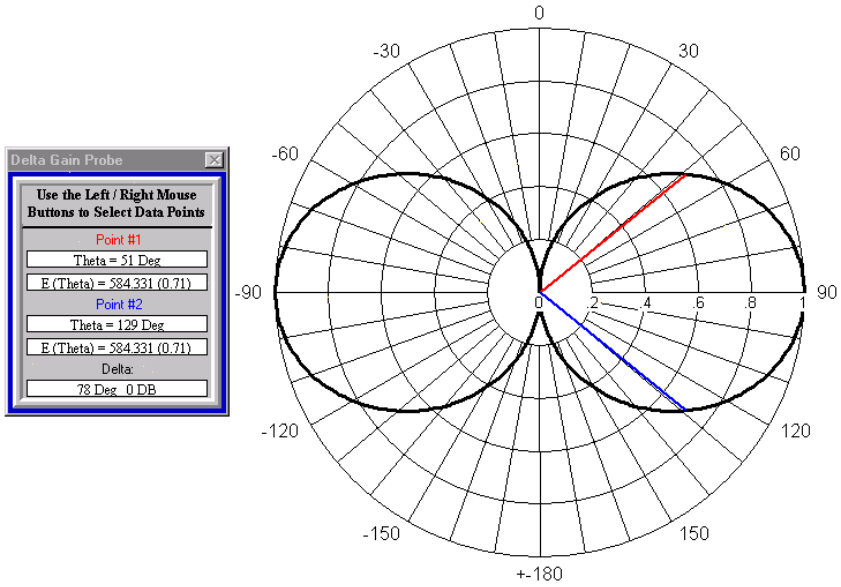


Figure no. 2.9 – *Radiation pattern of half wave cylindrical dipole in planes containing the dipole*

The *power density* of the radiated field in the radiation region is:

$$P_{\Sigma} = \frac{|\boldsymbol{E}|^2}{2\eta_0} = \frac{15I_0^2}{\pi r^2}\left[\frac{cos\left(\frac{\pi}{2}cos\,\theta\right)}{sin\,\theta}\right]^2 \qquad [\mathrm{W/m^2}] \qquad (2.94)$$

and the *radiation intensity* in the same region is:

$$P_{\Omega} = r^2 P_{\Sigma} = \frac{15 I_0^2}{\pi}\left[\frac{cos\left(\frac{\pi}{2}cos\,\theta\right)}{sin\,\theta}\right]^2 \qquad \text{[W/ste]} \qquad (2.95)$$

The *total radiated power* of a half wave dipole is:

$$P_{rad} = 30 I_0^2 \int_0^{\pi}\left[\frac{cos\left(\frac{\pi}{2}cos\,\theta\right)}{sin\,\theta}\right]^2 d\theta \qquad \text{[W]} \qquad (2.96)$$

The integral in the above expression has a very complex closed form, but we could obtain a simpler form by numerical approximation:

$$P_{rad} \approx 36.5648 I_0^2 \qquad \text{[W]} \qquad (2.97)$$

The *directivity* of the half wave dipole is:

$$D(\theta,\phi) \overset{\text{def}}{=} 4\pi \frac{P_{\Omega}(\theta,\phi)}{P_{rad}} \approx 1.64\left[\frac{cos\left(\frac{\pi}{2}cos\,\theta\right)}{sin\,\theta}\right]^2 \qquad (2.98)$$

The maximum value of the directivity is obtained in the direction of $\theta = \pi/2$ and is:

$$D_{max} \approx 1.64 \quad \text{or} \quad D_{max} \approx 10\,log\,1.64 \approx 2.15 \quad \text{[dBi]} \qquad (2.99)$$

The *radiation resistance* of the half wave dipole is:

$$R_{rad} = \frac{P_{rad}}{\frac{1}{2} I_0^2} \approx 73{,}13 \quad \text{[ohms]} \qquad (2.100)$$

Considering that the conducting losses in the antenna are negligible, the gain of the half wave dipole is equal to its directivity and we can compute its cross section area:

$$S_{ef} \overset{\text{def}}{=} \frac{\lambda_0^2}{4\pi} G_{max} \approx \frac{\lambda_0^2}{4\pi} D_{max} \approx 0.13\lambda_0^2 \qquad (2.101)$$

The *effective height* of the half wave dipole is:

$$\boldsymbol{h_{ef}} \approx -\frac{\lambda_0}{\pi}\frac{cos\left(\frac{\pi}{2}cos\,\theta\right)}{sin\,\theta}\widehat{\boldsymbol{\theta}} \qquad (2.102)$$

and it has a maximum modulus of λ_0/π in the direction $\theta = \pi/2$. Note that the maximum modulus of the *effective height is smaller than the physical length* of the half wave dipole, a property that is true for any antenna.

Biconical Dipole

The biconical dipole is a dipole having cones, instead of cylinders, as its arms (figure no. 2.10). The input impedance of a biconical dipole varies with the vertex angle $2\theta_0$ of the cones in accordance with the following relation:

$$Z_{in} = 120\,ln[cot(\theta_0/2)] \quad \text{[ohms]} \qquad (2.103)$$

This relation is presented graphically in figure no. 2.11. Usually, values of angle θ_0 between 30° and 60° are used, where the variation is almost linear.

The main advantages of the biconical dipole over the cylindrical one are the frequency bandwidth which is greater and the fact that input impedance is variable and could be set at the desired value by choosing an appropriate value for the angle θ_0 .

The radiation pattern of the biconical dipole is mainly given the length of its arms, but the vertex angle θ_0 directly influences the beamwidth of the main lobe. For instance, for $\theta_0 = 30°$, the main lobe beamwidth of the half wave biconical dipole is about 100° .

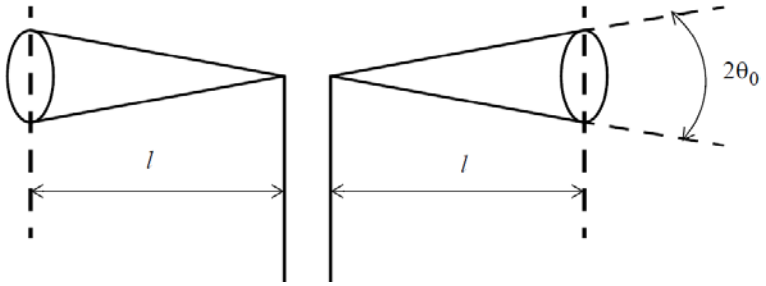


Figure no. 2.10 – *Geometry of biconical dipole*

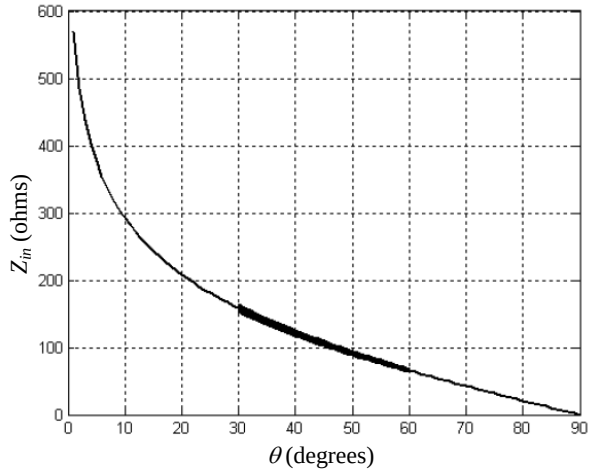


Figure no. 2.11 – *Input impedance versus cone vertex semi-angle*

Folded Dipole

A folded dipole (figure no. 2.12) is a group of cylindrical dipoles connected at both of their ends fed through the center of one of them. The cross section diameter of the dipoles could be different. The radiation pattern of the folded dipole is identical to the ones of the dipoles in the set, but the input impedance is the input impedance of a cylindrical dipole multiplied by the square of the number of dipoles. For instance, if half wave cylindrical dipoles are used to build a folded dipole, then the input impedance of the folded dipole is about $4 \times 75 = 300$ ohms when two dipoles are used, $9 \times 75 = 675$ ohms when three dipoles are used etc. The main advantage of a folded dipole over a cylindrical one is that can be connected directly, without an electrical isolator, to a supporting physical structure in the middle of a non-fed dipole, as this point is already at a zero potential: due to the interconnected at both of their ends, a standing wave voltage regime is established in the dipoles with maxima at the ends and zero value in the center.

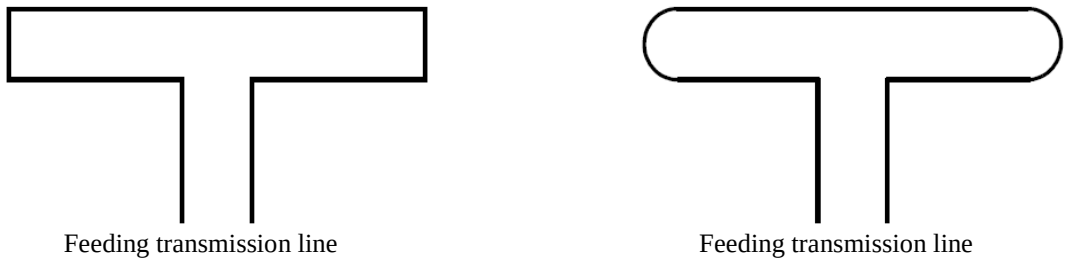


Figure no. 2.12 – *Geometry of folded dipole*

As we showed earlier, a two branch half wave folded dipole has an input impedance of about 300 ohms. In order to have matched interconnection between the folded dipole and the feeding transmission line we use a bifilar symmetrical transmission line with a characteristic impedance of 300 ohms. But, such a transmission line is prone to strong interferences, because the component conductors are not shielded. If a shielded transmission line, like a coaxial cable, should be used, then special circuitry should be introduced in order to realize a matched interconnection between the transmission line (characteristic impedance: 75 ohms) and the folded dipole (input impedance: 300 ohms). Also, this circuitry should realize a transition from the asymmetry of the coaxial cable output to the symmetrical input of the folded dipole. This circuitry, denoted as "balun" (from balanced-unbalanced), is in fact a half wave length coaxial stub connected as in figure no. 2.13: its shield connected with the shield of the feeding coaxial cable and one end of its center conductor connected to the center conductor of the feeding coaxial cable. The half wave length stub transforms the characteristic impedance of 75 ohms of the feeding coaxial cable into an impedance of 300 ohms at the output of the stub. Also, the output of this circuitry is symmetric as the input of the folded dipole is.

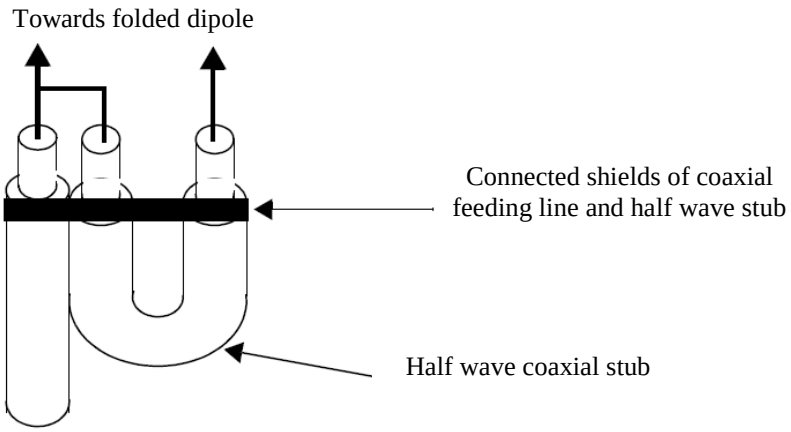


Figure no. 2.13 – *Half wave coaxial stub for matching and asymmetry/symmetry transformation*

Monopole Antenna

The monopole antenna is a wire antenna operating close to a ground plane, perpendicular to it, and fed through the end close to the ground (figure no. 2.14), not through its middle point as a dipole is. If the ground plane is a perfect electrical conductor and has infinite dimensions, then the ensemble of monopole antenna and its image is center fed dipole antenna. Consequently, the radiation pattern of a monopole antenna above the ground plane is identical to the one of a dipole antenna with double length, that is it contains a main lobe as torus with the antenna as axis of symmetry. Obviously, the monopole antenna does

not radiate below the ground plane. Actual ground planes are not perfect conductors and have finite dimensions and, as a result, the main lobe is no longer perpendicular to the antenna, but it approaches the antenna.

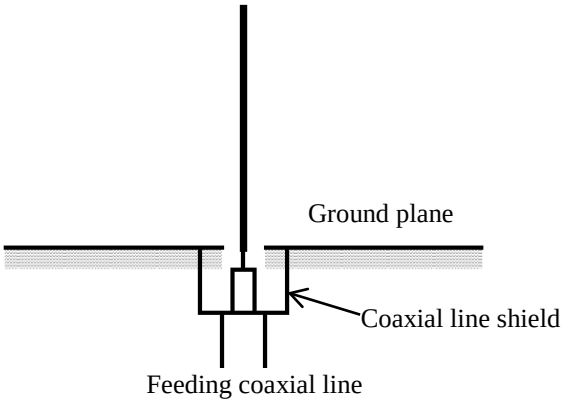
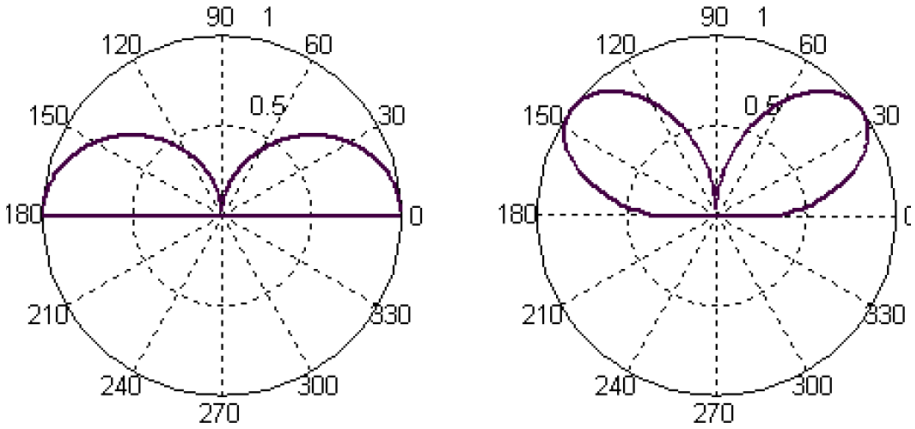


Figure no. 2.14 – *Geometry of monopole antenna*

Monopole antenna with the length equal to $\lambda_0/4$ above an infinite perfect conducting ground plane has the upper half torus of the equivalent dipole antenna as a radiation pattern: the direction of the lobe is $\theta_m = \pi/2$. If the ground plane extends only to a finite distance *d* from the antenna, then the direction of the lobe is smaller the $\pi/2$ (figure no. 2.15) and is given by the following formula:

$$sin\,\theta_m = 1 - \frac{3\lambda_0}{4d} \qquad (2.104)$$



Infinite size ground plane Finite size ground plane

Figure no. 2.15 – *Monopole antenna radiation pattern*

As long as $d \gg \lambda_0$ the field value at the ground plane level is the same fraction from the field in the main lobe direction, irrespective of the value of *d*:

$$\frac{E(\pi/2)}{E(\theta_m)} = 0.428 \qquad (2.105)$$

Disk-conical Antenna

The disk-conical antenna is a modified bi-conical antenna: one of the conical branches is replaced by a conducting disk. The presence of this disk makes the antenna input

impedance remain constant over a greater domain of frequency and thus the antenna frequency bandwidth increases. The disk diameter should be carefully chosen: if it is too small – the main lobe approaches too much the conical branch, if it is too big – the radiated field in the disk plane decreases too much.

Quadrant Antenna

The quadrant antenna is a set of two perpendicular cylindrical dipoles and it has an almost omnidirectional pattern in the plane containing the dipoles (figure no. 2.16). This property remains even if the branches of the dipoles are bended at $90°$. This bending allows for a simpler fixing to a supporting structure.

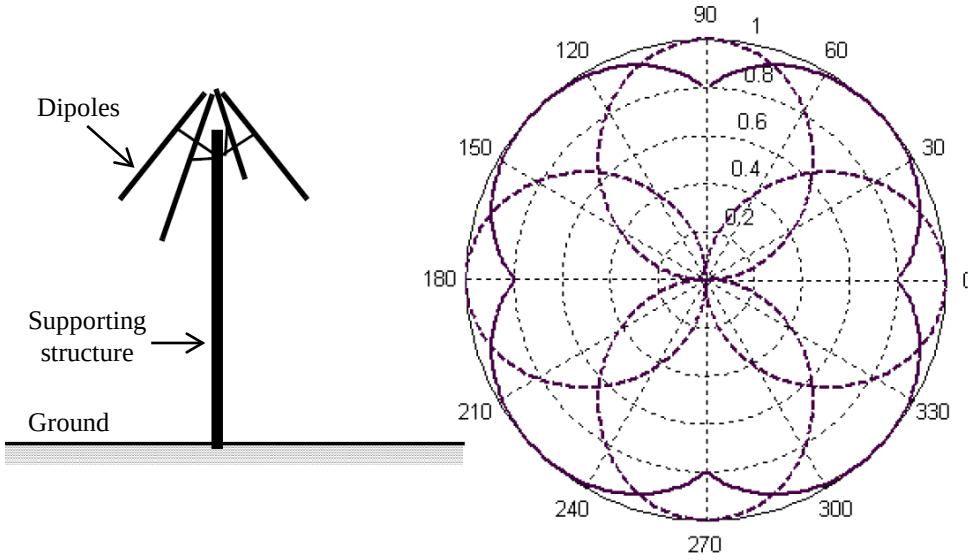


Figure no. 2.16 – *Quadrant antenna: geometry (left) and radiation pattern (right)*

Long-wire Antennas

Long-wire antennas are thin wire antennas with an electrical length *n* greater than 3. When radiating in free space and *n* is an integer, their radiation pattern has the following properties:

- the total number of lobes is equal to *n*;
- each of the lobe is a torus with the antenna as its axis of symmetry;
- for standing wave antennas the lobes are symmetrically distributed, the main lobes are the most close to the antenna and the level of the secondary lobes monotonically decreases with the angle separation from the main lobes until $\theta = \pi/2$; the angle between the main lobe direction and the antenna decreases with *n* is approximately given by the formula: $\cos\theta_{m1} = 1 - 0.371/n$;
- the traveling wave antennas has a single main lobe, the most close to the antenna, and the level of secondary lobes monotonically decreases with the angle of separation from the main lobe until $\theta = \pi$;

Some of the above properties are no longer valid when the electrical length *n* is not an integer.

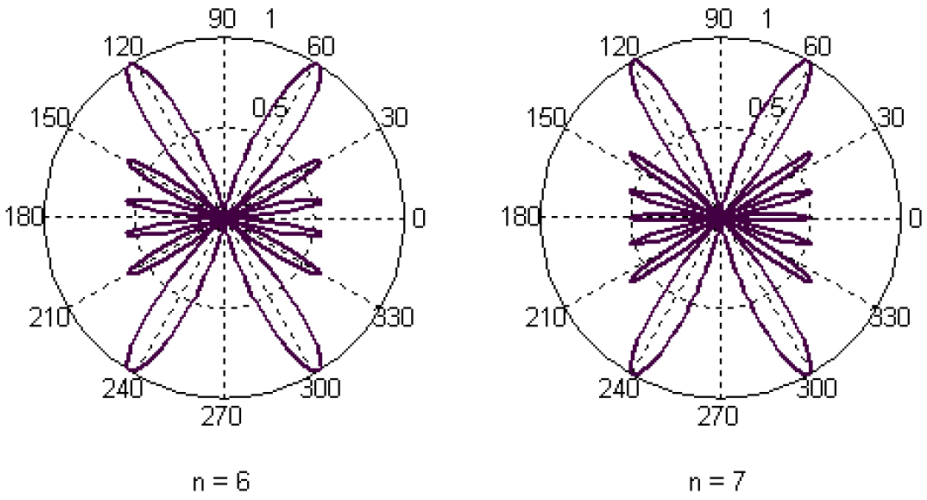


Figure no. 2.17 – *Radiation patterns of standing wave long-wire antenna*

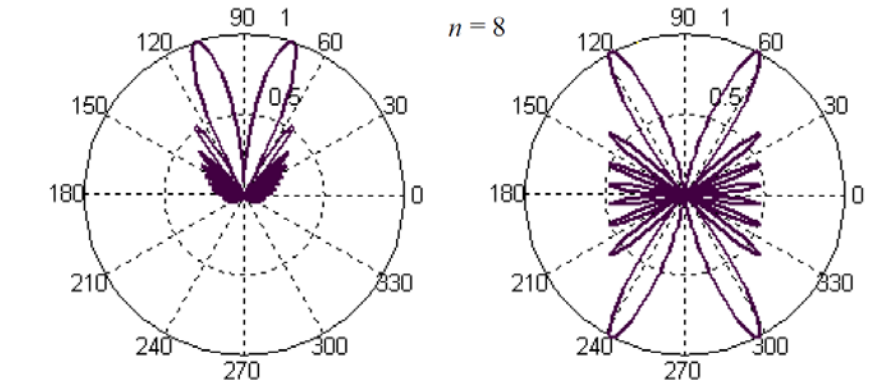


Figure no. 2.18 – *Radiation patterns of long-wire antenna*

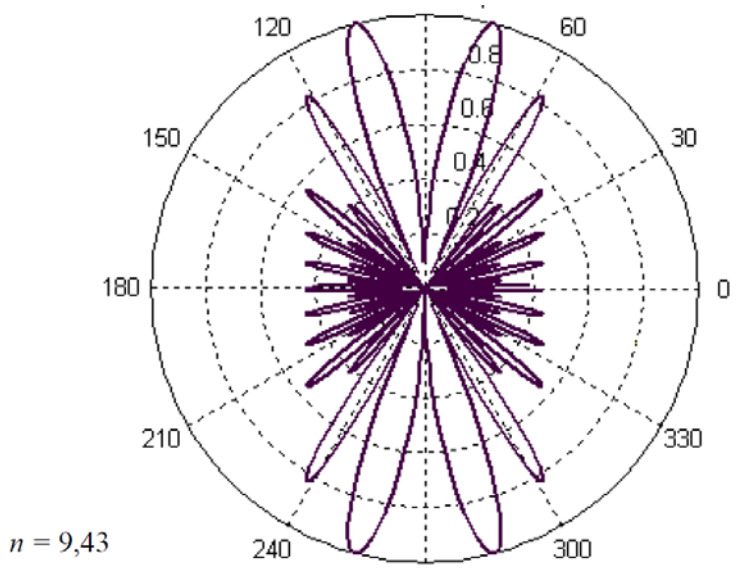


Figure no. 2.19 – *Radiation pattern of standing wave long-wire antenna*

Figure no. 2.17 illustrates the radiation pattern for standing wave long wire antennas with two electrical lengths, figure no. 2.18 illustrates the radiation pattern for standing wave long wire antennas and traveling wave long wire antenna, respectively, with same electrical length, and figure no. 2.19 illustrates the radiation pattern for a long wire antenna with a non-integer electrical length.

Practical long-wire antennas include a great variety of implementations.

Slanted Wire Antenna

The slanted wire antenna (figure no. 2.20) has the feeding end close to a ground plane and, usually, it has an angle tilting of $s = \theta_{m1}$ above the ground plane and thus its main radiation is at the ground plane level in two opposite directions. When $s > \theta_{m1}$ one of the main lobes is directed towards ionosphere, which could be a useful direction in some applications. In order to avoid interferences the tilting angle should be kept as small as possible, such that secondary lobes not be directed along or above the ground plane (figure no. 2.21).

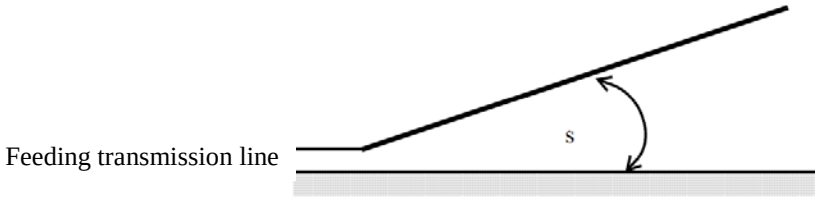
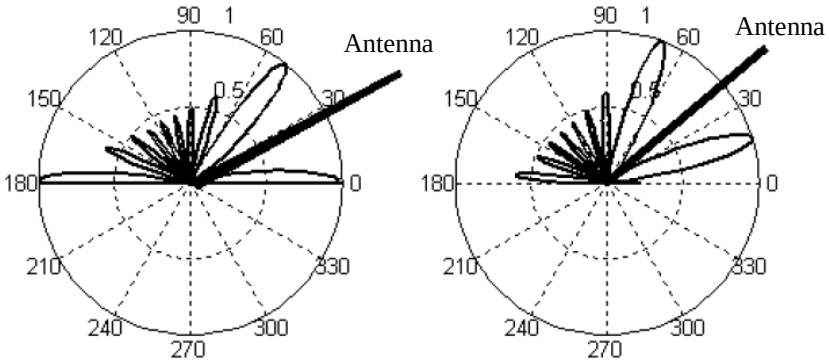


Figure no. 2.20 – *Geometry of slanted wire antenna*



$s = \theta_{m1}$ $s > \theta_{m1}$

Figure no. 2.21 – *Radiation pattern of slanted wire antenna*

Semirhombic Antenna

The semirhombic antenna is a series of two identical long-wire antennas, tilted by an appropriate angle, and radiating above a ground plane (figure no. 2.22). Its name is due to the fact that it constitutes a rhomb together with its image against the ground plane. Usually, it is terminated on matched impedance and a travelling wave current is established through it.

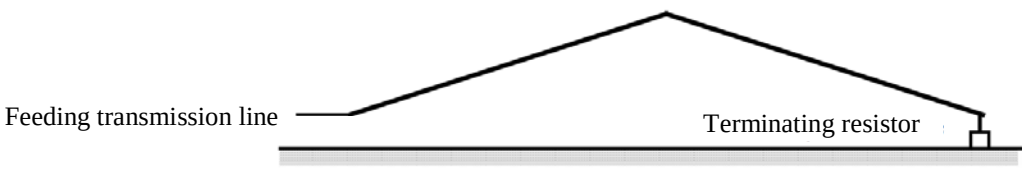


Figure no. 2.22 – *Geometry of semirhombic antenna*

Horizontal V Antenna

The horizontal V antenna is a group of two identical long-wire antennas in a horizontal plane, making an appropriate angle between them. It radiates in free space or close to a horizontal ground plane.

When no matching resistance is used a standing wave regime is established through the component antennas and a maximum directivity of:

Slanted V antenna

The tilted V antenna is a group of two identical long-wire antennas making an appropriate angle between them and placed in a tilted plane above the ground plane (figure no. 2.23). It is fed through the end at higher distance from the ground plane and terminated on matching resistance at the ends close to the ground plane. A traveling wave regime is installed through the component antennas.

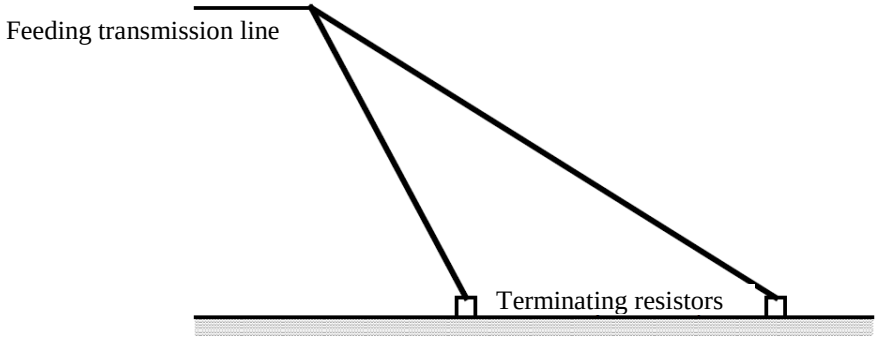


Figure no. 2.23 – *Geometry of slanted V antenna*

Rhombic Antenna

The rhombic antenna (figure no. 2.24) is one of the most used long-wire antennas. It operates both in free space and close to a ground plane. It could be considered as a series of two V antennas or as a symmetric transmission line whose conductors are separated by quite a large distance in order to radiate. The rhombic antenna is always terminated on a matched resistance and a traveling wave regime is established.

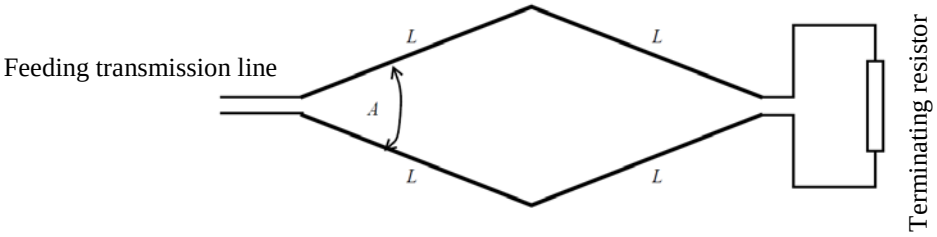


Figure no. 2.24 – *Geometry of rhombic antenna*

The basic constructive parameters are the length L (usually much greater than λ_0) of each of the branches and the acute angle A between the branches. If $A < 2\theta_{m1}$ the main lobe of the radiation pattern lies in the axial plane perpendicular to the antenna. If $A = 2\theta_{m1}$ the main lobe is along the rhombus great axis. The case with $A > 2\theta_{m1}$ is avoided as it allows for two main lobes in the plane of the antenna.

Because the angle θ_{m1} decreases when the electrical length L/λ_0 increases, the requirement $A \leq 2\theta_{m1}$ represents in fact a limitation of the maximum frequency of the bandwidth.

The level of the secondary lobes increases with the tilting of the main lobe above the plane of the antenna and this one increases when the electrical length decreases. The requirement to maintain the level of the secondary lobes under some desired threshold represents in fact a limitation of the minimum frequency of the bandwidth.

The *input impedance* of the rhombic antenna has a quite small reactive part, but a quite big active part: 750 – 800 ohms, impossible to match the characteristic impedance of usual transmission lines. Using 3 parallel wires for each of the branches decreases this value at about 600 ohms and the matching condition could be reached by using special transmission lines.

A matched *terminal resistance* is used at the end opposite to the feeding point for most of the practical rhombic antennas in order to dissipate the input power not radiated by the antenna and to establish a traveling wave regime through the antenna. The matching condition should be fulfilled with great accuracy, because the power reflected due to mismatching yields a standing wave component through the antenna and a disturbing back lobe appears in the radiation pattern. For moderate values of the input power a physical resistor is used, but for big values of the input power a physical resistor should present quite a large size. For this situation a lossy transmission line (usually made of stainless steel) or a combination of lossy transmission line and physical resistor could present a more convenient size. When the terminal resistance dissipates a big fraction of the input power, the efficiency of the rhombic antenna is very small. It could be increased by replacing the terminal resistance with an active circuit that could bring the fraction of power not radiated by the antenna to its input. Unfortunately, this active circuit is usually of limited bandwidth and it compromises one of the main parameter of the rhombic antenna: its large bandwidth as compared to most of the antennas of similar size.

Franklin Antenna

The Franklin antenna is a long-wire antenna with a modified current distribution, aiming at obtaining a pure omnidirectional radiation pattern without side lobes.

The necessary change in the distribution current results from noting that the number of lobes in the radiation pattern of a standing wave long-wire antenna equals the number of phase sign changes in its current distribution (figure no. 2.25a). Inserting current resonating circuit elements, which introduce a phase shift of π, in points where the phase sign changes, one could maintain the same current phase sign along all the antenna (figure no. 2.25b). These resonating circuit elements could be, for instance, short circuited transmission line stubs with length equal to $\lambda_0/4$.

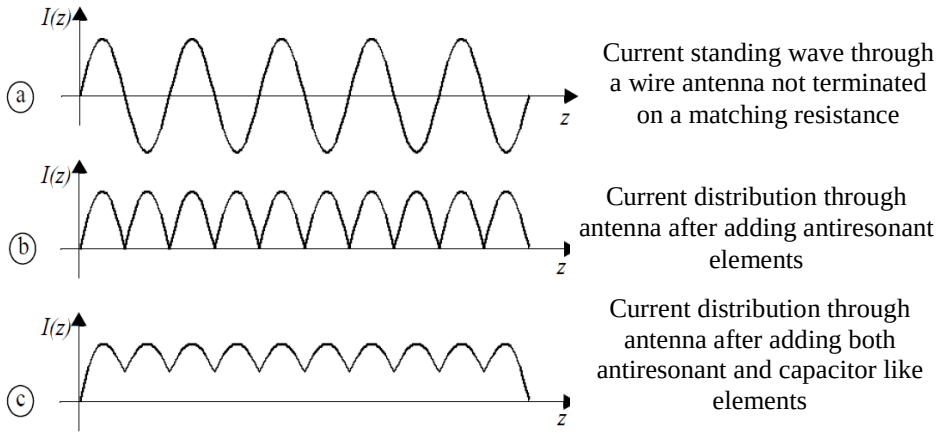


Figure no. 2.25 – *Modified current distribution through Franklin antenna*

The zero values in the current distribution could be eliminated by introducing energy storing circuit elements (like capacitors) in the same points. This way the current distribution along the antenna becomes more close to a desired constant wave (figure no. 2.25c) that generates an omnidirectional radiation pattern without side lobes.

A practical implementation of the concept of Franklin antenna is the *zig-zag antenna* (figure no. 2.26). The effect of both the anti-resonant circuit and the capacitor is obtained by placing the antenna segments at rigorously chosen distances. The *x*, *y*, and *z* segments' length are precisely computed based on the operating frequency.

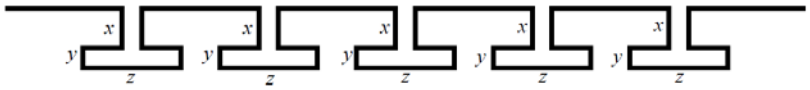


Figure no. 2.26 – *Geometry of zig-zag antenna*

The bandwidth of Franklin antenna is small because its operation is based on the resonant properties of the constituent elements.

2.7 – Loop Antenna and Helix Antenna

These antennas are practical implementations of the concept of magnetic dipole (as opposed to the antennas we previously talked about which implement in practice the concept of electric dipole).

Loop Antenna

The loop antenna is a radiative toroidal coil made of one or more turns in air or laid on a ferrite core. The ferrite core significantly increases the antenna gain as it increases the density of the magnetic field lines inside the coil. The radiation pattern of the loop antenna is identical to the one of an equivalent electric dipole that is perpendicular on the loop plane. The loop antenna is very useful when used in the receiving mode of operation because it is sensitive to the magnetic component of an electromagnetic field and this one is less prone to local interferences which are electric in nature in most of the cases.

Helix Antenna

A helix antenna is a helix on the surface of a dielectric cylinder and made of a cylindrical or a band conductor (figure no. 2.27). The cylindrical supporting surface could be a physical one, but it could miss if the helix is rigid enough to support itself. The feeding point is the end of the helix which is the closest to a ground surface; the other end of the helix is free.

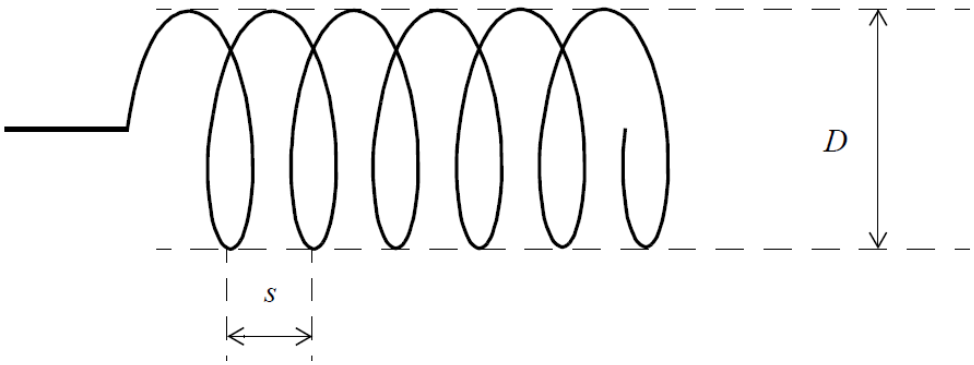


Figure no. 2.27 – *Geometry of helix antenna*

The geometry of a helix antenna is described by the following parameters:
- n – number of turns;
- D – diameter of the supporting cylinder;
- d – diameter of the cylindrical conductor or width of the conductive band;
- s – step of the spiral;
- $C = \pi D$ – circumference of the supporting cylinder;
- L – length of a turn;
- $\psi = \text{atan}(s/C)$ – slope of turns.

There are two fundamental modes of operation for the helix antenna: the normal mode and the axial mode.

The *normal mode* of operation establishes for $C/\lambda_0 < 0.5$ and it typical for low frequencies. In this mode the helix antenna is basically a shorted monopole antenna. The current distribution along the helix antenna is similar to the one along a monopole antenna with length *nL*, while the radiation pattern consists in a torus around the helix: the maximum radiation is in the plane perpendicular to the helix and a null occurs in the axial direction. The frequency bandwidth is smaller than the one of a monopole antenna with the same length, but the gain could be much greater if the resonance condition $nL \approx \lambda_0/4$ is fulfilled.

There are few applications of the helix antenna in its normal mode of operation because its parameters are similar to the ones of the monopole antenna with the same length, which is much simpler to build. An illustrative application is the helix antenna at transmission, when it operates in the presence of a metallic supporting cylinder not connected electrically to the antenna. Due to the interaction between the antenna and the close metallic support the torus of the radiation pattern is quite narrow and there is no side lobe.

The *axial mode* of operation establishes for a quite narrow frequency bandwidth where the condition $0.75 < C/\lambda_0 < 1.25$ is fulfilled and if the angle ψ remains in a specified domain. For an adequate design the radiated field is circularly polarized and, for an optimal length, the maximum gain occurs in the axial direction and remains constant in the entire frequency bandwidth. The input impedance is almost resistive and it remains constant in the entire frequency bandwidth.

There are few details in the literature about a complete design of a helix antenna and intense simulation studies should be conducted in order to obtain adequate results. But the antenna parameters do not significantly change for slight changes in the size of physical structure, so even a brief design could yield good results.

The following formulas could be used for obtaining the main parameters of a helix antenna:
- input impedance (almost resistive): $R_{in} = 140\, C/\lambda_0$ [ohms];
- 3 dB beamwidth: $\theta_{3dB} = \frac{52}{(C/\lambda_0)\sqrt{ns/\lambda_0}}$ [degrees];
- axial maximum directivity: $D_{max} = 15(C/\lambda_0)^2 ns/\lambda_0$;
- gain: 10 – 15 [dBi] (typically);
- side lobes' level: smaller than –10dB
- $12° < \psi_{optimal} < 14°$.

The helix antenna is a high bandwidth antenna. It radiates a circularly polarized field in the axial direction. A linear polarized field could be produced by using two identical helix antennas with opposite sense turns and fed in parallel or in series (figure no. 2.28).

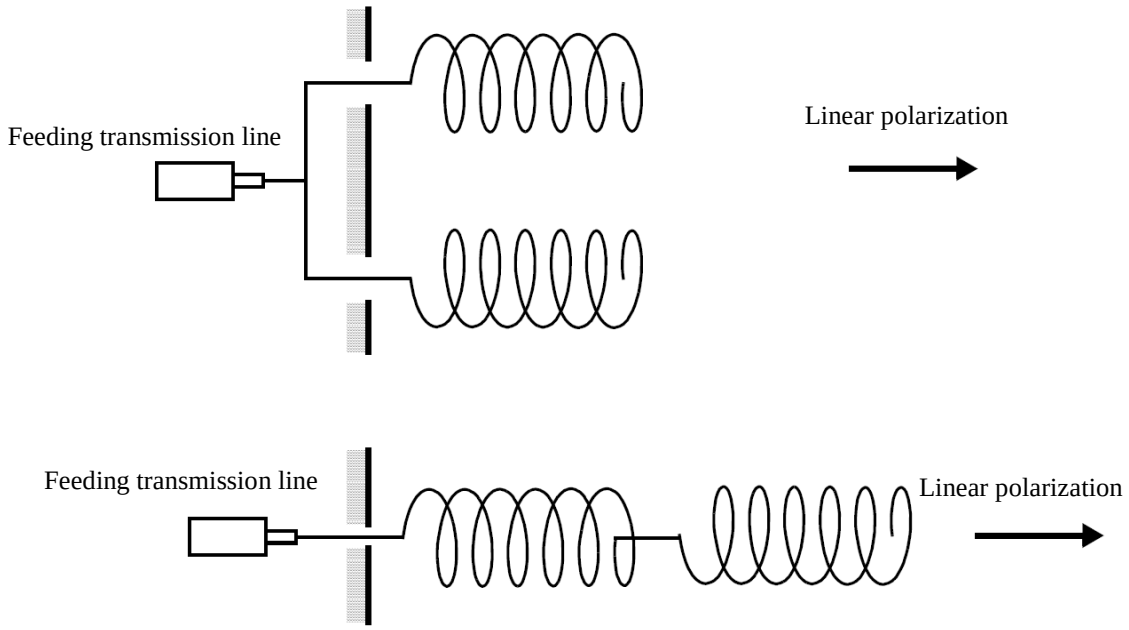


Figure no. 2.28 – *Groups of helix antennas radiating linear polarized field*

C h a p t e r I I I

RADIATION FROM APERTURES

An aperture is an opening (window) realized in surface. Radiation produced in free space by an existing electromagnetic field in the aperture depends on the field distribution throughout the aperture and on the shape and the electric properties of the surface the aperture belongs to. Simpler approximate formulas for the radiated field are obtained for aperture dimensions much greater than the field wavelength, because the field on the surface containing the aperture is almost zero, except for the aperture itself.

3.1 – Radiation of Rectangular Aperture in a Perfect Conducting Infinite Plane Surface

Let us consider a rectangular aperture of size $2a \times 2b$ situated in the *xOy* plane of Cartesian coordinate system. The plane is infinite in both of its dimensions and is considered to be a perfect conductor (conductivity is infinite). The symmetry center of the aperture is the center of the coordinate system (see figure no. 3.1). Let's denote by $(\boldsymbol{E_a}, \boldsymbol{H_a})$ the field distribution in the aperture. We shall show that the free space radiated field in the half space $z > 0$ is completely determined by the field distribution in the aperture.

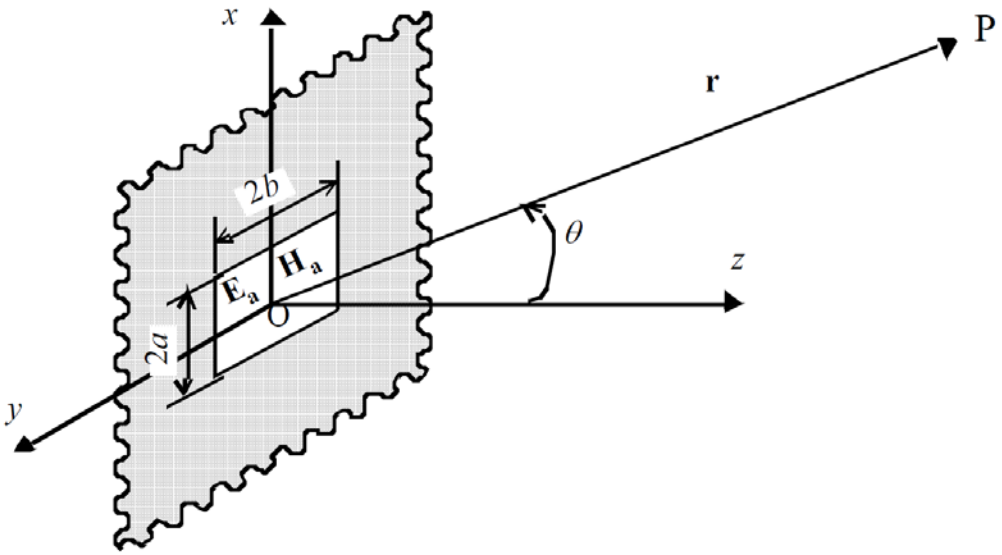


Figure no. 3.1 – *Rectangular aperture in perfect conducting infinite plane surface*

The current density $\boldsymbol{J}$ in the half space $z > 0$ is zero, because there is no radiation source besides the aperture field. Thus, the wave equation in this space region is a homogeneous one:

$$\nabla^2 \boldsymbol{E} + k_0^2 \boldsymbol{E} = 0 \qquad (3.1)$$

As we noted previously, the vectorial function $\boldsymbol{E}$ depends on time (t) and position in space ($\boldsymbol{r}$ – position vector), that is it depends on four independent variables: $\boldsymbol{E}(\boldsymbol{r}, t) = \boldsymbol{E}(x, y, z, t)$.

Also, the spatial electric charge density $\rho = 0$ in the half space $z > 0$ and so, the corresponding Maxwell equation is:

$$\nabla \cdot \boldsymbol{E} = 0 \tag{3.2}$$

Due to this condition the electric field is said to be solenoidal.

A solution for the wave equation (3.1) is easier to find in the bi-dimensional Fourier transform domain. Usually, in some specified conditions, an one-dimensional Fourier transformation associates a frequency dependent function $F(\omega)$ to a time dependent function $f(t)$, based on the following relation:

$$F(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} f(t)e^{j\omega t}dt \tag{3.3}$$

$F(\omega)$ is denoted as the image or the Fourier transform of $f(t)$, while $f(t)$ is denoted as the original function.

Similarly, a function $F(k_x, k_y)$ could be associated to a spatial variable dependent function $f(x, y)$, based on the relation:

$$F(k_x, k_y) = \int_{-\infty}^{\infty} f(x, y)e^{(jk_x x + jk_y y)}dxdy \tag{3.4}$$

The variables k_x, k_y are denoted as *spatial frequencies*.

Dropping the dependence on time an image function $\boldsymbol{E}(k_x, k_y, z)$ could be associated to the original function $\boldsymbol{E}(x, y, z)$ by means of a bi-dimensional Fourier transform:

$$\boldsymbol{E}(k_x, k_y, z) = \int_{-\infty}^{\infty} \boldsymbol{E}(x, y, z)e^{(jk_x x + jk_y y)}dxdy \tag{3.5}$$

Once the image function is determined, the original one is computed by inverse bi-dimensional Fourier transformation:

$$\boldsymbol{E}(x, y, z) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty} \boldsymbol{E}(k_x, k_y, z)e^{-(jk_x x + jk_y y)}dk_x dk_y \tag{3.6}$$

The bi-dimensional Fourier transformation enjoys similar properties as the one-dimensional one. For instance, the derivation after a spatial variable (x, for instance) in the original domain is replaced by multiplication with the corresponding spatial frequency (jk_x – in this case) in the Fourier domain. Based on this property, the wave equation in the space (3.1) becomes in the Fourier domain as follows:

$$-k_x^2\boldsymbol{E}(k_x, k_y, z) - k_y^2\boldsymbol{E}(k_x, k_y, z) + \frac{\partial^2 \boldsymbol{E}(k_x, k_y, z)}{\partial z^2} + k_0^2\boldsymbol{E}(k_x, k_y, z) = 0 \tag{3.7}$$

By using the following notation:

$$k_z^2 \triangleq k_0^2 - k_x^2 - k_y^2 \tag{3.8}$$

equation (3.7) could be written as:

$$\frac{\partial^2 \boldsymbol{E}(k_x, k_y, z)}{\partial z^2} + k_z^2\boldsymbol{E}(k_x, k_y, z) = 0 \tag{3.9}$$

When the variable k_z defined through the notation (3.8) is real, the equation (3.9) has a single solution that fulfills the Sommerfeld conditions and this is the following:

$$\boldsymbol{E}(k_x, k_y, z) = \boldsymbol{f}(k_x, k_y)e^{-jk_z z} \tag{3.10}$$

where $\boldsymbol{f}(k_x, k_y)$ is an integration constant (it should be a vector, because the unknown variable in the equation (3.9) is a vector !).

By using this solution, the formula (3.6) for the electric field in becomes as follows:

$$\boldsymbol{E}(x,y,z) = \frac{1}{(2\pi)^2}\int_{-\infty}^{\infty}\boldsymbol{f}(k_x,k_y)e^{-jk_z z}e^{-(jk_x x+jk_y y)}dk_x dk_y =$$

$$= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\boldsymbol{f}(k_x,k_y)e^{-(jk_x x+jk_y y+jk_z z)}dk_x dk_y =$$

$$= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\boldsymbol{f}(k_x,k_y)e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}dk_x dk_y \qquad (3.11)$$

This relation shows that the radiated electric field $\boldsymbol{E}(x,y,z)$ is a superposition of infinite plane waves of type $\boldsymbol{f}e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}$. We denoted by **k** a vector having the components (k_x, k_y, k_z) in the considered Cartesian coordinate system *Oxyz*. Due to the notation (3.8) the modulus of **k** is equal to k_0, which is the free space propagation constant. This is why **k** is denoted as *propagation vector*.

The requirement for variable k_z to be real means:

$$k_0^2 - k_x^2 - k_y^2 > 0 \quad \Rightarrow \quad k_x^2 + k_y^2 < k_0^2 \qquad (3.12)$$

This inequality defines a finite area in the domain of spatial frequencies (k_x, k_y), denoted as *visible domain*. The power radiated by the aperture field is associated only to spatial frequencies included in the visible domain.

The property (3.2) has the following consequence upon the integration constant $\boldsymbol{f}(k_x, k_y)$ (we drop the arguments for simplicity of the relations):

$$\nabla\cdot\boldsymbol{E} = 0 \quad \Rightarrow \quad \nabla\cdot\left(\boldsymbol{f}e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}\right) = 0 \quad \Rightarrow \quad e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}\nabla\cdot\boldsymbol{f} + \boldsymbol{f}\cdot\nabla e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} = 0 \quad \Rightarrow$$

$$\Rightarrow \quad \boldsymbol{f}\cdot\nabla e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} = 0 \quad \Rightarrow \quad \boldsymbol{f}\cdot(-\mathrm{j}\mathbf{k})e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} = 0 \Rightarrow \quad \boldsymbol{k}\cdot\boldsymbol{f} = 0 \Leftrightarrow \qquad (3.13)$$

$$\Leftrightarrow \quad k_x f_x + k_y f_y + k_z f_z = 0$$

In developing the above relations we took into account that $\boldsymbol{f}$ is dependent on variables (k_x, k_y) and, hence, it is independent from the variables (x, y, z) that define the derivation operator ∇. As a consequence $\nabla\cdot\boldsymbol{f} = 0$. Also, in the spherical coordinated system (r, θ, ϕ) associated to the Cartesian one *Oxyz*, the exponential function $e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}$ has a single argument (r) and, as a consequence, its gradient contains only one derivative $\nabla e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} = (-\mathrm{j}\mathbf{k})e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}$.

The result (3.13) shows that the components of $\boldsymbol{f}$ are not independent of each other and it substantiates our previous statement that *the radiated field is <u>completely</u> determined by the field distribution in the aperture.*

We denote the components of variables in the aperture plane *xOy* as tangential components. We have: $\boldsymbol{E_t} \triangleq E_x\hat{\boldsymbol{x}} + E_y\hat{\boldsymbol{y}}$, $\boldsymbol{f_t} \triangleq f_x\hat{\boldsymbol{x}} + f_y\hat{\boldsymbol{y}}$, and $\boldsymbol{k_t} \triangleq k_x\hat{\boldsymbol{x}} + k_y\hat{\boldsymbol{y}}$. Also, $\boldsymbol{E} = \boldsymbol{E_t} + E_z\hat{\boldsymbol{z}}$, $\boldsymbol{f} = \boldsymbol{f_t} + f_z\hat{\boldsymbol{z}}$, and $\boldsymbol{k} = \boldsymbol{k_t} + k_z\hat{\boldsymbol{z}}$.

For $z \to 0$ the radiated field (3.11) should reach the aperture field $\boldsymbol{E_a}$. So, we have:

$$\boldsymbol{E}(x,y,0) = \boldsymbol{E_t}(x,y) = \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\boldsymbol{f}(k_x,k_y)e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}dk_x dk_y\,|_{z=0} =$$

$$= \frac{1}{4\pi^2}\int_{-\infty}^{\infty}\boldsymbol{f_t}(k_x,k_y)e^{-(jk_x x+jk_y y)}dk_x dk_y = \boldsymbol{E_a}(x,y) \qquad (3.14)$$

The above relation shows that aperture field $\boldsymbol{E_a}$ is the bi-dimensional inverse Fourier transform of the tangential component $\boldsymbol{f_t}$ of the integration constant $\boldsymbol{f}$. Conversely, the tangential component $\boldsymbol{f_t}$ of the integration constant $\boldsymbol{f}$ is the bi-dimensional Fourier transform of the aperture field $\boldsymbol{E_a}$:

$$\boldsymbol{f_t}(k_x, k_y) = \int_{-\infty}^{\infty} \boldsymbol{E_a}(x,y) e^{(jk_x x + jk_y y)} dxdy \qquad (3.15)$$

The third component of the integration constant $\boldsymbol{f}$ is obtained from the condition (3.13):

$$f_z = -\boldsymbol{k_t} \cdot \boldsymbol{f_t}/k_z \qquad (3.16)$$

Summarizing: knowing the aperture field $\boldsymbol{E_a}$, the tangential component $\boldsymbol{f_t}$ of the integration constant is computed by means of bi-dimensional Fourier transformation (3.15), then the third component of $\boldsymbol{f}$ is computed with (3.16), and, finally, the radiated field $\boldsymbol{E}$ is computed by means of bi-dimensional inverse Fourier transformation (3.11).

The formula for the magnetic component $\boldsymbol{H}$ of the radiated field is obtained from the Maxwell equation, taking into account that there is no spatial electric charge in the half space $z > 0$. We have:

$$\nabla \times \boldsymbol{E} = -j\omega\mu_0 \boldsymbol{H} \quad \Rightarrow \quad \boldsymbol{H} = -\frac{1}{j\omega\mu_0} \nabla \times \boldsymbol{E} =$$

$$= -\frac{1}{j\omega\mu_0} \nabla \times \left[\frac{1}{4\pi^2} \int_{-\infty}^{\infty} \boldsymbol{f}(k_x, k_y) e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} dk_x dk_y \right] =$$

$$= \frac{1}{4\pi^2 j\omega\mu_0} \int_{-\infty}^{\infty} \nabla \times \boldsymbol{f}(k_x, k_y) e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} dk_x dk_y \qquad (3.17)$$

Interchanging the order of applying the operations in the above development was possible because the integration and the curl operation are defined in independent variable systems: (k_x, k_y, k_z) for integration and (x, y, z) for curl operator. For the same reason the vector $\boldsymbol{f}$ is a constant for curl operator. Thus:

$$\nabla \times \boldsymbol{f} e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} = e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} (\nabla \times \boldsymbol{f}) + \left(\nabla e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}\right) \times \boldsymbol{f} =$$

$$= \left(\nabla e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}\right) \times \boldsymbol{f} = \left(-j\boldsymbol{k} e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}}\right) \times \boldsymbol{f} = -j(\boldsymbol{k} \times \boldsymbol{f}) e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} \qquad (3.18)$$

and so:

$$\boldsymbol{H} = \frac{1}{4\pi^2 k_0 \eta_0} \int_{-\infty}^{\infty} (\boldsymbol{k} \times \boldsymbol{f}) e^{-\mathrm{j}\mathbf{k}\cdot\boldsymbol{r}} dk_x dk_y \qquad (3.19)$$

3.2 – Field Equivalence Principle

This principle allows for replacing a volumetric current distribution by a current distribution on the surface that limits that volume. *The equivalence is valid from the point of view of the electromagnetic field outside the volume*. Applying this principle does not imply that a solution is found for a particular situation, but it could suggest a possible way to find an approximate solution, extremely useful when a closed solution could not be found.

Let's consider a volumetric electric $\boldsymbol{J}$ and a magnetic $\boldsymbol{J_m}$ current distribution in a volume V limited by closed surface Σ (figure no. 3.2a). This current distribution radiates an electromagnetic field $(\boldsymbol{E}, \boldsymbol{H})$ satisfying the Maxwell equations.

We eliminate the radiating sources from the volume V and we postulate the existence of the same $(\boldsymbol{E}, \boldsymbol{H})$ outside the volume V and an electromagnetic field $(\boldsymbol{E_1}, \boldsymbol{H_1})$ inside the volume V (figure no. 3.2b). For this situation to comply with the Maxwell equations it is necessary to postulate, also, the existence of an electric and magnetic current distribution

$(\boldsymbol{J_s}, \boldsymbol{J_{ms}})$ on the surface Σ that accounts for the field discontinuity at this surface. According to equations (1.62) it is necessary that:

$$\boldsymbol{J_s} = \hat{\boldsymbol{n}} \times (\boldsymbol{H} - \boldsymbol{H_1})$$
$$\boldsymbol{J_{ms}} = -\hat{\boldsymbol{n}} \times (\boldsymbol{E} - \boldsymbol{E_1}) \qquad (3.20)$$

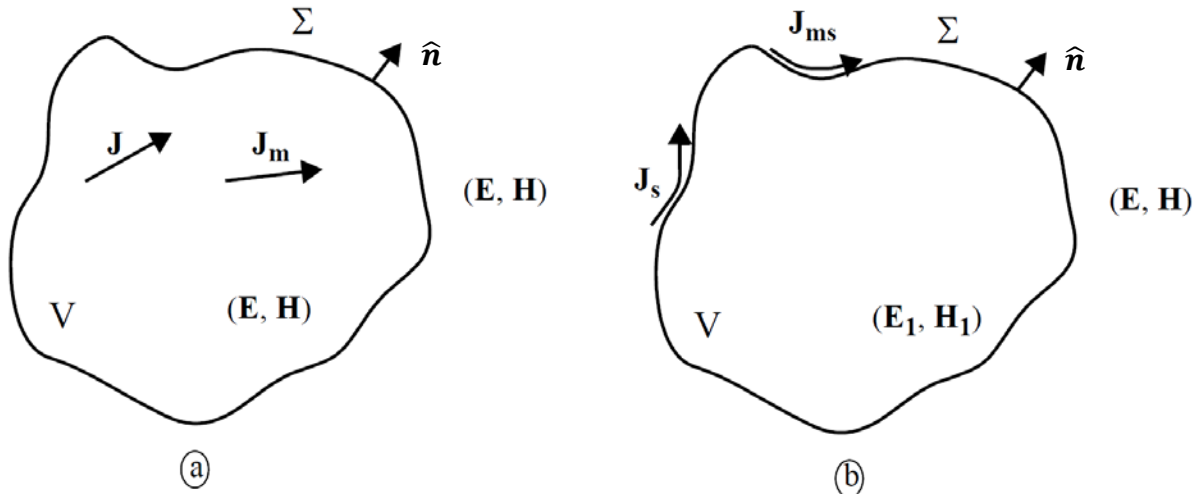


Figure no. 3.2 – *Field equivalence principle*

Remind that this surface current distribution is equivalent to the volumetric one, only from the point of view of the field distribution outside of the volume V. This limitation is not a restrictive requirement in antenna theory, because the region of interest is the one very far away from the radiation source.

The general form (3.20) of the field equivalence principle may take particular forms in particular situations. For instance, if we postulate a zero field distribution inside the volume V we obtain the *Love principle*:

$$\boldsymbol{J_s} = \hat{\boldsymbol{n}} \times \boldsymbol{H}$$
$$\boldsymbol{J_{ms}} = -\hat{\boldsymbol{n}} \times \boldsymbol{E} \qquad (3.21)$$

Love principle could be further simplified. Because the electric current distribution inside the volume V, we could introduce a perfect electric conducting surface infinitesimally close to the surface Σ without modifying the electric field distribution inside the volume V. But, in the presence of this perfect electric surface, the electric current density $\boldsymbol{J}$ is zero and the volumetric electric and magnetic current density inside the volume V is equivalent only to the surface magnetic current density:

$$\boldsymbol{J_{ms}} = -\hat{\boldsymbol{n}} \times \boldsymbol{E} \qquad (3.22)$$

On the other side, because the magnetic current distribution inside the volume V, we could introduce a perfect magnetic conducting surface closed to the surface Σ without modifying the magnetic field distribution inside the volume V. But, in the presence of this perfect magnetic surface, the magnetic current density $\boldsymbol{J_m}$ is zero and the volumetric electric and magnetic current density inside the volume V is equivalent only to the surface electric current density:

$$\boldsymbol{J_s} = \hat{\boldsymbol{n}} \times \boldsymbol{H} \qquad (3.23)$$

3.3 – Applying Field Equivalence Principle to Aperture Radiation

Taking into account the radiating aperture configuration in figure no. 3.1, we notice the electromagnetic field in the half space $z < 0$ is zero. Thus, we could introduce a perfect electric conducting surface infinitesimally close to the aperture and, according to the Love principle, the field distribution inside the rectangular aperture is equivalent to a magnetic current distribution on this surface given by:

$$\boldsymbol{J_{ms}} = -\hat{\boldsymbol{z}} \times \boldsymbol{E_a} \tag{3.24}$$

By adding this perfect conducting surface, the entire plane *xOy* becomes a perfect conducting surface. According to the image principle, the radiation of a magnetic current density $\boldsymbol{J_{ms}}$ infinitesimally close to this surface is equivalent, in the half plane $z > 0$, with the radiation of a magnetic current density $2\boldsymbol{J_{ms}}$ in free space.

We proved in paragraph (2.3) that an arbitrary electric current density $\boldsymbol{J}(\boldsymbol{r'})$ inside a volume V radiates in the radiation region an electromagnetic field with a magnetic component:

$$\boldsymbol{H}(\boldsymbol{r}) = \frac{jk_0}{4\pi}\frac{e^{-jk_0 r}}{r}\int_V (\boldsymbol{J}(\boldsymbol{r'}) \times \hat{\boldsymbol{r}}) e^{jk_0 \hat{\boldsymbol{r}}\cdot\boldsymbol{r'}} dV \tag{3.25}$$

in the radiation region. According to the duality theorem, a magnetic current density $2\boldsymbol{J_{ms}}$ radiates in the radiation region an electromagnetic field with an electric component:

$$\boldsymbol{E}(\boldsymbol{r}) = -\frac{jk_0}{4\pi}\frac{e^{-jk_0 r}}{r}\int_V (2\boldsymbol{J_{ms}}(\boldsymbol{r'}) \times \hat{\boldsymbol{r}}) e^{jk_0 \hat{\boldsymbol{r}}\cdot\boldsymbol{r'}} dV \tag{3.26}$$

Taking into account the particular shape of the half space $z < 0$, we have:

$$\begin{aligned}\boldsymbol{E}(\boldsymbol{r}) &= -\frac{jk_0}{4\pi}\frac{e^{-jk_0 r}}{r}\int_{S_a} (2\boldsymbol{J_{ms}}(\boldsymbol{r'}) \times \hat{\boldsymbol{r}}) e^{jk_0 \hat{\boldsymbol{r}}\cdot\boldsymbol{r'}} d\sigma = \\ &= \frac{jk_0}{2\pi}\frac{e^{-jk_0 r}}{r}\int_{S_a} [(\hat{\boldsymbol{z}} \times \boldsymbol{E_a}) \times \hat{\boldsymbol{r}}] e^{jk_0 \hat{\boldsymbol{r}}\cdot\boldsymbol{r'}} d\sigma\end{aligned} \tag{3.27}$$

But:

$$(\hat{\boldsymbol{z}} \times \boldsymbol{E_a}) \times \hat{\boldsymbol{r}} = (\hat{\boldsymbol{r}} \cdot \hat{\boldsymbol{z}})\boldsymbol{E_a} - (\hat{\boldsymbol{r}} \cdot \boldsymbol{E_a})\hat{\boldsymbol{z}} = \boldsymbol{E_a}\cos\theta - (\hat{\boldsymbol{r}} \cdot \boldsymbol{E_a})\hat{\boldsymbol{z}} \tag{3.28}$$

while:

$$\hat{\boldsymbol{r}} \cdot \boldsymbol{r'} = r' \sin\theta \tag{3.29}$$

Hence:

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{jk_0}{2\pi}\frac{e^{-jk_0 r}}{r}\left[\cos\theta \int_{S_a} \boldsymbol{E_a} e^{jk_0 r' \sin\theta} dxdy - \left(\hat{\boldsymbol{r}} \cdot \int_{S_a} \boldsymbol{E_a} e^{jk_0 r' \sin\theta} dxdy\right)\hat{\boldsymbol{z}}\right] \tag{3.30}$$

The surface integral in the above relation could be extended over the entire plane *xOy* without changing the result, because the electric field is zero outside the aperture S_a. After this extension to infinite the integral becomes similar to the bi-dimensional Fourier transform definition (3.5), except for the transformation kernel $e^{-(jk_x x + jk_y y)}$ that is replaced by $e^{jk_0 r' \sin\theta}$. This means that we could the result of the integration as the bi-dimensional transform of the aperture electric field $\boldsymbol{E_a}$. We proved earlier that this transform is the tangential component $\boldsymbol{f_t}$ of the integration (vectorial) constant $\boldsymbol{f}$. Hence:

$$\boldsymbol{E}(\boldsymbol{r}) = \frac{jk_0}{2\pi}\frac{e^{-jk_0 r}}{r}[\cos\theta\, \boldsymbol{f_t}(k_0 r' \sin\theta) - \hat{\boldsymbol{r}} \cdot \boldsymbol{f_t}(k_0 r' \sin\theta)]\hat{\boldsymbol{z}} \tag{3.31}$$

This expression could be simplified further if we write it in spherical coordinate system. We know that (see figure no. 3.3):

$$\begin{aligned}\hat{\boldsymbol{x}} &= \sin\theta\cos\phi\,\hat{\boldsymbol{r}} + \cos\theta\cos\phi\,\hat{\boldsymbol{\theta}} - \sin\phi\,\hat{\boldsymbol{\phi}}\\ \hat{\boldsymbol{y}} &= \sin\theta\sin\phi\,\hat{\boldsymbol{r}} + \cos\theta\sin\phi\,\hat{\boldsymbol{\theta}} + \cos\phi\,\hat{\boldsymbol{\phi}}\\ \hat{\boldsymbol{z}} &= \cos\theta\,\hat{\boldsymbol{r}} - \sin\theta\,\hat{\boldsymbol{\theta}}\end{aligned} \tag{3.32}$$

and, also:

$$\begin{aligned}\hat{\boldsymbol{r}} &= \sin\theta\cos\phi\,\hat{\boldsymbol{x}} + \sin\theta\sin\phi\,\hat{\boldsymbol{y}} + \cos\theta\,\hat{\boldsymbol{z}}\\ \hat{\boldsymbol{\theta}} &= \cos\theta\cos\phi\,\hat{\boldsymbol{x}} + \cos\theta\sin\phi\,\hat{\boldsymbol{y}} - \sin\theta\,\hat{\boldsymbol{z}}\\ \hat{\boldsymbol{\phi}} &= -\sin\phi\,\hat{\boldsymbol{x}} + \cos\phi\,\hat{\boldsymbol{y}}\end{aligned} \tag{3.33}$$

Hence:

$$\begin{aligned}\hat{\boldsymbol{r}}\cdot\boldsymbol{f_t} &= (\sin\theta\cos\phi\,\hat{\boldsymbol{x}} + \sin\theta\sin\phi\,\hat{\boldsymbol{y}} + \cos\theta\,\hat{\boldsymbol{z}})\cdot\left(f_x\hat{\boldsymbol{x}} + f_y\hat{\boldsymbol{y}}\right) =\\ &= \sin\theta\left(f_x\cos\phi + f_y\sin\phi\right)\end{aligned} \tag{3.34}$$

and so:

$$\begin{aligned}&\cos\theta\,\boldsymbol{f_t}(k_0r'\sin\theta) - \hat{\boldsymbol{r}}\cdot\boldsymbol{f_t}(k_0r'\sin\theta)\hat{\boldsymbol{z}} =\\ &= \cos\theta\left(f_x\hat{\boldsymbol{x}} + f_y\hat{\boldsymbol{y}}\right) - \sin\theta\left(f_x\cos\phi + f_y\sin\phi\right)\hat{\boldsymbol{z}} =\\ &= \cos\theta\,f_x\left(\sin\theta\cos\phi\,\hat{\boldsymbol{r}} + \cos\theta\cos\phi\,\hat{\boldsymbol{\theta}} - \sin\phi\,\hat{\boldsymbol{\phi}}\right) +\\ &+\cos\theta\,f_y\left(\sin\theta\sin\phi\,\hat{\boldsymbol{r}} + \cos\theta\sin\phi\,\hat{\boldsymbol{\theta}} + \cos\phi\,\hat{\boldsymbol{\phi}}\right) -\\ &- \sin\theta\left(f_x\cos\phi + f_y\sin\phi\right)\left(\cos\theta\,\hat{\boldsymbol{r}} - \sin\theta\,\hat{\boldsymbol{\theta}}\right) =\\ &= \left(f_x\cos\phi + f_y\sin\phi\right)\hat{\boldsymbol{\theta}} + \cos\theta\left(-f_x\sin\phi + f_y\cos\phi\right)\hat{\boldsymbol{\phi}}\end{aligned} \tag{3.35}$$

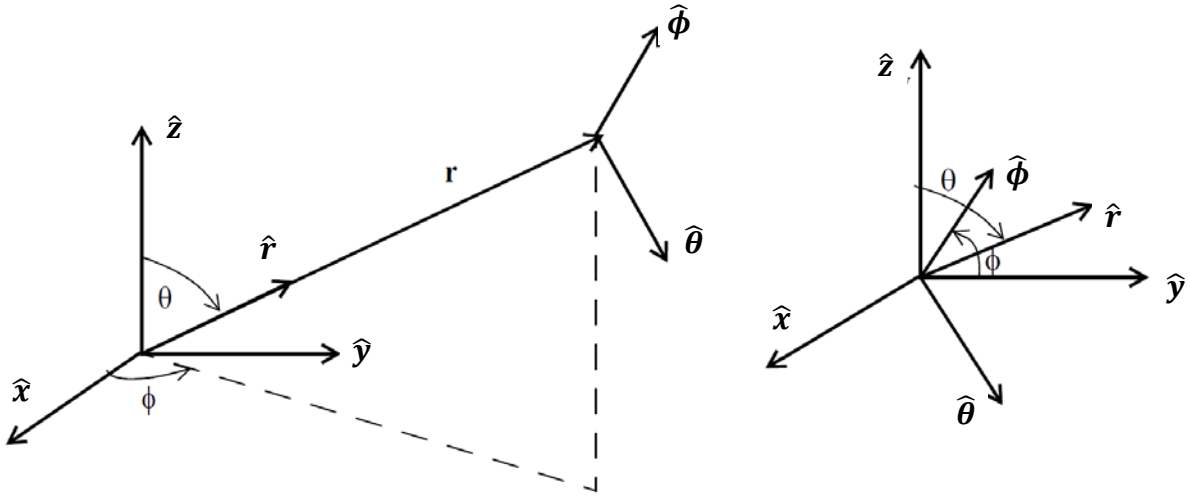


Figure no. 3.3 – *Versors of Cartesian and spherical coordinate systems*

By using these results we could write that:

$$\boldsymbol{E}(r,\theta,\phi) = \frac{jk_0}{2\pi}\frac{e^{-jk_0r}}{r}\left[\left(f_x\cos\phi + f_y\sin\phi\right)\hat{\boldsymbol{\theta}} + \cos\theta\left(-f_x\sin\phi + f_y\cos\phi\right)\hat{\boldsymbol{\phi}}\right] \tag{3.36}$$

We note that:

$$E_r = 0$$

$$E_\theta = \frac{jk_0}{2\pi}\frac{e^{-jk_0 r}}{r}\left(f_x \cos\phi + f_y \sin\phi\right) \tag{3.37}$$

$$E_\phi = \frac{jk_0}{2\pi}\frac{e^{-jk_0 r}}{r}\cos\theta\left(-f_x \sin\phi + f_y \cos\phi\right)$$

The above result validates our conclusion when studying the radiation of an arbitrary current distribution that the electromagnetic field in the radiation region is a transversal wave.

3.4 – Radiation from Apertures with Typical Field Distributions

Uniform Field Distribution

Let us consider that a constant field $\boldsymbol{E_a}$ exists in the rectangular aperture studied in the previous sub-chapters. A field oriented along the positive direction of the *Ox* axis that has a constant modulus E_0 is analytically described by the following relation:

$$\boldsymbol{E_a} = E_0\hat{\boldsymbol{x}} \tag{3.38}$$

According to (3.15) the tangential component $\boldsymbol{f_t}$ of the integration constant $\boldsymbol{f}$ is:

$$\boldsymbol{f_t} = \int_{-\infty}^{\infty}\boldsymbol{E_a}e^{\left(jk_x x + jk_y y\right)}dxdy = E_0\hat{\boldsymbol{x}}\left(\int_{-a}^{a}e^{jk_x x}dx\right)\left(\int_{-b}^{b}e^{jk_y y}dy\right) =$$

$$= \frac{4E_0}{k_x k_y}\sin(k_x a)\sin\left(k_y b\right)\hat{\boldsymbol{x}} \tag{3.39}$$

The spatial frequencies k_x, k_y are the components along *Ox* and *Oy*, respectively, of the propagation vector $\boldsymbol{k}$, whose modulus is k_0. By using (3.32) we obtain that:

$$k_x = k_0 \sin\theta\cos\phi \qquad k_y = k_0 \sin\theta \sin\phi \tag{3.40}$$

and so:

$$\boldsymbol{f_t} = 4abE_0\frac{\sin(k_0 a\sin\theta\cos\phi)}{k_0 a\sin\theta\cos\phi}\frac{\sin(k_0 b\sin\theta\sin\phi)}{k_0 b\sin\theta\sin\phi}\hat{\boldsymbol{x}} \tag{3.41}$$

We note from the above relation that:

$$f_x = 4abE_0\frac{\sin(k_0 a\sin\theta\cos\phi)}{k_0 a\sin\theta\cos\phi}\frac{\sin(k_0 b\sin\theta\sin\phi)}{k_0 b\sin\theta\sin\phi} \quad \text{and} \quad f_y = 0 \tag{3.42}$$

Using this result from (3.37) we get:

$$E_\theta = \frac{2jabk_0E_0}{\pi}\frac{e^{-jk_0 r}}{r}\frac{\sin(k_0 a\sin\theta\cos\phi)}{k_0 a\sin\theta\cos\phi}\frac{\sin(k_0 b\sin\theta\sin\phi)}{k_0 b\sin\theta\sin\phi}\cos\phi$$

$$E_\phi = -\frac{2jabk_0E_0}{\pi}\frac{e^{-jk_0 r}}{r}\frac{\sin(k_0 a\sin\theta\cos\phi)}{k_0 a\sin\theta\cos\phi}\frac{\sin(k_0 b\sin\theta\sin\phi)}{k_0 b\sin\theta\sin\phi}\cos\theta\sin\phi \tag{3.43}$$

These relations become less complex in two particular planes:

- in *xOz* plane ($\phi = 0$):

$$E_\theta = \frac{2jabk_0E_0}{\pi}\frac{e^{-jk_0 r}}{r}\frac{\sin(k_0 a\sin\theta)}{k_0 a\sin\theta} \quad \text{and} \quad E_\phi = 0 \tag{3.44}$$

- while in *yOz* plane ($\phi = \pi/2$):

$$E_\theta = 0 \quad \text{and} \quad E_\phi = -\frac{2jabk_0E_0}{\pi}\frac{e^{-jk_0 r}}{r}\frac{\sin(k_0 b\sin\theta)}{k_0 b\sin\theta}\cos\theta \tag{3.45}$$

The modulus of the electric component of the radiated field has a typical variation of $\left|\frac{\sin x}{x}\right|$ in both of the planes. Its graphical representation for the *xOz* plane is presented in figure no. 3.4.

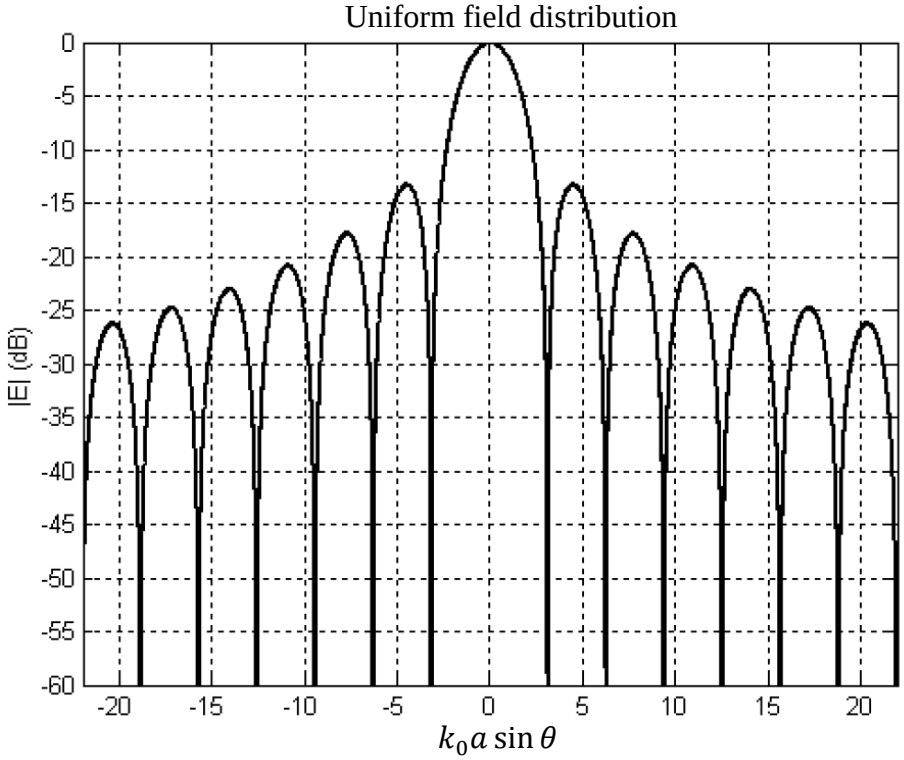


Figure no. 3.4 – *Radiation pattern for apertures with uniform field distribution*

Physically, the angle θ varies from $-\pi/2$ to $\pi/2$ and, as a consequence, the variable $k_x = k_0 a \sin\theta$ varies from $-k_0 a$ to $k_0 a$. The interval $[-k_0 a, k_0 a]$ is denoted as the *visible domain* of the variable k_x .

The first null in the radiation pattern of the aperture in *xOz* plane appears for (see 3.44):

$$k_0 a \sin\theta = \pi \qquad \Rightarrow \qquad \sin\theta = \frac{\pi}{k_0 a} = \frac{\lambda_0}{2a} \tag{3.46}$$

Remind that $2a$ is the aperture dimension along the *Ox* axis and that we supposed that the aperture dimensions are much greater than the wavelength of the radiated field. Hence, $\lambda_0/(2a) \ll 1$, the first null is very close to the main lobe and this makes the main lobe beamwidth to be small: $\approx 0.88\,\lambda_0/(2a)$. The greatest side lobe is the one closest to the main lobe and its relative level is ≈ 0.217 or $\approx -13.6\,dB$. The maximum directivity is $16\pi ab/\lambda_0^2$ and it is obtained in the direction $\theta = 0$, that is perpendicular to the aperture.

Tapered Field Distributions

We use the above results as a reference, but we should be aware that apertures with uniform field distribution could not be found in real world because discontinuities appear at the aperture edges where the field modulus changes from E_0 inside the aperture to 0 outside it. Real apertures have continuous transition of field at its edges: the electric field modulus has a maximum in the aperture center and monotonically decreases towards 0 at its edges. Because the electric field is null outside the aperture it is obvious that the transition at the aperture edges is continuous. This type of field distribution is denoted as tapered.

We assume that the radiated field in a plane perpendicular to aperture depends only on the aperture field distribution type in that plane. This is why, we study only one dimensional distribution laws in the followings.

A *cosine field distribution* law of the aperture field along the Ox axis is analytically described by the relation:

$$\boldsymbol{E_a}(x) = E_0 \cos\left(\frac{\pi x}{2a}\right)\hat{\boldsymbol{x}} \qquad -a \le x \le a \tag{3.47}$$

Note that this relation yields a maximum modulus in the center of the aperture ($x = 0$) and a zero value at its edges ($x = \pm a$).

According to (3.15):

$$\boldsymbol{f_t} = f_x\hat{\boldsymbol{x}} = \int_{-\infty}^{\infty} \boldsymbol{E_a} e^{jk_x x} dx = E_0\hat{\boldsymbol{x}} \int_{-a}^{a} \cos\left(\frac{\pi x}{2a}\right) e^{jk_0 x \sin\theta} dx =$$

$$= \frac{\pi a E_0 \cos(k_0 a \sin\theta)}{\left(\frac{\pi}{2}\right)^2 - (k_0 a \sin\theta)^2}\hat{\boldsymbol{x}} \tag{3.48}$$

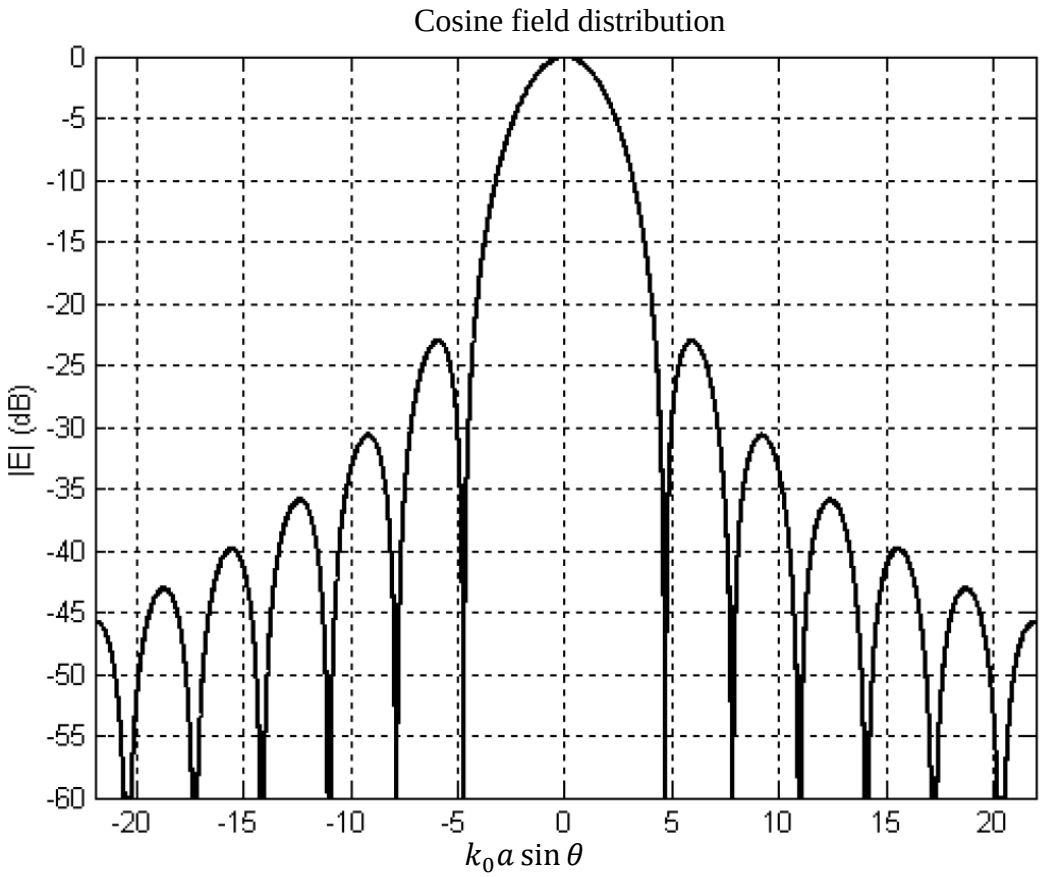


Figure no. 3.5 – *Radiation pattern for apertures with cosine field distribution*

The radiation pattern presented in figure no. 3.5 has a main lobe directed towards $\theta = 0$, identical to the uniform field distribution, but the first null appears for:

$$k_0 a \sin\theta = \frac{3\pi}{2} \qquad \Rightarrow \qquad \sin\theta = \frac{3\pi}{2k_0 a} = \frac{3\lambda_0}{4a} \tag{3.49}$$

This value yields a beamwidth of the main lobe of $\approx 1.2\lambda_0/(2a)$, (greater than the one for the uniform field distribution) and a relative level of the maximum side lobe of ≈ 0.07 or $\approx -23\ dB$ (smaller by 10 dB than the one for the uniform field distribution).

A *triangular field distribution* is obtained when there is a linear variation of the field modulus from a maximum in the aperture center to zero at its edges. Analytically:

$$\boldsymbol{E_a}(x) = E_0\left(1 - \frac{|x|}{a}\right)\hat{\boldsymbol{x}} \qquad -a \le x \le a \tag{3.50}$$

According to (3.15):

$$\boldsymbol{f_t} = f_x\hat{\boldsymbol{x}} = \int_{-\infty}^{\infty} \boldsymbol{E_a} e^{jk_x x}dx = E_0\hat{\boldsymbol{x}}\int_{-a}^{a}\left(1 - \frac{|x|}{a}\right)e^{jk_0 x\sin\theta}dx =$$

$$= aE_0\left(\frac{\sin\frac{k_0 a\sin\theta}{2}}{\frac{k_0 a\sin\theta}{2}}\right)^2\hat{\boldsymbol{x}} \tag{3.51}$$

The direction of the main lobe is also at $\theta = 0$, but the first null in the radiation pattern (figure no. 3.6) appears for:

$$\frac{k_0 a\sin\theta}{2} = \pi \qquad \Rightarrow \qquad \sin\theta = \frac{2\pi}{k_0 a} = \frac{\lambda_0}{a} \tag{3.52}$$

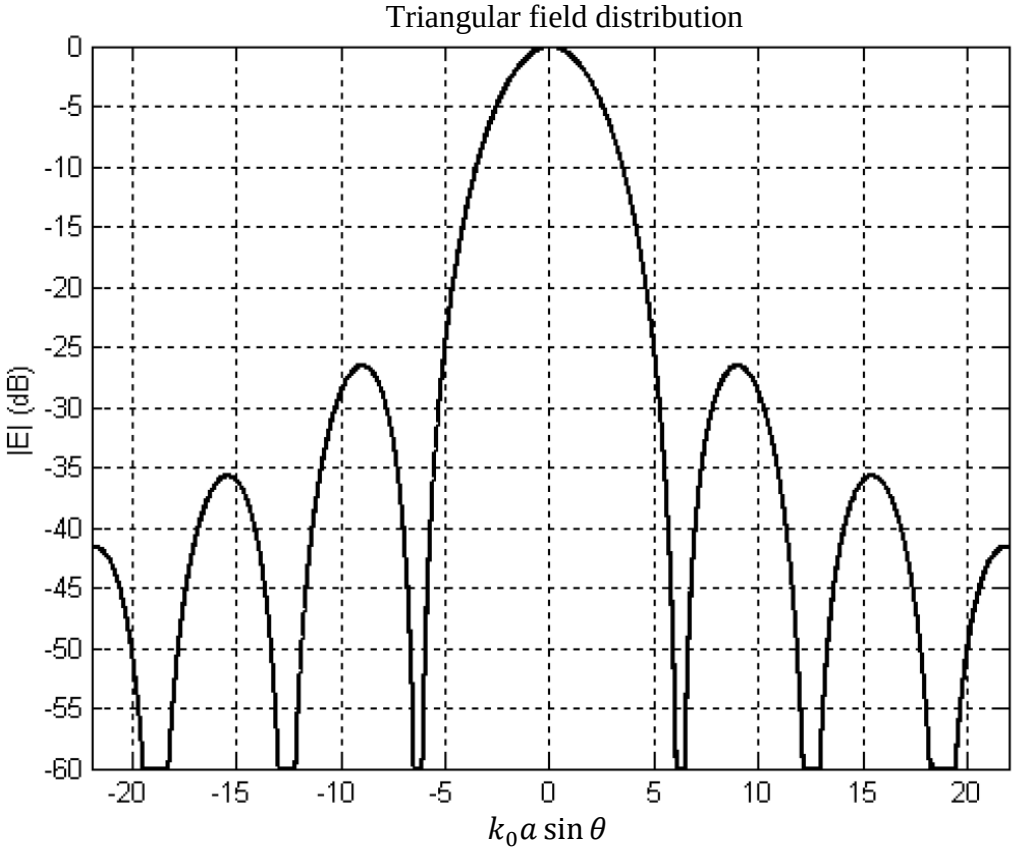


Figure no. 3.6 – *Radiation pattern for apertures with triangular field distribution*

The main lobe beamwidth ($\approx 1.28\lambda_0/2a$) is slightly greater than the one of the cosine distribution, while the relative level of the greatest side lobe decreases significantly and it becomes ≈ 0.047 or $\approx -24.6\,dB$.

In conclusion, *tapered field distributions, which are the distributions' type met in the real world, yield greater beamwidth for the main lobe and smaller levels for the side lobes than the reference ideal uniform distribution.*

Linear Phase Field Distributions

The above studies took into account only the modulus of the field in the aperture. By default, the phase of the field is assumed to be the same all over the aperture (constant phase). Assuming that phase of the field varies, we expect that some changes in the radiation pattern appear.

Let us assume that the phase of the field varies linearly with distance:

$$\boldsymbol{E}_{\boldsymbol{a}}' = \boldsymbol{E}_{\boldsymbol{a}} e^{-jk_0 x \sin\theta_0} \tag{3.53}$$

The Fourier transform of this field is:

$$\boldsymbol{f}_{\boldsymbol{t}}'(k_0 a \sin\theta) = f_x'\hat{\boldsymbol{x}} = \int_{-a}^{a} \boldsymbol{E}_{\boldsymbol{a}}' e^{jk_x x} dx = \int_{-a}^{a} \boldsymbol{E}_{\boldsymbol{a}} e^{-jk_0 x \sin\theta_0} e^{jk_0 x \sin\theta} dx =$$

$$= \int_{-a}^{a} \boldsymbol{E}_{\boldsymbol{a}} e^{jk_0 x(\sin\theta - \sin\theta_0)} dx = \boldsymbol{f}_{\boldsymbol{t}}[k_0 a(\sin\theta - \sin\theta_0)] \tag{3.54}$$

where $\boldsymbol{f}_{\boldsymbol{t}}$ is the Fourier transform of the constant phase aperture field $\boldsymbol{E}_{\boldsymbol{a}}$.

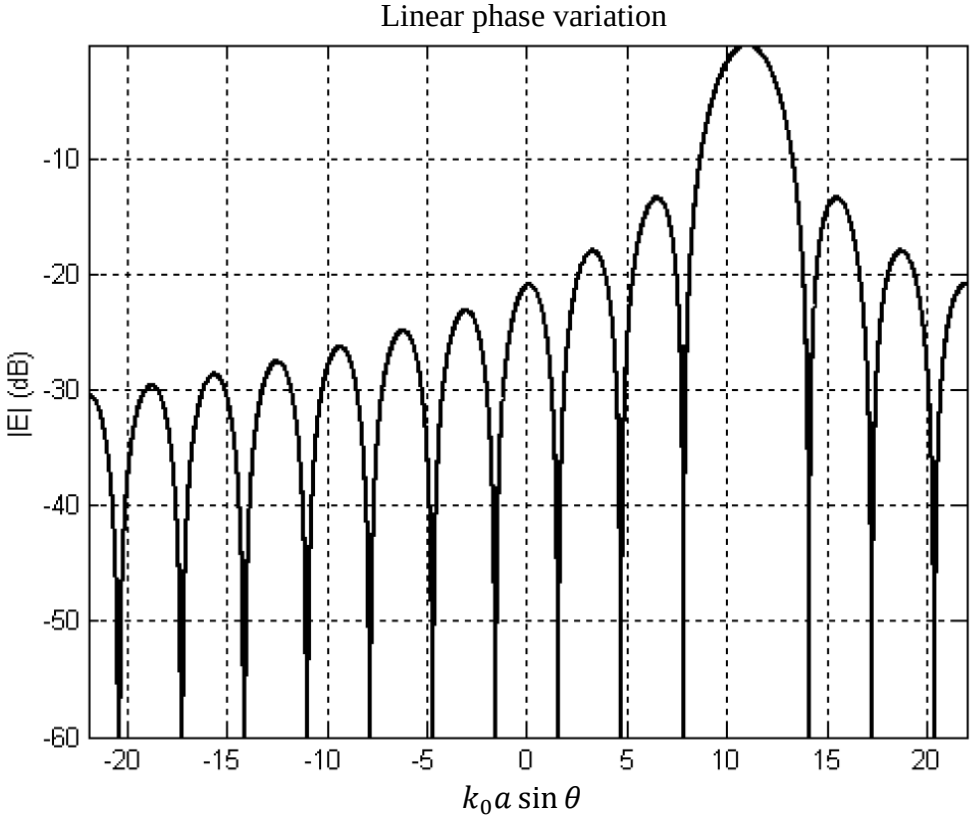


Figure no. 3.7 – *Radiation pattern for apertures with linear phase variation*

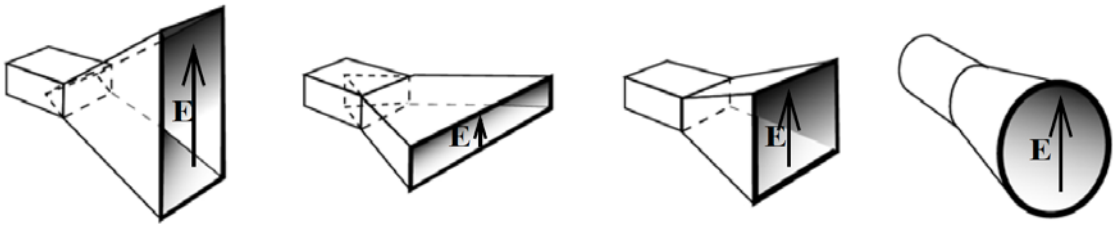
The above relation shows that the direction of the main lobe of the radiation pattern is modified by the linear phase variation from $\theta = 0$ to $\theta = \theta_0$ (see figure no. 3.7). Due to this tilt, the main lobe beamwidth increases by $1/\cos\theta_0$ times. The structure of the side lobes and their relative levels do not change.

3.6 – Horn Antenna

The typical implementation of the theoretical concept of aperture is the horn antenna. A horn antenna, usually fed by a rectangular waveguide, is a truncated pyramid having the smaller basis identical to the waveguide cross section. It allows for a smooth transition from the waveguide medium, where a limited number of propagation modes are possible, to the free space, where an infinite number of propagation modes are possible. Horn antenna is a

large bandwidth antenna. Antenna dimensions for a particular application are chosen such that one of the following parameters to be optimum: gain, bandwidth, input impedance or radiation pattern shape.

The most common type of horn antenna is the *pyramidal* one. When one of the outer cross section dimensions is identical to the corresponding dimension of the feeding waveguide the antenna is said to be a *sectoral horn*. If the feeding waveguide is circular, the horn antenna is a truncated cone and it is denoted as a *conical horn* (figure no. 3.8).



E sectoral horn H sectoral horn Pyramidal horn Conical horn

Figure no. 3.8 – *Typical shapes of horn antennas*

Sectoral Horn Antenna

When the larger dimension of the outer cross section is along the magnetic component of the electromagnetic field in the aperture, the antenna is denoted as *H sectoral horn.* When the larger dimension of the outer cross section is along the electric component of the electromagnetic field in the aperture, the antenna is denoted as *E sectoral horn.*

Obviously, the radiation pattern of a sectoral horn antenna in the plane of the horn larger dimension is quite similar to the one of an aperture antenna. The main dissimilarity is that the radiated field in the nulls' directions is not zero, only much less than the one in the neighboring directions. This is due to the fact that the phase of the field in the horn outer cross section is not the same. The field propagates through horn from the waveguide to the free space by means of spherical wave fronts and it reaches the center of the outer cross section earlier than other points of this cross section, so that the phase of the field in the center leads the one in other points. The phase difference increases as the horn length increases.

The H sectoral horn has larger main lobe beamwidth smaller side lobes' level than the E sectoral horn with identical dimensions. This is because the distribution field in the horn outer cross section approaches a cosine law for the H sectoral horn, while the distribution field in the horn outer cross section is almost uniform for the E sectoral horn.

The *gain* of a sectoral horn antenna depends on the larger cross section dimension (a or b) and on the horn length l, as related to the radiated wavelength (that is, a/λ, b/λ, and l/λ, respectively). It increases monotonically with l/λ, and it has a maximum for some optimum value of a/λ (or b/λ).

Input impedance. The benefits of horn antenna large bandwidth could be exploited only when the antenna is perfectly matched at both ends. The matching between the horn and the feeding waveguide is obtained by means of reactive slots in the waveguide walls, close to the junction waveguide-horn. The matching between the horn and the free space is obtained by covering the inner surface of the horn by an appropriately chosen dielectric.

Pyramidal Horn Antenna

The radiation pattern of a pyramidal horn antenna in planes parallel to its cross section edges is identical to one of the sectoral horn with the same size. It is used in applications that require a tight control of the radiation pattern in both of the planes.

Pyramidal horn antenna is often used as a reference antenna for measuring the gain of other antennas because its gain is very precisely computed from its physical dimensions (error is smaller than one tenth of a dB):

$$G[dB] = 10\left[1.008 + \log\left(\frac{a}{\lambda}\frac{b}{\lambda}\right)\right] - L_e - L_h \tag{3.55}$$

where L_e and L_h are correction factors dependent on the horn length l/λ; their values are given in the literature.

When the outer cross section dimensions equal the optimal values that yields the maximum gain for the corresponding sectoral horn, that is:

$$\frac{a}{\lambda} = \sqrt{3l} \qquad \frac{b}{\lambda} = \sqrt{2l} \tag{3.56}$$

then the pyramidal horn gain is maximum:

$$G_{optim}[dB] = 10\left[0.008 + \log\left(\frac{a}{\lambda}\frac{b}{\lambda}\right)\right] \tag{3.57}$$

and the antenna is said to be an *optimal pyramidal horn*. This antenna is rarely used because it is a narrow bandwidth antenna.

Conical Horn Antenna

Conical horn antenna is a truncated cone fed by a circular waveguide; it is used for applications requiring symmetrical radiation pattern around the cone axis. The side lobes have greater relative levels than the ones of the pyramidal horn of the same size.

The gain of the conical horn antenna is:

$$G[dB] = 20\log\frac{C}{\lambda} - L \tag{3.58}$$

where C is the circumference of the outer cross section, while L is a correction factor dependent on the horn length l.

There is also an *optimal conical horn antenna*, obtained for $C = \pi\sqrt{3l\lambda}$ that has:

$$G_{optim}[dB] = 20\log\frac{C}{\lambda} - 2.82 \tag{3.59}$$

C h a p t e r I V

RECEIVING ANTENNA

The electromagnetic wave received by an antenna is characterized by the frequency band it covers, by the spatial distribution of directions of arrival, by spatial and temporal variations of its amplitude, phase or/and polarization, and by possible correlations between these parameters. In order to establish a reference for our study we consider in the followings that the received wave is monochromatic (it has a single frequency) and it arrives from a single fixed point source.

The antenna feeds its load by means of a transmission line. For a maximal power transfer, matching conditions should be met both at the junction of antenna with the transmission line and at the connection of the transmission line with the load. As transmission medium a common symmetrical/unsymmetrical transmission line or a waveguide could be used. Any of them could be equivalent to an unsymmetrical transmission line (coaxial cable). This is the reason why only unsymmetrical transmission line is considered in the followings.

4.1 – Reciprocity Principle for Antennas

Let us consider two arbitrary antennas **a** and **b** fed by two separated shielded sources $\mathbf{g_a}$ and $\mathbf{g_b}$, respectively (figure no. 4.1). Let us assume that the considered antennas are placed in a lossless medium and that they are separated by a large distance, so that each antenna is placed in the radiation region of the other one. Denote by Z_{i1} and Z_{in1}, the impedance seen from the reference plane $\boldsymbol{S_1}$ towards the generator $\mathbf{g_a}$ and towards the antenna **a**, respectively. Also, denote by Z_{i2} and Z_{in2}, the impedance seen from the reference plane $\boldsymbol{S_2}$ towards the generator $\mathbf{g_b}$ and towards the antenna **b**, respectively.

When $\mathbf{g_a}$ is connected and $\mathbf{g_b}$ is unconnected (that is, when antenna **a** is transmitting and antenna **b** is receiving), the field created by the antenna **a** has the following components:

- in plane $\boldsymbol{S_1}$:

$$E_r = \frac{V_{1a}}{r\ln(r_{e1}/r_{i1})}, \qquad H_\phi = \frac{I_{1a}}{2\pi r} \qquad \text{with} \quad \frac{V_{1a}}{I_{1a}} = Z_{in1} \tag{4.1}$$

- in plane $\boldsymbol{S_2}$:

$$E_r = \frac{V_{2a}}{r\ln(r_{e2}/r_{i2})}, \qquad H_\phi = -\frac{I_{2a}}{2\pi r} \qquad \text{with} \quad \frac{V_{2a}}{I_{2a}} = Z_{i2} \tag{4.2}$$

When $\mathbf{g_a}$ is unconnected and $\mathbf{g_b}$ is connected (that is, when antenna **a** is receiving and antenna **b** is transmitting), the field created by the antenna **b** has the following components:

- in plane $\boldsymbol{S_1}$:

$$E_r = \frac{V_{1b}}{r\ln(r_{e1}/r_{i1})}, \qquad H_\phi = -\frac{I_{1b}}{2\pi r} \qquad \text{with} \quad \frac{V_{1b}}{I_{1b}} = Z_{i1} \tag{4.3}$$

- in plane $\boldsymbol{S_2}$:

$$E_r = \frac{V_{2b}}{r\ln(r_{e2}/r_{i2})}, \qquad H_\phi = \frac{I_{2b}}{2\pi r} \qquad \text{with} \quad \frac{V_{2b}}{I_{2b}} = Z_{in2} \tag{4.4}$$

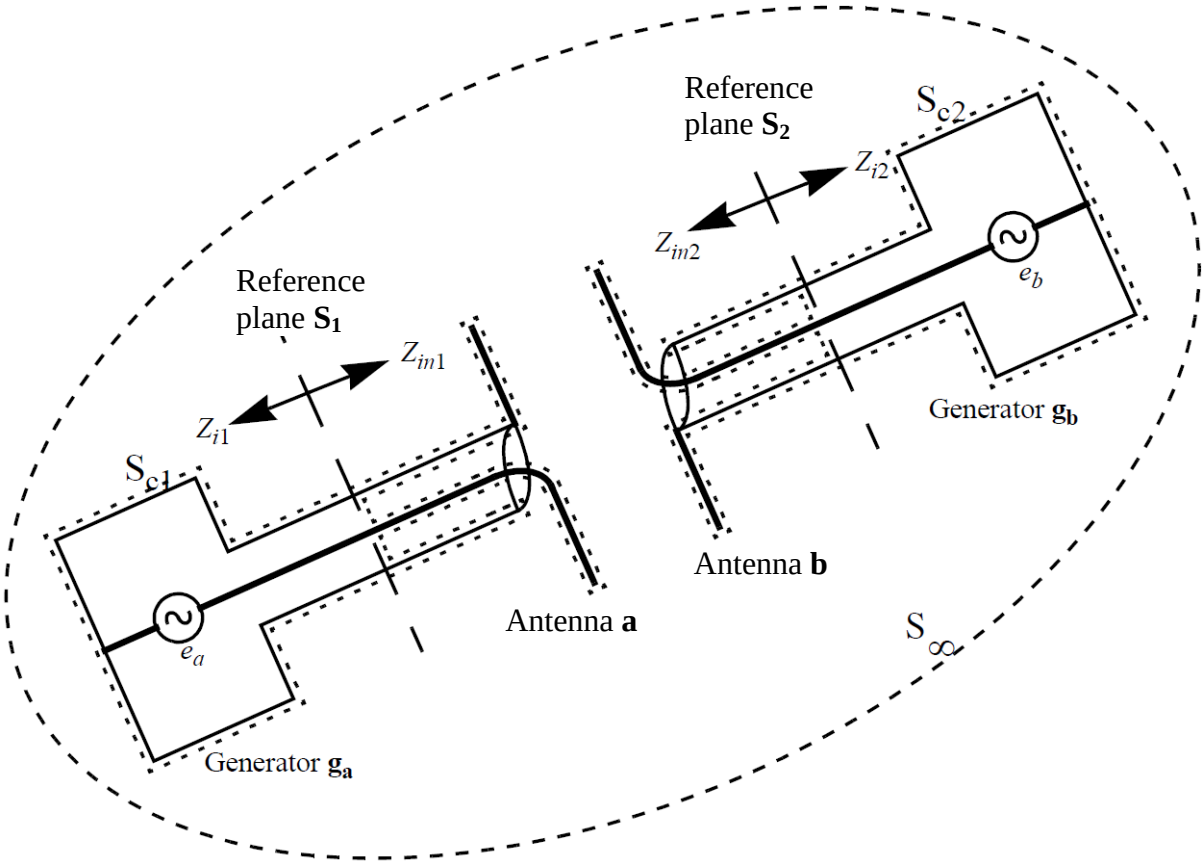


Figure no. 4.1 – *System of two arbitrary antennas*

In the above relations V_{1a} is the voltage created in the feeding coaxial cable in the reference plane $\boldsymbol{S_1}$ by the antenna **a** when it transmits, V_{1b} is the voltage created in the feeding coaxial cable in the reference plane $\boldsymbol{S_1}$ by the antenna **b** when it transmits, and so on. Also, r_e and r_i are the exterior radius and interior radius, respectively, of the dielectric in the feeding coaxial cable. (r, θ, ϕ) are the spherical variables associated to Cartesian coordinate systems centered in planes $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$, respectively, and having the inner conductor of the coaxial cables as their axis *Oz* with positive directions towards the antennas.

Based on the equation (1.70), expressing the general principle of reciprocity, we could write that:

$$\int_\Sigma (\boldsymbol{E_a} \times \boldsymbol{H_b} - \boldsymbol{E_b} \times \boldsymbol{H_a}) \cdot d\boldsymbol{\sigma} =$$

$$= \int_V [(\boldsymbol{E_b} \cdot \boldsymbol{J_a} - \boldsymbol{E_a} \cdot \boldsymbol{J_b}) - (\boldsymbol{H_b} \cdot \boldsymbol{J_{ma}} - \boldsymbol{H_a} \cdot \boldsymbol{J_{mb}})] dV \tag{4.5}$$

We define the surface Σ as the reunion of the following surfaces:

- $S_c = S_{c1} \cup S_{c2}$ – the surface infinitesimally close to the shield of the sources $\mathbf{g_a}$ and $\mathbf{g_b}$;
- $S' = S_1' \cup S_2'$ – areas of the intersections of the reference planes $\mathbf{S_1}$ and $\mathbf{S_2}$ with the dielectric of the feeding coaxial cables;
- S_∞ – the surface of the sphere with infinite radius.

The volume V closed by the surface Σ does not include any electric or magnetic source. Thus, the argument of the volume integral is zero and the result of the integration is, obviously, zero.

The surface of the sphere with infinite radius is placed in the radiation region of both of the antennas, so the radiated field is a transversal electromagnetic wave with:

$$\boldsymbol{H_a} = \frac{1}{\eta_0}\hat{\boldsymbol{r}} \times \boldsymbol{E_a} \quad , \quad \boldsymbol{H_b} = \frac{1}{\eta_0}\hat{\boldsymbol{r}} \times \boldsymbol{E_b} \tag{4.6}$$

Also, the tangential components of the vectors $\boldsymbol{E}$ and $\boldsymbol{H}$ on the surface S_c are related by the same type of expressions:

$$\boldsymbol{H_{at}} = \frac{1}{\eta_m}\hat{\boldsymbol{n}} \times \boldsymbol{E_{at}} \quad , \quad \boldsymbol{H_{bt}} = \frac{1}{\eta_m}\hat{\boldsymbol{n}} \times \boldsymbol{E_{bt}} \tag{4.7}$$

where $\hat{\boldsymbol{n}}$ is a versor normal to the surface S_c in each of its points, while η_m is the characteristic impedance of the shielding material.

Due to the above relations, the surface integral in equation (4.5) is zero on surfaces S_∞ and S_c. Thus, the equation (4.5) becomes simply:

$$\int_{S_1'} (\boldsymbol{E_a} \times \boldsymbol{H_b} - \boldsymbol{E_b} \times \boldsymbol{H_a}) \cdot d\boldsymbol{\sigma} + \int_{S_2'} (\boldsymbol{E_a} \times \boldsymbol{H_b} - \boldsymbol{E_b} \times \boldsymbol{H_a}) \cdot d\boldsymbol{\sigma} = 0 \tag{4.8}$$

On the surface S_1' we have:

$$\begin{gathered} \boldsymbol{E_a} \times \boldsymbol{H_b} - \boldsymbol{E_b} \times \boldsymbol{H_a} = \\ = \frac{V_{1a}}{r \ln(r_{e1}/r_{i1})}\hat{\boldsymbol{r}} \times \left(-\frac{I_{1b}}{2\pi r}\hat{\boldsymbol{\phi}}\right) - \frac{V_{1b}}{r \ln(r_{e1}/r_{i1})}\hat{\boldsymbol{r}} \times \frac{I_{1a}}{2\pi r}\hat{\boldsymbol{\phi}} = \\ = \frac{V_{1a}I_{1b} + V_{1b}I_{1a}}{2\pi r^2 \ln(r_{e1}/r_{i1})}\hat{\boldsymbol{\theta}} \end{gathered} \tag{4.9}$$

On the surface S_2' we have:

$$\begin{gathered} \boldsymbol{E_a} \times \boldsymbol{H_b} - \boldsymbol{E_b} \times \boldsymbol{H_a} = \\ = \frac{V_{2a}}{r \ln(r_{e2}/r_{i2})}\hat{\boldsymbol{r}} \times \frac{I_{2b}}{2\pi r}\hat{\boldsymbol{\phi}} - \frac{V_{2b}}{r \ln(r_{e2}/r_{i2})}\hat{\boldsymbol{r}} \times \left(-\frac{I_{2a}}{2\pi r}\right)\hat{\boldsymbol{\phi}} = \\ = -\frac{V_{2a}I_{2b} + V_{2b}I_{2a}}{2\pi r^2 \ln(r_{e2}/r_{i2})}\hat{\boldsymbol{\theta}} \end{gathered} \tag{4.10}$$

Thus, equation (4.8) becomes:

$$\frac{V_{1a}I_{1b} + V_{1b}I_{1a}}{2\pi \ln(r_{e1}/r_{i1})}\hat{\boldsymbol{\theta}} \int_{S_1'} \frac{d\boldsymbol{\sigma}}{r^2} = \frac{V_{2a}I_{2b} + V_{2b}I_{2a}}{2\pi \ln(r_{e2}/r_{i2})}\hat{\boldsymbol{\theta}} \int_{S_2'} \frac{d\boldsymbol{\sigma}}{r^2} \tag{4.11}$$

The surfaces S_1' and S_2' are cross sections of the dielectric tube that separates the inner conductor and the shield in a coaxial cable. In order to evaluate the surface integral from equation (4.11) we define the vectorial surface infinitesimal element $d\boldsymbol{\sigma}$ as a vector normal to the surface with a modulus equal to the area between the circles with radius r and $r + dr$, respectively (figure no. 4.2):

$$d\boldsymbol{\sigma} = 2\pi r dr \hat{\boldsymbol{\theta}} \tag{4.12}$$

Thus:

$$\int_{S_1'} \frac{d\boldsymbol{\sigma}}{r^2} = \int_{r_{i1}}^{r_{e1}} \frac{2\pi r dr}{r^2} \widehat{\boldsymbol{\theta}} = 2\pi \ln(r_{e1}/r_{i1})\, \widehat{\boldsymbol{\theta}}$$

$$\int_{S_2'} \frac{d\boldsymbol{\sigma}}{r^2} = \int_{r_{i2}}^{r_{e2}} \frac{2\pi r dr}{r^2} \widehat{\boldsymbol{\theta}} = 2\pi \ln(r_{e2}/r_{i2})\, \widehat{\boldsymbol{\theta}} \qquad (4.13)$$

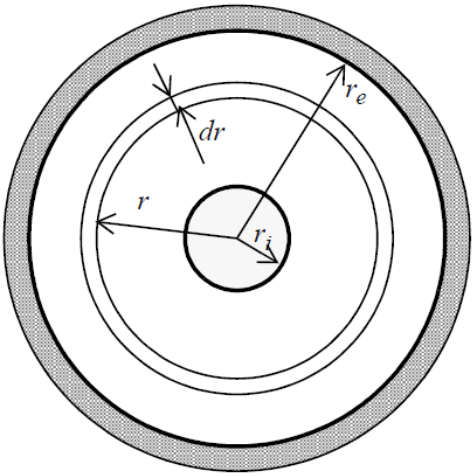


Figure no. 4.2 – *Definition of the vectorial surface infinitesimal element* $d\boldsymbol{\sigma}$

Based on these results, the equation (4.11) becomes:

$$V_{1a}I_{1b} + V_{1b}I_{1a} = V_{2a}I_{2b} + V_{2b}I_{2a} \qquad (4.14)$$

and this is a formula of the *general reciprocity principle for antennas*.

This formula could be simplified further if the electromotive voltages e_a and e_b of the generators $\mathbf{g_a}$ and $\mathbf{g_b}$ are taken into account. Figure no. 4.3 illustrates the equivalent electric circuits of the two antenna system from figure no. 4.1 for the two operational cases we discussed earlier:

a) $\mathbf{g_a}$ is connected and $\mathbf{g_b}$ is unconnected (antenna **a** is transmitting and antenna **b** is receiving);

b) $\mathbf{g_b}$ is connected and $\mathbf{g_a}$ is unconnected (antenna **b** is transmitting and antenna **a** is receiving).

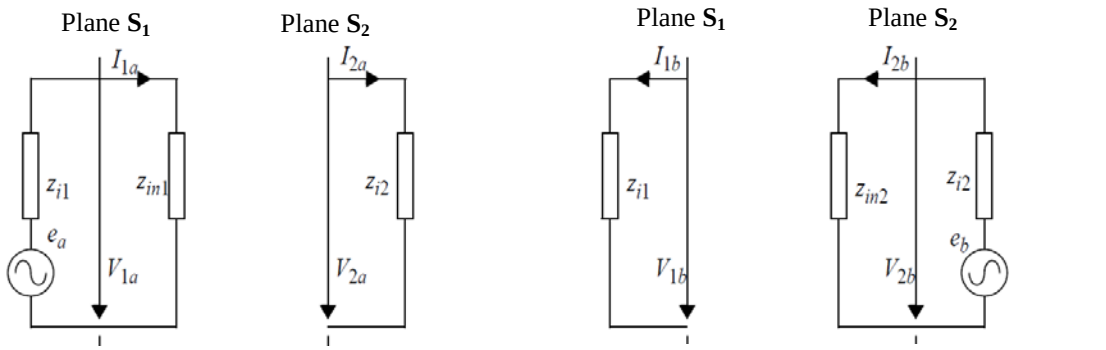


Figure no. 4.3 – *Equivalent circuits for the two antenna system*

Based on these equivalent circuits we could write:

$$(a)\begin{cases} V_{1a} = e_a - Z_{i1}I_{1a} \\ V_{2a} = Z_{i2}I_{2a} \end{cases} \qquad (b)\begin{cases} V_{1b} = Z_{i1}I_{1b} \\ V_{2b} = e_b - Z_{i2}I_{2b} \end{cases} \qquad (4.15)$$

Replacing these expressions of the voltages in (4.14) we get:

$$(e_a - Z_{i1}I_{1a})I_{1b} + Z_{i1}I_{1b}I_{1a} = Z_{i2}I_{2a}I_{2b} + (e_b - Z_{i2}I_{2b})I_{2a} \qquad (4.16)$$

or:

$$e_a I_{1b} = e_b I_{2a} \qquad (4.17)$$

which is another formula for the *general reciprocity principle for antennas*.

4.2 – The Equivalent Circuit of Two Antenna System

When the two generators in figure no. 4.1 are simultaneously connected, there are voltages and currents in the reference planes $\boldsymbol{S_1}$ and $\boldsymbol{S_2}$ produced by both of the generators. Based on the assumption that each antenna is placed in the radiation region of the other one, the electric variables are not influenced by each other and the total currents and voltages in each of the planes are the algebraic sum of the respective variables when only one of the generators is connected. Hence (see figure no. 4.4):

$$\begin{cases} V_1 = V_{1a} + V_{1b} \\ I_1 = I_{1a} - I_{1b} \end{cases} \qquad \begin{cases} V_2 = V_{2a} + V_{2b} \\ I_2 = -I_{2a} + I_{2b} \end{cases} \qquad (4.18)$$

Figure no. 4.4 – *Total electric variables in the reference planes*

Viewing the free space between the two antennas as an electric circuit, it looks like a two-port circuit. Describing it by means of the *z* parameters we could write:

$$\begin{cases} V_1 = z_{11}I_1 + z_{12}I_2 \\ V_2 = z_{21}I_1 + z_{22}I_2 \end{cases} \qquad (4.19)$$

These generic equations get specific forms for the two operational cases of the two antenna systems. Thus, when antenna **a** is transmitting ($e_a \neq 0$) and antenna **b** is receiving ($e_b = 0$):

$$\begin{cases} V_{1a} = z_{11}I_{1a} - z_{12}I_{2a} \\ V_{2a} = z_{21}I_{1a} - z_{22}I_{2a} \end{cases} \qquad (4.20)$$

while when antenna **b** is transmitting ($e_b \neq 0$) and antenna **a** is receiving ($e_a = 0$):

$$\begin{cases} V_{1b} = -z_{11}I_{1b} + z_{12}I_{2b} \\ V_{2b} = -z_{21}I_{1b} + z_{22}I_{2b} \end{cases} \qquad (4.21)$$

Replacing the above expressions of voltages in the first formula (4.14) of the general principle of reciprocity for antennas we get:

$$(z_{11}I_{1a} - z_{12}I_{2a})I_{1b} + (-z_{11}I_{1b} + z_{12}I_{2b})I_{1a} = \\ = (z_{21}I_{1a} - z_{22}I_{2a})I_{2b} + (-z_{21}I_{1b} + z_{22}I_{2b})I_{2a} \qquad (4.22)$$

or:

$$(I_{1a}I_{2b} - I_{1b}I_{2a})z_{12} = (I_{1a}I_{2b} - I_{1b}I_{2a})z_{21} \tag{4.23}$$

and:

$$(I_{1a}I_{2b} - I_{1b}I_{2a})(z_{12} - z_{21}) = 0 \tag{4.24}$$

The generators $\mathbf{g_a}$ and $\mathbf{g_b}$ are independent and, thus, the currents they generate are independent of each other. As a consequence, the first factor in the above equation could not be zero and, for the equation to be fulfilled, the second factor *must* be zero. This means that:

$$z_{12} \equiv z_{21} \tag{4.25}$$

The above relation shows that *the two-port equivalent circuit for the space between the two antennas is a reciprocal one*. This equivalent circuit is illustrated in figure no. 4.5. It includes also equivalent circuits for the characteristic impedance of the two generators, which are, in general, T-impedances. The circuit in figure 4.5 is *the complete electric model for a two antenna system*.

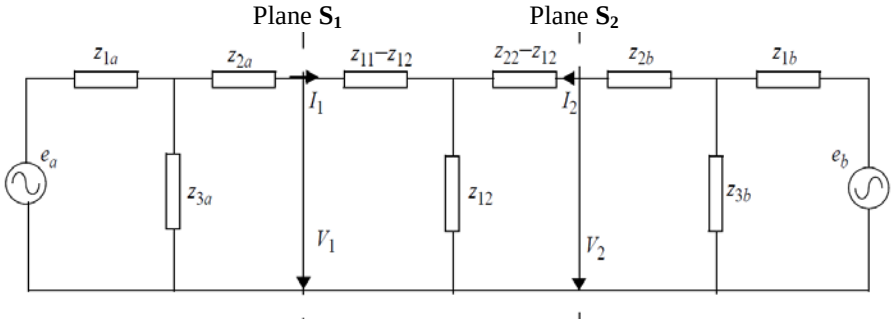


Figure no. 4.5 – *Reciprocal equivalent circuit for a two arbitrary antenna system*

4.3 – Directive Properties of Antennas

Let us consider the two antenna system presented in the figure no. 4.6. Assume that each of them is matched with its load. Without restraining the generality, but simplifying the involved relations, let us assume that the antenna input impedances are purely resistive: $Z_{in1} = R_a$ and $Z_{in2} = R_b$. We consider that antenna **a** is fixed, while antenna **b** moves such that its distance from the antenna **a** and the solid angle (θ_b, ϕ_b) are constant. These assumptions make the position of antenna **b** relative to antenna **a** be completely described by the solid angle (θ_a, ϕ_a).

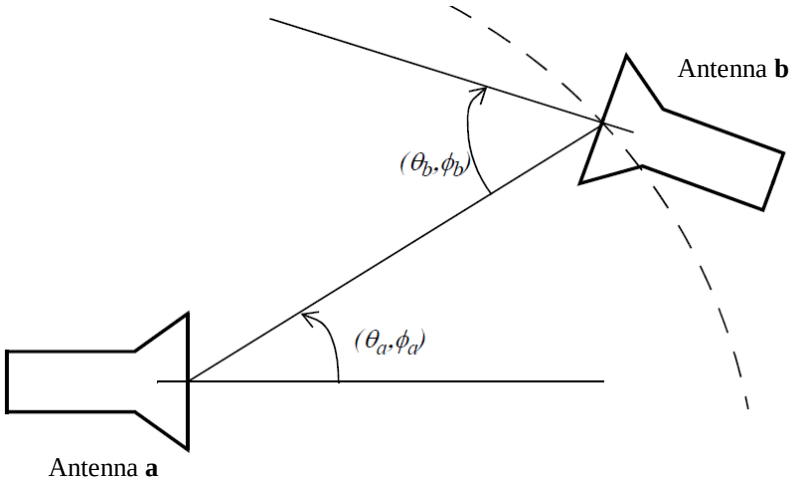


Figure no. 4.6 – *Two arbitrary antenna system:* **a** *– fixed and* **b** *– moving*

Let consider the case when antenna **a** is transmitting ($e_a \neq 0$) and antenna **b** is receiving ($e_b = 0$). If antenna **b** is placed in the radiation region of antenna **a** and it does not modify the electromagnetic field at the receiving point, then the power received by antenna **b** while moving on the sphere surface centered in **a** is a measure of the directive properties of antenna **a**.

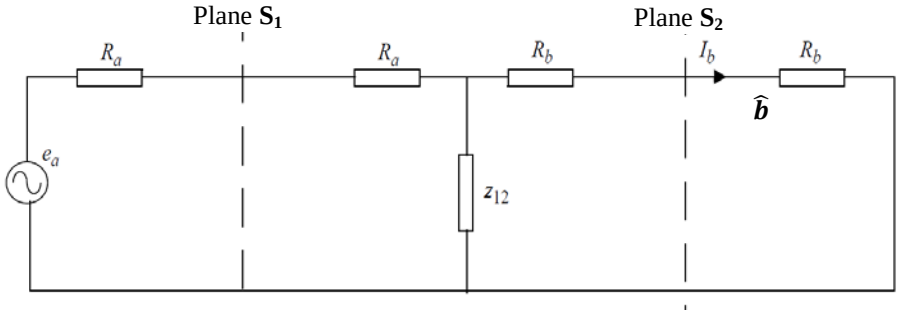


Figure no. 4.7 – *Equivalent circuit of the system from figure no. 4.6*

The power received by antenna **b** could be evaluated by means of the electric equivalent circuit in figure no. 4.5. This circuit is redrawn in figure no. 4.7 in simplified form due to the above assumptions about the antenna input impedance being purely resistive and their matching with loads connected at their ports. Based on these assumptions we have $Z_{i1} = z_{11} = R_a$ and $Z_{i2} = z_{22} = R_b$. Assume that the modulus of the transfer impedance z_{12} is much smaller than input and output impedances of the two-port circuit, that is $|z_{12}| \ll R_a$ and $|z_{12}| \ll R_b$. The power received by antenna **b** is the power dissipated by the characteristic impedance $Z_{i2} = R_b$ of generator $\mathbf{g_b}$ which is the load of the antenna **b**. Based on the equivalent circuit from figure no. 4.7 and taking into account the previous assumptions this power is given by the following formula:

$$P_b = \frac{1}{2} R_b |I_b|^2 = \frac{1}{2} R_b \left| \frac{e_a}{2R_a + z_{12} \| (2R_b)} \right|^2 \left| \frac{z_{12}}{2R_b + z_{12}} \right|^2 =$$

$$= \frac{1}{2} R_b \left| \frac{e_a}{2R_a} \right|^2 \left| \frac{z_{12}}{2R_b} \right|^2 = \frac{|e_a|^2}{32 R_a^2 R_b} |z_{12}|^2 \qquad (4.26)$$

The power received by antenna ***b*** in a point on the sphere surface it moves on is proportional with the power density of the electromagnetic field in that point and this one is proportional with gain of antenna ***a*** in the direction (θ_a, ϕ_a) where antenna ***b*** is placed. Hence:

$$P_b \sim G_a(\theta_a, \phi_a) \qquad (4.27)$$

The only factor in the formula (4.26) of P_b that could depend on the gain of antenna **a** is $|z_{12}|^2$; so:

$$|z_{12}|^2 \sim G_a(\theta_a, \phi_a) \qquad (4.28)$$

Following a similar rationale we get the formula for the power P_a received by antenna **a** when it is receiving ($e_a = 0$), while antenna **b** is transmitting ($e_b \neq 0$):

$$P_a = \frac{|e_b|^2}{32 R_b^2 R_a} |z_{12}|^2 \qquad (4.29)$$

Hence:

$$P_a \sim |z_{12}|^2 \sim G_a(\theta_a, \phi_a) \qquad (4.30)$$

that is, *the power received by antenna* **a** *from a direction is proportional with the its gain in that direction*. We made no restriction on antenna **a** used in the above developments, so it is an arbitrary antenna. The conclusion is that *any antenna has the same radiation pattern irrespective of its operation mode as a transmitter or as a receiver*. In other words, *any antenna is a reciprocal device*.

This property of reciprocity is obtained thanks to the reciprocal ($z_{12} = z_{21}$) two-port model of the space between two antennas, that is, *antenna reciprocity is a consequence of the medium it operates in of being isotropic*. If the medium is anisotropic or if the antenna contains non reciprocal devices, reciprocity does not hold anymore.

It could be proved that the antenna reciprocity holds even when the matching condition is not fulfilled. The main requirement is that the transfer impedance of the equivalent two-port circuit be much smaller than its input/output impedances: $|z_{12}| \ll |z_{11}|$ and $|z_{12}| \ll |z_{22}|$.

4.4 – Antenna Receiving Cross Section

Receiving cross section or *effective area* is a global parameter that characterizes the receiving properties of an antenna. Its usefulness is given by the observation that the power transferred to the load by an antenna it is not influenced by the properties of the transmitting antenna that creates the electromagnetic field in the receiving point, but only by the properties of the receiving antenna and its relationship with the electromagnetic field it receives. We show in the followings that antenna cross section is a direct consequence of its directive properties.

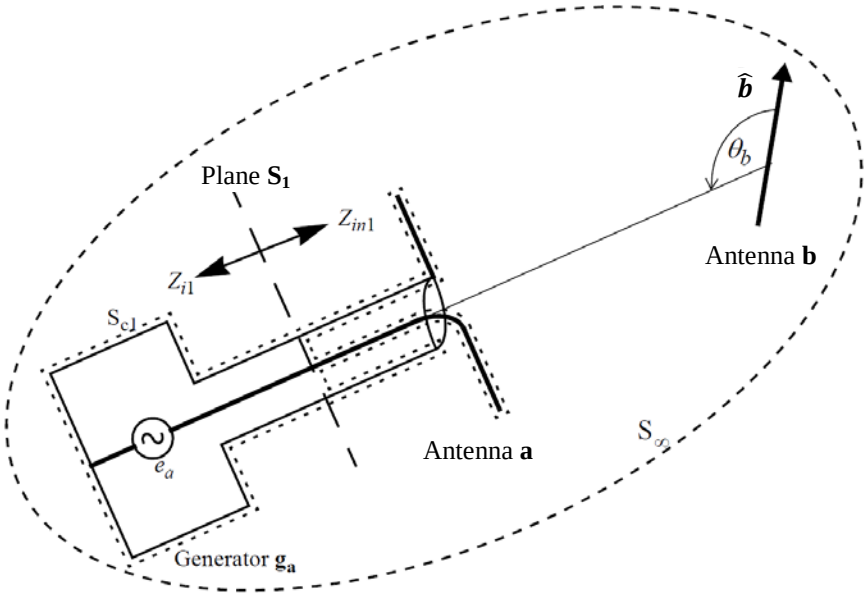


Figure no. 4.8 – *System containing an arbitrary antenna* **a** *and a current element antenna* **b**

Let us consider the two antenna system in figure no. 4.8: antenna **a** is an arbitrary antenna, while antenna **b** is a rectilinear element of length l carrying a constant current I. We assume that each of the antennas is placed in the radiation region of the other one.

We consider two operational cases:

a) antenna **a** is transmitting and antenna **b** is receiving and

b) antenna **b** is transmitting and antenna **a** is receiving.

For the first case the variables created in the reference plane $\boldsymbol{S_1}$ are given by the equations (4.1), while for the second case – they are given by the equations (4.3).

We apply the general reciprocity principle (equation 1.70) for a volume V delimited by the surface Σ defined as the reunion of the following surfaces:

- S_c – the surface infinitesimally close to the shield of the source $\mathbf{g_a}$;
- S_1' – area of the intersection of the reference plane $\mathbf{S_1}$ with the dielectric of the feeding coaxial cable;
- S_∞ – the surface of the sphere with infinite radius.

We have:

$$\int_\Sigma (\boldsymbol{E_a} \times \boldsymbol{H_b} - \boldsymbol{E_b} \times \boldsymbol{H_a}) \cdot d\boldsymbol{\sigma} =$$

$$= \int_V [(\boldsymbol{E_b} \cdot \boldsymbol{J_a} - \boldsymbol{E_a} \cdot \boldsymbol{J_b}) - (\boldsymbol{H_b} \cdot \boldsymbol{J_{ma}} - \boldsymbol{H_a} \cdot \boldsymbol{J_{mb}})]dV \qquad (4.31)$$

We proved previously that:

$$\int_{S_1'} (\boldsymbol{E_a} \times \boldsymbol{H_b} - \boldsymbol{E_b} \times \boldsymbol{H_a}) \cdot d\boldsymbol{\sigma} = V_{1a}I_{1b} + V_{1b}I_{1a} \qquad (4.32)$$

The magnetic current densities $\boldsymbol{J_{ma}}$ and $\boldsymbol{J_{mb}}$ are actually zero as they derive from a fictitious magnetic charge. Also, the electric current density $\boldsymbol{J_a}$ is zero in the volume V, because the source $\mathbf{g_a}$ is not placed inside the volume V. Thus, the volume integral in (4.31) restrains to integration on the volume occupied by antenna **b** and this one becomes a line integral due to the particular shape of the antenna **b**. Hence:

$$\int_V [(\boldsymbol{E_b} \cdot \boldsymbol{J_a} - \boldsymbol{E_a} \cdot \boldsymbol{J_b}) - (\boldsymbol{H_b} \cdot \boldsymbol{J_{ma}} - \boldsymbol{H_a} \cdot \boldsymbol{J_{mb}})]dV = -\int_V \boldsymbol{E_a} \cdot \boldsymbol{J_b} dV =$$

$$= -\int_0^l \boldsymbol{E_a} \cdot \hat{\boldsymbol{b}} Idl = -\boldsymbol{E_a} \cdot \hat{\boldsymbol{b}} Il = -|\boldsymbol{E_a}| Il \cos\varphi \qquad (4.33)$$

where φ is the angle between the vector $\boldsymbol{E_a}$ and the versor $\hat{\boldsymbol{b}}$ along the rectilinear element antenna.

Hence:

$$V_{1a}I_{1b} + V_{1b}I_{1a} = -\boldsymbol{E_a} \cdot \hat{\boldsymbol{b}} Il = -|\boldsymbol{E_a}| Il \cos\varphi \qquad (4.34)$$

After some tedious mathematical manipulations we obtain that the cross section area of an arbitrary antenna **a** is:

$$S_{ef,a} = \frac{\lambda_0^2}{4\pi} G_a(\theta_a, \phi_a) \frac{(\cos\varphi)^2}{(\sin\theta_b)^2} \frac{4\mathrm{Re}(Z_{i1})\mathrm{Re}(Z_{in1})}{|Z_{i1}+Z_{in1}|^2} \qquad (4.35)$$

where θ_b is the angle between the positive direction of the rectilinear element antenna and the line joining it with antenna **a**, while λ_0 is the wavelength of the radiated electromagnetic field.

The maximum value yielded by formula (4.35) is considered as the cross section area of an antenna. In order to obtain it the rectilinear element antenna should be oriented such that its maximum gain be oriented towards antenna **a** and this implies that $\theta_b = \pi/2$. Also, the polarization properties of the rectilinear element antenna should match the polarization properties of the field it receives. This implies that $\varphi = 0$. If these requirements are met:

$$\frac{(\cos\varphi)^2}{(\sin\theta_b)^2} = 1 \qquad (4.36)$$

The maximum power transfer between an antenna and its load is obtained when matching conditions are met, that is:

$$\mathrm{Re}(Z_{i1}) = \mathrm{Re}(Z_{in1}) \text{ and } \mathrm{Im}(Z_{i1}) = -\mathrm{Im}(Z_{in1}) \qquad (4.37)$$

These equations imply:

$$\frac{4\mathrm{Re}(Z_{i1})\mathrm{Re}(Z_{in1})}{|Z_{i1}+Z_{in1}|^2}=1 \tag{4.38}$$

In conclusion, when matching conditions between antenna and its load are met, when the antenna polarization properties match the ones of the received electromagnetic field, and the antennas are relatively oriented such that they are in the maximum gain direction of each other, then the receiving antenna cross section has a maximum value and this is:

$$S_{ef,a}=\frac{\lambda_0^2}{4\pi}G_a(\theta_a,\phi_a) \tag{4.39}$$

This result proves our previous statement that an *antenna cross section is a direct consequence of its directive properties*.

When load is not matched with antenna input impedance the value of the antenna cross section decreases:

$$S_{ef,a}=(1-|\Gamma|^2)\frac{\lambda_0^2}{4\pi}G_a(\theta_a,\phi_a) \tag{4.40}$$

where:

$$\Gamma \overset{\text{def}}{=} \frac{Z_{in1}-Z_c}{Z_{in1}+Z_c} \tag{4.41}$$

is the reflection coefficient at the antenna input port, while Z_c is the characteristic impedance of the transmission line between antenna and its load.

4.5 – Reception of Completely Polarized Waves

Based on their polarization state the electromagnetic waves are grouped into three categories:

- *completely polarized waves* – waves with a fixed polarization state;
- *random polarized waves* – waves with permanently changing polarization state;
- *partial polarized state* – waves having both completely and random polarized components.

The completely polarized waves are specific for technical applications and are the subject of this subchapter. The other two categories of waves are met in the radio astronomy domain and are not discussed here.

We show in the followings that an antenna receives the maximum power when its polarization properties match the polarization state of the received electromagnetic field.

It results from (4.34) that the open circuit voltage at the antenna output port is:

$$V_0=V_{1b}|_{I_{1b}=0}=-\frac{\boldsymbol{E_a}\cdot\widehat{\boldsymbol{b}}}{I_{1a}}Il \tag{4.42}$$

When positioning antenna **b** such that antenna **a** receive the maximum power the vectors $\boldsymbol{E_a}$ and $\widehat{\boldsymbol{b}}$ lie in the same plane normal to the line between antennas **a** and **b** that contains the versor $\hat{\boldsymbol{r}}$. As a consequence the vectors $\boldsymbol{E_a}$ and $\widehat{\boldsymbol{b}}$ have components only along the versors $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\phi}}$ and:

$$\boldsymbol{E_a}\cdot\widehat{\boldsymbol{b}}=E_\theta b_\theta+E_\phi b_\phi \tag{4.43}$$

So, based on (4.42) and (4.43) we could write that:

$$V_0\sim E_\theta b_\theta+E_\phi b_\phi \tag{4.44}$$

The power received by antenna **a** is proportional with the square of the open circuit voltage:

$$P_{rec,a} \sim |V_0|^2 \sim \left|E_\theta b_\theta + E_\phi b_\phi\right|^2 \tag{4.45}$$

We showed that the elliptical polarization is the most general polarization state of an electromagnetic field. This state is analytically described by the relation:

$$E_\phi = \tau e^{j\beta} E_\theta \qquad \text{with} \qquad 0 \le \tau \le 1 \tag{4.46}$$

where E_θ is a real positive variable. For $\beta = 0$ the elliptical polarization degenerates into a linear one, while for $\beta = \pm\pi/2$ and $\tau = 1$ the elliptical polarization degenerates into a circular one.

Based on (4.46) the expression (4.45) becomes:

$$\begin{aligned} P_{rec,a} &\sim \left|E_\theta b_\theta + \tau(\cos\beta + j\sin\beta)E_\theta b_\phi\right|^2 = \\ &= \left(b_\theta^2 + \tau^2 b_\phi^2 + 2\tau b_\theta b_\phi \cos\beta\right)E_\theta^2 \end{aligned} \tag{4.47}$$

The above expression has a maximum for $\beta = 0$, so:

$$P_{rec,a}^{max} \sim \left(b_\theta + \tau b_\phi\right)^2 E_\theta^2 \tag{4.48}$$

This maximum value could be maximized further by choosing particular values for the components b_θ, b_ϕ of the versor $\hat{\boldsymbol{b}}$. These components are not independent, but they are related by the relation:

$$b_\theta^2 + b_\phi^2 = 1 \qquad \Rightarrow \qquad b_\phi = \sqrt{1 - b_\theta^2} \tag{4.49}$$

Note that due to the above relation the expression (4.48) is a function of a single variable b_θ and its maximum is obtained for a value of b_θ for which its derivative is zero. We have:

$$\frac{dP_{rec,a}^{max}}{db_\theta} \sim \left(b_\theta + \tau b_\phi\right)\left(1 - \tau\frac{b_\theta}{b_\phi}\right)E_\theta^2 \tag{4.50}$$

Requiring that this derivative be zero we obtain:

$$\frac{b_\theta}{b_\phi} = \frac{1}{\tau} \tag{4.51}$$

Note that for the above condition of $\beta = 0$ the relation (4.46) yields:

$$\frac{E_\theta}{E_\phi} = \frac{1}{\tau} \tag{4.52}$$

Summarizing, the power received by antenna **a** has a maximum value when:

$$\beta = 0 \quad \text{and} \quad \frac{b_\theta}{b_\phi} = \frac{E_\theta}{E_\phi} \tag{4.53}$$

In other words, the power received by antenna **a** has a maximum value when its polarization is linear ($\beta = 0$), identical to the polarization of the rectilinear element transmitting antenna, and it is orientated in the same direction as the rectilinear element transmitting antenna $\left(\frac{b_\theta}{b_\phi} = \frac{E_\theta}{E_\phi}\right)$.

In general, *a receiving antenna transfers the maximum power to its load when its polarization properties match the polarization properties of the transmitting antenna*.

By using rather complex mathematical derivation the polarization state of an antenna could be described by means of three variables denoted as *Stoke's parameters*. These variables define an unique point on the surface of a sphere with radius one and denoted as

Poincare sphere (figure no. 4.9). The equation of this sphere represents the linear polarization, the North pole represents the left-hand circular polarization, the South pole represents the right-hand circular polarization, and the other points on the sphere's surface represent different types of elliptical polarization. A conjugate polarization state is represented by the symmetric point of the representative point relative to the equatorial plane.

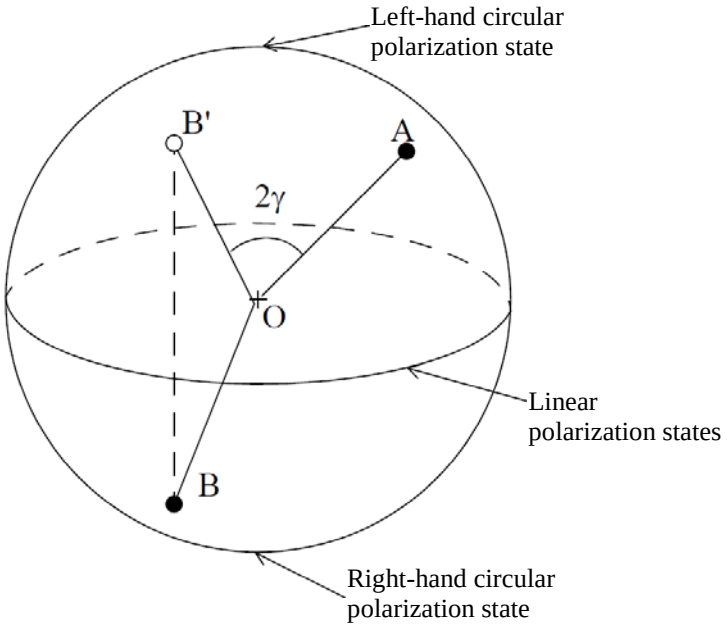


Figure no. 4.9 – *Poincare sphere*

It is proved that the power transfer between two antennas, from the point of view of polarization properties, is dependent on the angle 2γ between the radius corresponding to the polarization state of an antenna and the one corresponding to the conjugate state of the other antenna. Specifically:

$$P_{rec} \sim \cos^2 \gamma \qquad (4.54)$$

4.6 – Noise in Antennas

The minimum level of the signal that can be processed by an electronic system depends on the level of the noise that is associated with it. This noise could be introduced by the electronic blocks of the processing system or it could be added by the antenna that receives it together with the useful signal. The noise signal an antenna receives is produced by man-made sources or it originates from natural sources (Sun, stars, planets etc.). Man-made noise could be easily minimized by an appropriate placing and orienting the receiving antenna. The following presentation aims at natural noise minimization.

Let us consider an isolated system consisting in an antenna and a black body (figure no. 4.10). We assume that the antenna feeds a load R_L by a lossless transmission line with characteristic impedance Z_c and that the temperature of the black body is T. Let us denote by $d\Omega$ the angle in which is seen the black body from the point where antenna is placed.

The power radiated by the black body at temperature T on a unit surface, in a unit solid angle, and in a unit frequency band is denoted as brightness and is given by the Planck formula:

$$B = \frac{2K}{\lambda^2} T \qquad \left[\frac{W}{m^2 ste \cdot cycle}\right] \qquad (4.55)$$

where K is the Boltzmann's constant.

The noise power received by the antenna in a frequency bandwidth Δf is:

$$P_{rec} = (1 - |\Gamma|^2) B S_{ef} \frac{\Delta f}{2} d\Omega = (1 - |\Gamma|^2) \frac{2K}{\lambda^2} T \frac{\lambda^2}{4\pi} G \frac{\Delta f}{2} d\Omega =$$

$$= (1 - |\Gamma|^2) KT\Delta f \frac{G d\Omega}{4\pi} \qquad (4.56)$$

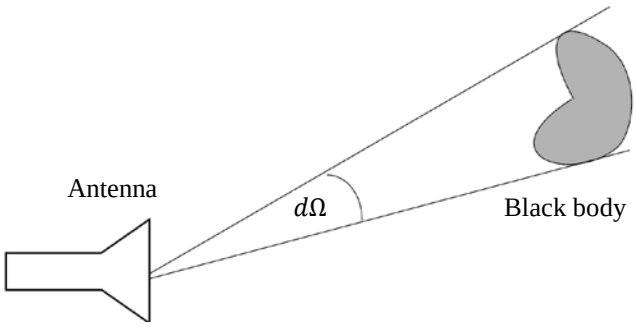


Figure no. 4.10 – *Antenna-Black body system*

At the thermodynamic equilibrium the power transmitted by the antenna should equals the power absorbed by the black body.

When the black body is not a perfect one and it absorbs only a fraction α of the incident power, the equilibrium is established for a receiving power level of:

$$P_{rec} = \alpha (1 - |\Gamma|^2) KT\Delta f \frac{G d\Omega}{4\pi} \qquad (4.57)$$

Actually, a receiving antenna is surrounded by a multitude bodies with different equivalent noise temperatures and with different absorption coefficients. Then, the power exchanged by the antenna with these bodies should be:

$$P_{rec} = (1 - |\Gamma|^2) K\Delta f \int_{4\pi} \alpha(\theta, \phi) T(\theta, \phi) G(\theta, \phi) \frac{d\Omega}{4\pi} \qquad (4.58)$$

The above expression should be modified if we take into consideration the losses in the transmission line. If the transmission line has a loss coefficient α_0 and an equivalent noise temperature T_0, then the power exchanged by the antenna with the surrounding bodies becomes:

$$P_{rec} = (1 - |\Gamma|^2) K\Delta f \int_{4\pi} \alpha(\theta, \phi) T(\theta, \phi) G(\theta, \phi) \frac{d\Omega}{4\pi} + (1 - |\Gamma|^2) \alpha_0 K\Delta f T_0 \qquad (4.59)$$

If the antenna and its surrounding bodies have the same temperature T_a and the antenna exchanges the same power P_{rec} with them, then the expression of this power would be:

$$P_{rec} = (1 - |\Gamma|^2) K\Delta f T_a \qquad (4.60)$$

Equating the last two expressions of P_{rec} we obtain that:

$$T_a = \alpha_0 T_0 + \int_{4\pi} \alpha(\theta, \phi) T(\theta, \phi) G(\theta, \phi) \frac{d\Omega}{4\pi} \qquad (4.61)$$

which represents the *equivalent noise temperature* of the antenna.

For an antenna to have a small equivalent noise temperature it is necessary that the transmission line connecting the antenna with the load have small losses and that the lobes of the radiation pattern (especially, the main lobe) not to be directed towards surrounding bodies with high noise temperature (the product $T(\theta, \phi) G(\theta, \phi)$ should be kept as small as

possible). Common high power noise sources are the Earth, the Sun, and some celestial bodies. Their positions are very well known and high gain antennas for special applications should be placed such that they avoid orienting their lobes towards these bodies. For common applications attention should be paid to large natural obstacles (hills, buildings, etc.) in the neighborhood.

C h a p t e r V

ANTENNA ARRAYS

Many applications require radiation patterns that cannot be realized by using a single antenna, but by a group of many antennas appropriately positioned relative to each other and fed by currents with amplitudes and phases adequately computed. They are denoted as *antenna arrays*.

5.1 – Factorization

The expression of the electromagnetic field radiated by an antenna array is found out by using the one previously obtained for the field radiated by an arbitrary current distribution. We proved in Chapter 2 that the electric vector potential produced in the radiation region by an arbitrary current distribution $\boldsymbol{J}(\boldsymbol{r}')$ is:

$$\boldsymbol{A}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\frac{e^{-jk_0 r}}{r}\int_V \boldsymbol{J}(\boldsymbol{r}')e^{jk_0\hat{\boldsymbol{r}}\cdot\boldsymbol{r}'}dV \qquad (5.1)$$

where V is the volume occupied by this current distribution.

Let us consider an array of n identical antennas, identically oriented having similar current distributions that differs from each other only by multiplicative complex coefficients. We define the *reference antenna* as an antenna identical to the ones in the array, identically oriented with them, with a similar current distribution and *placed in the origin of the coordinate system.*

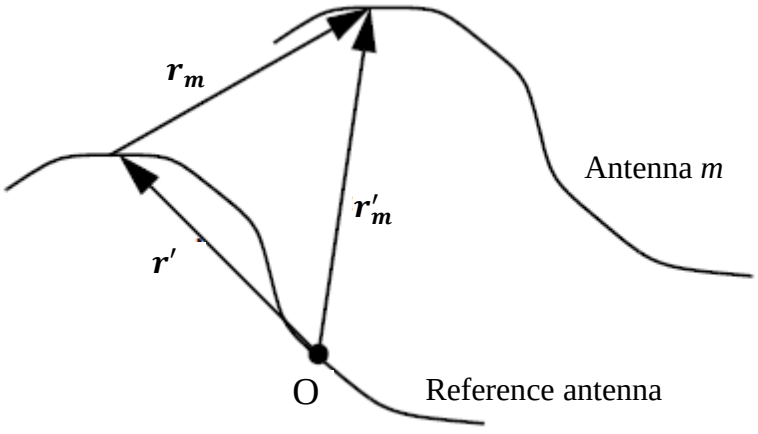


Figure no. 5.1 – *Geometry of antenna array*

According to the above definition the position vector of the reference antenna is zero and the phase of its current distribution is also zero.

We conclude from definitions of the antenna array and of the reference antenna that any antenna m of the array can be regarded as a translation of the reference antenna by a translation vector $\boldsymbol{r_m}$, while its current distribution is the reference antenna current distribution multiplied by a complex coefficient a_m. Based on the notations in figure no. 5.1 we can write:

$$\boldsymbol{r'_m} = \boldsymbol{r'} + \boldsymbol{r_m} \quad , \quad \boldsymbol{J}(\boldsymbol{r'_m}) = a_m \boldsymbol{J_{ref}}(\boldsymbol{r'}) \tag{5.2}$$

According to (5.1) the vector potential created in the radiation region by an arbitrary antenna m of the array is:

$$\boldsymbol{A_m}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\frac{e^{-jk_0 r}}{r}\int_{V_m} \boldsymbol{J}(\boldsymbol{r'_m}) e^{jk_0 \hat{\boldsymbol{r}}\cdot \boldsymbol{r'_m}} dV_m \tag{5.3}$$

where V_m is the volume occupied by the antenna m.

Assuming that the field radiated by an individual antenna of the array is not influenced by the other antennas of the array we can write the vector potential created by the whole array in the radiation region as the sum of the vector potentials created by each of the antennas:

$$\boldsymbol{A}(\boldsymbol{r}) = \sum_{m=1}^{n} \boldsymbol{A_m}(\boldsymbol{r}) = \frac{\mu_0}{4\pi}\frac{e^{-jk_0 r}}{r}\sum_{m=1}^{n}\left[\int_{V_m} \boldsymbol{J}(\boldsymbol{r'_m}) e^{jk_0 \hat{\boldsymbol{r}}\cdot \boldsymbol{r'_m}} dV_m\right] \tag{5.4}$$

Extending the integral in the above relation to the whole volume V occupied by the array does not modify the result of the integration because the current density $\boldsymbol{J}(\boldsymbol{r'_m})$ is zero outside the volume V_m occupied by the antenna m. Thus:

$$\begin{aligned}
\boldsymbol{A}(\boldsymbol{r}) &= \frac{\mu_0}{4\pi}\frac{e^{-jk_0 r}}{r}\sum_{m=1}^{n}\left[\int_{V} \boldsymbol{J}(\boldsymbol{r'_m}) e^{jk_0 \hat{\boldsymbol{r}}\cdot \boldsymbol{r'_m}} dV\right] = \\
&= \frac{\mu_0}{4\pi}\frac{e^{-jk_0 r}}{r}\sum_{m=1}^{n}\left[\int_{V} a_m \boldsymbol{J_{ref}}(\boldsymbol{r'}) e^{jk_0 \hat{\boldsymbol{r}}\cdot (\boldsymbol{r'}+\boldsymbol{r_m})} dV\right] = \\
&= \left[\frac{\mu_0}{4\pi}\frac{e^{-jk_0 r}}{r}\int_{V} \boldsymbol{J_{ref}}(\boldsymbol{r'}) e^{jk_0 \hat{\boldsymbol{r}}\cdot \boldsymbol{r'}} dV\right]\left[\sum_{m=1}^{n} a_m e^{jk_0 \hat{\boldsymbol{r}}\cdot \boldsymbol{r_m}}\right]
\end{aligned} \tag{5.5}$$

In developing the above relation we took into account that the summation operator is independent from the integration one and they could be interchanged as order of application to the same argument. Also, we regrouped the variables and highlighted two factors with significant meaning:

- the first one describes the vector potential created in the radiation region by the reference antenna $\boldsymbol{A_{ref}}(\boldsymbol{r})$. As the reference antenna is identical to all the antennas in the array, this factor characterizes the radiation of any individual antenna (array element) and it is denoted as the *element factor*.
- the second factor describes the relative positioning ($\boldsymbol{r_m}$) of the individual antennas inside the array and the relative amplitudes (a_m) and the relative phases ($k_0 \hat{\boldsymbol{r}} \cdot \boldsymbol{r_m}$) of their current distributions; it is denoted as the *array factor*. Note that it does not depend on the position vector module r, but only on its direction in space (through $\hat{\boldsymbol{r}}$). To highlight this, the usual notation for the array factor is $f(\theta, \phi)$:

$$f(\theta, \phi) = \sum_{m=1}^{n} a_m e^{jk_0 \hat{\boldsymbol{r}}\cdot \boldsymbol{r_m}} \tag{5.6}$$

The relation (5.5) could be rewritten shortly as:

$$\boldsymbol{A}(\boldsymbol{r}) = \boldsymbol{A_{ref}}(\boldsymbol{r}) \cdot f(\theta, \phi) \tag{5.7}$$

It could be proved that the expressions of the field intensity vectors in the radiation region could be written, also, as:

$$\boldsymbol{E}(\boldsymbol{r}) = \boldsymbol{E_{ref}}(\boldsymbol{r}) \cdot f(\theta, \phi)$$
$$\boldsymbol{H}(\boldsymbol{r}) = \boldsymbol{H_{ref}}(\boldsymbol{r}) \cdot f(\theta, \phi) \qquad (5.8)$$

The above expressions suggest that the radiation pattern of an antenna array is the radiation pattern of its reference antenna multiplied by $|f(\theta, \phi)|$; this is why $|f(\theta, \phi)|$ is regarded as the *array radiation pattern*.

We underline that the array radiation pattern, as defined above, depends only on the relative positioning of its elements and on the relative modules and relative phases of their current distributions.

Note, also, that the *total radiation pattern* of an antenna array is the radiation pattern of its reference antenna ($\left|\boldsymbol{E_{ref}}(\boldsymbol{r})\right|$) multiplied by the array radiation pattern as defined above ($|f(\theta, \phi)|$).

The elements of an array could be distributed along a straight line (linear array), on the circumference of a circle (circular array), on a plane surface (planar array) or inside a volume (3D array).

5.2 – Uniform Linear Arrays

The uniform linear array (ULA) has, according with the general definition of an array: identical antennas, identically oriented and similar current distributions through the antennas, and, supplementary: *antennas distributed regularly along a straight line (the same separation distance between two neighbor antennas), equal amplitudes of the current distributions, and current phases that vary with same value from an antenna to the next one*.

For an n element array the module of the current distributions is usually normalized to $1/n$, the value δ with which the current phases vary from an element to the other is denoted as array *phase constant*, while the separation distance d between any two neighbor elements is denoted as the array *distance constant*.

Assuming that an ULA is laid along the *Oz* axis of a Cartesian coordinate system and that the first element (no. 1) of the array is chosen as the reference antenna, then the antenna no. m has:

$$a_m = \frac{1}{n} e^{j(m-1)\delta} \quad \text{and} \quad \boldsymbol{r_m} = (m-1)d\hat{\boldsymbol{z}} \qquad (5.9)$$

It results that:

$$f(\theta, \phi) = \sum_{m=1}^{n} \left[\frac{1}{n} e^{j(m-1)\delta} e^{jk_0 \hat{\boldsymbol{r}} \cdot (m-1)d\hat{\boldsymbol{z}}}\right] = \frac{1}{n} \sum_{m=1}^{n} \left[e^{j(m-1)(\delta + k_0 d \cos\theta)}\right] \qquad (5.10)$$

Note that for ULAs the array factor $f(\theta, \phi)$ is only a function of θ.

Making the notation:

$$\gamma \triangleq \delta + k_0 d \cos\theta \qquad (5.11)$$

the array factor could be written in a more compact form:

$$f(\theta, \phi) = \frac{1}{n} \sum_{m=1}^{n} e^{j(m-1)\gamma} = \frac{1}{n} \frac{e^{jn\gamma} - 1}{e^{j\gamma} - 1} = e^{j\frac{(n-1)\gamma}{2}} \frac{\sin\frac{n\gamma}{2}}{n \sin\frac{\gamma}{2}} \qquad (5.12)$$

Hence, the array radiation pattern is:

$$|f(\theta, \phi)| = \left|\frac{\sin\frac{n\gamma}{2}}{n \sin\frac{\gamma}{2}}\right| \qquad (5.13)$$

Given the particular expression of the argument γ (see 5.11), the graphical representation of the radiation pattern consists in a *polar diagram* construction by taking the following steps (see figure no. 5.2):

- make a Cartesian representation of $|f(\gamma)|$; the horizontal *axis γ coincides with the axis Oz* on which the array elements are positioned;
- redraw the horizontal axis γ some distance lower;
- draw a semicircle with radius $k_0 d$ centered in the point A on the new axis γ at a distance δ from the origin;
- draw a segment AD on every radius of the semicircle with a length equal to the value of $|f(\gamma)|$ (segment MN) in the direction θ (θ is the angle between the positive direction of axis γ and the considered radius); *the set of all points D represents the array radiation pattern* in a plane containing the axis *Oz*.

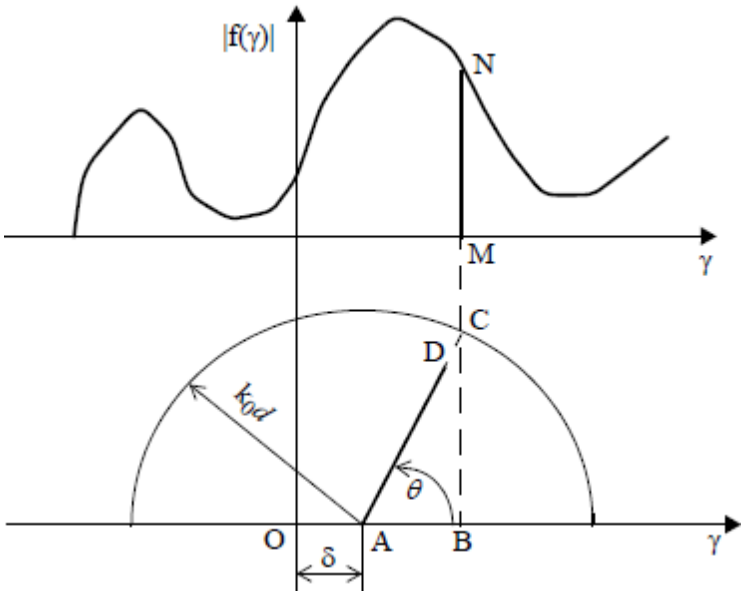


Figure no. 5.2 – *Polar diagram principle*

Because the array factor does not depend on ϕ the radiation pattern has revolution symmetry around the axis *Oz* and, thus, by building the symmetric of the previously obtained curve relative to axis Oz we get the array radiation pattern in the paper plane. The tri dimensional array radiation pattern is obtained by rotating this curve around the axis Oz.

The nominal variation interval of θ is $[0, \pi]$; this implies that $\cos\theta \in [-1,1]$ and that $\gamma \in [\delta - k_0 d, \delta + k_0 d]$. This interval is denoted as the *visible domain* of the variable γ. Its length is $2k_0 d$ and is directly related to the array length through the array distance constant, while its position on the axis γ is given by the array phase constant (the visible domain is centered in δ).

Figures no. 5.3 – 5.6 illustrate the building of the array radiation pattern for a two element array ($n = 2$) using the above presented polar diagram construction, for three values of the phase constant and four values of the distance constant. In each case the array factor:

$$|f(\theta, \phi)| = \left|\frac{\sin\frac{2\gamma}{2}}{2\sin\frac{\gamma}{2}}\right| = \left|\cos\frac{\gamma}{2}\right| \qquad (5.14)$$

is drawn only for the visible domain of γ.

In figure no. 5.3 the visible domain has the length: $2\frac{2\pi}{\lambda}\frac{\lambda}{4} = \pi$.

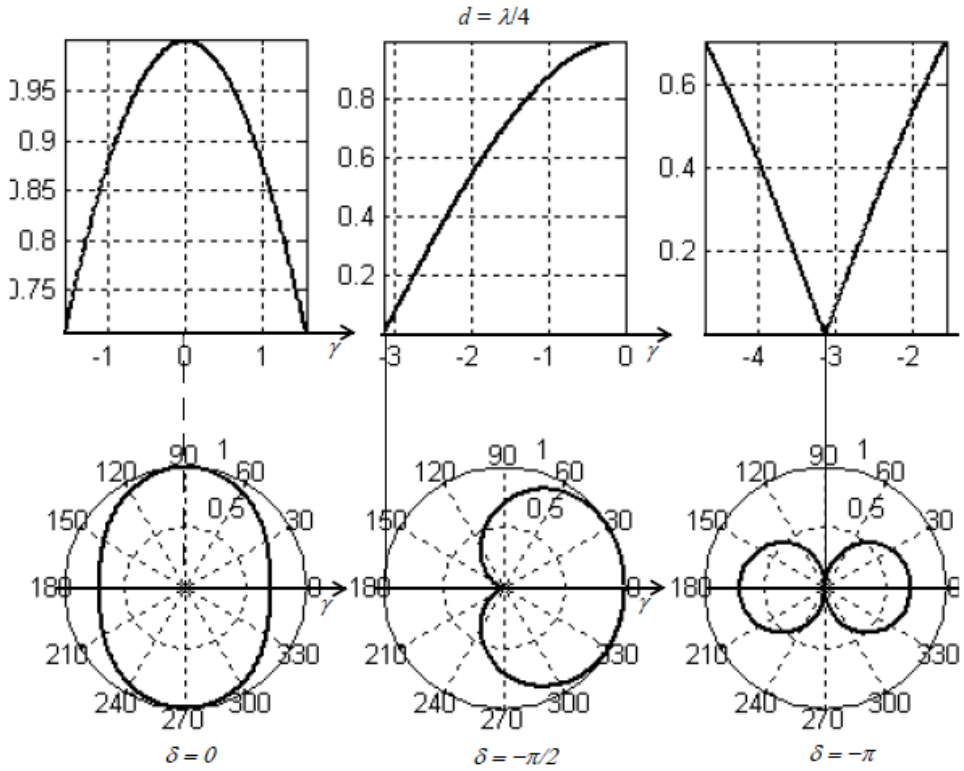


Figure no. 5.3 – *Radiation patterns for two element uniform array*

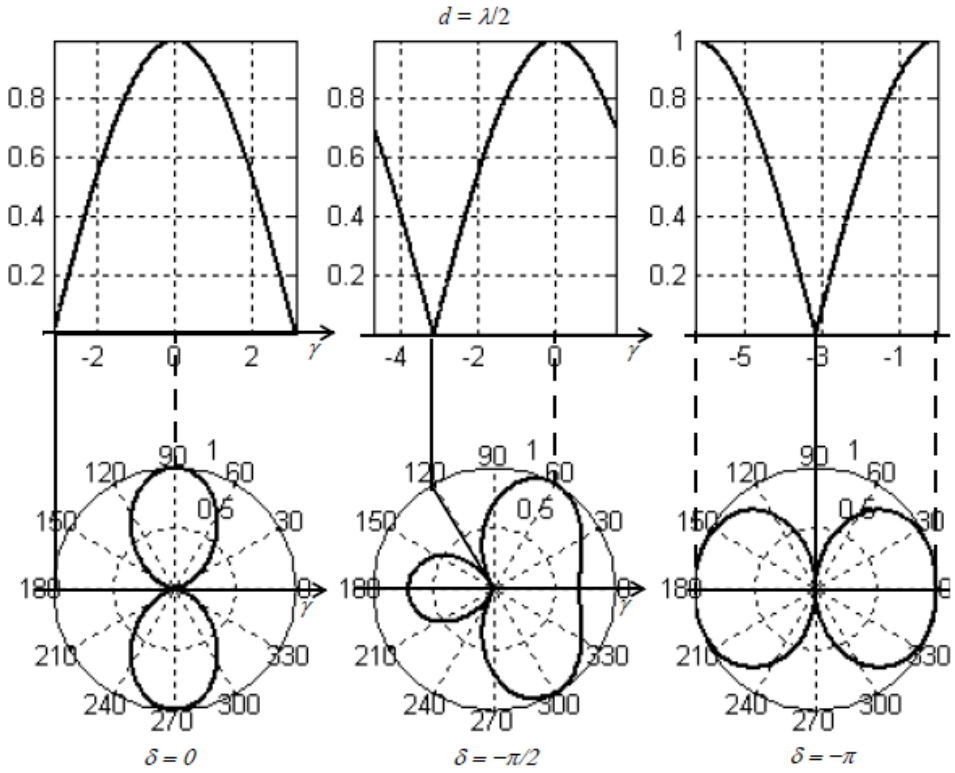


Figure no. 5.4 – *Radiation patterns for two element uniform array*

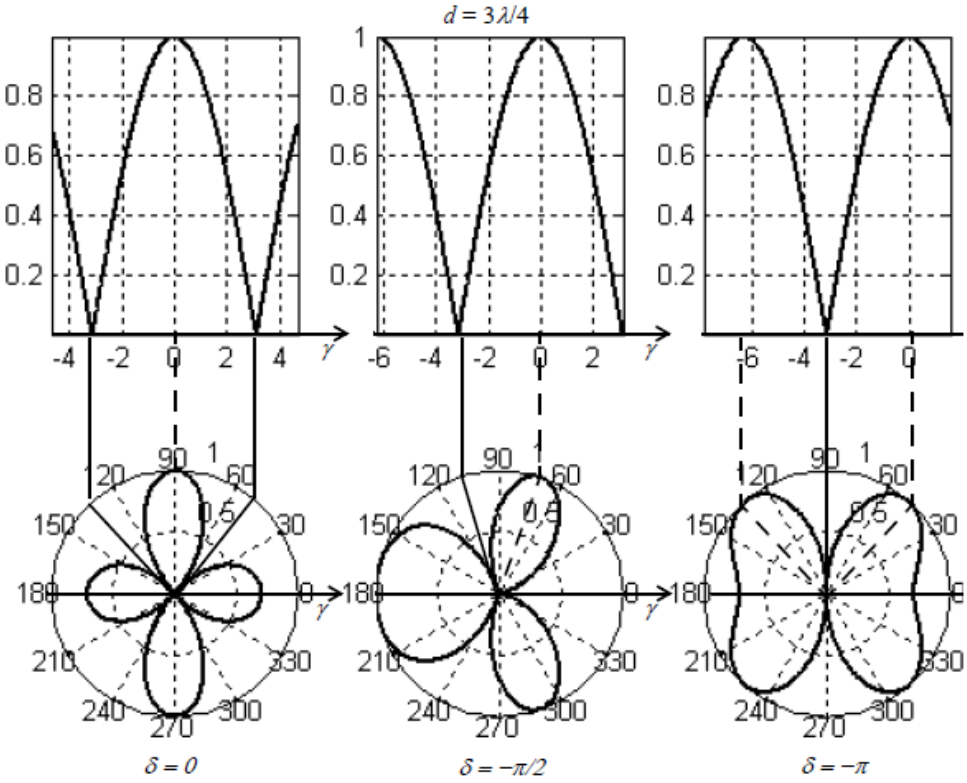


Figure no. 5.5 – *Radiation patterns for two element uniform array*

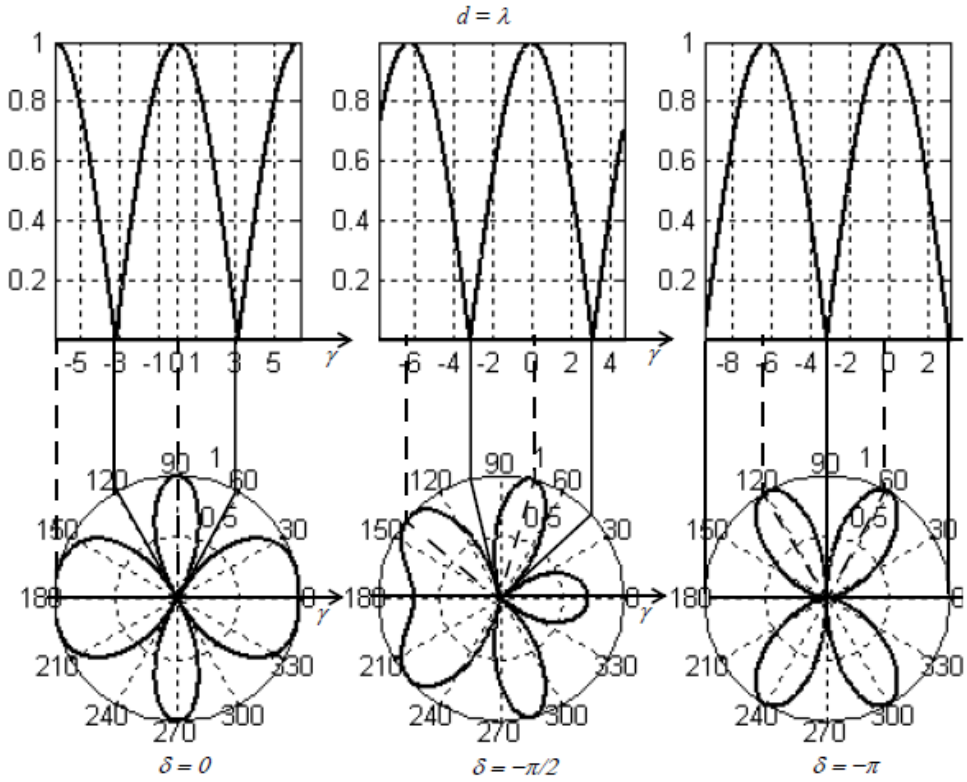


Figure no. 5.6 – *Radiation patterns for two element uniform array*

For $\delta = 0$ it covers the interval $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. When $\theta = 0$ the array factor is $\frac{1}{\sqrt{2}} \approx 0.707$, it increases monotonically while θ increases towards $\frac{\pi}{2}$, it reaches the maximum value of 1 when $\theta = \frac{\pi}{2}$, it decreases monotonically for θ increasing to π, and it reaches again the starting value of $\frac{1}{\sqrt{2}}$ when $\theta = \pi$. Correspondingly, the array radiation pattern is an ellipsis.

For $\delta = -\frac{\pi}{2}$ the visible domain covers the interval $[-\pi, 0]$. When $\theta = 0$ the array factor is 1, it decreases monotonically for θ increasing to π, and it reaches the value of 0 when $\theta = \pi$. Correspondingly, the array radiation pattern is a cardioid.

For $\delta = -\pi$ the visible domain covers the interval $\left[-\frac{3\pi}{2}, -\frac{\pi}{2}\right]$. When $\theta = 0$ the array factor is $\frac{1}{\sqrt{2}} \approx 0.707$, it decreases monotonically while θ increases towards $\frac{\pi}{2}$, it reaches the minimum value of 0 when $\theta = \frac{\pi}{2}$, it increases monotonically for θ increasing to π, and it reaches again the starting value of $\frac{1}{\sqrt{2}}$ when $\theta = \pi$. Correspondingly, the array radiation pattern comprises two opposite lobes with the same level.

Similar rules apply for building the array radiation pattern in figures no. 5.4 – 5.6.

The following features, common for ULAs, are obvious while inspecting the above figures:

- for $\delta = 0$ the main lobe is in the direction $\theta = \frac{\pi}{2}$, that is – perpendicular on the axis along which the elements of the array are positioned; the array with such a radiation pattern is denoted as a *broadside array*;
- for $\delta + k_0 d = 2p\pi$, where p is an arbitrary integer, the main lobe is in the direction $\theta = 0$, that is – in the direction of the axis along which the elements of the array are positioned; the array with such a radiation pattern is denoted as an *end-fire array*. For $p = 0$ the above condition becomes $\delta = -k_0 d$, which is the *basic condition* for an ULA to be and end-fire array.
- for a fixed value of δ when the array distance constant increases, the number of lobes of the radiation pattern also increases. It is proved that for a fixed d the number of lobes increases for increasing n; because $(n-1)d$ is the physical length of an array, the obvious conclusion is that the *number of lobes increases when the array length increases*.

Going back to the ULA array factor expression (5.13), note that it is a periodic function with the period 2π, it is symmetrical around $\gamma = 0$, it has $n - 1$ zeros for:

$$\frac{n\gamma}{2} = p\pi \quad \Longrightarrow \quad \gamma = p\frac{2\pi}{n}, \quad p = 1, 2, \ldots\ldots\ldots, n-1 \qquad (5.15)$$

it has a global maximum equal to 1, and it has local maxima less than 1 approximately in the middle of the intervals between zeros.

Figure no. 5.7 illustrates the array factor for $n = 2 \div 10$ and $0 \leq \gamma \leq \pi$.

Figures no. 5.8 and 5.9 illustrate the radiation pattern of an ULA with $n = 6$ elements for $\delta = 0$ (broadside array) and $\delta = -k_0 d$ (end-fire array), respectively. Note that the main lobe of the broadside array is narrower than the one of the end-fire array. This property holds for any ULA.

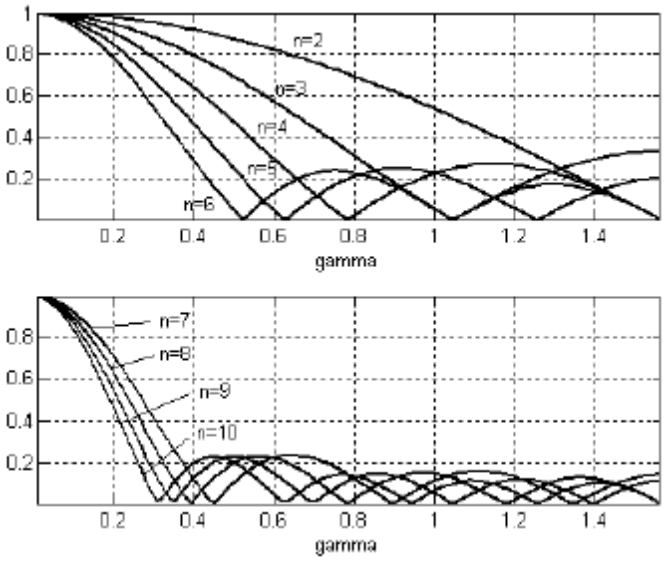


Figure no. 5.7 – *Array factor for 2 to 10 element arrays*

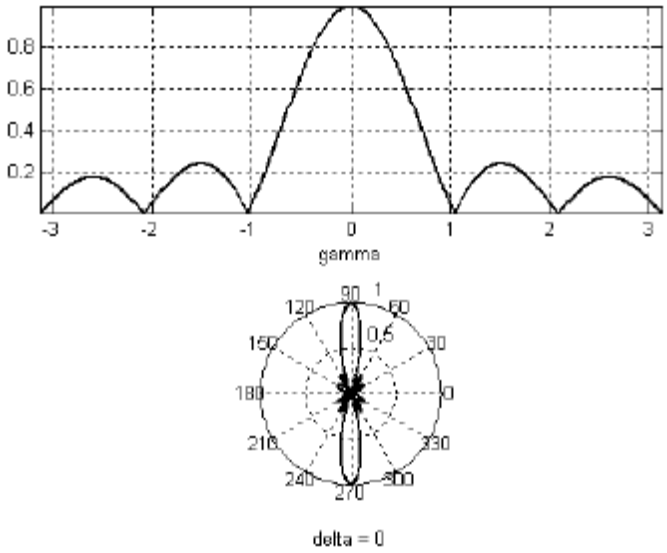


Figure no. 5.8 – *Radiation pattern for a broadside array*

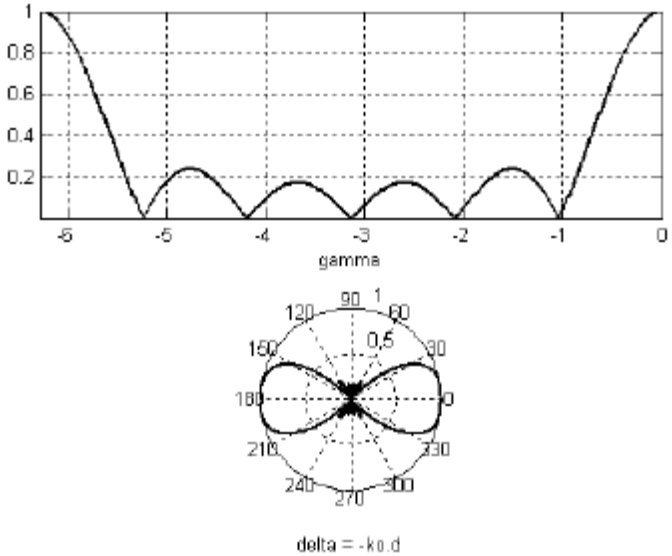


Figure no. 5.9 – *Radiation pattern for a end-fire array*

5.2 – Directive Properties of the Uniform Linear Arrays

Main Lobe Beamwidth

Figure 5.10 illustrates the radiation pattern of an ULA with an arbitrary number of elements and an arbitrary phase constant. Direction of the main lobe is θ_0, while the closest nulls to the main lobe are directed towards θ_1 and θ_2. Let's define $2\alpha = \theta_2 - \theta_1$ as the main lobe beamwidth (note that this definition differs from the standard one, but it is very useful here as it simplifies the relations; the qualitative result of this analysis remains true when using the standard definition).

For big values of n, α is small and the following approximations hold:

$$\theta_1 \approx \theta_0 - \alpha \quad , \quad \theta_2 \approx \theta_0 + \alpha \tag{5.16}$$

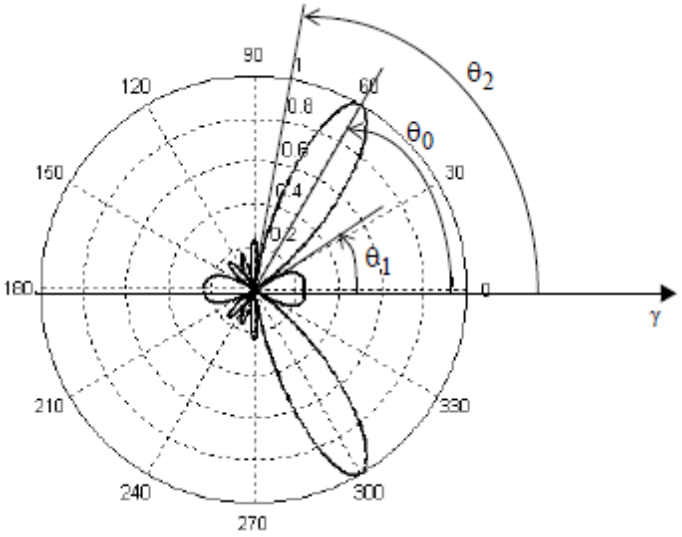


Figure no. 5.10 – *Radiation pattern for an arbitrary ULA*

From (5.15) we see that the closest nulls to the main lobe are obtained for $\gamma = \pm\frac{2\pi}{n}$; hence:

$$\delta + k_0 d \cos\theta_2 = -\frac{2\pi}{n} \tag{5.17}$$

The direction of the main lobe is always obtained for $\gamma = 0$; hence:

$$\delta + k_0 d \cos\theta_0 = 0 \tag{5.18}$$

The last two relations yield:

$$\cos\theta_2 - \cos\theta_0 = -\frac{2\pi}{n}\frac{1}{k_0 d} \tag{5.19}$$

For very small values of α, $\cos\theta_2$ is very well approximate by the first three terms of its Taylor series:

$$\cos\theta_2 = \cos(\theta_0 + \alpha) \approx \cos\theta_0 - \alpha\sin\theta_0 - \frac{\alpha^2}{2}\cos\theta_0 \tag{5.20}$$

Based on this approximation and replacing $k_0 = \frac{2\pi}{\lambda_0}$, the relation (5.19) becomes:

$$2\alpha\sin\theta_0 + \alpha^2\cos\theta_0 = \frac{2\lambda_0}{nd} \tag{5.21}$$

This relation yields for a broadside array $\left(\theta_0 = \frac{\pi}{2}\right)$ a main lobe beamwidth of:

$$2\alpha_{broadside} = \frac{2\lambda_0}{nd} \tag{5.22}$$

while for an end-fire array ($\theta_0 = 0$) the main lobe width is:

$$2\alpha_{end-fire} = 2\sqrt{\frac{2\lambda_0}{nd}} \tag{5.23}$$

As we considered an ULA with a big number of elements, then its length $(n-1)d$ is greater than $2\lambda_0$, and also, $nd > 2\lambda_0$ or $\frac{2\lambda_0}{nd} < 1$. Hence:

$$2\alpha_{broadside} = \frac{2\lambda_0}{nd} < 2\frac{2\lambda_0}{nd} < 2\sqrt{\frac{2\lambda_0}{nd}} = 2\alpha_{end-fire} \tag{5.24}$$

In words: *the radiation pattern of an ULA operating as a broadside array has a main lobe narrower than the one when the same ULA operates as an end-fire array.*

Side Lobe Maximum Level

An antenna is supposed to concentrate the radiated power in the desired space region. Actually, part of the radiated power is radiated outside this area. A measure of the antenna capability of focusing its radiation in the desired solid angle is the relative levels of its side lobes, mainly the maximum of them.

For an ULA the side lobe with the maximum level is the one closest to the main lobe (see figure no. 5.7). For big values of the number n of the array elements we could approximate the direction of this lobe with the average of the directions of the adjacent nulls. From (5.13) we see that the two successive nulls closest to the main lobe are obtained for $\gamma_1 = 2\pi/n$ and $\gamma_2 = 2 \cdot 2\pi/n$, thus the direction of the closest side lobe to the main lobe is obtained for $\gamma' \approx 3\pi/n$ and its level is (see 5.13):

$$r \approx \left|\frac{\sin\left(n\frac{3\pi/n}{2}\right)}{n\sin\left(\frac{3\pi/n}{2}\right)}\right| = \left|\frac{1}{n\sin\left(\frac{3\pi}{2n}\right)}\right| \approx \frac{2}{3\pi} \tag{5.25}$$

Hence, the maximum side lobe level decreases as the number of elements increases and it reaches in the limit (for $n \to \infty$) a value of $\approx 2/3\pi$ or ≈ -13.5 dB.

Maximum Directivity

The maximum directivity of an antenna or of an antenna array is one of the main parameters of evaluating its capability of focusing the radiated power in a desired solid angle.

For *broadside* ULAs ($\delta = 0$), assuming that the distance constant d is much smaller than the radiation wavelength λ_0, the array factor could be approximated by:

$$f(\theta,\phi) = \left.\frac{\sin\frac{n\gamma}{2}}{n\sin\frac{\gamma}{2}}\right|_{\delta=0, d\ll\lambda_0} \approx \frac{\sin\left(\frac{nk_0 d\cos\theta}{2}\right)}{\frac{nk_0 d\cos\theta}{2}} = \frac{\sin z}{z} \tag{5.26}$$

where $z \stackrel{\text{def}}{=} \frac{nk_0 d\cos\theta}{2}$.

The radiation intensity is:

$$P_\Omega(\theta,\phi) = [f(\theta,\phi)]^2 = \left(\frac{\sin z}{z}\right)^2 \tag{5.27}$$

and it has a maximum value of 1 for $z = 0$ (that is, for $\theta = \pi/2$), as expected.

The total radiated power is:

$$P_{rad} = \int_0^{2\pi}\int_0^{\pi}\left(\frac{\sin z}{z}\right)^2 \sin\theta\, d\theta d\phi = \frac{4\pi}{nk_0 d}\int_{-z_m}^{z_m}\left(\frac{\sin z}{z}\right)^2 dz \qquad (5.28)$$

where $z_m \stackrel{\text{def}}{=} \frac{nk_0 d}{2}$.

For n very big, the limits of the integration could be extended towards ∞, without significantly modifying the result of the integration. Hence:

$$P_{rad} \approx \frac{4\pi}{nk_0 d}\int_{-\infty}^{\infty}\left(\frac{\sin z}{z}\right)^2 dz = \frac{4\pi^2}{nk_0 d} \qquad (5.29)$$

It results that:

$$D_{max,broadside} \stackrel{\text{def}}{=} 4\pi\frac{P_{\Omega,max}}{P_{rad}} \approx \frac{nk_0 d}{\pi} = 2n\frac{d}{\lambda_0} \qquad (5.30)$$

Knowing that the physical length of an ULA is $L = (n-1)d$ we obtain that:

$$D_{max,broadside} \approx 2\left(1+\frac{L}{d}\right)\frac{d}{\lambda_0}\Big|_{L\gg d} \approx 2\frac{L}{\lambda_0} \qquad (5.31)$$

For *end-fire* ULAs ($\delta = -k_0 d$), assuming that the distance constant d is much smaller than the radiation wavelength λ_0, the array factor could be approximated by:

$$f(\theta,\phi) = \frac{\sin\frac{n\gamma}{2}}{n\sin\frac{\gamma}{2}}\Big|_{\delta=-k_0 d, d\ll\lambda_0} \approx \frac{\sin\left(\frac{nk_0 d(\cos\theta-1)}{2}\right)}{\frac{nk_0 d(\cos\theta-1)}{2}} = \frac{\sin z}{z} \qquad (5.32)$$

where $z \stackrel{\text{def}}{=} \frac{nk_0 d(\cos\theta-1)}{2}$.

The radiation intensity is:

$$P_\Omega(\theta,\phi) = [f(\theta,\phi)]^2 = \left(\frac{\sin z}{z}\right)^2 \qquad (5.33)$$

and it has a maximum value of 1 for $z = 0$ (that is, for $\theta = 0$), as expected.

The total radiated power is:

$$P_{rad} = \int_0^{2\pi}\int_0^{\pi}\left(\frac{\sin z}{z}\right)^2 \sin\theta\, d\theta d\phi = \frac{4\pi}{nk_0 d}\int_0^{2z_m}\left(\frac{\sin z}{z}\right)^2 dz \qquad (5.34)$$

where $z_m \stackrel{\text{def}}{=} \frac{nk_0 d}{2}$.

For n very big, the upper limit of the integration could be extended towards ∞, without significantly modifying the result of the integration. Hence:

$$P_{rad} \approx \frac{4\pi}{nk_0 d}\int_0^{\infty}\left(\frac{\sin z}{z}\right)^2 dz = \frac{2\pi^2}{nk_0 d} \qquad (5.35)$$

It results that:

$$D_{max,end-fire} \stackrel{\text{def}}{=} 4\pi\frac{P_{\Omega,max}}{P_{rad}} \approx \frac{2nk_0 d}{\pi} = 4n\frac{d}{\lambda_0} \approx 4\left(1+\frac{L}{d}\right)\frac{d}{\lambda_0} \approx 4\frac{L}{\lambda_0} \qquad (5.36)$$

Based on the above results we conclude that the *maximum directivity of an ULA increases with the array length increase* and that *it almost doubles for end-fire operation mode as compared to the broadside operation mode*.

Hansen-Woodyard condition

By using graphical techniques, proved analytically afterwards, it was found out that the maximum directivity of an end-fire ULA can be increased beyond the limit (5.36) if the

basic condition ($\delta = -k_0 d$) for end-fire operation is replaced by the Hansen-Woodyard condition:

$$\delta = -k_0 d - \frac{2.94}{n} \approx -k_0 d - \frac{\pi}{n} \tag{5.37}$$

where n is the number of elements in the array. If this condition holds:

$$D_{max} \approx 1.805 \left(4n \frac{d}{\lambda_0}\right) = 7.22 \frac{L}{\lambda_0} \tag{5.38}$$

while the beamwidth of the main lobe is:

$$\theta_{-3dB} \approx 2 \arccos\left(1 - 0.1398 \frac{\lambda_0}{nd}\right) \tag{5.39}$$

5.3 – Linear Arrays with Tapered Current Distributions

Considering the line along which the elements of a linear array are positioned as an axis of a Cartesian coordinate system and associating to every point on this axis were an element exists a complex number representing the coefficient of its current distribution we obtain a discrete complex function dependent on a real variable. A continuous function built as a natural extension of the above discrete one can be associated to the antenna array. This continuous complex function is equivalent with two real functions: one for current modules and the second for current phases.

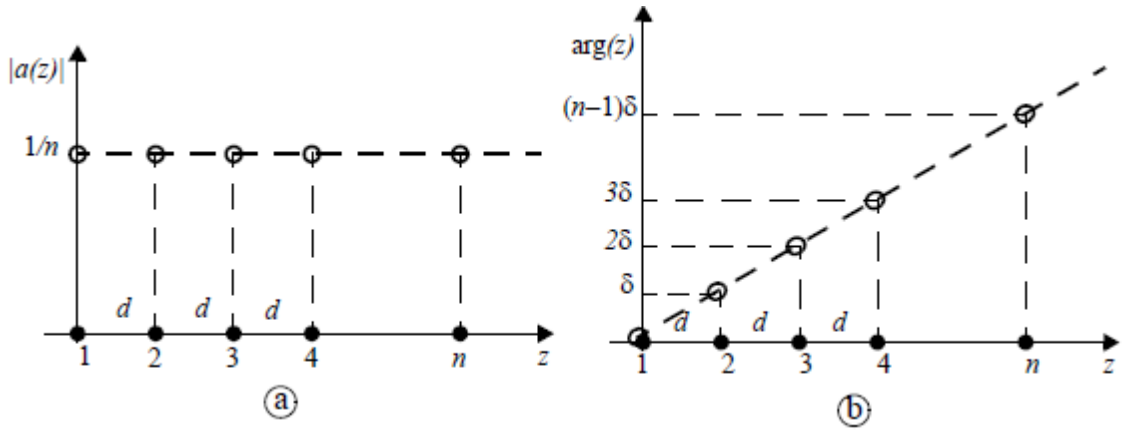


Figure no. 5.11 – *Real functions of modules (a) and phases (b) associated to an ULA*

According to the definition, the real functions illustrated in figure no. 5.11 are the most natural associations for an ULA: a horizontal line for the modules and a line passing through the origin for the phases.

The associated function for phases remain true for almost all practical antenna arrays. But, as the modules are regarding, many of practical arrays associate with different functions than the one of ULAs. Usually, the associated functions for the current modules have a maximum in the middle of the array and monotonically decrease towards the array edges. These current distributions are denoted as *tapered distributions*.

Two methods are used for the analysis of the arrays with tapered current distributions: *polynomial* method and *Z-transform* method.

Polynomial Method

When the modules $|a_m|$ of the current distribution coefficients are not equal with each other, but the phases of the current distributions remain to change with the same

quantity δ from an element to the next one, then the expression (5.12) of the array factor for a linear array with n equidistant elements is:

$$f(\theta,\phi) = \sum_{m=1}^{n} |a_m| e^{j(m-1)\gamma} \qquad (5.40)$$

where the variable γ is the same notation (5.11).

By renumbering of the elements' array from 0 to $n-1$, the above expression is rewritten as:

$$f(\theta,\phi) = \sum_{i=0}^{n-1} |a_i| e^{ji\gamma} = \sum_{i=0}^{n-1} |a_i| \eta^i \qquad (5.41)$$

where we made the notation $\eta \stackrel{\text{def}}{=} e^{j\gamma}$.

The expression in η of the array factor is, in fact, an $n-1$ order polynomial with real coefficients, denoted as Shelkunoff polynomial. When the angle θ varies from 0 to π, the variable η moves in the complex plane on the circle with radius 1 centered in the origin (figure no. 5.12) describing an arc of length $2k_0 d$ from $e^{j(\delta+k_0 d)}$ to $e^{j(\delta-k_0 d)}$. This arc is denoted as the *visible domain* of the variable η.

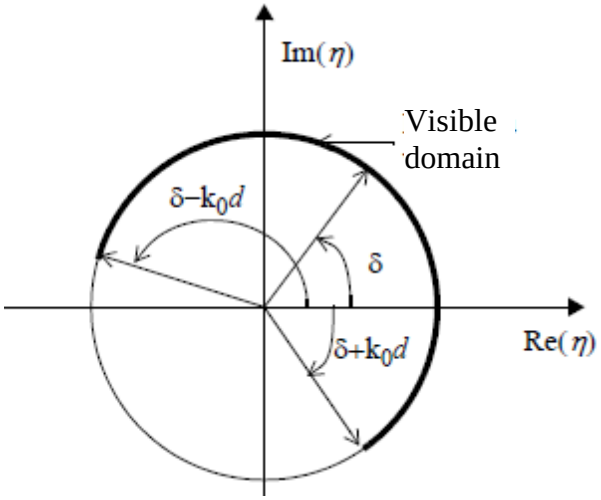


Figure no. 5.12 – *The complex plane of the variable* η

According to the fundamental theorem of algebra, the polynomial array factor (5.41) has $n-1$ complex roots and it can be written as:

$$f(\theta,\phi) = |a_{n-1}|(\eta-\eta_1)(\eta-\eta_2)\ldots\ldots\ldots(\eta-\eta_{n-1}) \qquad (5.42)$$

Hence, the modulus of the array factor in a given direction θ is given, except for a multiplicative factor, by the product of distances of the corresponding point in the visible domain to all the roots of the polynomial. Obviously, the nulls of the radiation pattern correspond to those roots that are in the visible domain of the variable η.

Example. Consider an antenna array with $n = 4$ elements, distance constant $d = \lambda_0/2 \Longrightarrow k_0 d = \pi$, phase constant $\delta = 0$, and current distributions' modules: $|a_0| = 1$, $|a_1| = 3$, $|a_2| = 3$, $|a_3| = 1$. According to (5.41), the array factor in η is:

$$f(\theta,\phi) = 1 + 3\eta + 3\eta^2 + \eta^3 = (1+\eta)^3 \qquad (5.43)$$

The above polynomial has a triple root equal to -1. The module of the array factor in a direction θ is proportional to the third power of the distance Δ between the corresponding point P in the visible domain and the point $A(-1,0)$, where the triple root is located (see figure no. 5.13a). The visible domain of η is the arc on the unit circle between $e^{j(\delta+k_0 d)} = e^{j\pi}$ and $e^{j(\delta-k_0 d)} = e^{-j\pi}$, that is the entire unit circle circumference.

Graphically, the array factor is proportional to Δ^3; its value is 0 for $\theta = 0$ and $\theta = \pi$ and it reaches the maximum value for $\theta = \frac{\pi}{2}$. Due to the absence of variable ϕ in the expression of η the array factor is independent of ϕ and the radiation pattern has revolution symmetry around the horizontal axis. The normalized array radiation pattern in figure no. 5.13b is built based on these observations.

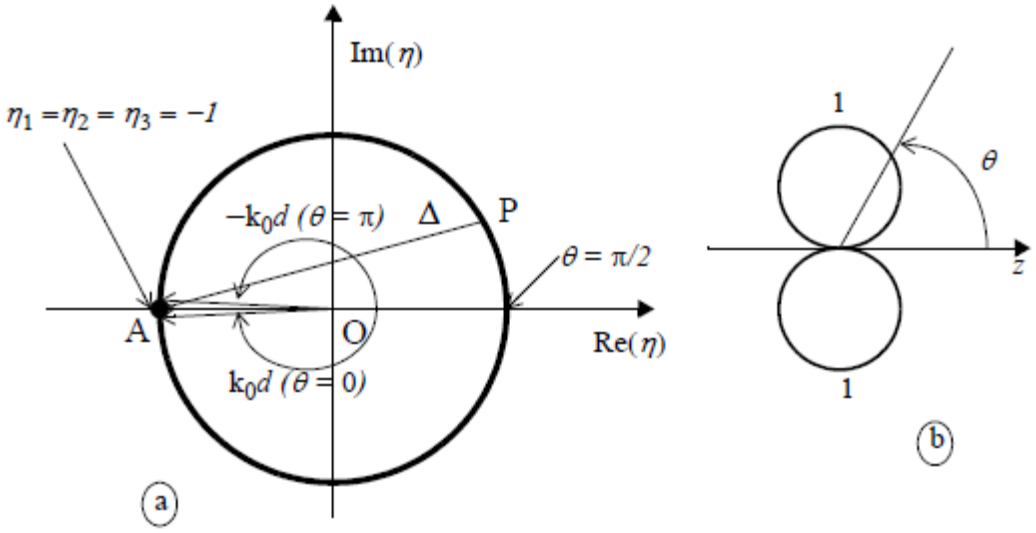


Figure no. 5.13 – *The visible domain of the variable η (a) and the radiation pattern (b) for a 4 element array with tapered current distributions*

Z Transform Method

The Z transform of real continuous function $f(x)$ is defined based on its samples at points spaced uniformly at distance d and it has the following expression:

$$\mathrm{F(z)} = \sum_{i=0}^{\infty} f(id) z^{-i} \tag{5.44}$$

The expression (5.41) of an antenna array with n equidistant elements can be rewritten as:

$$f(\theta, \phi) = \sum_{i=0}^{n-1} |a_i| e^{ji\gamma} = \sum_{i=0}^{n-1} |a_i| z^{-i} \tag{5.45}$$

if we make the notation $z \stackrel{\text{def}}{=} e^{-j\gamma}$.

This way, the array factor looks like the Z transform of the associated function to the modules of the current distributions, except for the upper limit of summation which is not ∞. We can extend this limit to ∞ by replacing the function $|a_i|$ with the product $|a_i| g(z)$, where $g(z)$ is the unit window function:

$$g(z) = \begin{cases} 0, & \text{for } z < 0 \\ 1, & \text{for } 0 \le z \le (n-1)d \\ 0, & \text{for } z > (n-1)d \end{cases} \tag{5.46}$$

Hence, the array factor is the Z transform of the function $|a_i| g(z)$:

$$f(\theta, \phi) = \sum_{i=0}^{\infty} |a_i| g(z) z^{-i}$$

The Z transform method is very useful in finding out the array factor for sophisticated distributions of coefficients $|a_i|$. For instance, for $a_i(z) = \sin(hz)$, where h is an arbitrary real constant, this method offers a simple way to arrive at the following expression of the array factor:

$$f(\theta,\phi)=\frac{e^{-j(n-2)\gamma}+e^{-j\gamma}}{1-2e^{-j\gamma}\cos(hd)+e^{-j2\gamma}}\sin(hd) \tag{5.47}$$

5.5 – Circular Arrays

A circular antenna array has its elements uniformly distributed on the circumference of a circle. If n is the number of the elements and a is the radius of the circle centered in the origin of a Cartesian coordinate system, then, considering that the reference antenna is situated in the center of the circle, we have (see figure no. 5.14):

$$v_m=\frac{2\pi}{n}m$$

$$\boldsymbol{r_m}=\hat{\boldsymbol{x}}a\cos v_m+\hat{\boldsymbol{y}}a\sin v_m \tag{5.48}$$

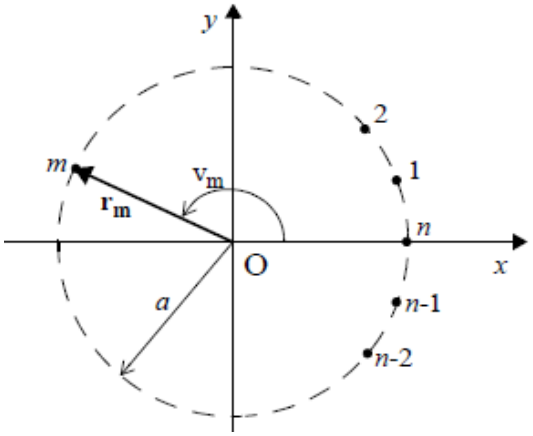


Figure no. 5.14 – *Geometry of a circular array*

According to (5.6) the array factor is:

$$f(\theta,\phi)\overset{\text{def}}{=}\sum_{m=1}^{n}a_m e^{jk_0\hat{\boldsymbol{r}}\cdot\boldsymbol{r_m}}=$$

$$=\sum_{m=1}^{n}a_m e^{jk_0a(\cos v_m\sin\theta\cos\phi+\sin v_m\sin\theta\sin\phi)}=$$

$$=\sum_{m=1}^{n}|a_m|e^{j[\delta_m+k_0a(\cos v_m\sin\theta\cos\phi+\sin v_m\sin\theta\sin\phi)]}=$$

$$=\sum_{m=1}^{n}|a_m|e^{j[\delta_m+k_0a\sin\theta\cos(\phi-v_m)]} \tag{5.49}$$

where δ_m is the phase of the coefficient a_m.

Note that the array factor is dependent on the variable ϕ and this means that the radiation pattern of a circular array does not have revolution symmetry.

Circular arrays could be implemented as directive or as omnidirectional arrays.

Directive circular array is obtained when radiated fields of component antennas combine constructively in the desired spatial direction (θ_0,ϕ_0). This requirement is fulfilled if:

$$\delta_m=-k_0a\sin\theta\cos(\phi-v_m) \tag{5.50}$$

Note that for a directive circular array to be obtained, a nonlinear distribution of current distribution phases should be used.

Theoretically, the only omnidirectional circular array is the one with infinite number of elements. But a good approximate of an omnidirectional radiation pattern could be obtained if the number of elements is greater than k_0a.

The *uniform circular arrays* are required to have identical modules for the current distributions and current phases varying with the same quantity δ form an element to the next one. The total variation of phase along the array should be a multiple p of 2π, because a phase change with the same quantity δ should exist from the last element to the first one (they are neighbors !). Thus:

$$a_m = \frac{1}{n} e^{-j\frac{p2\pi}{n}n} = \frac{1}{n} e^{-jpv_m}, \quad p - \text{integer} \tag{5.50}$$

and:

$$f(\theta, \phi) = \sum_{m=1}^{n} \frac{1}{n} e^{-j[pv_m - k_0 a \sin\theta \cos(\phi - v_m)]} \tag{5.51}$$

5.6 – Arrays of Arrays

The analysis made until now regarding the properties of antenna arrays did not refer to the *type* of the antennas used as array elements. Thus, it is possible for these antennas to be array of antennas themselves. This way we obtain arrays of antenna arrays.

The antenna elements are denoted as *subarrays* in this context, while the resulting array is denoted as *superarray*.

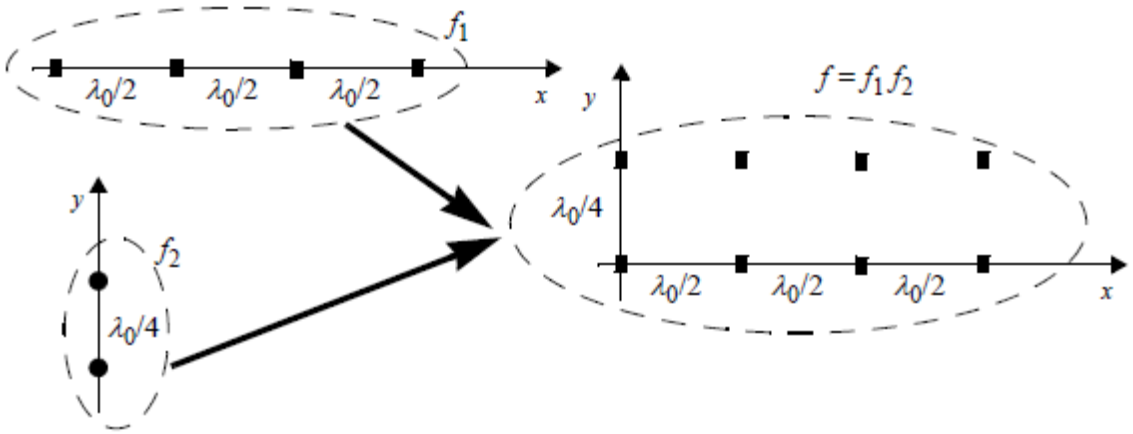


Figure no. 5.15 – *Building an array of arrays*

Let us consider an array of m antennas, with element no. p as the reference antenna; denote by f_1 its array factor. Let us build a (super)array by using a number of n above defined (sub)arrays and choose the subarray no. q as its reference antenna; denote by f_2 its array factor (see figure no. 5.15). Looking at this superarray as an array of $m \times n$ antennas, with the antenna p in the subarray q as its reference antenna, then its array factor is $f = f_1 \cdot f_2$. *Thus, the radiation pattern of the array with $m \times n$ antennas is obtained by multiplying the two radiation patterns.* This property is denoted as the *characteristic multiplication*.

Example. Let us consider as subarray an uniform linear array with $m = 2$ elements, distance constant $= \lambda_0/2$, and phase constant $\delta = 0$. We build a superarray as an uniform linear array by using $n = 2$ subarrays with distance constant $d = \lambda_0/4$ and phase constant $\delta = \pi/2$. It is illustrated in the figure no. 5.16 the resulting radiation pattern of the array of $4 \times 2 = 8$ antennas as graphical product of the two radiation patterns.

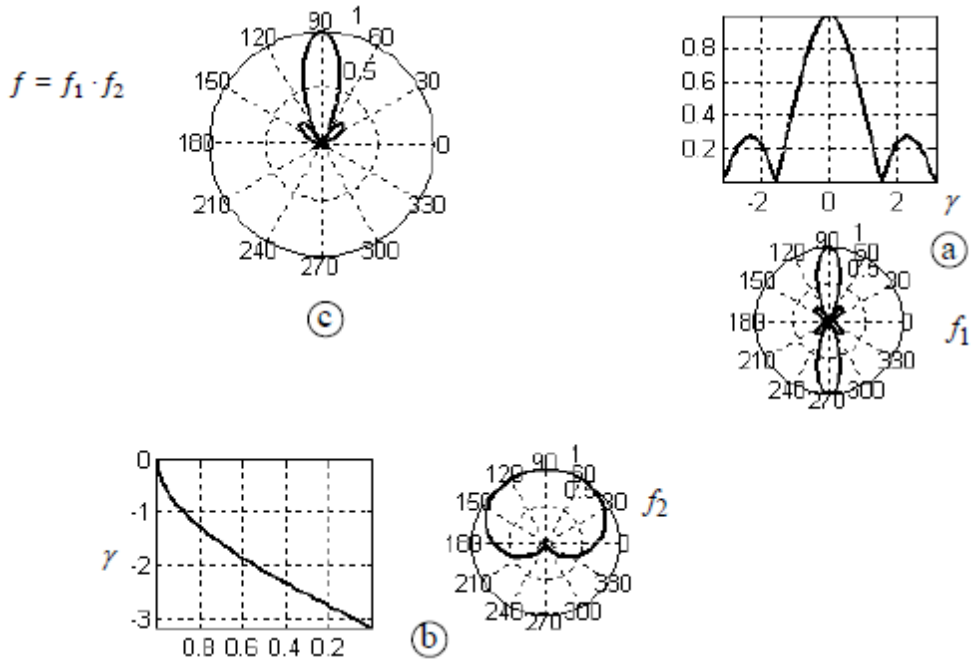


Figure no. 5.16 – *Example of characteristic multiplication*

Binomial array.

Let us consider a linear array of 2 elements with distance constant d, phase constant δ, and equal modules of the current distributions: $|a_1| = |a_2| = 1$. According to (5.10) its array factor is:

$$f(\theta, \phi) = 1 + e^{j\gamma} \tag{5.52}$$

We build an array with 3 elements by the superposition of the above array with its replica obtained by moving it to the right with a distance d and increasing the current phases with δ (see figure no. 5.17a). The array factor of this new array is:

$$f(\theta, \phi) = 1 + 2e^{j\gamma} + e^{j2\gamma} = \left(1 + e^{j\gamma}\right)^2 \tag{5.53}$$

According to (5.15) the array factor is zero for $\gamma = p\pi,\ \forall p \in \mathbb{Z}\backslash\{0\}$.

By repeating n times this process we obtain an array with $n + 1$ elements having the following array factor:

$$f(\theta, \phi) = \left(1 + e^{j\gamma}\right)^n \tag{5.54}$$

This array factor has the same zeros as the initial 2 element array. When the visible domain of γ does not include multiples of π, the array factor has not zeros and, consequently, the radiation pattern has not side lobes (see figure no. 5.17b). This property is very useful in some applications (radar, for instance).

The array is denoted after its array factor particular expression.

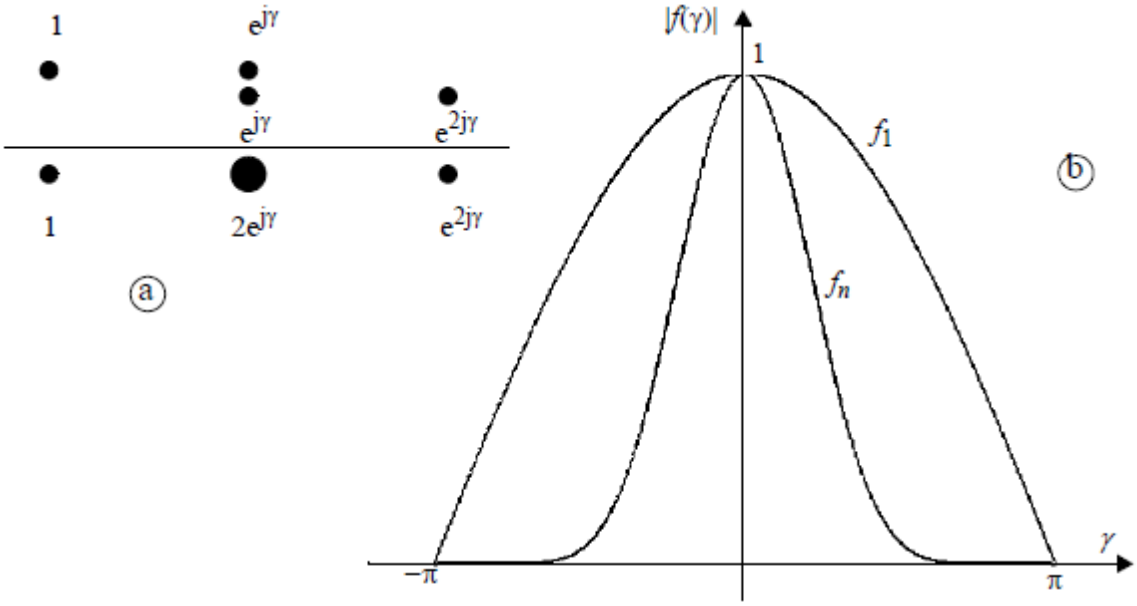


Figure no. 5.17 – *Building a binomial array (a) and its radiation pattern (b)*

Array with triangular current distribution.

The property of characteristic multiplication could be used to derive the expression of the array factor for complex current distributions.

Let us consider an uniform linear array with n elements, distance constant d, phase constant δ, and modules of all current distributions equal to 1 (figure no. 5.18). Taking the rightmost element as the reference antenna, the array factor is:

$$f(\theta,\phi) = e^{-j\frac{n\gamma}{2}}\frac{\sin\frac{n\gamma}{2}}{\sin\frac{\gamma}{2}} \qquad (5.55)$$

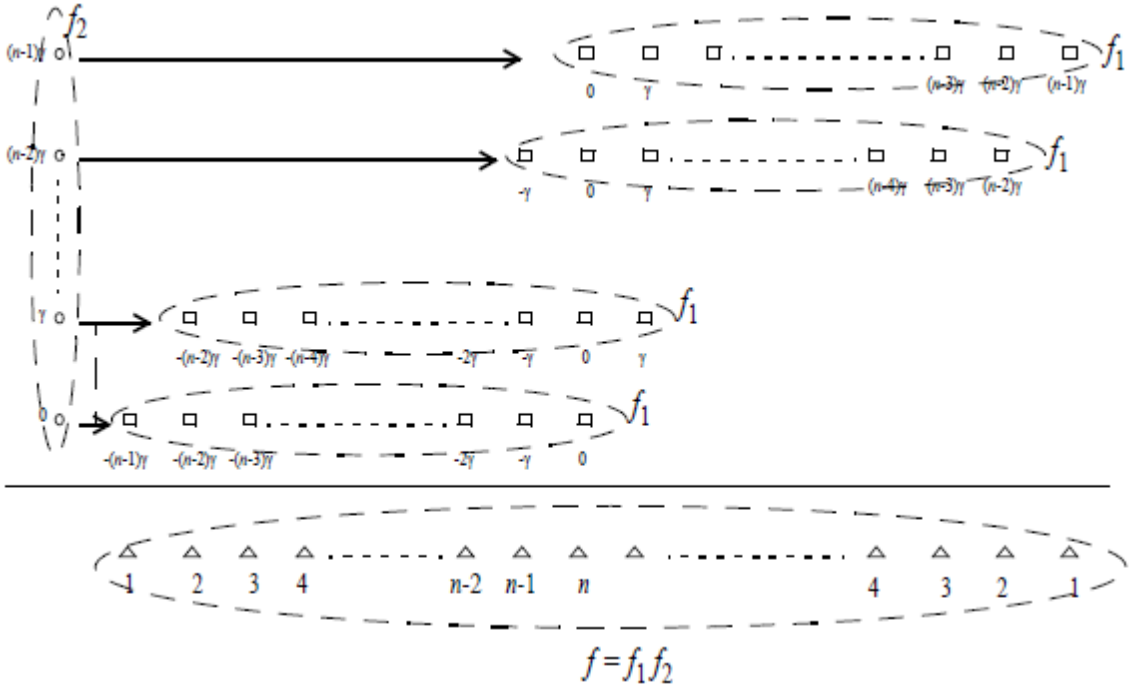


Figure no. 5.18 – *Building an array with triangular current distribution*

We build a new array by translating the above array by a distance constant d to the right and increasing the current phases by a constant phase δ. After $n-1$ translations we obtain $n-1$ arrays, each of them being the translation of the previous translated one with a position to the right. These arrays and the initial one form a superarray of n subarrays, the leftmost one being the reference subarray. The array factor of the superarray is:

$$f(\theta,\phi) = e^{j\frac{n\gamma}{2}} \frac{\sin\frac{n\gamma}{2}}{\sin\frac{\gamma}{2}} \tag{5.56}$$

By overlapping the elements on each of the positions, we obtain an array with $2n+1$ elements with the same distance constant and the same phase constant as the initial array, but with central element having the greatest current distribution module and linearly decreasing modules towards the edges for the rest of the elements. The current distribution is triangular.

By using the characteristic multiplication property, the array factor of the $2n+1$ element array is the product of the last two expressions, that is:

$$f(\theta,\phi) = \left(\frac{\sin\frac{n\gamma}{2}}{\sin\frac{\gamma}{2}}\right)^2 \tag{5.57}$$

5.7 – Arrays' Optimization

The above analysis revealed that, for a given array, trying to decrease the side lobe levels of its radiation pattern makes the main lobe beamwidth to increase or, conversely, decreasing the main lobe beamwidth makes the side lobe levels to increase. This correlation between the two parameters maintains for individual antennas, too. Practical implementations should make a compromise.

An *optimal array* has current distributions' coefficients a_m that allows for a minimum value of the side lobe levels for a given main lobe beamwidth or the minimum main lobe beamwidth for a given level of the side lobe levels. Dolph proved that the optimum is reached when the coefficients a_m are coefficients of Chebyshev polynomials. This is why optimal arrays are denoted as *Dolph-Chebyshev arrays*.

Chebyshev Polynomials

There are different definitions for the Chebyshev polynomials, but the most convenient for the arrays' optimization problem is the following one:

$$T_n(x) = \begin{cases} (-1)^n \cosh(n \operatorname{arccosh}|x|), & x < -1 \\ \cos(n \arccos x), & |x| \le 1 \\ \cosh(n \operatorname{arccosh} x), & x > 1 \end{cases} \tag{5.58}$$

where n is the degree of the polynomial.

Although the above definition appears as a complicated formula, the expressions of the Chebyshev polynomials are quite simple. For instance:

$$\begin{aligned} T_0(x) &= 1 \\ T_1(x) &= x \\ T_2(x) &= 2x^2 - 1 \\ T_3(x) &= 4x^3 - 3x \\ T_4(x) &= 8x^4 - 8x^2 + 1 \\ T_5(x) &= 16x^5 - 20x^3 + 5x \end{aligned} \tag{5.59}$$

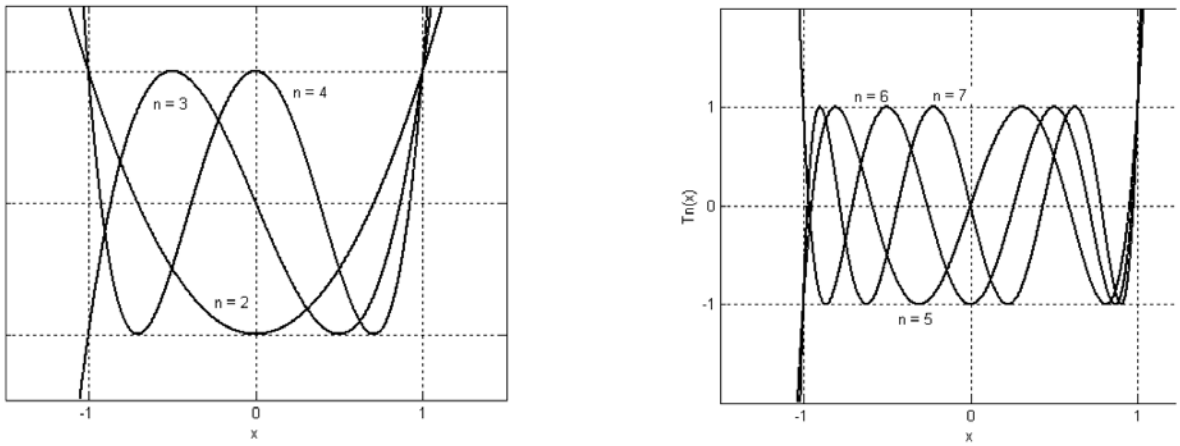


Figure no. 5.19 – *Chebyshev polynomials*

There exist two recurrent relations that allow for a fast computation of Chebyshev polynomials of higher degree:

$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$$
$$T_{mn}(x) = T_m\big(T_n(x)\big) = T_n\big(T_m(x)\big) \tag{5.60}$$

The definition, the expressions (5.59) and the graphical representations in figure no. 5.19 allow us to extract the following properties of the Chebyshev polynomials:

- the polynomials of even degree contain only even powers of the argument, while the polynomials of odd degree contain only odd powers of the argument;
- the polynomial of degree n includes points $(-1,(-1)^n)$ and $(1,1)$;
- for $-1 \le x \le 1,\ \ -1 \le T_n(x) \le 1,\ \forall n$;
- all the roots of a polynomial are real and lie in the interval $(-1,1)$; for a polynomial of degree n they are given by the following relation:

$$x = \cos\left[(2p+1)\frac{\pi}{2n}\right], \quad p = 0, 1, \dots\dots\dots, n-1 \tag{5.61}$$

- all the local extremes of a polynomial have modules equal to 1 and appear in the interval $(-1,1)$; for a polynomial of degree n their positions on the abscissa are given by the following relation:

$$x = \cos\left(p\frac{\pi}{n}\right), \quad p = 1,\ 2, \dots\dots\dots, n-1 \tag{5.62}$$

A useful *theorem* for the arrays' optimization problem is the following:

If $P_n(x)$ is an arbitrary polynomial of degree n with $P_n(x_0) = R > 1,\ \ x_0 > 1$, and $P_n(x_2) = 0$, for $x_1 \le x_2 < x_0$, where x_1 is the greatest root of the Chebyshev polynomial of the same degree, than $|P_n(x)| > 1$ for at least part of the interval $-1 < x < 1$.

In other words, a Chebyshev polynomial has the smallest module on the whole interval $-1 < x < 1$, among all the polynomials of the same degree that pass through the point (x_0, R) and have a root in the interval $[x_1, x_0)$.

A *reciprocal* of the above theorem is, also, useful for arrays' optimization:

The Chebyshev polynomial has the greatest root x_1 closest to x_0, among all the polynomials of the same degree that pass through the point (x_0, R) and maintain their module smaller than 1 on the whole interval $-1 < x < 1$.

In other words, a Chebyshev polynomial minimizes the distance $|x_0 - x_2|$ between its greatest root x_2 and the point with abscissa x_0.

Array Factor Equivalence with a Chebyshev Polynomial

Let us consider an antenna linear array with *odd* number n of elements positioned at equal distance d between each other and with current phases changing with same quantity δ from an element to the next one. We define the element in the middle of the array as the reference antenna and renumber the elements as illustrated in figure no. 5.20. By applying (5.13) we obtain that:

$$f(\theta,\phi) = |a_0|e^{j0} + \sum_{m=1}^{\frac{n-1}{2}} |a_m| e^{jm\gamma} + \sum_{m=1}^{\frac{n-1}{2}} \left| a_{m+\frac{n-1}{2}} \right| e^{-jm\gamma} \qquad (5.63)$$

If we require that:

$$|a_m| = \left| a_{m+\frac{n-1}{2}} \right|, \quad \forall m \qquad (5.64)$$

then:

$$f(\theta,\phi) = |a_0|e^{j0} + 2\sum_{m=1}^{\frac{n-1}{2}} |a_m| \cos(m\gamma) \qquad (5.65)$$

Based on common trigonometric formulas $\cos(m\gamma), m - \text{integer}$, could be written as a polynomial of degree m with argument $\cos\gamma$. Then, knowing that $\cos\gamma = 2\cos^2\left(\frac{\gamma}{2}\right) - 1$, $\cos(m\gamma)$ could be written as a polynomial of degree $2m$ with argument $\cos\left(\frac{\gamma}{2}\right)$ that contains only even powers of the argument. Thus:

$$f(\theta,\phi) = b_0 + \sum_{m=1}^{\frac{n-1}{2}} b_m \cos^{2m}\left(\frac{\gamma}{2}\right) \qquad (5.66)$$

that is *the array factor of an antenna array with n ($n - odd$) elements is a polynomial of degree $n-1$ with argument $cos\left(\frac{\gamma}{2}\right)$ that contains only even powers of the argument.*

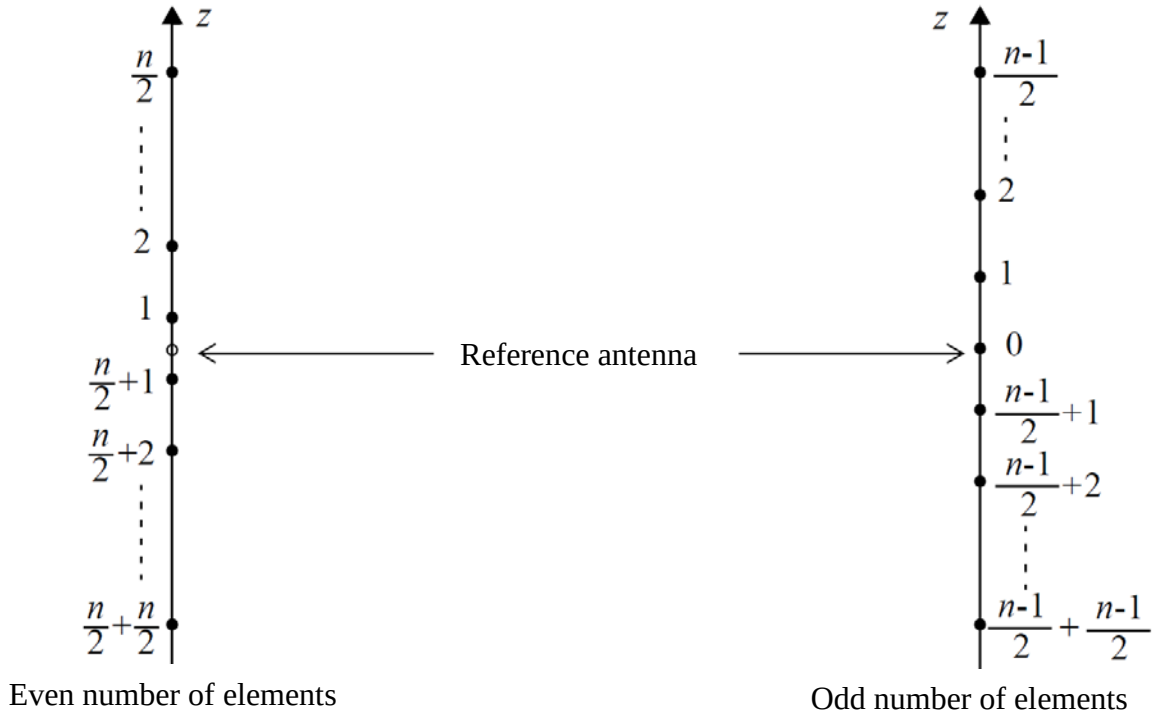


Figure no. 5.20 – *Renumbering the elements of array*

Let us consider now an antenna linear array with *even* number n of elements positioned at equal distance d between each other and with current phases changing with same quantity δ from an element to the next one. We define a fictitious element in the middle of the array as the reference antenna and renumber the elements as illustrated in figure no. 5.20. By applying (5.13) we obtain that:

$$f(\theta,\phi) = \sum_{m=1}^{\frac{n}{2}} |a_m| e^{j(2m-1)\frac{\gamma}{2}} + \sum_{m=1}^{\frac{n}{2}} \left|a_{m+\frac{n}{2}}\right| e^{-j(2m-1)\frac{\gamma}{2}} \qquad (5.67)$$

If we require that:

$$|a_m| = \left|a_{m+\frac{n}{2}}\right|, \quad \forall m \qquad (5.68)$$

then:

$$f(\theta,\phi) = 2\sum_{m=1}^{\frac{n}{2}} |a_m| \cos\left[(2m-1)\frac{\gamma}{2}\right] \qquad (5.69)$$

Based on common trigonometric formulas $\cos\left[(2m-1)\frac{\gamma}{2}\right], m - \text{integer}$, could be written as a polynomial of degree $2m-1$ with argument $\cos\left(\frac{\gamma}{2}\right)$ that contains only odd powers of the argument. Thus:

$$f(\theta,\phi) = \sum_{m=1}^{\frac{n}{2}} b'_m \cos^{2m-1}\left(\frac{\gamma}{2}\right) \qquad (5.66)$$

that is *the array factor of an antenna array with n ($n-even$) elements is a polynomial of degree $n-1$ with argument $\cos\left(\frac{\gamma}{2}\right)$ that contains only odd powers of the argument*.

The final conclusion is that ***the array factor for an antenna array with equidistant elements and current phases varying with same quantity from an element to the next one could be written as a Chebyshev polynomial when it uses equal modules for the current distributions of the elements symmetrically positioned in the array*** (see conditions 5.64 and 5.68). The degree of the equivalent Chebyshev polynomial is one unit smaller than the number of elements in the array.

In order to apply this conclusion we replace the argument x of a Chebyshev polynomial by:

$$x = b\cos\frac{\gamma}{2} \qquad (5.67)$$

where b is an arbitrary real constant and thus:

$$f(\theta,\phi) = T_{n-1}\left(b\cos\frac{\gamma}{2}\right) \qquad (5.68)$$

Polar Diagram for Dolph-Chebyshev Arrays

We illustrate this construction by an example. Let us consider a $n=5$ element array, with distance constant $d = \lambda_0/2 \implies k_0 d = \pi$. For phase constant $\delta = 0$, when θ varies from 0 to $\pi/2$, $\gamma = \delta + k_0 d\cos\theta$ varies from π to 0, while $x = b\cos(\gamma/2)$ varies from 0 to b (from point A to point B in the figure no. 5.21). The representative point on the 4 degree Chebyshev polynomial graph varies from point A' to point B'. When θ varies from $\pi/2$ to π, γ varies from 0 to $-\pi$, while x varies from b to 0 (from point B to point C in the figure no. 5.21). The representative point on the 4 degree Chebyshev polynomial graph varies from point B' to point $C' \equiv A'$.

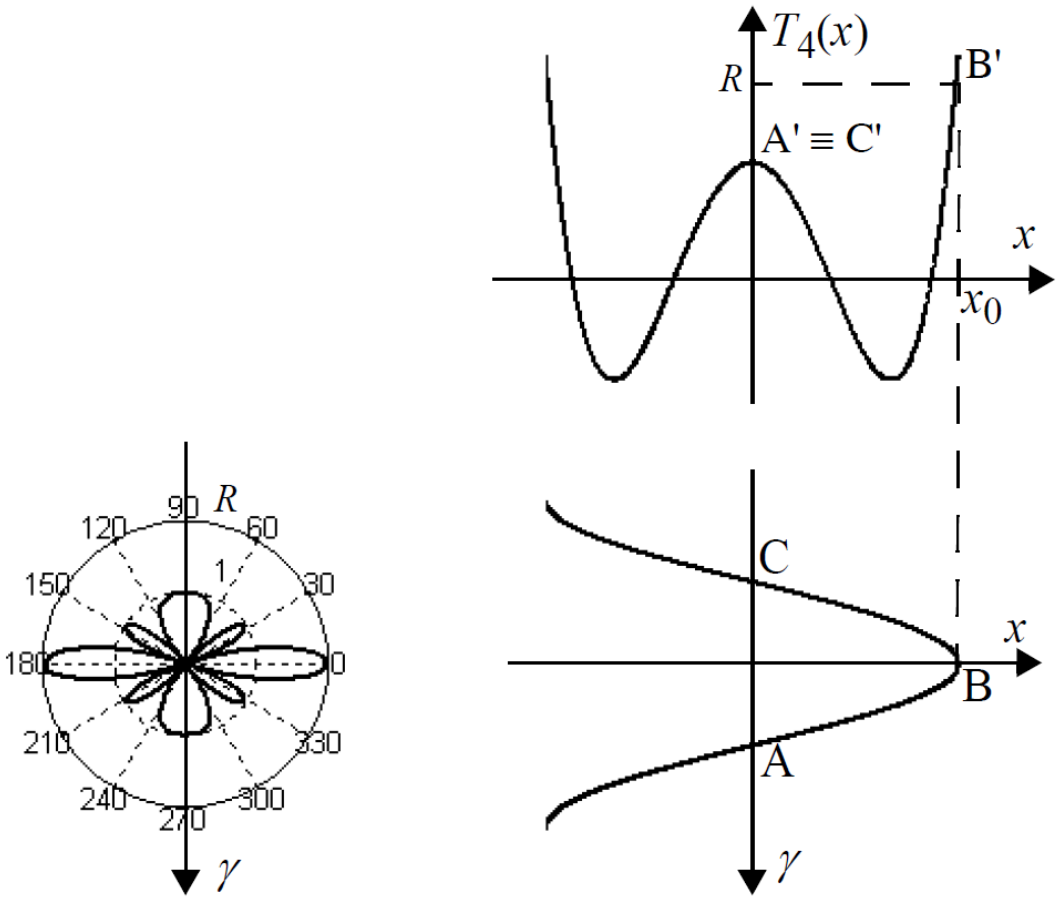


Figure no. 5.21 – *Polar diagram for an optimal broadside array*

In accordance with this excursion of the representative point on the 4 degree Chebyshev polynomial graph, the array factor reaches a global maximum $f(\theta = \pi/2, \phi) = T_4(x_0 = b) = R$ and 3 local maxima, all of them equal to 1. The radiation pattern has a main lobe with level R perpendicular to the axis on which the elements of the array are positioned (broadside array, as expected) and three side lobes with level 1; the relative level of the side lobes is $1/R$. The radiation pattern has revolution symmetry around the axis γ.

Following similar rules the array radiation pattern for $\delta = -\pi$ is built in figure no. 5.22. There are also three side lobes with relative level $1/R$. The main lobes are directed along the axis γ (end-fire array).

Design of Dolph-Chebyshev Arrays

Let us denote by 2α the main lobe beamwidth between its adjacent nulls.

Broadside arrays

For a broadside array (main lobe direction $\theta_0 = \pi/2$) we should choose $\delta = 0$. The main lobe adjacent nulls correspond to the greatest root of the Chebyshev polynomial (see figure no. 5.21) and this one is obtained from 5.61 for $p = 0$:

$$x_1 = \cos\frac{\pi}{2(n-1)} \tag{5.69}$$

(*Note that we used* $n-1$ *instead of* n *at the denominator, because the array factor for an* n *element optimal array is a* $n-1$ *degree Chebyshev polynomial*)

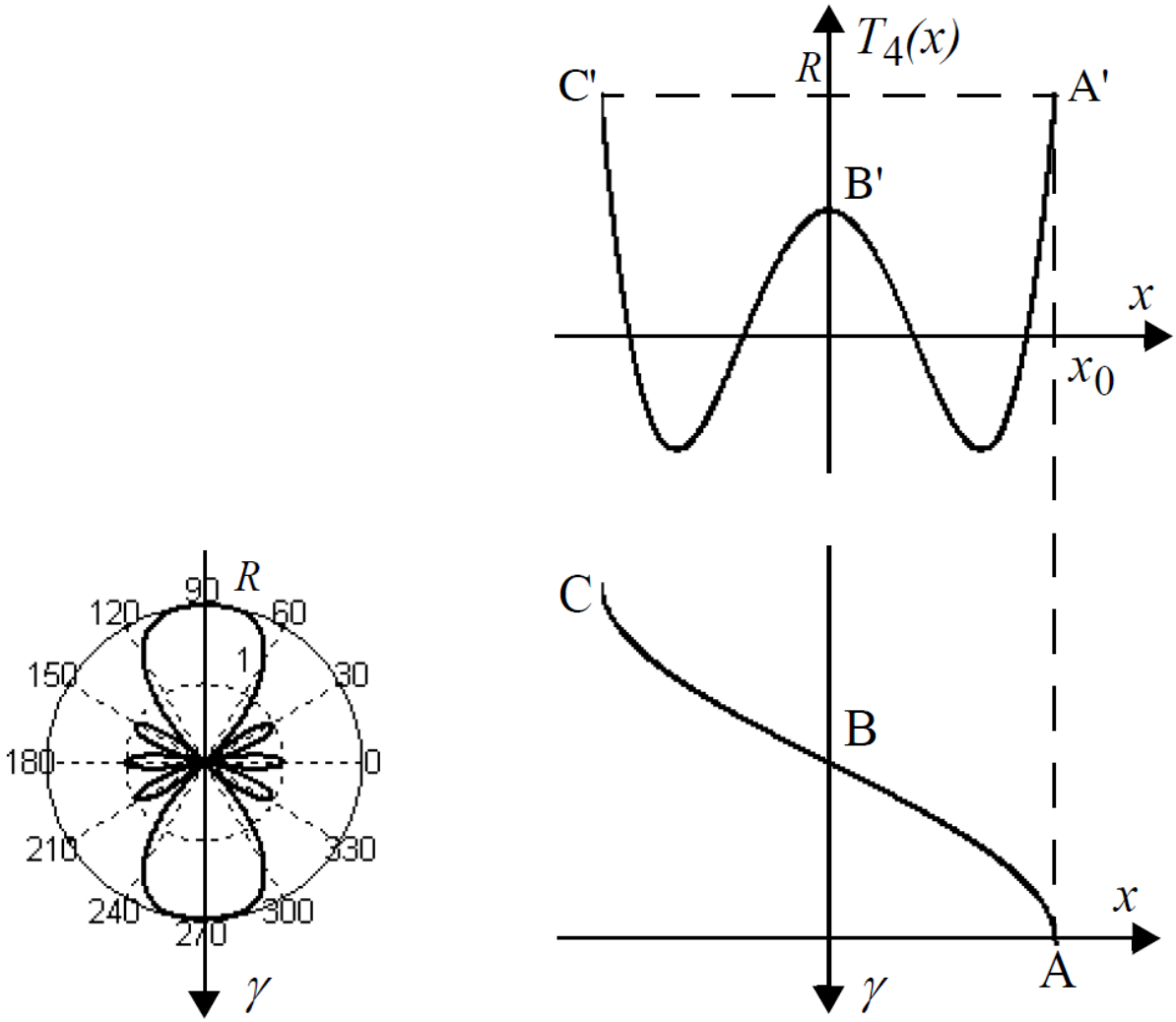


Figure no. 5.22 – *Polar diagram for an optimal end-fire array*

Requiring that this root corresponds to the direction $\theta = \pi/2 - \alpha$ of the main lobe adjacent null we obtain the equation (see 5.67):

$$\cos\frac{\pi}{2(n-1)} = b\cos\left(\frac{k_0 d\sin\alpha}{2}\right) \tag{5.70}$$

The main lobe level is the value of the Chebyshev polynomial for $\theta_0 = \pi/2$, for which the argument of the Chebyshev polynomial is $x_0 = b\cos(\pi/2) = b$. Thus:

$$R = T_{n-1}(x_0) = \cosh[(n-1)\operatorname{arccosh} b] \tag{5.71}$$

Case 1. When a value of 2α is required for the main lobe beamwidth, we compute parameter b with equation (5.70), then we find out the main lobe level R with equation (5.71). The relative level $1/R$ of the side lobes is the smallest possible to obtain because the optimum theorem previously presented states that the Chebyshev polynomial is the only polynomial that maintains its module less than 1 on the whole interval $[-1,1]$.

Case 2. When a maximum relative value of $1/R$ is required for the side lobes' level, we compute parameter b with equation (5.71), then we find out the value of α with equation (5.70). The main lobe beamwidth 2α is the smallest possible to obtain because the reciprocal

of the optimum theorem states that, among all the polynomials of the same degree, the Chebyshev polynomial has its greatest root x_1 the closest to x_0.

End-fire arrays

The main lobe direction for an end-fire array is $\theta_0 = 0$. For this value of θ_0 we obtain from (5.67) that:

$$x_0 = b\cos\left(\frac{\delta + k_0 d}{2}\right) \tag{5.72}$$

The main lobe level is:

$$R = T_{n-1}(x_0) = \cosh[(n-1)\operatorname{arccosh}(x_0)] \tag{5.73}$$

The adjacent null of the main lobe corresponds to $\theta = \alpha$ (see figure no. 5.22). Imposing that the greatest root x_1 of the Chebyshev polynomial corresponds to the direction of the main lobe adjacent null we obtain the equation (see 5.67):

$$\cos\frac{\pi}{2(n-1)} = b\cos\left(\frac{\delta + k_0 d\cos\alpha}{2}\right) \tag{5.74}$$

In order to make use of the optimal properties of the Chebyshev polynomial, we should require that its argument x spans the whole interval $[-1,1]$, this meaning that $x = -1$ for $\theta = \pi$ or:

$$b\cos\left(\frac{\delta - k_0 d}{2}\right) = -1 \tag{5.75}$$

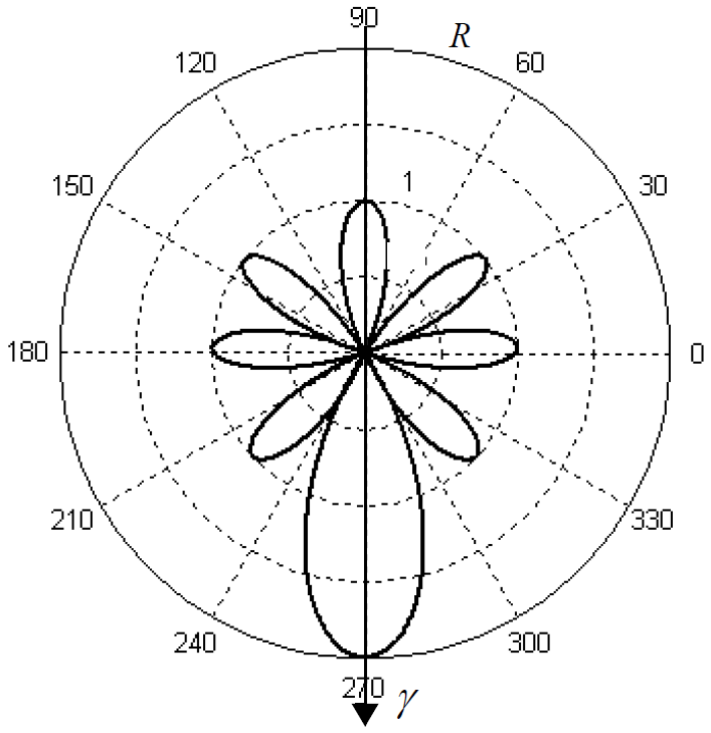


Figure no. 5.23 – *Polar diagram for an optimal end-fire array*

Case 1. When a value of 2α is required for the main lobe beamwidth, we compute parameters b and δ from equations (5.74) and (5.75), then we compute x_0 from equation (5.72) and, finally, we find out R from equation (5.73). The relative level $1/R$ of the side lobes is the smallest possible to obtain because the optimum theorem previously presented

states that the Chebyshev polynomial is the only polynomial that maintains its module less than 1 on the whole interval $[-1,1]$.

Case 2. When a maximum relative value of $1/R$ is required for the side lobes' level, we compute parameter x_0 with equation (5.73), then we compute b and δ with equations (5.72) and (5.75) and, finally, we find out the value of α with equation (5.74). The main lobe beamwidth 2α is the smallest possible to obtain because the reciprocal of the optimum theorem states that, among all the polynomials of the same degree, the Chebyshev polynomial has its greatest root x_1 the closest to x_0.

Figure no. 5.23 illustrates the radiation pattern of an end-fire array as resulted from the above design: the side lobes have the same relative level of $1/R$.

5.8 – Parasitic Antenna Arrays

Parasitic antenna array is an array containing elements, denoted as *parasitic antennas*, that are not fed by means of a transmission line, but by the electromagnetic field radiated by the other elements of the array that are fed, as usually, by means of a transmission line. Parasitic antenna arrays comprising dipole antennas are denoted as Yagi-Uda arrays, Yagi-Uda antennas or Yagi antennas. The name is due to the fact that they appeared as a result of extensive experimental research conducted by a group coordinated by professor Shintaro Uda at Tohoku University, Sendai, Japan, while the research results were firstly published in English by Hidetsugu Yagi, one of the group members. Yagi antennas are frequently used because they are constructively simple and have a relative big gain.

Simulation studies show that the radiation pattern of a set of dipoles can be modified by changing the dipoles' dimensions.

For instance, for a set of two identical half wave dipoles with an extremely small separation distance and fed with currents with the same modules, but with opposite phases, the radiation pattern is a torus with the dipoles as its symmetry axis. The radiation pattern in a plane containing the dipoles has two main lobes in opposite directions (figure no. 5.24).

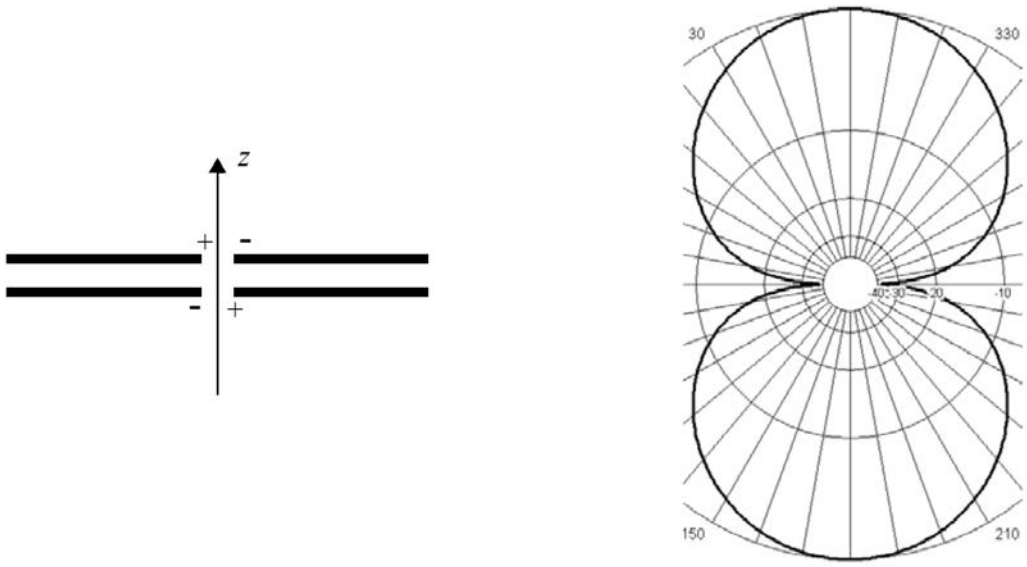


Figure no. 5.24 – *Radiation pattern of 2 dipoles fed with opposite phase currents*

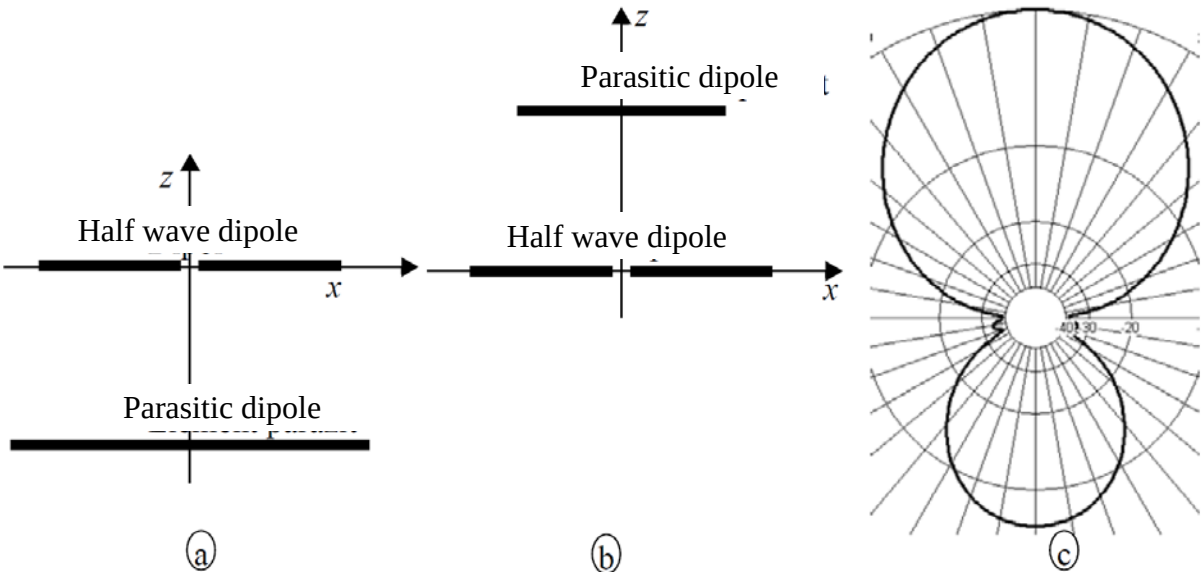


Figure no. 5.25 – *Modifying a dipole radiation pattern by means of a parasitic dipole*

By increasing the length of a dipole (figure no. 5.25a) or by decreasing the length of the other dipole (figure no. 5.25b) and maintaining the feeding only for the unmodified dipole results in change of the radiation pattern: the lobe towards the shorter dipole in the set has a greater level than the other one (figure no. 5.25c). When using more parasitic dipoles the difference between the levels of the two lobes becomes greater.

Because the fed dipole remains with the length equal to the half wavelength it is denoted as a *vibrator* as it resonates on the radiation frequency. The longer dipoles are denoted as *reflectors* because the lobe in their direction has a smaller level and it looks like they reflect the wave radiated by the vibrator. The shorter dipoles are denoted as *directors* because the lobe in their direction has a greater level and it looks like they direct towards them the wave radiated by the resonator.

Basic Yagi antenna comprises three elements: a transmission line fed vibrator dipole, a parasitic reflector dipole and a parasitic director dipole. Its radiation pattern in the dipoles' plane is presented in figure no. 5.26.

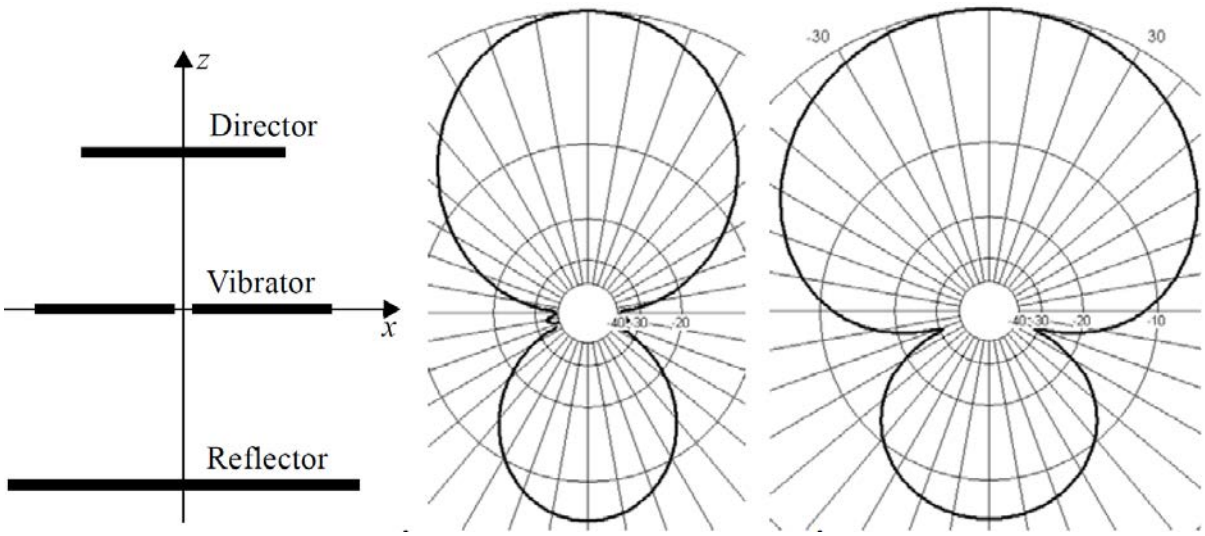


Figure no. 5.26 – *Radiation pattern for a 3 element Yagi antenna*

The effect of using more than one reflector upon the radiation pattern is negligible small, so practical Yage antennas use only one reflector. The effect of using more than one director upon the radiation pattern is very strong, so common practical Yagi antennas use up to 5 directors, but there were reported implementations with 40 directors.

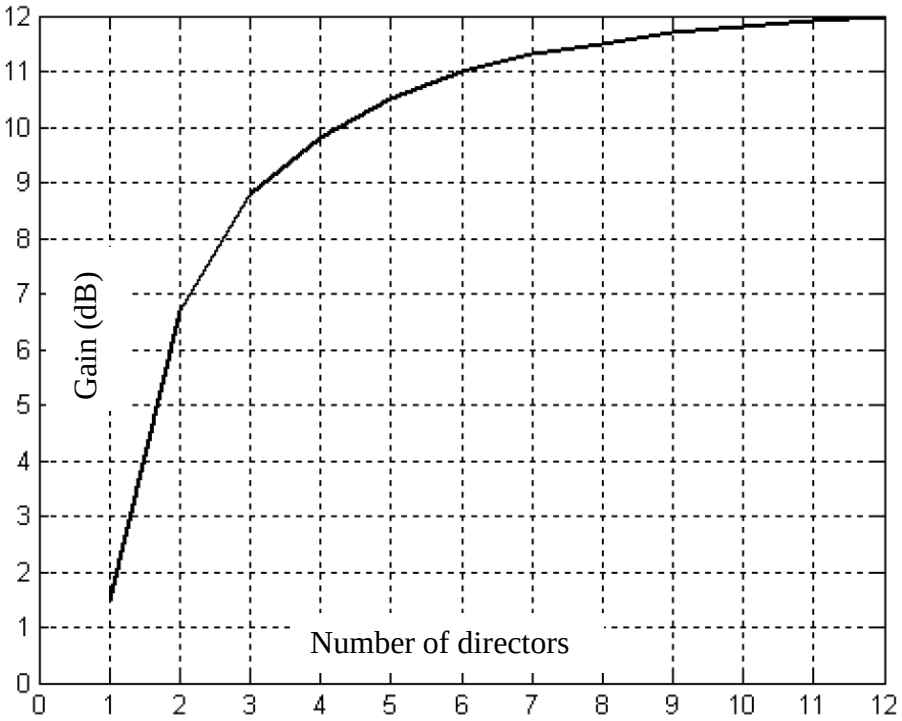


Figure no. 5.27 – *Yagi antenna gain versus number of directors*

Figure no. 5.27 illustrates the Yagi antenna versus the number of directors and it reveals that the gain increase is slower at high numbers. For instance, increasing the number of the directors from 2 to 3 makes the antenna gain to increase with more than 2 dB, while increasing the number of the directors from 9 to 10 makes the antenna gain to increase with about 0.2 dB. This is one reason for limiting the number of directors in practical implementations. The other reason is that the antenna dimensions and weight reach high values and it is difficult to physically support it.

The side lobe opposite to the main lobe is the highest one and this is why the front-to-back ratio (F/B) is one of the most important Yagi antenna parameters.

The design of a Yagi antenna is a very difficult task due to the little knowledge about the current distribution along the individual dipoles. Usually, numerical integration techniques or experimental measurements allow for finding out dipoles' optimal length and separation distance starting with an initial solution. Generally, dipoles' length decreases when going farther from the vibrator, while the interspacing follows a quasi-logarithmic law.

There are published tables with precise values for dipoles' length and interspacing for a given number of elements and a given operating frequency. Fortunately, small errors in implementing the tabulated dimensions do not influence strongly the antenna parameters and, thus, practical implementation is not a critical undertaking. It is very important that the dipoles' diameter should remain much smaller than the operating wavelength and that the material they are made of should have good conducting properties.

Table no. 5.1 – *Parameters of Yagi antennas with equidistant elements and identical length directors*

No. of elements	d/λ	l_R/λ	l_V/λ	l_D/λ	*G (dB)*	*F/B (dB)*	Z_{in} *(ohms)*	θ_E *(deg)*	θ_H *(deg)*
3	0.25	0.479	0.453	0.451	9.4	5.6	22.3 + 15j	84	66
4	0.15	0.486	0.459	0.453	9.7	8.2	36.7 + 9.6j	84	66
	0.2	0.503	0.474	0.463	9.3	7.5	5.6 + 20.7j	64	54
	0.25	0.486	0.463	0.456	10.4	6.0	10.3 + 23.5j	60	52
	0.3	0.475	0.453	0.446	10.7	5.2	25.8 + 23.2j	64	56
5	0.15	0.505	0.476	0.456	10	13.1	9.6 + 13j	76	62
	0.2	0.486	0.462	0.449	11	9.4	18.4 + 17.6j	68	58
	0.25	0.477	0.451	0.442	11	7.4	53.3 + 6.2j	66	58
	0.3	0.482	0.459	0.451	9.3	2.9	19.3 + 39.4j	42	40
6	0.2	0.482	0.456	0.437	11.2	9.2	51.3 + 1.9j	68	58
	0.25	0.484	0.459	0.446	11.9	9.4	23.2 + 21j	56	50
	0.3	0.472	0.449	0.437	11.6	6.7	61.2 + 7.7j	56	52
7	0.2	0.489	0.463	0.444	11.8	12.6	20.6 + 16.8j	58	52
	0.25	0.477	0.454	0.434	12	8.7	57.2 + 1.9j	58	52
	0.3	0.475	0.455	0.439	12.7	8.7	35.9 + 21.7j	50	46

Table no. 5.1 presents the dimensions and the electric characteristics of special Yagi antennas with equidistant elements and identical length directors.

Arrays of 2 or 4 Yagi antennas could be used for significant gain increase.

C h a p t e r V I

FREQUENCY INDEPENDENT ANTENNAS

6.1 – Operation Principles

A frequency independent antenna has not an infinite frequency bandwidth as its name suggests, but a very large value, much greater than the bandwidth of usual antennas. They are built according with one of the following two principles.

The first principle exploits the dependence of antenna parameters not on their physical dimensions, but on their *electric dimensions*, that is on the ratio of their physical dimensions to the wavelength at the working frequency. If we have the possibility to modify antenna dimensions when the working frequency changes, such that its electric dimensions remains constant, then the antenna parameters do not change and it could be efficiently used at the new working frequency. The physical structures exhibiting such a property have shapes that are completely defined by angles, not by lengths. These antennas satisfy an *angle condition*. *Conical antennas* and *equiangular antennas* are typical representatives of this category.

The second principles states that if a physical structure maps into itself when scaling by an arbitrary coefficient τ, then an antenna based on this structure has identical properties at frequencies f and τf. Actually, antenna properties vary periodically with the period $\log \tau$ and, for $\tau \approx 1$, the period is extremely small and the antenna properties remain constant for a large frequency bandwidth. It is noticed experimentally that the antenna properties remain constant even for $\tau > 1$. The antennas in this category are denoted as *log-periodic antennas*.

Based on both principles we get infinite size physical structure starting from a point (usually the origin of a coordinate system). Actual antennas are parts of these physical structures obtained by cutting them at distances r_1 and r_2 ($r_1 < r_2$) from the origin.

The distance r_1 limits the upper limit of the frequency bandwidth because the antenna smaller dimension should remain much greater than the feeding line and its coupling circuit to the antenna (the antenna is fed at its smaller end) for all operating frequencies. Practically, the dimension of the coupling circuit between the antenna and its feeding line is the limiting factor of the bandwidth upper limit.

The distance r_2 limits the lower limit of the frequency bandwidth due to the followings. The current along the antenna decreases from the feeding point towards the opposite end. The infinite size physical structure should be cut at a sufficiently great distance from the origin such that the current value to be negligibly small at the cutting point and the current distribution along the finite size antenna resembles very well the one in the infinite structure. The current decreases slower for low frequency and the distance r_2 becomes greater as the frequency decreases. The conclusion is that we could obtain any bandwidth lower limit, as long as we can afford the great dimension that results for the antenna. The speed of the current decreasing along the antenna depends on its particular shape, too. So, different antennas have different sizes for the same frequency bandwidth.

6.2 – Typical Frequency Independent Antennas

Equiangular Antenna

An equiangular antenna has the shape of a logarithmic spiral on a conical surface (figure no. 6.1). If $2\theta_0$ is the cone vertex angle, then, for a Cartesian coordinate system centered in the cone vertex and having the cone symmetry axis as its *Oz* axis, the spiral is described by the following equation system:

$$\begin{cases} r = r_0 e^{a\phi} \\ \theta = \theta_0 \end{cases} \qquad a \in \mathcal{R} - \text{constant} \tag{6.1}$$

The angle φ between the spiral and the position vector has the same value in any of the spiral points:

$$\varphi = \operatorname{atan} \frac{\sin \theta_0}{a} = \text{constant} \tag{6.2}$$

and this is the angle condition that the equiangular antenna fulfills.

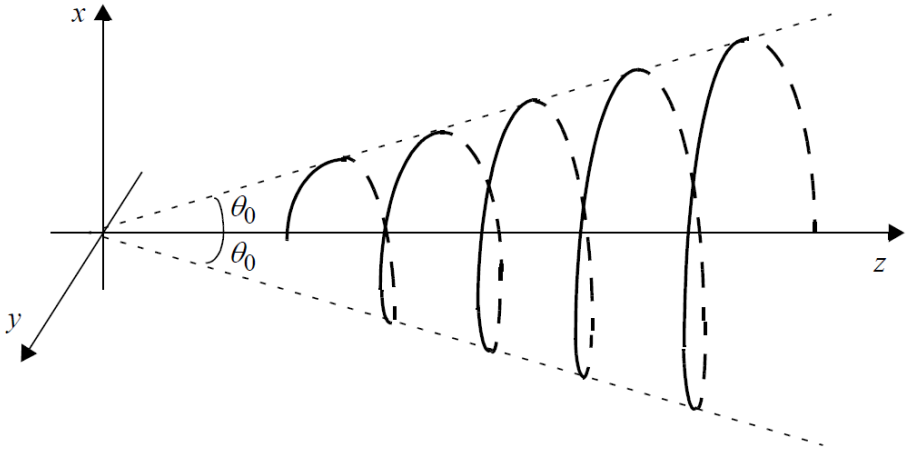


Figure no. 6.1 – *Geometry of equiangular antenna*

The electric component of the electromagnetic field created in its radiation region by the equiangular antenna is:

$$E(\theta, \phi) = \frac{e^{-jk_0 r}}{r} \left[P_\theta(\theta, \phi)\hat{\boldsymbol{\theta}} + P_\phi(\theta, \phi)\hat{\boldsymbol{\phi}} \right] \tag{6.3}$$

where the functions $P(\theta, \phi)$ are Legendre polynomials that have the following property:

$$P\left(\theta, \phi, \frac{f}{\tau}\right) = P\left(\theta, \phi - \frac{\ln \tau}{a}, f\right) \tag{6.4}$$

which shows that the radiation pattern rotates with an angle $\frac{\ln \tau}{a}$ around *Oz* axis when the operation frequency changes from $\frac{f}{\tau}$ to f. Practically, the radiation pattern remains unchanged when the frequency changes because it has revolution symmetry around *Oz* axis.

The main lobe of the equiangular antenna is oriented along its symmetry axis towards its smaller dimension. The main lobe beamwidth depends on the cone vertex angle (figure no. 6.2).

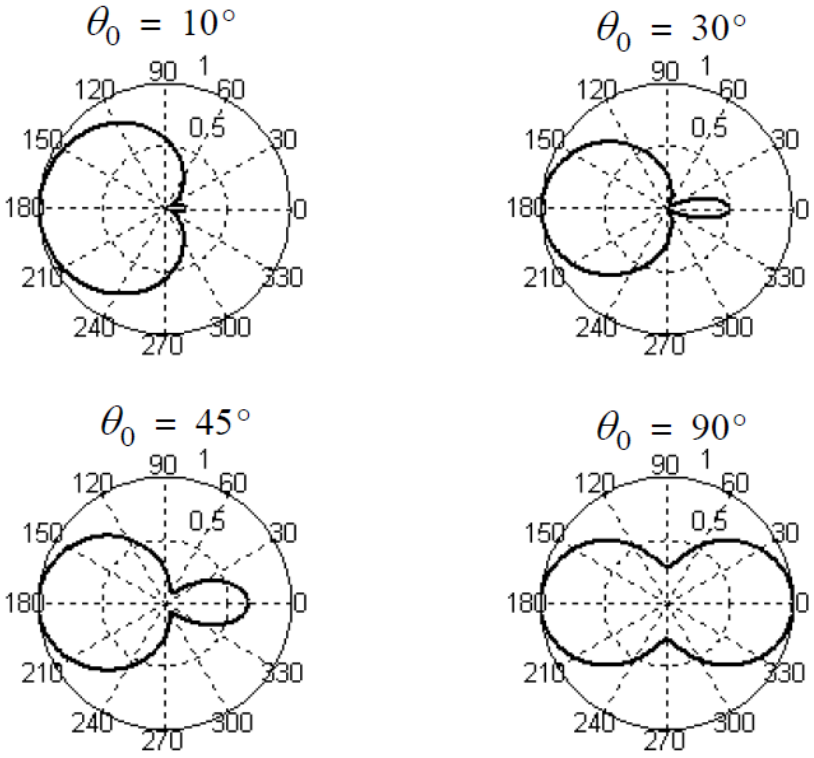


Figure no. 6.2 – *Radiation patterns of equiangular antenna*

For $\theta_0 = \pi/2$ the equiangular antenna lies in a plane and it comprises two symmetrical branches (figure no. 6.3). According to the theory the branches' cross section should increase with the distance from the vertex (as in figure no. 6.3), but very good results are obtained for constant cross section branches, too.

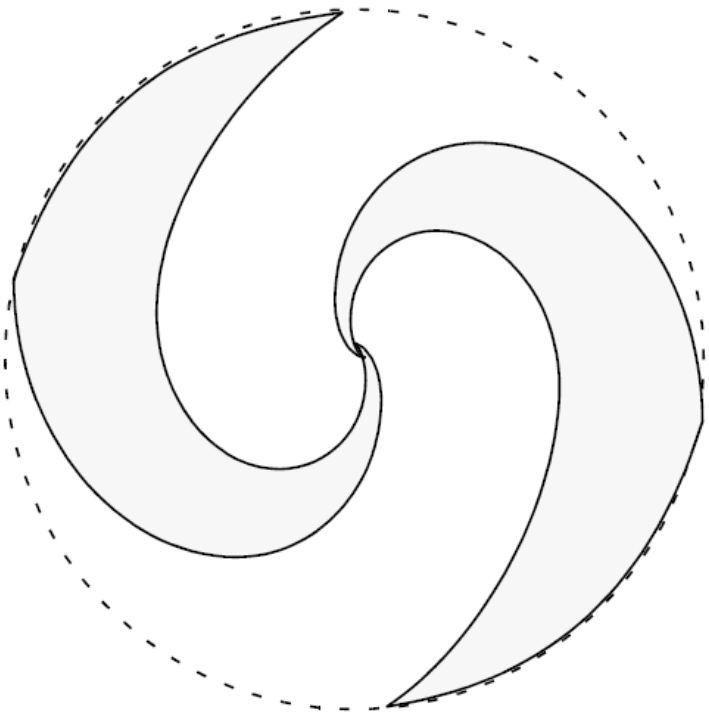


Figure no. 6.3 – *Plane equiangular antenna*

Log-periodic Antennas

There is a large variety of log-periodic antennas. Figure no. 6.4 illustrates a *log-periodic antenna with trapezoidal teeth* mostly used as transmitting antenna in 6-30 MHz band. The two branches are anti-symmetrically deployed and the teeth's length and the width increase with the distance from the vertex. In accordance with the previously stated operational principle the following relations should be fulfilled:

$$\frac{R_{n+1}}{R_n} = \tau - \text{constant}, \qquad \forall n \tag{6.5}$$

The input impedance varies periodically with frequency, but the variation limits could be kept sufficiently close if an appropriate value is chosen for τ.

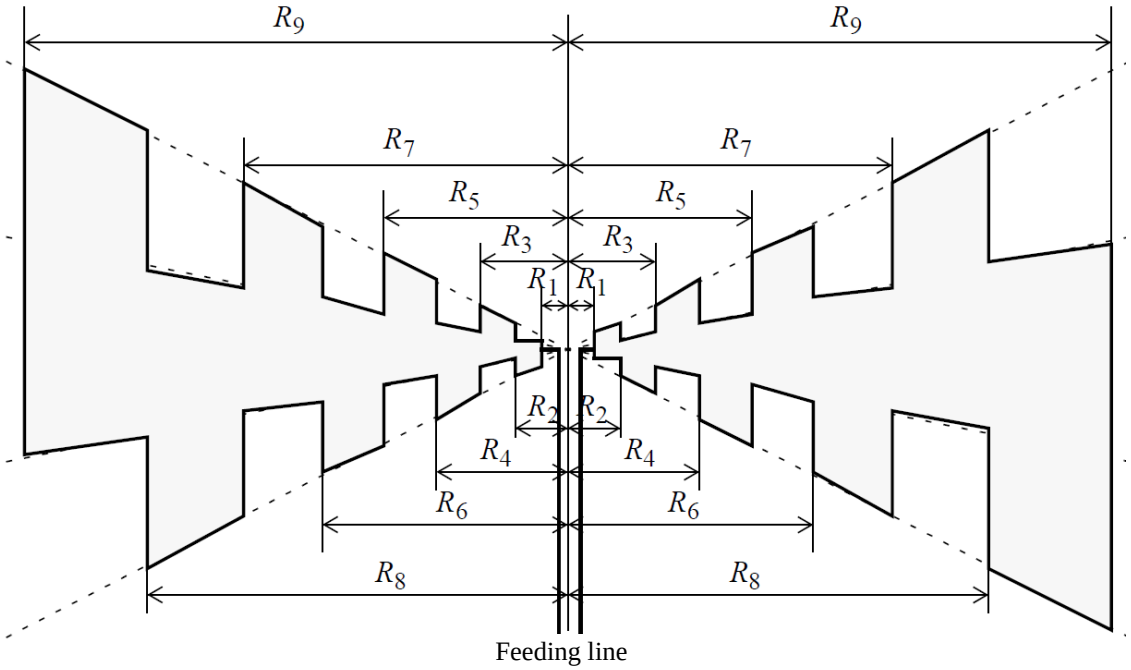


Figure no. 6.4 – *Geometry of a log-periodic antenna with trapezoidal teeth*

Figure no. 6.5 illustrates a *dipole logarithmic array* that comprises cylindrical dipoles with lengths and positions that fulfill the following relations:

$$\frac{R_{n+1}}{R_n} = \frac{l_{n+1}}{l_n} = \tau - \text{constant}, \qquad \forall n \tag{6.6}$$

According to the operational principle the dipoles' diameter should fulfill a similar relation, but good results are obtained even when the diameter is the same for all dipoles. If bi-conical dipoles are used instead of the cylindrical ones, the array active zone includes a greater number of neighboring dipoles and the radiation efficiency increases. The array active zone is a group of neighboring dipoles that radiate the most part of the total radiated power.

The feeding current phase should change by π from a dipole to another. For this, a flexible feeding transmission line is twisted (as in figure no. 6.5), while for a rigid one, the dipole branches are positioned alternately on the line branches (as in figure no. 6.6). The branches of a rigid feeding line could be metallic tubes and their connection to the source (that should be made at the array small dimension edge) can be realized by conductors laid inside the tubes. This way the coupling of the source to the array has a weak influence on the radiation pattern.

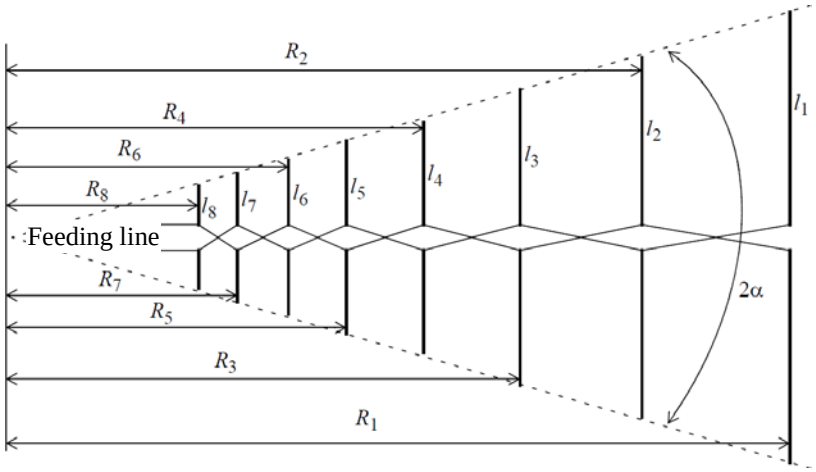


Figure no. 6.5 – *Geometry of the dipole log-periodic array*

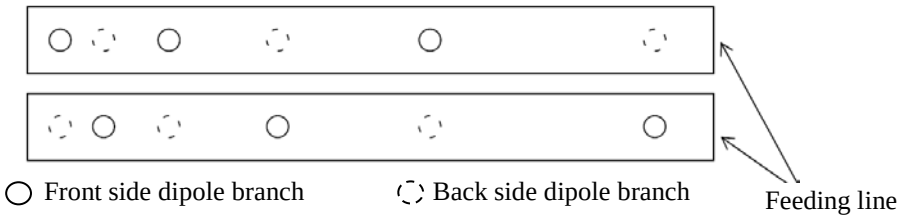


Figure no. 6.6 – *Dipole logarithmic array with rigid feeding line*

The behavior as a frequency independent antenna of the dipole logarithmic array is explained by the fact that the active zone comprises the neighboring dipoles whose lengths are close to the half wavelength and its position changes with the working frequency.

As any other frequency independent antenna, the dipole logarithmic array radiation pattern has a main lobe along its symmetry axis towards its smaller dimension. Its beamwidth slightly increases with frequency.

Due to its performance and its simple shape, dipole logarithmic arrays are frequently used for radio broadcasting in the short wave band and in antenna measurement domain.

The limits of the frequency band for a dipole logarithmic array could reach a ratio of 30:1, while its gain varies between 6 and 10.5 dBd.

The *design of a dipole logarithmic array* begins by setting a desired value for its gain. Then, appropriate values for the coefficients τ and σ in figure no. 6.7 are chosen, followed by the computation of the vertex semi angle α and of the number N of the dipoles:

$$\alpha = \operatorname{atan}\frac{1-\tau}{4\sigma} \tag{6.7}$$

$$N = 1 + \frac{\ln B_s}{\ln\frac{1}{\tau}} \tag{6.8}$$

A slightly greater value B_s is used for the frequency bandwidth in the design, as compared to the required one B, in order for the active zone to remain inside the array length, even for the extreme frequencies in the band:

$$B_s = [1.1 + 7.7(1-\tau)^2 \cot\alpha]B \tag{6.9}$$

The longest dipole has a length equal with:

$$l_1 = \frac{\lambda_{max}}{2} \tag{6.10}$$

and is positioned at the distance:

$$R_1 = \frac{l_1}{2}\cot\alpha \tag{6.11}$$

The other dimensions are computed by means of relations (6.6).

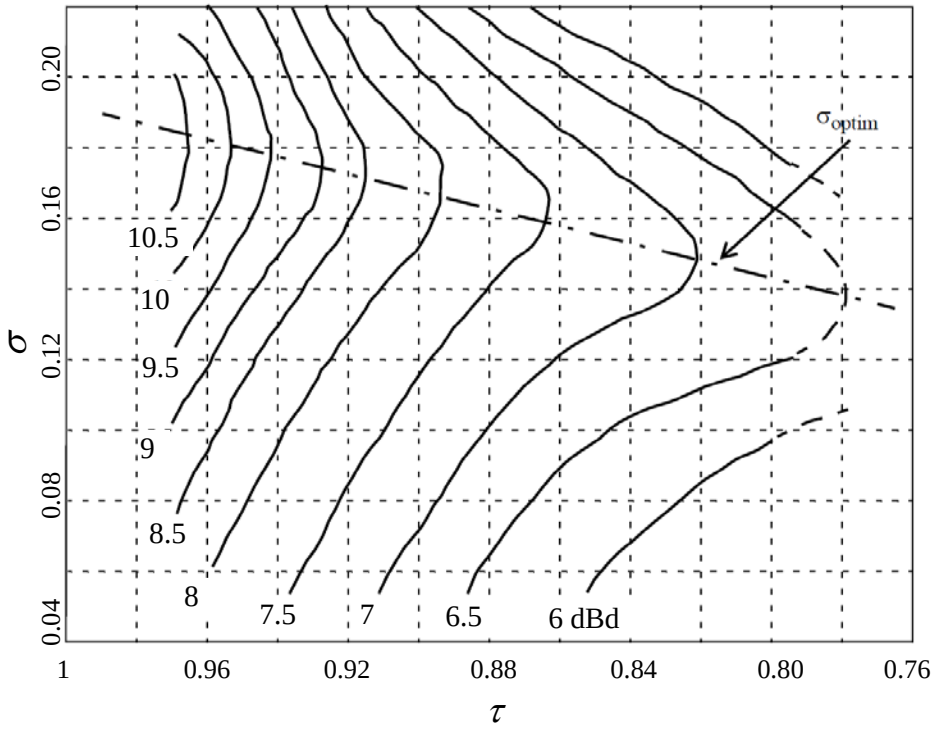


Figure no. 6.8 – *Constant gain curves for dipole logarithmic array*

The values of the coefficients τ and σ should be carefully chosen in order to avoid the appearance of a second active zone. A second active zone appears at frequencies for which a dipole with length $\frac{3\lambda}{2}$ exists. The existence of two active zones negatively influence the array input impedance and radiation pattern.

C h a p t e r V I I

REFLECTOR ANTENNA

Reflector antennas were used even since the electromagnetic waves appeared (Heinrich Hertz, 1886), although they were denoted as such. But their extensive use began with the spectacular development of radar techniques during the WWII, when precise analysis and design methods were developed. After the war new applications were implemented in astronomy, microwave communications, and satellite communications that asked for more sophisticated reflectors to be imagined and new analytical and experimental tools were envisaged in order to obtain efficiently illuminated optimal shaped reflectors, maximizing the overall gain of the antenna plus reflector combination. Frequent use of reflector antennas for deep space communications and space exploration, for Moon based research centers, for satellite direct broadcasting systems, and for multiple satellite connections to Internet made reflector antennas a common device. There are a large number of shapes and dimensions of reflectors, but the most used ones are the plane reflector, the corner reflector, and the parabolic reflector.

In this ensemble antenna + reflector, the antenna is usually denoted as *primary antenna*, while the reflector is denoted as *secondary antenna*.

Generally, the reflecting surface is made of conducting materials with finite conductivity that increases with frequency. Also, its size is much greater than the antenna size. In order to simplify the reflector analysis, its conductivity and size are considered to be infinite. Finally, the obtained results suffer some corrections to take into account the effects of the finite values for conductivity and size.

The *plane reflector antenna* analysis and design are simply realized by means of the image theorem that allows for replacing the reflector antenna with the antenna plus its image with respect to the reflector and removing the plane reflector. The results are valid only for the half space containing the antenna, the radiated field behind the reflector being zero. The plane reflector is the simplest implementation of a reflector, but we still could control the antenna radiation pattern, input impedance, and gain by adequately placing the antenna in front of the reflector.

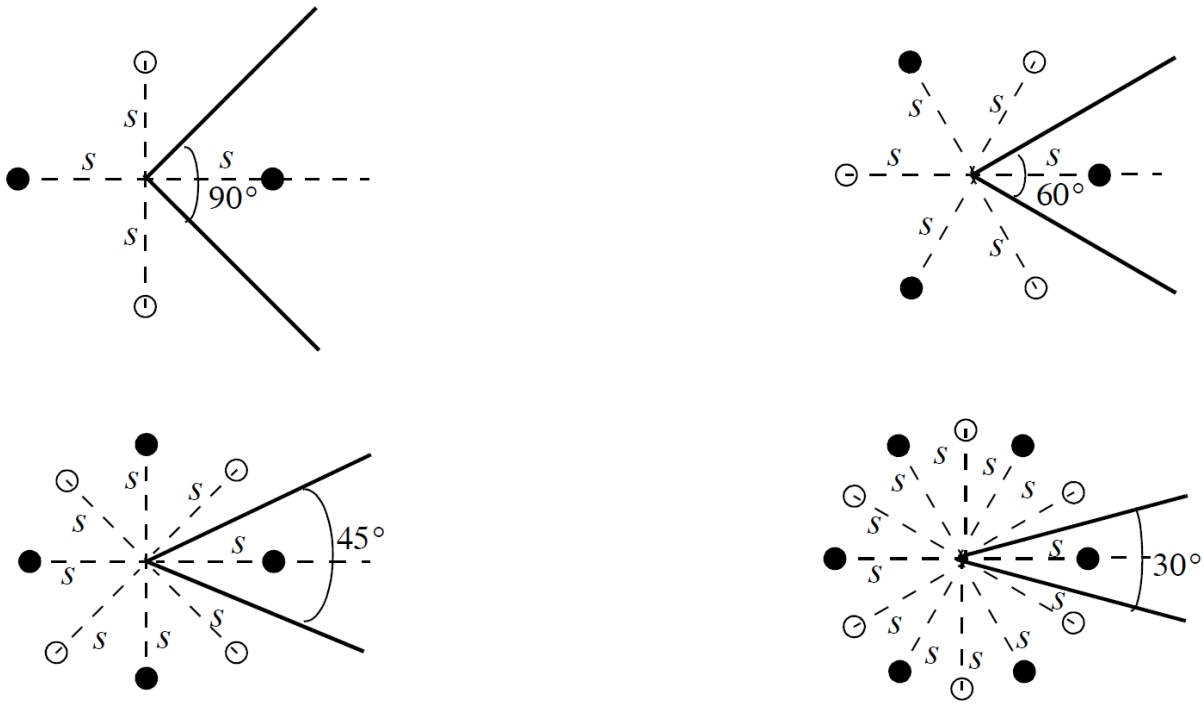
7.1 – Corner Reflector Antenna

In order to concentrate the antenna radiation in narrower solid angles more complex reflector shapes should be used. A quite simple implementation is obtained by combining more plane surfaces: *corner reflector antennas* are thus obtained. They have numerous applications. For instance, *passive radar* uses a 90° corner reflector antenna that has the property of reflecting an electromagnetic wave precisely in the same direction it arrives from (retrodirectivity). Also, *terrestrial retransmission stations* for satellite direct broadcasting systems frequently use corner reflector antennas.

Most of the corner reflector antennas use 90° plane surfaces, but there are used other angles between the reflecting plane surfaces, too: 60°, 45°, and 30°. As the angle becomes smaller, the antenna distance with respect to reflector edge should increase, if good radiation efficiency is to be kept. For infinite size reflectors the antenna gain is greater at smaller angles between planes, but for actual finite size reflectors this property is no longer valid for too small angles. So, decreasing the angle between planes under a threshold value is not recommended. In the following analysis, only the dimension of the two planes along their common edge is considered to be infinite.

Usually, corner reflector antennas are single or multiple cylindrical or conical dipoles. For large sizes, reflecting planes are made of mesh wires or wires parallel to the planes' common edge in order to maintain their weight at acceptable values. The distance between the wires should not be greater than $\lambda/10$.

Usually, the width of the reflecting planes is twice the distance between the antenna and the planes' common edge, with greater values for smaller angles between the planes. A too great width values do not improve much antenna directivity and main lobe beamwidth, but greater values for antenna frequency bandwidth and input impedance are obtained. The distance between the antenna and the reflector edge is chosen between $\lambda/3$ and $2\lambda/3$. There is an optimal value: if it is too small, the antenna input impedance is small and its radiation efficiency is small, too; if it is too great, numerous side lobes appear and antenna directivity is small. The theoretical infinite length of the planes' common edge is actually approximated by values of about 1.2 to 1.5 times the antenna length.



NOTE: Filled circle antennas and empty circle antennas are fed with opposite phase currents

Figure no. 7.1 – *Antenna images for different angles between plane reflectors*

The analysis of corner reflector antennas is easier for angles between the planes having values of type π/n, $n-$ integer, because the image theorem could be used. Figure no. 7.1 presents the images of the actual antenna for several angle values. Figure no. 7.2 illustrates the details of building antenna images for 90° angle.

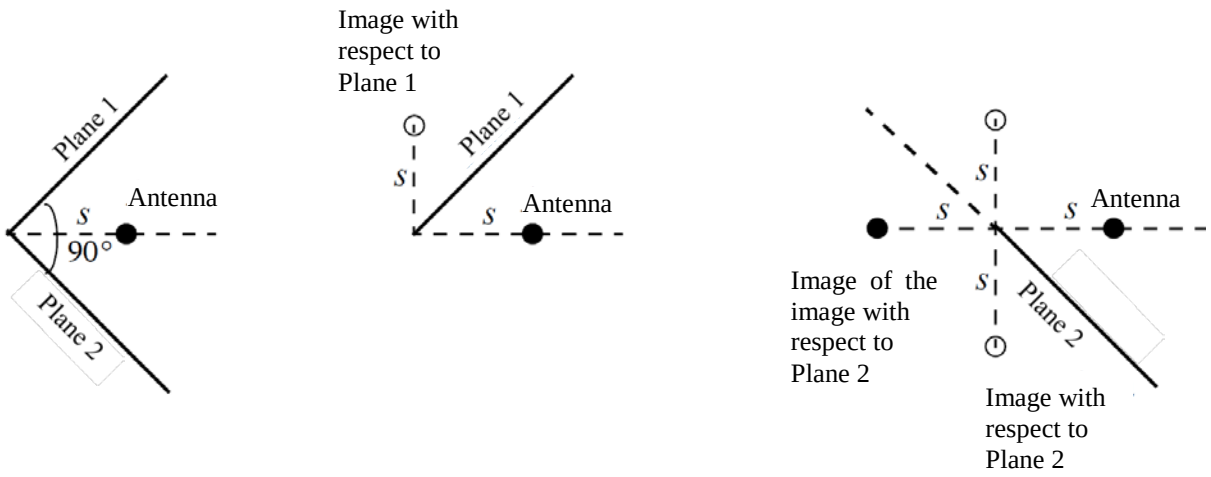


Figure no. 7.2 – *Building of antenna images for* 90° *corner reflector*

Analysis of* 90° *Corner Reflector Antenna

The total radiated field of the antenna in the presence of the plane reflectors is identical to the one created by the actual antenna together with its images, but with plane reflectors removed (image theorem), as depicted in figure no. 7.3. This is true for the space region between the plane reflectors containing the actual antenna; theoretically, outside this region the radiated field is zero.

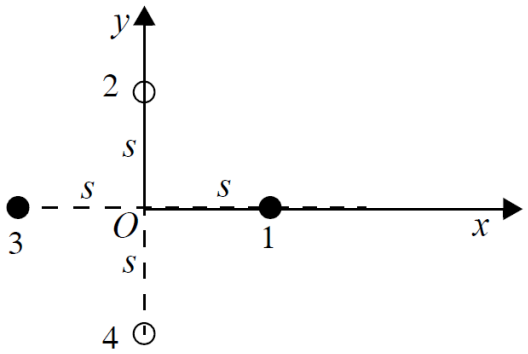


Figure no. 7.3 – *Distribution of the equivalent sources for* 90° *corner reflector antenna*

Based on the antennas' positions and their current phases as depicted in figure no. 7.3, we could write that the array factor is:

$$f(\theta,\phi) \overset{\text{def}}{=} \sum_{m=1}^{4} a_m e^{jk_0\hat{\mathbf{r}}\cdot\mathbf{r}'} = e^{jk_0 s\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}} - e^{jk_0 s\hat{\mathbf{r}}\cdot\hat{\mathbf{y}}} + e^{-jk_0 s\hat{\mathbf{r}}\cdot\hat{\mathbf{x}}} - e^{-jk_0 s\hat{\mathbf{r}}\cdot\hat{\mathbf{y}}} =$$
$$= 2[\cos(k_0 s\sin\theta\cos\phi) - \cos(k_0 s\sin\theta\sin\phi)] \qquad (7.1)$$

where s is the distance of the actual antenna and its images to the reflectors' common edge. This edge is considered infinitely long and laid along the *Oz* axis. According to the above considerations, this relation is valid for:

$$\phi \in \left[0,\frac{\pi}{4}\right] \cup \left[\frac{3\pi}{4},2\pi\right] \text{ and } \theta \in [0,\pi] \qquad (7.2)$$

The array factor in the plane normal to the reflector edge and containing the primary antenna ($\theta = \pi/2$) is:

$$f(\theta,\phi)|_{\theta=\pi/2} = 2[\cos(k_0 s\cos\phi) - \cos(k_0 s\sin\phi)] \qquad (7.3)$$

and it is graphically presented in figure no. 7.4. We note that for small values of s the array radiation pattern has one single lobe – the main lobe, while for bigger values of s, side lobes or multiple main lobes appear.

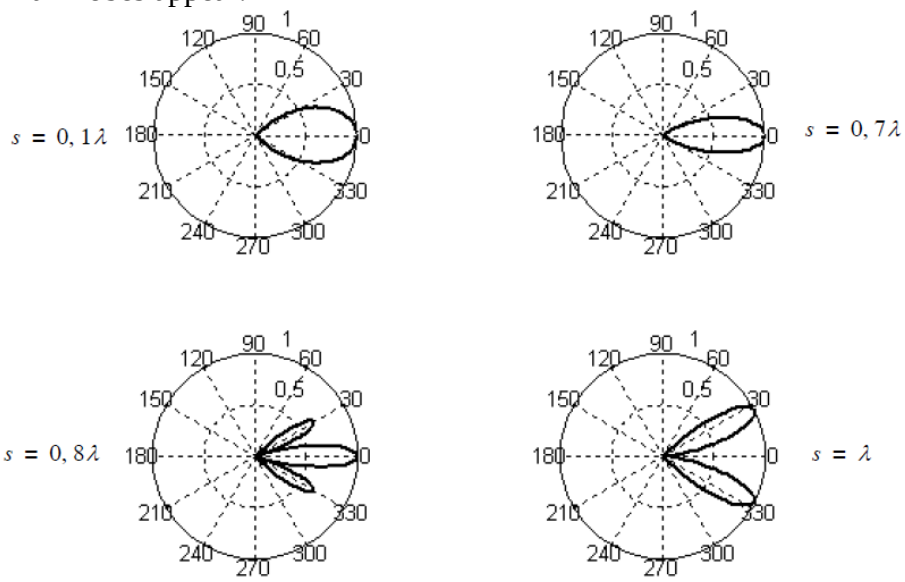


Figure no. 7.4 – *Array factor of the equivalent array for* 90° *corner reflector antenna*

The array gain depends on the distance s between the antenna and the reflector edge. Figure no. 7.5 illustrates the gain variation with s in the azimuthal plane (the plane containing the reflector edge and the antenna): gain varies periodically with period equal with λ, all maximum values are 4, and the first one appears for $s = \lambda/2$.

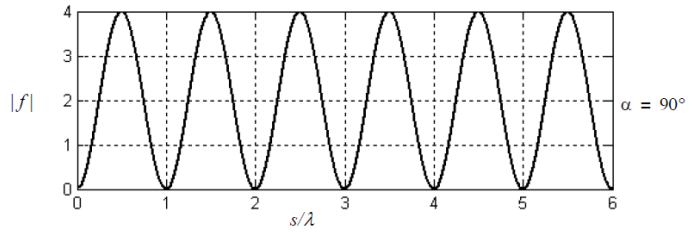


Figure no. 7.5 – *Array factor in the azimuthal plane vs* s/λ *for* 90° *corner reflector antenna*

***Corner Reflector Antennas with Angles Smaller than* 90°**

For reflector angles $\alpha = \pi/n$, $n -$ integer, an equivalent antenna array is built by means of the image theorem and its array factor is easily computed. Thus:

- for $\alpha = \pi/3$ (60°):

$$f(\theta,\phi) = 4\sin\frac{x}{2}\left[\cos\frac{x}{2} - \cos\left(\sqrt{3}\frac{y}{2}\right)\right] \tag{7.4}$$

- for $\alpha = \pi/4$ (45°):

$$f(\theta,\phi) = 2\left[\cos x + \cos y - 2\cos\left(\sqrt{2}\frac{x}{2}\right)\cos\left(\sqrt{2}\frac{y}{2}\right)\right] \tag{7.5}$$

- for $\alpha = \pi/6$ (30°):

$$f(\theta,\phi) = 2\left[\cos x + 2\cos\left(\sqrt{3}\frac{x}{2}\right)\cos\frac{y}{2} - \cos y - 2\cos\frac{x}{2}\cos\left(\sqrt{3}\frac{y}{2}\right)\right] \tag{7.6}$$

where $x \triangleq k_0 s\sin\theta\cos\phi$ and $y \triangleq k_0 s\sin\theta\sin\phi$.

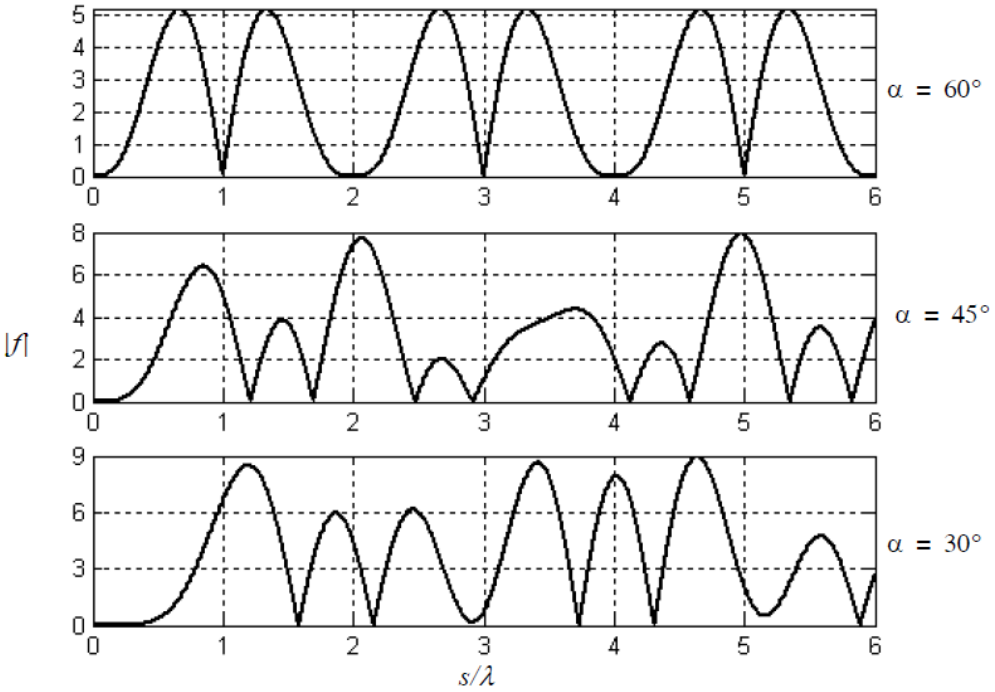


Figure no. 7.6 – *Array factor in the azimuthal plane vs s/λ for corner reflector antenna*

Figure no. 7.6 illustrates the dependence on s of these array factors in the azimuthal plane. The array gain varies periodically with period equal with 2λ for $\alpha = 60°$, 16.69λ for $\alpha = 45°$, and 30λ for $\alpha = 30°$, but the relative maxima are no longer equal. The greatest relative maximum is 5.2 for $\alpha = 60°$, 8 for $\alpha = 45°$, and 9 for $\alpha = 30°$. The first maximum (not necessarily the greatest one) appears for $s = 0.65\lambda$ for $\alpha = 60°$, $s = 0.85\lambda$ for $\alpha = 45°$, and $s = 1.2\lambda$ for $\alpha = 30°$.

Like for 90° corner reflector antennas, for angles smaller than 90° the array factor in the plane normal to the reflector edge and containing the primary antenna ($\theta = \pi/2$) has a single lobe for small values of s, but side lobes appear for bigger values of s. It is proved that side lobes appear for values of s over a threshold. This threshold is $s_{th} = 0.95\lambda$ for $\alpha = 60°$, $s_{th} = 1.2\lambda$ for $\alpha = 45°$, and $s_{th} = 2.5\lambda$ for $\alpha = 30°$.

7.2 – Parabolic Reflector Antenna

A parabolic reflector allows for a main lobe beamwidth smaller than 15° and simultaneously side lobes with very small levels. Usually, the gain of parabolic reflector antennas is between 20 dBi and 40 dBi, but for some applications even values of 60 dBi are reported. In most of the cases, revolution symmetry for the total radiation pattern is required and this asks for a similar symmetry for the radiation pattern of the primary antenna. For less critical application this requirement could be relaxed by asking for equal beamwidth of the main lobe in two reciprocally perpendicular planes; this allows for using a less sophisticated primary antenna like pyramidal horn.

Geometrical properties

The paraboloid is the geometrical locus of points in space at equal distance from a fixed plane and a fixed point, denoted as *focus*, not included in the respective plane. The

paraboloid cross the line segment drawn perpendicular to the plane from the focus in the middle point of the segment; this point is denoted as the paraboloid *vertex*.

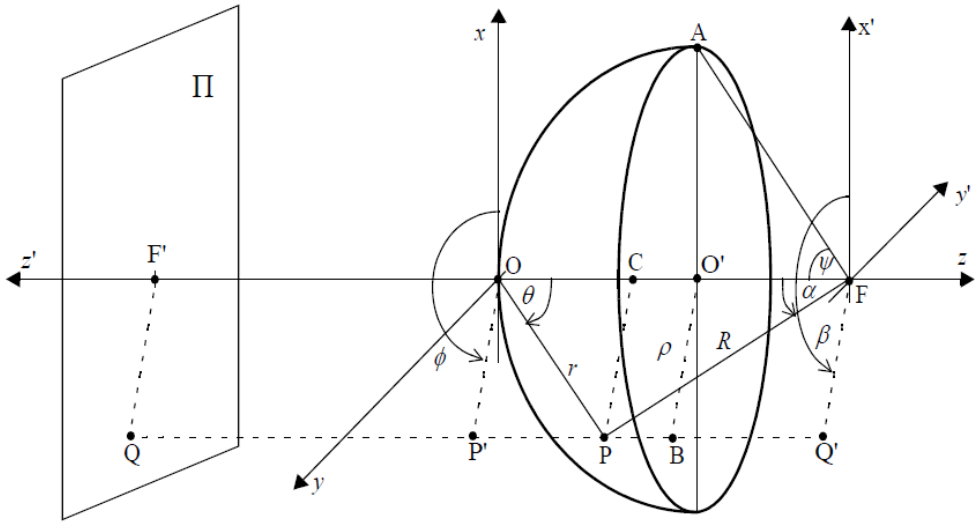


Figure no. 7.7 – *Geometry of paraboloid*

In figure no. 7.7 the fixed plane is denoted as Π, the focus as F, the vertex as O (because it is the center of a Cartesian coordinate system with OF line as its *Oz* axis), and the intersection of the *Oz* axis with plane Π as F′. The points F and F′ are symmetrical with respect to vertex O. The distance f between focus and vertex is denoted as the paraboloid *focal distance*.

If $P(x, y, z)$ is an arbitrary point on the paraboloid surface the condition PQ = PF asked by the definition to be fulfilled yields:

$$z + f = \sqrt{(f - z)^2 + x^2 + y^2} \quad \Longrightarrow \quad x^2 + y^2 = 4fz \tag{7.7}$$

which represents the equation of the paraboloid in the considered Cartesian coordinate system.

Using the spherical variables (r, θ, ϕ) associated to the above Cartesian system the equation (7.7) is written as:

$$(r \sin\theta \cos\phi)^2 + (r \sin\theta \sin\phi)^2 = 4fr \cos\theta \implies r = 4f \frac{\cos\theta}{\sin^2\theta} \tag{7.8}$$

Using the cylindrical variables (ρ, z, ϕ) associated to the above Cartesian coordinate system the equation (7.7) is written as:

$$(\rho \cos\phi)^2 + (\rho \sin\phi)^2 = 4fz \implies \rho^2 = 4fz \tag{7.9}$$

Finally, the paraboloid surface should be described in a Cartesian system centered in F, because the primary antenna is located here. The FO line is chosen as its *Fz'* axis (note that *Fz'* and *Oz* axes are in opposite directions). Also, the *Oy* and *Fy'* axes are in opposite directions, but *Ox* and *Fx'* axes are identically oriented. Using the spherical variables (R, α, β) associated to this coordinate system (see figure no.7.7), equation (7.7) is written as:

$$(R \sin\alpha \cos\beta)^2 + (R \sin\alpha \sin\beta)^2 = 4f(f - R\cos\alpha) \implies R = \frac{f}{\left(\cos\frac{\alpha}{2}\right)^2} \tag{7.10}$$

The surface resulted applying the paraboloid definition has revolution symmetry around the line OF and extends towards infinite. A parabolic reflector is a finite size body obtained by cutting the infinite surface at an appropriate distance from the vertex. Usually

the cutting plane is perpendicular to the paraboloid axis of symmetry (as is the one presented in figure no. 7.7). The cut has a circular shape and it is denoted as the *great circle* of the parabolic reflector.

The parabolic shape has some remarkable properties:

- the intersection of the paraboloid with a plane parallel with axis of symmetry is a parabola having a focal distance equal with the one of the paraboloid. This allows for checking the shape of the practical parabolic reflector by means of a reference plane parabola.
- the path length of ray starting in the focal point, reflecting on the parabolic reflector and arriving in the great circle plane is constant, irrespective of the reflection point. This property allows for a constant phase electromagnetic field in plane of the great circle.
- the reflected rays are parallel with paraboloid axis of symmetry. The incident and the reflected rays in a point make equal angles with the normal to the parabolic surface in that point and both of them are half the angle in the focal point between the incident ray and the paraboloid axis of symmetry (see figure no. 7.8).
- the angle ψ under which is seen from the focal point the diameter D of the great circle (figure no. 7.7) is uniquely determined by the ratio f/D. Indeed, in the triangle AFO' we write:

$$\tan\psi = \frac{AO'}{O'F} = \frac{AO'}{OF-OO'} = \frac{D/2}{z_0} \tag{7.11}$$

Based on relation (7.9) it results from figure no. 7.7 that:

$$\left(\frac{D}{2}\right)^2 = 4fz_0 \implies z_0 = \frac{D^2}{16f} \tag{7.12}$$

Hence:

$$\tan\psi = \frac{f/D}{2(f/D)^2-1/8} \tag{7.13}$$

q.e.d.

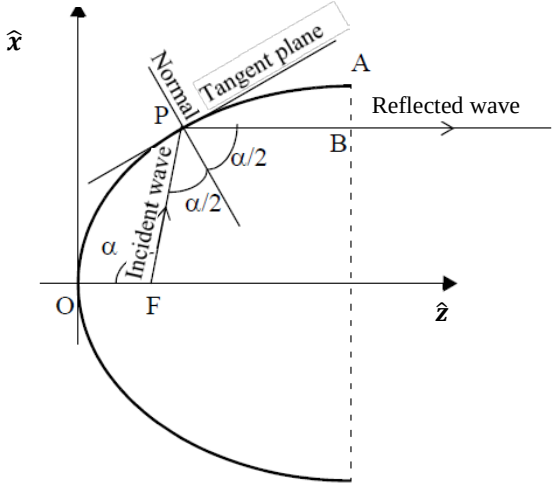


Figure no. 7.8 – *Reflected wave trajectory*

Operational Structures

In order to exploit the property of parabolic reflector of concentrating in its focal point the rays arriving parallel with its symmetry axis or, conversely, to reflect parallel with

its symmetry axis the rays arriving from its focal point it is mandatory that the primary antenna be placed in the focal point. This operational structure is denoted as *front end* configuration.

When used in the transmitting mode the front end configuration has the main drawback of blocking the rays arriving from region around the reflector vertex with two negative effects: decreasing the reflector radiation efficiency and influencing primary antenna parameters (because the blocked rays arrive from the direction of antenna maximum gain).

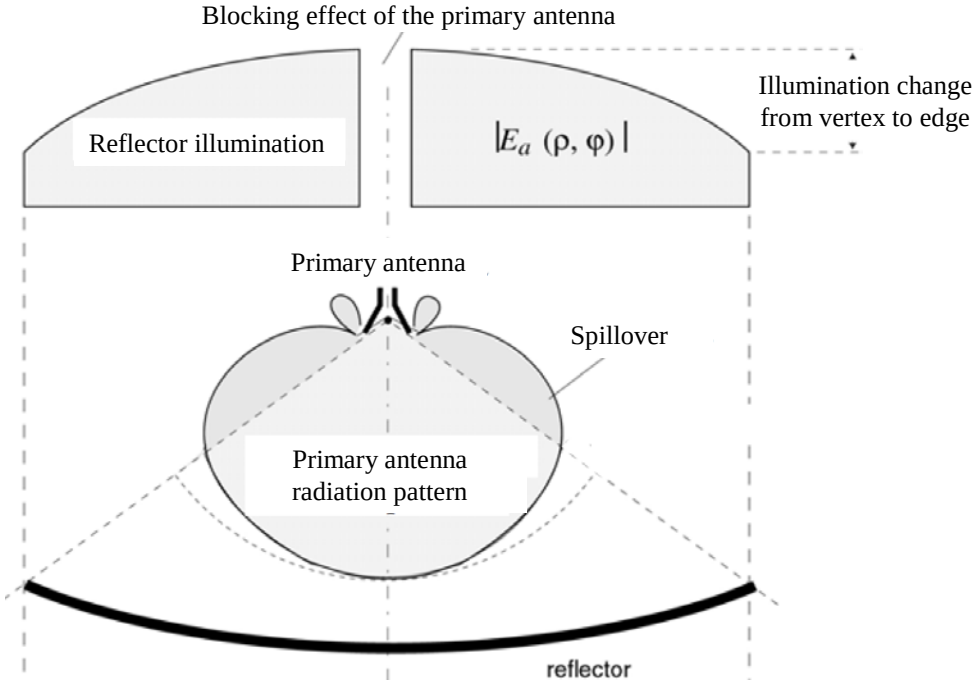


Figure no. 7.9 – *Spillover and blocking effect*

When used in the receiving mode a long transmission line is needed between the primary antenna and the receiving equipment, which is placed usually behind the reflector. An important fraction of the received waves are blocked and the receiving conditions worsen (figure no.7.9). The blocking effect is even more important when the receiving equipment is placed in the focal point together with the primary antenna.

The *double reflector* configuration is one solution for minimizing the blocking effect (figure no.7.10): a secondary reflector is placed in the focal point of the parabolic reflector (denoted in the context as the primary reflector), while the antenna is placed in the vertex of the paraboloid, behind the reflective surface. The secondary reflector is smaller than the antenna and, so, the blocking effect is smaller. When the secondary reflector is a convex hyperbolical surface this configuration is denoted as *Cassegrain antenna* (Cassegrain is a French astronomer who used this configuration for the first time). When the secondary reflector is a concave ellipsoidal surface this configuration is denoted as *McGregor* or *Gregorian antenna*.

An alternative solution for minimizing the blocking effect is using an *offset reflector*, whose surface is obtained by cutting the paraboloid infinite surface with a plane not perpendicular to the symmetry axis: the resulted parabolic surface is no longer symmetrical relative to the paraboloid symmetry axis. The antenna should be placed in the focal point, which is not placed on the new symmetry axis of the reflector. The antenna maximum gain is directed towards the new symmetry center and, thus, it no longer receives the reflected

waves from the direction of its maximum gain; thus, the antenna parameters are less influenced by the reflected waves. The blocking effect is completely eliminated if the offset reflector does not include the paraboloid vertex (figure no. 7.11).

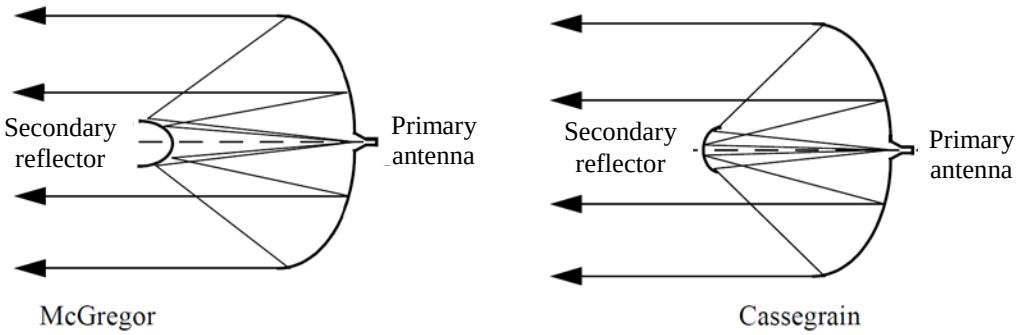


Figure no. 7.10 – *Double reflector configurations*

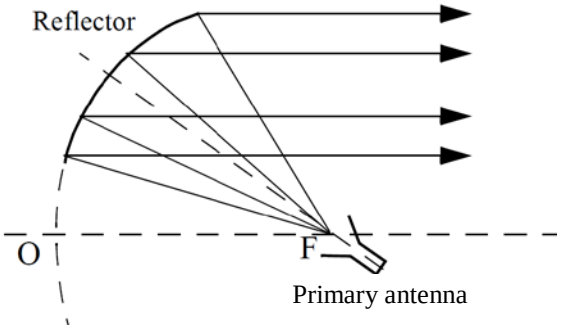


Figure no. 7.11 – *Offset parabolic reflector*

Symmetric parabolic reflector allows for both main lobe minimum beamwidth and small level side lobes. It is mostly used in radio astronomical researches, being a cost effective solution. The Cassegrain antenna is mainly used for terrestrial stations in satellite communication links due to its good radiation efficiency and to tight control of the radiation pattern.

As primary antenna conical horns are preferred as they have revolution symmetry radiation pattern. Corrugated horns are present in modern applications which allows for polarization control of the transmitted/received wave.

7.3 – Parabolic Reflector Antenna Radiation Pattern

The expression of the field radiated by the parabolic reflector antenna could be determined by means of two variables:

a. the electromagnetic field distribution in the area of the great circle of the parabolic reflector;
b. the current distribution on the parabolic surface.

In the first case we have the advantages of a uniform field distribution (due to the above presented remarkable properties of a parabolic surface) and the simple shape of the surface (a circle) upon which we have to integrate the wave equation, but little information about the radiation pattern side lobe structure is obtained. In the second case we obtain detailed information about the whole radiation pattern, but the integration process is more

complex because the surface upon which we have to carry the integration is not so simple and the current distribution is not uniform.

Irrespective of the chosen variable, simplifying hypotheses should be assumed, if we aim at obtaining less complex formulas. The usual assumed hypotheses are the following:

- the current on the back surface of the parabolic reflector is zero;
- the current discontinuity at the parabolic reflector edge is negligible small;
- the contribution of the primary antenna back lobe is negligible small;
- the blocking effect of the primary antenna is negligible small.

The results show that there is an important transversal component (normal to the symmetry axis of the paraboloid) of the current that controls the main lobe region of the radiation pattern and, also, a current longitudinal component which is responsible for the structure of the side lobe region.

7.4 – Parabolic Reflector Antenna Gain

There are three factors that contribute to the final value of the parabolic reflector antenna gain:

a) the influence of the primary antenna radiation pattern upon the optimal value of the angle ψ under which the radius of the parabolic great circle is seen from the focal point;
b) the interference of the parabolic reflector main lobe with the primary antenna back lobe;
c) the aperture efficiency of the parabolic reflector surface.

a) The existence of an *optimal value of the angle* ψ is explained as follows: for small values of ψ, the parabolic surface spans a small fraction of the primary antenna main lobe and most of the waves starting from the focal point do not intercept the reflector, a great amount of power radiated by the primary antenna is wasted due to the spill over phenomenon and the reflector gain is small. For big values of ψ, the parabolic surface spans a large portion of the primary antenna main lobe and, thus, the gain towards the edges of the parabolic reflector is much smaller than the one towards its vertex. As a result, the induced current density on the parabolic surface is far from being uniformly distributed and this prevents the reflector gain from having high gain. The conclusion is that there is an intermediary value of ψ for which the spill over is moderate and the induced current density does not differ significantly from being uniform and, thus, a maximum gain value is obtained.

The optimal value of ψ is strongly influenced by the primary antenna main lobe beamwidth. Indeed, it can be shown after some mathematical manipulations that, assuming revolution symmetry of the primary antenna main lobe, the maximum gain of the parabolic reflector is given by:

$$G_{max} = \left(\frac{\pi D}{\lambda}\right)^2 \varepsilon_{ap} \tag{7.14}$$

where (see notations in figure no. 7.7):

$$\varepsilon_{ap} = \left(\cot\frac{\psi}{2}\right)^2 \left(\int_0^{\psi} \sqrt{G(\alpha)} \tan\frac{\alpha}{2} d\alpha\right)^2 \tag{7.15}$$

is the parabolic surface radiation efficiency.

A closed form of the above expression could be obtained if an appropriate analytical expression is used for the gain $G(\alpha)$. A typical one is the following:

$$G(\alpha) = \begin{cases} G_0(\cos\alpha)^n, & 0 \le \alpha \le \pi/2 \\ 0, & \pi/2 \le \alpha \le \pi \end{cases} \tag{7.16}$$

For small values of the exponent n big values of the main lobe beamwidth are obtained and, conversely, big values of the exponent n yield small values of the main lobe beamwidth. After a normalization process we obtain that $G_0 = 2(n+1)$.

So:

$$G(\alpha) = \begin{cases} 2(n+1)(\cos\alpha)^n, & 0 \le \alpha \le \pi/2 \\ 0, & \pi/2 \le \alpha \le \pi \end{cases} \tag{7.17}$$

and:

$$\varepsilon_{ap} = 2(n+1)\left(\cot\frac{\psi}{2}\right)^2 \left(\int_0^{\psi} \sqrt{(\cos\alpha)^n}\tan\frac{\alpha}{2}d\alpha\right)^2 \tag{7.18}$$

For even integer values of n we could solve easily the integral in the above relation and obtain:

$n = 2$: $$\varepsilon_{ap} = 24\cot^2\frac{\psi}{2}\left[\sin^2\frac{\psi}{2} + \ln\left(\cos\frac{\psi}{2}\right)\right]^2 \tag{7.19}$$

$n = 4$: $$\varepsilon_{ap} = 40\cot^2\frac{\psi}{2}\left[\sin^4\frac{\psi}{2} + \ln\left(\cos\frac{\psi}{2}\right)\right]^2 \tag{7.20}$$

$n = 6$: $$\varepsilon_{ap} = 14\cot^2\frac{\psi}{2}\left[\frac{1}{2}\sin^2\psi + \ln\left(\cos\frac{\psi}{2}\right) + \frac{1}{3}(1-\cos\psi)^3\right]^2 \tag{7.21}$$

$n = 8$: $$\varepsilon_{ap} = 18\cot^2\frac{\psi}{2}\left[\frac{1}{4}(1-\cos^4\psi) - \frac{1}{2}\sin^2\psi - 2\ln\left(\cos\frac{\psi}{2}\right) + \frac{1}{3}(1-\cos\psi)^3\right]^2 \tag{7.22}$$

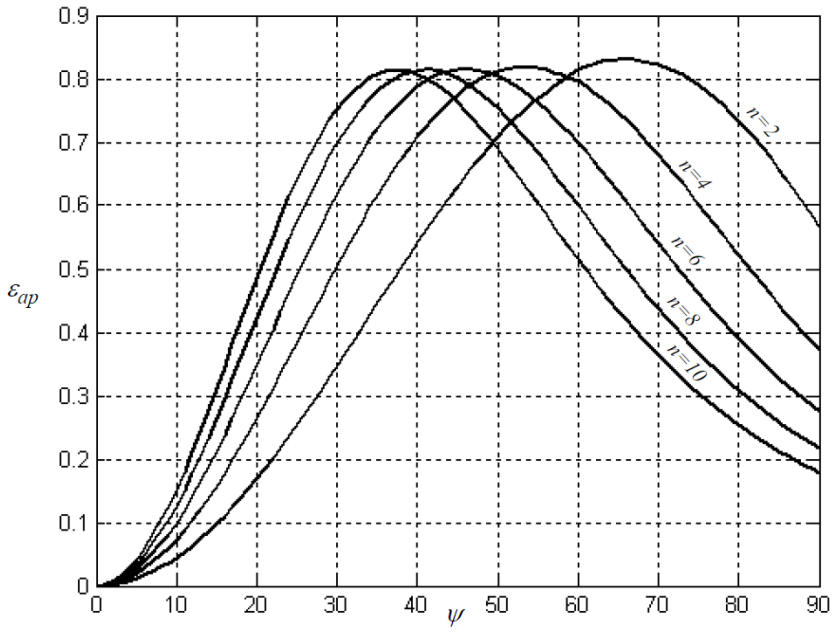


Figure no. 7.12 – *Parabolic reflector radiation efficiency versus angle ψ*

The above expressions of the parabolic reflector radiation efficiency are graphically presented in figure no. 7.12. We note that:

- it has a maximum value of about $0.82 - 0.83$, irrespective of the value of n, that is irrespective of the primary antenna main lobe beamwidth;

- the maximum is obtained for decreasing values of ψ when n increases, that is the ratio f/D should be small (see 7.13) when primary antenna main lobe beamwidth is small;

Choosing the appropriate value for the ratio f/D when designing a parabolic reflector in order to obtain its maximum efficiency requires finding out the value of the exponent n that allows for the best fitting of its main lobe region with the law (7.17); this task could prove to be extremely difficult. A simpler approach could be obtained if we use the expression (7.17) and plot the difference ΔG between the maximum gain (obtained for $\alpha = 0$) and the gain in the direction $\alpha = \psi_{opt}$. We note from figure no. 7.13 that this difference remains around 9 dB for a large domain of the exponent n values (from 2 to 10), that is for large changes of primary antenna main lobe beamwidth. Hence, the practical design method: *for a maximum radiation efficiency choose a value of ψ for which the gain of the primary antenna is about* 9 dB *smaller than the maximum value and find out the ratio* f/D *from equation (7.17).*

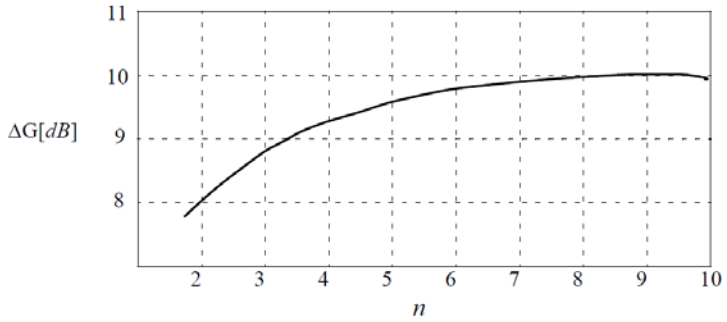


Figure no. 7.13 – *Gain difference* ΔG *versus exponent* n

For some applications like point-to-point links the main concern is the level of the side lobes which has to be maintained under a threshold value in order to limit interference with other communication systems. Based on similar considerations it can be proved that a parabolic reflector exhibits small level side lobes of its radiation pattern when its size allows for a decrease with 20 dB of the primary antenna gain towards the reflector edge as compared to the one towards the reflector vertex.

b) The primary antenna could radiate an important fraction of its power in a direction opposed to the main lobe (back lobe). The power radiated through the *primary antenna back lobe interferes with the one radiated by the reflector through its main lobe* as both of them are directed towards the reflector symmetry axis. The overall gain of the parabolic reflector antenna could be modified significantly by this interference. When the field created by the primary antenna back lobe is in phase with the one created by the parabolic reflector main lobe the resulting field intensity is increased and the interference is beneficial. But when the field created by the primary antenna back lobe is out of phase with the one created by the parabolic reflector main lobe the resulting field intensity is decreased and the interference is no longer beneficial. We have to avoid this situation.

The phase difference of the fields radiated by the primary antenna through its main lobe and through its back lobe could have only two values: 0 (radiated fields are in phase) or π (radiated fields have opposite phases). Two additional phases add to this difference:

- phase $k_0 \cdot 2f$ (f is the focal distance) introduced by the supplementary path traveled by the reflected wave from the focal point to the reflector surface and back to the focal plane;

- phase (of about π, in most of the cases) introduced in the reflection point.

The focal distance should have an appropriate value such that the field radiated by the primary antenna through its main lobe to arrive in the focal plane with same phase as the one of the field radiated by the primary antenna through its back lobe.

The interference with the back lobe can modify the reflector antenna gain by about 3%. The change is important especially for small size reflectors that have small gain.

c) the *aperture efficiency* of the parabolic reflector surface is defined as the ratio between its receiving cross section (see definition in chapter 2) and the physical area of the parabolic surface. It is computed as a product of the following partial efficiencies:

- the ratio of the power radiated through its main lobe by the primary antenna that is intercepted by the reflector (the complement of the spill over power) and the primary antenna total radiated power;
- effect of the deviation from the uniformity of the current distribution on the reflector surface (theoretically, the maximum gain is obtained for a uniform current distribution, as modulus and phase);
- effect of deviation of the electromagnetic field phase in the great circle aperture (theoretically, the maximum gain is obtained when there is no phase variation of the field);
- effect of non-uniform distribution of the field polarization state in the great circle aperture;
- ratio between the power blocked by the primary antenna and the total reflected power;
- effect of the deviation of the reflector shape from its ideal parabolic shape.

There are approximate formulas for each of these partial efficiencies. One surprising conclusion offered by these formulas is that a non-ideal parabolic reflector has maximum gain at a particular radiation frequency.

7.5 – Design of Parabolic Reflector Antenna

It results from the above considerations that the design of a parabolic reflector antenna consists in the following successive phases:

1. Choose the optimal value for ψ, which is equal to the primary antenna main lobe beamwidth at 9 or 10 dB.
2. Compute the shape factor f/D of the parabolic reflector from eq. (7.13).
3. Choose the value of the focal distance f as a multiple of $\lambda/2$ if phase of the field radiated by the primary antenna through its back lobe is opposed to the one in the main lobe. Choose the value of the focal distance f as an odd multiple of $\lambda/4$ if the field radiated by the primary antenna through its back lobe is in-phase with the one in the main lobe. *In both of the cases the field radiated by the primary antenna through its back lobe is in-phase with the one reflected by the parabolic surface*.
4. If the focal distance could not meet *exactly* the above condition, *place the primary antenna on the symmetry axis, as closed as possible to the focal point*, such that its distance from the vertex meets the above condition.
5. Compute the paraboloid great circle diameter D based on the above computed values of f/D and f.
6. Check for the phase errors of the field in the great circle aperture; desirable, they should be smaller than $\pi/4$, but in no case should they be greater than $\pi/2$.

C h a p t e r V I I I

ANTENNA MEASUREMENT

8.1 – Introduction

An antenna is an interface between a hardware and software signal processing equipment and free space where electromagnetic waves propagate. This is why antenna performance evaluation – the subject of this chapter – uses a complex combination of electrical, mechanical, and optical devices and techniques. Antenna measurement is an unique domain of science and technology using complex computing systems and sophisticated data processing techniques in order to deal with the complexity of antennas as electrical devices and with the strict requirements they have to fulfill. It is not unusual to ask for thousands of measurement points in order to define a radiation pattern or to simultaneously measure power levels spread in 50 dB range or more. Using an automated testing tool is a must.

An antenna is asked to spatially distribute its radiated power in a desired manner and this is checked by finding out its radiation pattern. This one includes a spatial region with big power density – the main lobe; a 10% error in antenna gain measurement yields a 20% error in the evaluated radiated power. So, the measurement equipment should measure big power levels with very good accuracy. In order to avoid interference with other communication systems an efficient antenna radiates extremely low power in the side lobe region of its radiation pattern. The measurement equipment should measure accurately these extremely low power levels in order to have a good evaluation of interferences. Moreover, the measurement space should be carefully prepared in order to prevent the reception of any reflected wave.

Besides gain, other parameters should be measured or computed from measurement data: radiated wave polarization, direction of the main lobe, direction and depth of nulls, reflection coefficient or voltage standing wave ration (VSWR), input impedance etc.

Most of the parameters are easily found out form a graphical representation of the radiation pattern. This could be done in a polar coordinate system as in figure no. 8.1 or in a Cartesian coordinate system as in figure 8.2. In the first case the direction of the main lobe and its beamwidth are accurately read, but the region of the side lobes is concentrated around the origin and it is difficult to read. In the latter case the side lobe region occupies a greater space and it is easily read. There are other representation modes: constant level curves (figure no. 8.3), (θ, ϕ) Cartesian system, and 3D (figure no. 8.4).

8.2 – Standard Definition of Terms

Gain – ratio of the radiation intensity to the antenna input power multiplied by 4π. Gain value is decreased by impedance or polarization mismatching between antenna and feeding line; Gain equals Directivity at matching.

Absolute Gain – Gain value in the direction of the main lobe (usually referred as antenna gain).

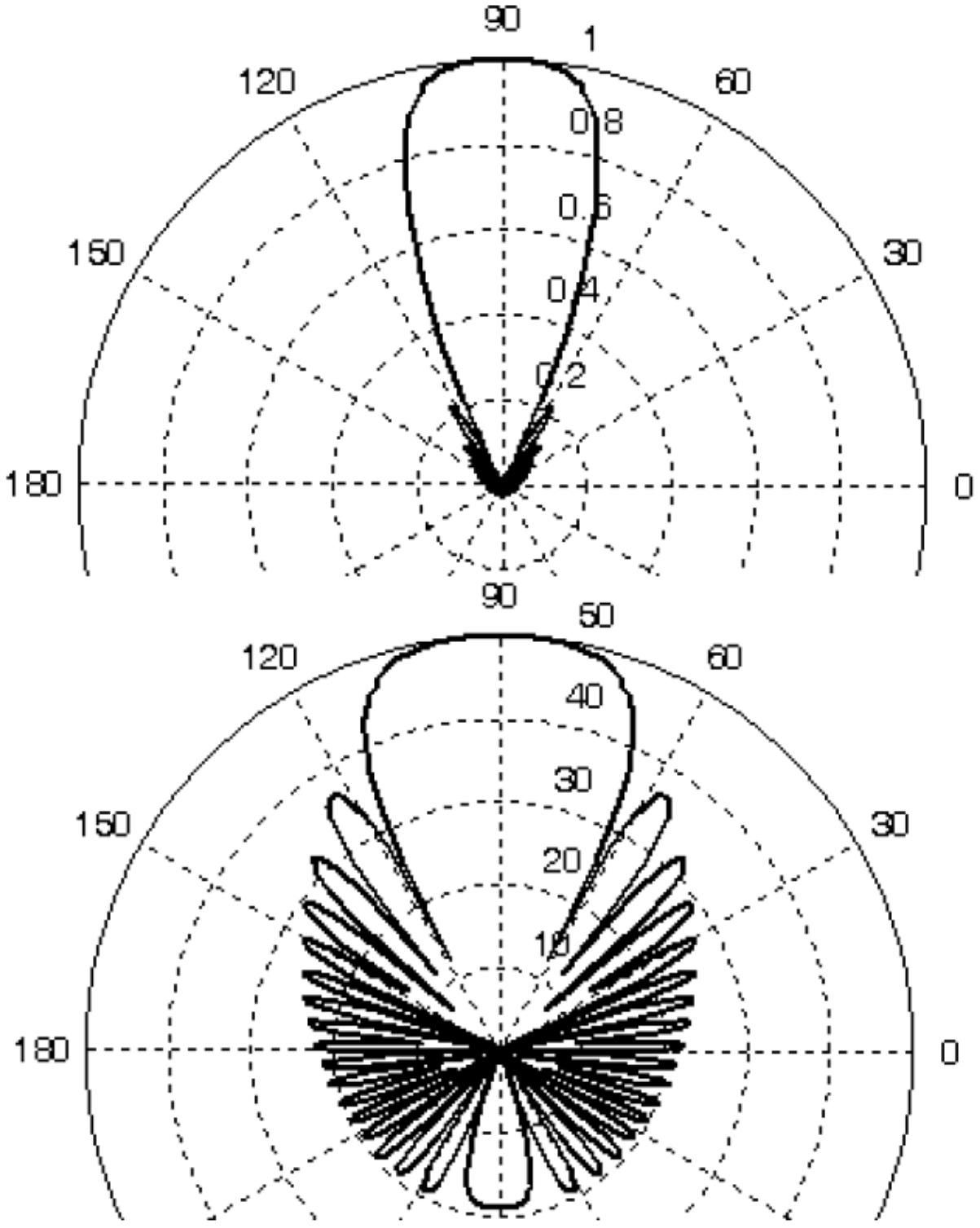


Figure no. 8.1 – *Normalized polar representation of radiation pattern: in relative values (upper) and in dB (lower)*

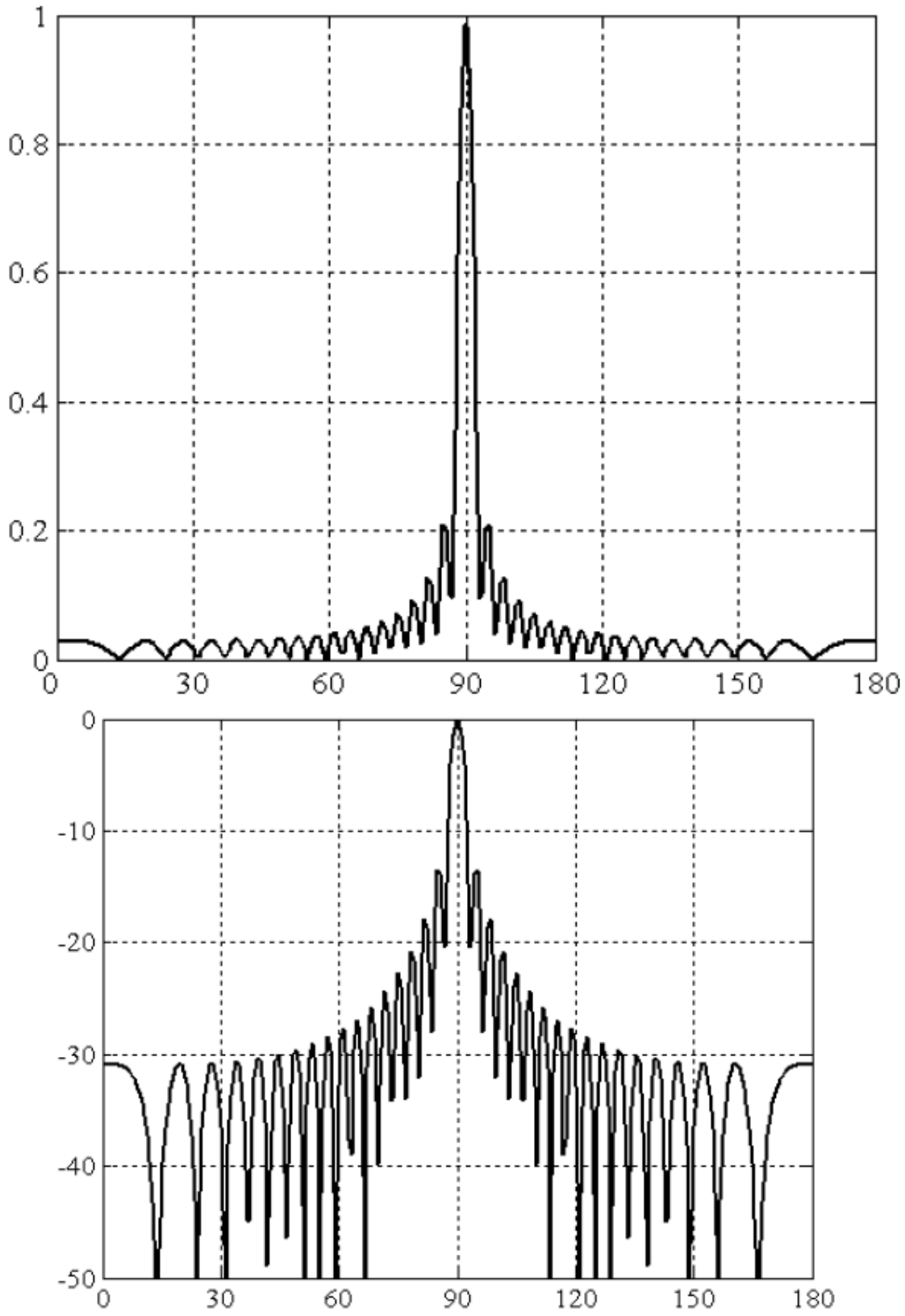


Figure no. 8.2 – *Normalized Cartesian representation of radiation pattern: in relative values (upper) and in dB (lower)*

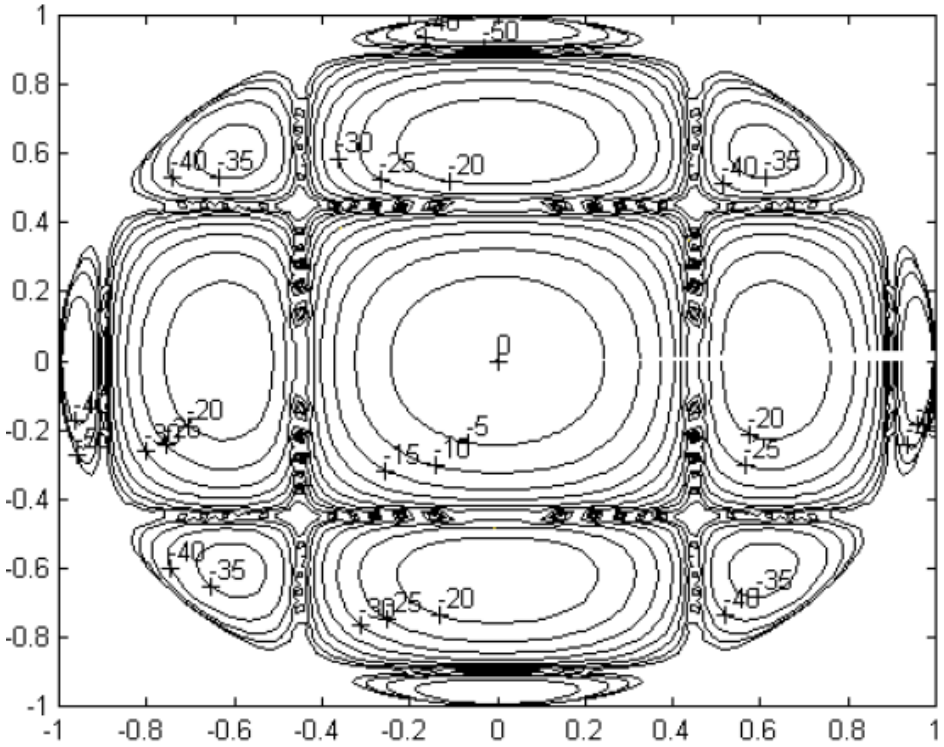


Figure no. 8.3 – *Constant level curves representation of radiation pattern (in dB)*

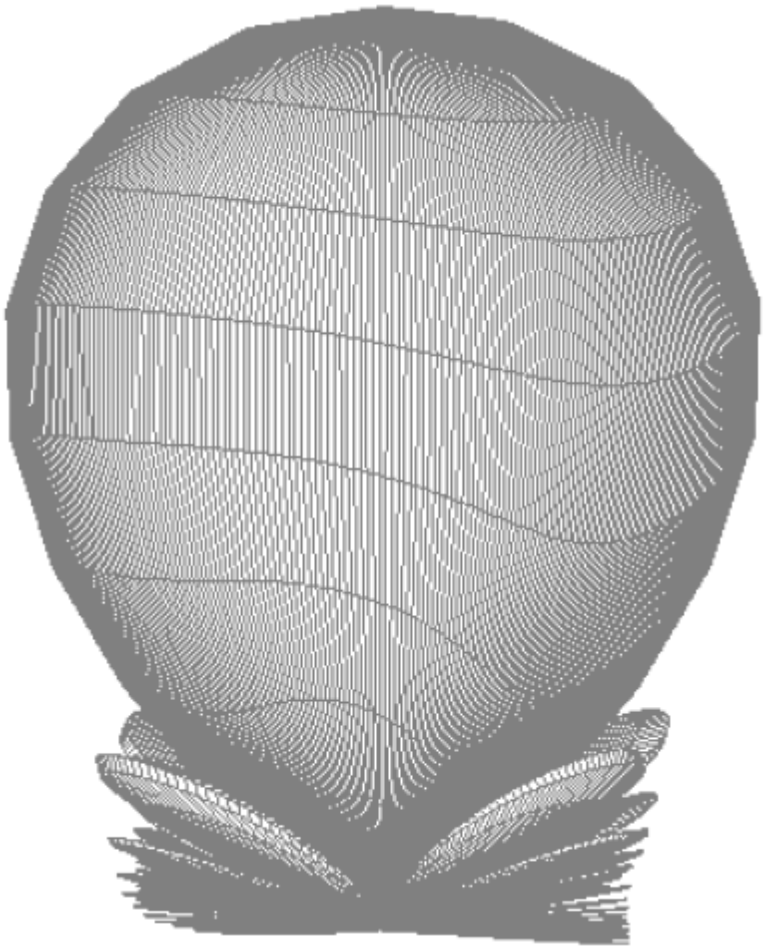


Figure no. 8.4 – *3D representation of radiation pattern (in dB)*

Partial Gain – Gain value for a single polarization component; Gain is the sum of Partial Gains for any two orthogonal components.

Relative Gain – ratio between the Gain in an arbitrary direction and the Gain in the direction of the main lobe.

Phase Center – point associated to an antenna such that the electromagnetic field has the same phase on the surface in the radiation region of a sphere centered in this point. Some antennas have multiple Phase Centers.

Co-polar component – electromagnetic wave (or one of its component) with desired polarization.

Cross-polar component – electromagnetic wave (or one of its component) with polarization orthogonal to the desired one.

Directivity – ratio of the radiation intensity to the total radiated power multiplied by 4π. According to the definition Directivity can be computed for an arbitrary direction in space, but its usual meaning is the maximum possible value for an antenna, which is obtained in the direction of the main lobe.

Partial Directivity – Directivity for a specified polarization component. Antenna Directivity is the sum of Directivity values for any two orthogonal components.

Efficiency (for an aperture antenna) – ratio of the effective antenna area to the antenna geometrical area.

Merit Factor – ratio of the Absolute Gain to the equivalent noise temperature at the antenna input.

Array Factor – antenna array radiation pattern considering that the antennas in the array are isotropic antennas. For an uniform linear array, the array factor multiplied by the element antenna radiation pattern is the total array radiation pattern.

Phase (for a circular polarized field) – angle of the field phasor with a reference direction in the polarization plane; angle has a positive sign in the sense the phasor is rotating.

Active Impedance (of an array element) – ratio of the voltage to the current at the feeding port when all the other elements are fed.

Isolated (Free Air) Impedance (of an array element) – input impedance of an array element when all the other array elements are eliminated.

Mutual Impedance (between two array elements) – ratio of the open circuit input voltage of an array element to the input current of the fed array element when all the other array elements are in the open circuit state.

Auxiliary (Side or Secondary) Lobe – any lobe of the radiation pattern besides the main one.

Main Lobe – lobe of a radiation pattern containing the maximum directivity direction. Some antennas have multiple main lobes.

Average Side Lobe Level – relative average radiated power in a given solid angle that does not contain the main lobe; the power maximum density is used as the reference value.

Polarization Plane – plane containing the polarization ellipsis; for linear polarization, the one perpendicular to the radiation direction is considered (among the infinite number possible planes).

Cardinal Plane – any symmetry plane perpendicular on a planar array and parallel to the sides of the elementary cell network the array elements are positioned.

Inter-Cardinal Plane – any plane between two successive cardinal planes containing their intersection line.

Elliptical Polarization – polarization of an electromagnetic wave (and of the antenna that radiated it) whose electric component vertex follows an ellipsis in fixed point in space. Geometrically, straight line and circle are particular shapes of an ellipsis, so linear and

circular polarizations are elliptical polarizations, too. But, usually, elliptical polarizations are considered the ones that are not linear and nor circular.

Isotropic Radiator – ideal lossless antenna radiating the same density power in all directions in space.

Noise Temperature (of an antenna) – temperature of resistor that generates a noise power density in unity frequency bandwidth equal to the noise power density at the antenna input at a specified frequency; it depends on the antenna coupling with the surrounding bodies and on the noise generated by the antenna itself.

8.3 – Measurement Techniques

There are two great categories of antenna measurement techniques: *far-field* measurement techniques and *near-field* measurement techniques.

In the first case the electromagnetic field is measured at very big distance from the antenna (the radiation region) where the field is a plane wave. The measurement accuracy depends on the accuracy of meeting the condition of a plane wave and the condition of constant phase in all of the measurement points. The antenna under test can operate as a transmitter or as a receiver. The measurements are made in free space (far-field range) or in closed spaces (compact range) where reflected waves are drastically attenuated. For compact ranges the plane wave condition is met at a smaller distance from the antenna by using parabolic reflectors, plane array of antennas or lens antennas.

Near-field measurement techniques use complex computing systems and software tools that allow for finding out the field radiated by the antenna in its radiation region based on measured radiated field (amplitude and phase) close to the antenna. Thousands or tens of thousands measurement points are needed for a good accuracy. Measurement costs are drastically reduced. The final accuracy depends on the number of the measurement points, on the surface they are distributed on, and on the performance of the data processing algorithm. The measurement points can be distributed on a sphere centered on the antenna under test, on a cylinder around the antenna or on a plane surface close to the antenna.

8.4 – Far-field Measurements

Although near-field measurements offer big cuts in measurement costs, most of the present day antenna measurements are still done in far-field ranges. The reason is the huge costs of building such far-field ranges and the need of the manufacturing companies to recover the initial investments. Moreover, even smaller as they are, the costs of near-field measurement equipment are still high and most of the companies cannot afford them.

There are three categories of far-field measurement ranges:

- with antennas placed at *big height from the ground* (figure no. 8.5). Antennas are situated on special built high masts such that ground reflected waves do not reach receiving antenna.
- with antennas placed *close to the ground* and carefully setting up the space between the two antennas as a perfect flat surface (figure no. 8.6). Ground reflected waves reach receiving antenna, but their contribution to the total received power can be precisely computed and extracted, due to the surface flatness. The departure of the ground shape from the flat condition should be kept extremely small in order for the measurement errors to remain under a given threshold.
- by using *compact ranges*. A plane electromagnetic wave can be created at small distance from the transmitting antenna by specific means like: an intermediary parabolic reflector, a special horn with covered interior walls or a planar array transmitting antenna. This way, the distance between the transmitting antenna

and the receiving one is reduced very much and they could be placed in a closed space with special treated walls, completely attenuating the reflected waves.

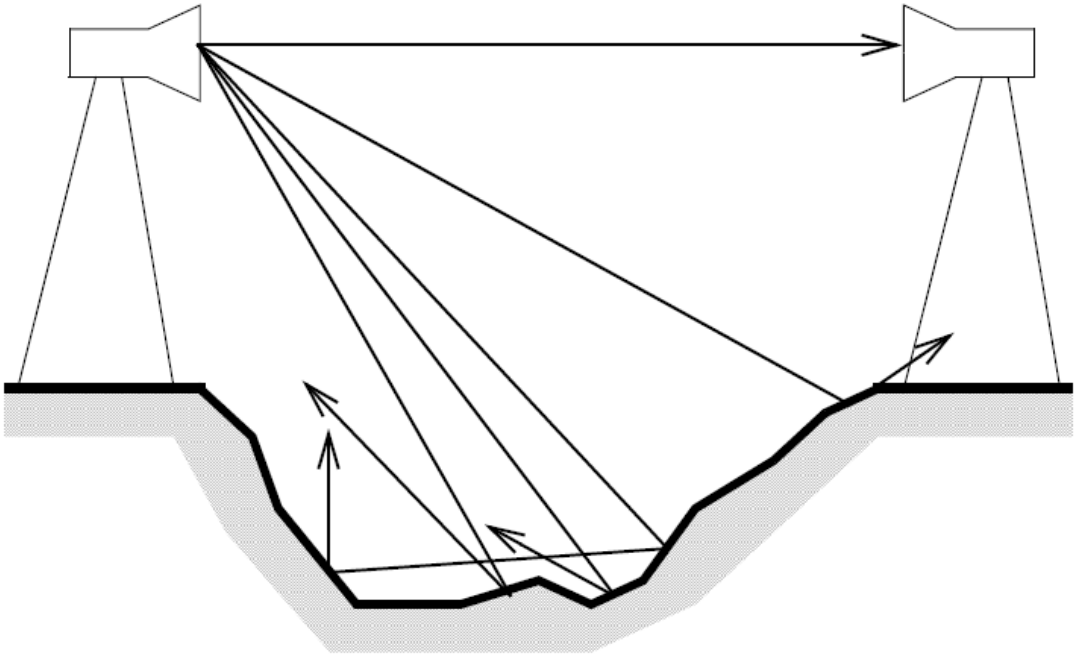


Figure no. 8.5 – *Placing antennas at big height from the ground*

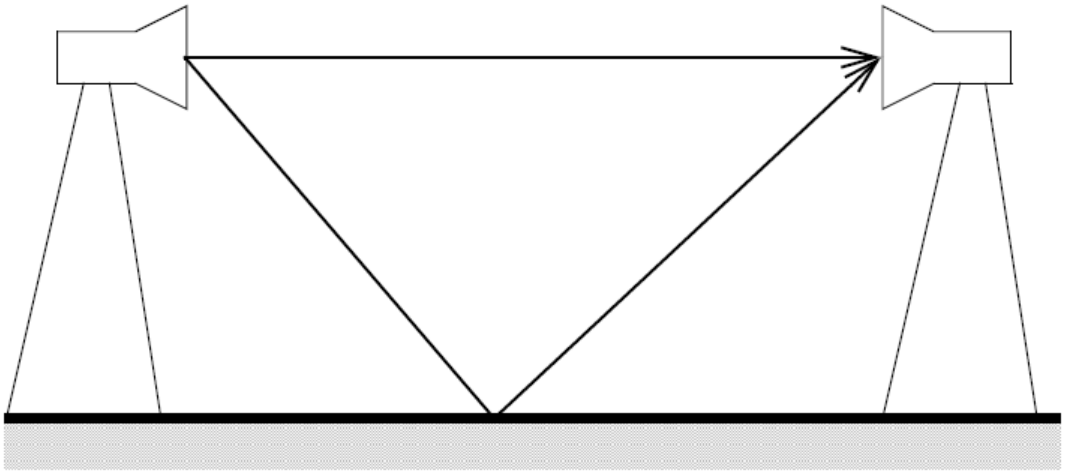


Figure no. 8.6 – *Placing antennas close to the ground*

When the antenna under test is in receiving mode, the following requirements should be met in order to reach good measurement accuracy:

- receiver noise factor should be as small as possible such that the thermal noise be smaller than the smallest signal delivered by the antenna;
- interference from other communication systems should be absent;
- measurement equipment should be able to accurately measure signal levels varying in broad range;
- the relative angle between the receiving antenna and the transmitting one must be precisely modified;
- measured field amplitude and phase should remain constant in all of the space occupied by the receiving antenna.

This last requirement is extremely important and big efforts are made in order to fulfill it. The regularity of the field distribution is the main parameter in characterizing the quality of antenna measurement range.

There is a great diversity of socio-economic factors that influence the positioning of antenna measurement range. Some of them are:

- domain of testing frequency;
- maximum size of tested antennas;
- side lobe least level that should be measured;
- possibility of big vehicles transporting the antennas to access the range;
- secure storing of antennas and testing equipment;
- other communication systems' interference level;
- existence of radio or cable communication links.

Open-space Range Size

Range size is primarily given by the requirement o realizing a very good regularity of the field distribution over the space occupied by the receiving antenna. Rigorously, the wave front at the receiving point is spherical. If the receiving antenna is placed on the Oz axis of a Cartesian coordinate system centered in the transmitting antenna phase center, at distance $OA = R$ from the origin (figure no. 8.7), then the phase difference of the field in point $A(0,0,R)$ with respect to the one in an arbitrary point $M(x,y,R)$ situated in a plane parallel to xOy and containing point A is:

$$\gamma = k_0(OA - OM) = \frac{2\pi}{\lambda_0}\left(R - \sqrt{x^2 + y^2 + R^2}\right) \tag{8.1}$$

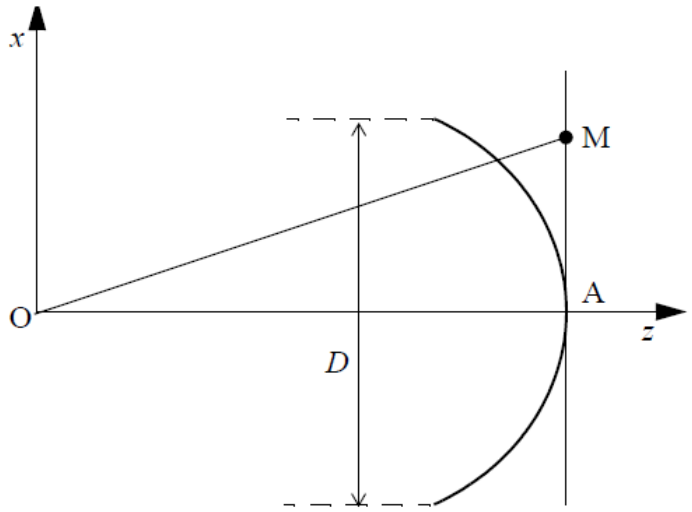


Figure no. 8.7 – *Field phase variation across the receiving antenna space*

Point A is the closest point to the transmitting antenna (point O) and its field phase is taken as reference. Thus, the field phase in M relative to the one in A is:

$$\gamma = -\frac{2\pi}{\lambda_0}\sqrt{x^2 + y^2 + R^2} = -2\pi\sqrt{\left(\frac{x}{\lambda_0}\right)^2 + \left(\frac{y}{\lambda_0}\right)^2 + \left(\frac{R}{\lambda_0}\right)^2} \tag{8.2}$$

When point M is situated in the xOz plane, its coordinates are $(x,0,R)$; for $x \ll \lambda_0$ its relative phase could be approximated as:

$$\gamma_x = -2\pi\sqrt{\left(\frac{x}{\lambda_0}\right)^2 + \left(\frac{R}{\lambda_0}\right)^2} \approx -\pi\frac{x^2}{\lambda_0 R} \tag{8.3}$$

If D is the maximum receiving antenna size along the Ox axis, then the maximum field phase difference across the antenna along the Ox axis is:

$$\gamma_{x,max} \approx \pi \frac{(D/2)^2}{\lambda_0 R} = \frac{\pi}{4} \frac{D^2}{\lambda_0 R} \tag{8.4}$$

Usually, it is considered that the radiation region of an antenna of size D is beyond the distance $R = 2D^2/\lambda_0$ from it. Thus, the field phase difference across the antenna at its radiation region limit is:

$$\gamma_{x,max} \approx \frac{\pi}{4} \frac{D^2}{\lambda_0 (2D^2/\lambda_0)} = \frac{\pi}{8} \quad \text{or} \quad 22{,}5° \tag{8.5}$$

This limit obtained for a field to be considered as having a constant phase seems too great (and it is really too great for some applications !), but it accepted in most of the practical cases due to cost reasons. Big values of the field phase difference across the receiving antenna yield significant errors in measured nulls' depth and in measured levels of the side lobes close to the main lobe. The measuring errors in the main lobe region remain small. This measuring error levels are accepted for most of the situation because trying to reduce them by 1 dB asks for doubling the range size and, consequently, increasing 4 times the total costs.

Conclusion is that the *range length should be at least* $2D^2/\lambda$, where D is the maximum antenna size to be measured, while λ is the minimum wavelength in the measuring frequency domain.

As the field amplitude uniformity across the receiving antenna is regarding, usually, the maximum accepted amplitude difference is $0.25\ dB$. Field amplitude variation across the receiving antenna is given by the wave front curvature and this one is great for narrow transmitting antenna main beam. Roughly speaking, big size antennas yield narrow beams, so limiting field amplitude variations finally translates into limiting the size of transmitting antenna. As a rule of thumb, a transmitting antenna is appropriate to use in measurements if the angle between the nulls adjacent to the main lobe is 2.4 times greater than the $-3dB$ main lobe beamwidth or 8.3 times greater than the $-0.25dB$ main lobe beamwidth.

The *range width* is inferior limited by the requirement that the transmitting antenna main lobe remain inside the range, because the reflected waves should be kept as small as possible. The range width could result smaller than this if the range is limited by large water surfaces or special walls are built at the range edges that absorb the incident waves. As a rule of thumb, without special measures, the range width should be 10 times greater than the greatest size of tested antennas.

Compact Ranges

Most part of the far-field measurement error made in open-space range is due to departure from the uniformity requirement of the field amplitude or/and phase across the receiving antenna area. This, in turn, is the effect of the finite range length, which is limited by costs associated with its implementation and maintenance.

A solution for reducing measurement error is the generation of plane front wave at smaller distance from the transmitting antenna by using special means. As the needed distance between transmitting and receiving antennas is much smaller the measurement space can be cost-efficiently closed by walls covered with conical or pyramidal bodies that destroy incident waves through multiple reflections and diffractions. This is a compact range. The generated front wave is almost plane at quite small distance from the transmitting antenna and the reflected waves are extremely weak.

The compact range has the following advantages:

- occupied space is small;

- it can be used irrespective of the weather conditions;
- measuring equipment is securely placed;
- there is no interference with other communication systems;
- tested antenna can switch simply between transmitting and receiving modes;
- testing set-up is fast and simple.

The main drawback of a compact range is its narrow testing frequency bandwidth. The maximum frequency is limited by the errors in manufactured conical or plane surfaces of the reflecting bodies. At high frequency these errors are comparable with the wave front wavelength and the associated diffraction compromises the absorbing effect of the incident wave. The minimum frequency is limited by the absorbing bodies themselves, whose size becomes comparable with the wave front wavelength and the absorbing effect is no longer efficient.

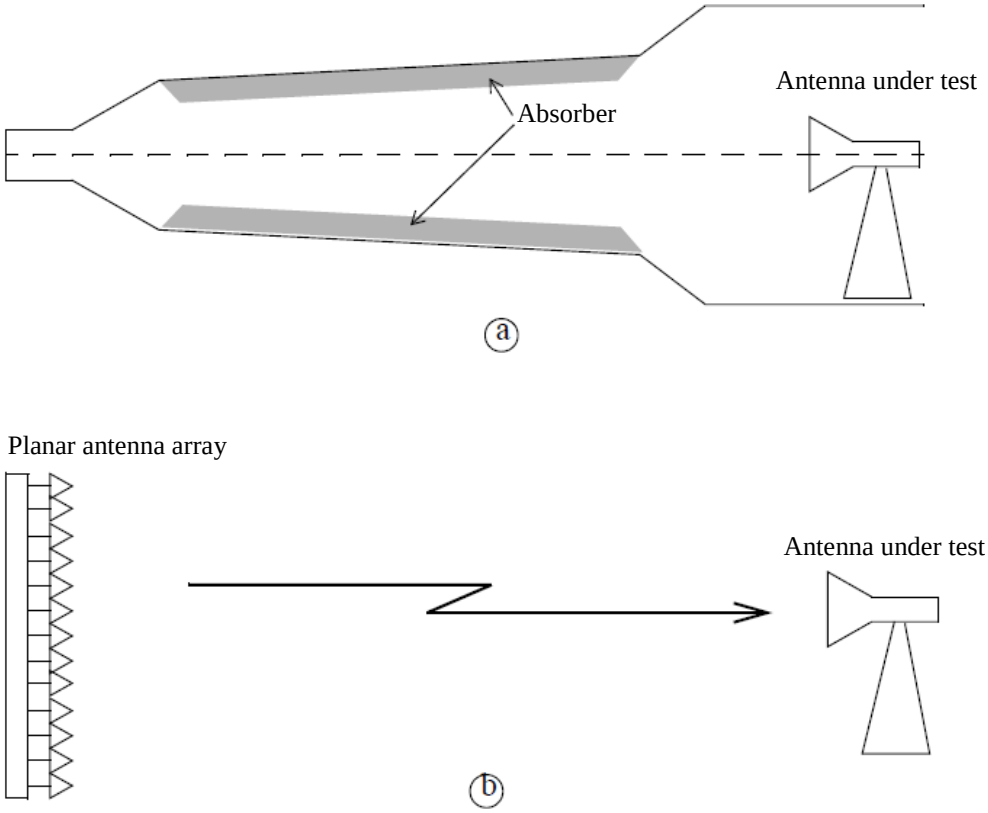


Figure no. 8.8 – *Compact ranges: a) special shaped horn and b) planar antenna array*

The means used to generate a plane wave front are the following:

a) *use of a special shaped horn* (figure no. 8.8a). The last part of the horn inner surface is covered with absorbing type materials. Small error measurements are possible if tested antenna is placed inside the "quite zone"; this is a cone around the horn symmetry axis with vertex at the end of the uncovered inner part and cross-section diameter equal to $\sqrt{R\lambda/2}$, R – distance from the vertex, λ – wavelength at the testing frequency. Note that, unlike other techniques, the quite zone size is greater at smaller frequencies.

b) *use of a big size transmitting planar array* (figure no. 8.8b). For inter-element spacing smaller than $\lambda/2$ (to keep array size at reasonable values), the quite zone begins at some small multiples of λ in front of the array. This technique is used mostly at small testing frequency.

c) *use of a horn antenna and two parabolic-cylindrical reflectors*. This technique is less costly, but its performance is modest, so it is rarely used.
d) *use of a horn antenna and an offset parabolic reflector* (figure no. 8.9). It is the most used technique. The offset reflector avoids the blocking effect and the reflected wave front near the paraboloid symmetry axis is a plane wave with uniform distribution of field amplitude and phase in planes normal to the symmetry axis.

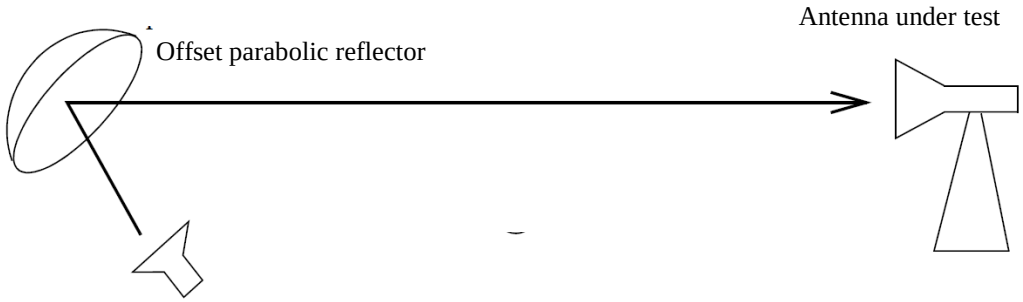


Figure no. 8.9 – *Compact ranges: Offset parabolic reflector*

Gain Measurement

a) ***Absolute Gain***. Absolute gain of an antenna is its *maximum gain* as defined by standard: the ratio of the transmitted power density in the main direction of radiation to the total antenna input power multiplied by 4π.

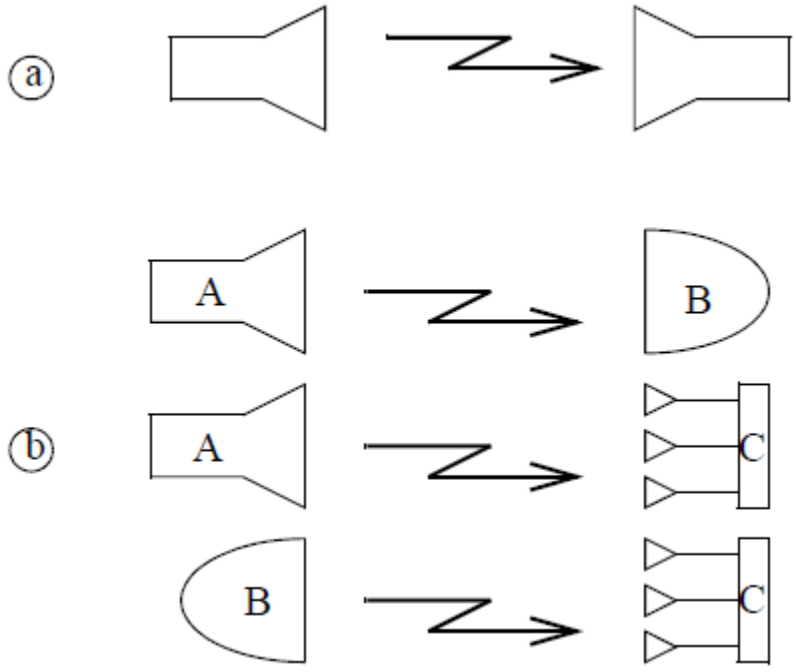


Figure no. 8.10 – *Absolute gain measuring techniques: a) using two identical antennas and b) using three different antennas*

There are three methods (figure no. 8.10):
- use of two identical antennas;
- use of three different antennas;
- use of natural radio sources.

When using *two identical antennas*, a transmission is realized between two identical antennas, each of them placed in the radiation region of the other and working in matched conditions. It is supposed that the two antennas have the same maximum gain G_a and that the each antenna is in the main lobe direction of each other.

The measured power at the receiving antenna output is:

$$P_{rec} = S_{ef} P_{\Sigma} \tag{8.6}$$

where:

$$S_{ef} = \frac{\lambda_0^2}{4\pi} G_a \tag{8.7}$$

is the antenna cross section (effective) area, λ_0 is the wavelength at the measuring frequency, while:

$$P_{\Sigma} = \frac{P_{tr}}{4\pi R^2} G_a \tag{8.8}$$

is the density power of the field created by transmission antenna in the receiving point, P_{tr} is the total transmitting power, and R is the distance between the antennas.

It results from the above equations that:

$$G_a = \frac{4\pi R}{\lambda_0} \sqrt{\frac{P_{rec}}{P_{tr}}} \tag{8.9}$$

which is the parameter we need to measure. The above expression is well known in the antenna domain and is denoted as the *Friis formula*.

When using *three different antennas* a number of three successive measurements are made. It is supposed that for each of the measurement: transmitting and receiving antennas are separated by the same distance R, they are in the radiation region of each other, they operate in matching conditions, transmitting power P_0 is the same, and the same testing frequency is used (for which the wavelength is λ_0). By using equations 8.6, 8.7, and 8.8 we obtain that:

- the power received by antenna B when antenna A transmits is:

$$P_{rAB} = \frac{\lambda_0^2}{4\pi} G_B \frac{P_0}{4\pi R^2} G_A \tag{8.10}$$

- the power received by antenna C when antenna A transmits is:

$$P_{rAC} = \frac{\lambda_0^2}{4\pi} G_C \frac{P_0}{4\pi R^2} G_A \tag{8.11}$$

- the power received by antenna C when antenna B transmits is:

$$P_{rBC} = \frac{\lambda_0^2}{4\pi} G_C \frac{P_0}{4\pi R^2} G_B \tag{8.12}$$

Thus, we can determine the gains of the three antennas:

$$G_A = \frac{4\pi R}{\lambda_0} \sqrt{\frac{P_{rAB} P_{rAC}}{P_0 P_{rBC}}} \quad G_B = \frac{4\pi R}{\lambda_0} \sqrt{\frac{P_{rAB} P_{rBC}}{P_0 P_{rAC}}} \quad G_C = \frac{4\pi R}{\lambda_0} \sqrt{\frac{P_{rAB} P_{rBC}}{P_0 P_{rAB}}} \tag{8.13}$$

Measuring technique by *using natural radio sources* is specific for large size antennas that cannot be moved from the operation site. It is supposed that power density of the radio source is known, as well as, receiver noise factor, receiver noise bandwidth, and free air attenuation in actual weather conditions (temperature, humidity, etc.). Reasonable

measuring errors are obtained when tested antenna has high gain (more than $40\ dB$) and radio source power density is much greater than local background noise level.

b) ***Relative (Transfer) Gain***. The (maximum) gain of the antenna under test is determined by comparison with the gain of a reference (probe) antenna. As probe antenna the half-wave dipole is used for frequencies below $1\ GHz$; pyramidal horn is used for frequencies greater than $1\ GHz$. Gain and polarization properties of both of the antennas are known with good accuracy.

The measuring set up is presented in figure no. 8.12. In the first phase receiver is connected to antenna under test, which is rotated in azimuth and elevation such that it receives maximum power. Also, perfect impedance and polarization matching is performed. In the second phase receiver is connected to probe antenna, for which the same settings (rotation and matching) are made. The gain of antenna under test is:

$$G_a = \frac{P_a}{P_s} G_s \qquad (8.14)$$

where P_a is the power received by antenna under test, P_s is the power received by probe antenna, while G_s is the gain of the probe antenna.

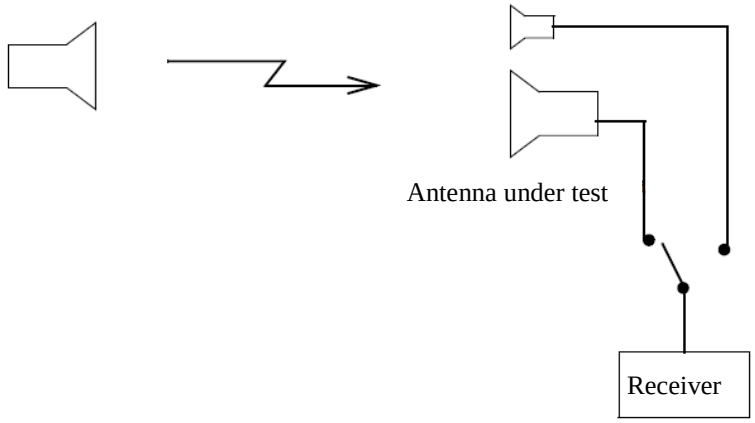


Figure no. 8.12 – *Transfer gain measuring set up*

The measuring error is given by positioning error of antennas, error in evaluating transmission line losses, and error due to receiver inaccuracy. Big error appears if the probe main lobe is much broader than the tested antenna main lobe. In order to reduce it, average of the measurements for different positions of the probe around the antenna under test should be considered.

Radiation Pattern Measurement

Measuring an antenna radiation pattern consists in measuring its gain in a number of directions, big enough to reveal its main parameters: main lobe direction, main lobe beamwidth, main lobe cross section shape, direction and depth of nulls, direction and level of side lobes, etc. Main lobe parameters show the accuracy of antenna orientation to a desired direction and its capability of rejecting undesired signals from directions close to the desired one. The side lobe structure characterizes the antenna efficiency in rejecting undesired signals from all the other directions and the relative level of interference it introduces at the receiver input.

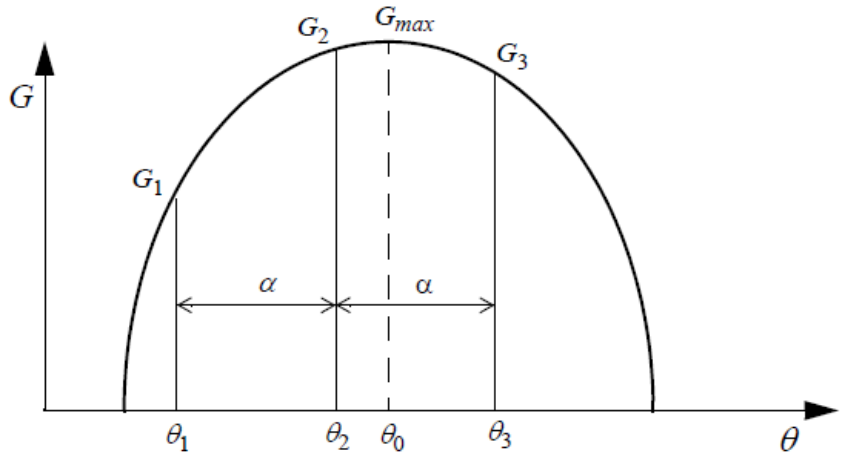


Figure no. 8.13 – *Gain measuring in parabolic shaped lobe*

Giving the huge volume of needed measurement data, the measuring process is completely automated and the direction of measurement is changed with a constant step. Thus, it is possible that measurements be made in directions close to the main lobe direction, but not precisely in this direction. The main lobe direction is computed based on the assumption that gain follows a parabolic law around main lobe vertex. If measuring direction is automatically modified with step α, then (see figure no. 8.13):

$$\theta_2 - \theta_1 = \theta_3 - \theta_2 = \alpha \tag{8.15}$$

For the parabolic law case:

$$G = a + b\theta + c\theta^2 \tag{8.16}$$

and we can write (measure unit for θ and α is *radian*):

$$G_1 = a + b(\theta_2 - \alpha) + c(\theta_2 - \alpha)^2$$

$$G_2 = a + b\theta_2 + c\theta_2^2 \tag{8.17}$$

$$G_3 = a + b(\theta_2 + \alpha) + c(\theta_2 + \alpha)^2 \tag{8.18}$$

which yield:

$$a = G_2 + \frac{G_1 - G_3}{2}\frac{\theta_2}{\alpha} + \frac{G_1 - 2G_2 + G_3}{2}\left(\frac{\theta_2}{\alpha}\right)^2$$

$$b = -\frac{G_1 - G_3}{2\alpha}\frac{\theta_2}{\alpha} - \frac{G_1 - 2G_2 + G_3}{2\alpha}\frac{\theta_2}{\alpha} \tag{8.19}$$

$$c = \frac{G_1 - 2G_2 + G_3}{2\alpha^2}$$

The coordinates (θ_0, G_0) of the parabola vertex are main lobe direction and antenna maximum gain, respectively. That is:

$$\theta_0 \triangleq -\frac{b}{2c} = \theta_2 + \frac{\alpha}{2}\frac{G_1 - G_3}{8(G_1 - 2G_2 + G_3)} \tag{8.20}$$

$$G_0 \triangleq -\frac{b^2 - 4ac}{4c} = G_2 - \frac{(G_1 - G_3)^2}{8(G_1 - 2G_2 + G_3)} \tag{8.21}$$

This algorithm can be used to find out direction and level of side lobes if the measuring step size is small enough, such that the three necessary measurement directions are close to the lobe vertex, where parabolic law variation of gain approximates well the actual gain dependence on the radiation direction.

Measurement Errors

Main causes of the far-field measurement errors are the following:

a) *Non-uniformity of the field amplitude across the receiving antenna*. Assuming that the size of projection of the area occupied by the receiving antenna on a plane normal to transmitting antenna main lobe direction (figure no. 8.14) is D and is symmetrical, the maximum relative field variation across the projection is:

$$\Delta E = \frac{E_A - E_B}{E_A} \tag{8.22}$$

For a $-3dB$ beamwidth of the transmitting antenna main lobe equal with α and a parabolic law variation of gain, by simple mathematical manipulations, we get:

$$\frac{E(\theta)}{E_A} = 1 - \left(1 - \frac{1}{\sqrt{2}}\right)\left(\frac{2\theta}{\alpha}\right)^2 \tag{8.23}$$

where θ is the angle in radians measured from the main lobe direction.

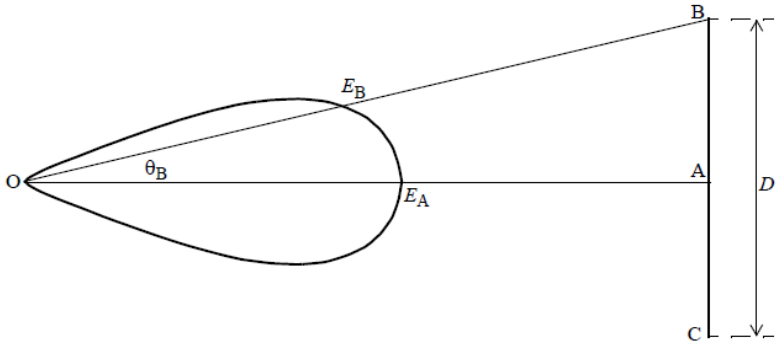


Figure no. 8.14 – *Field amplitude non-uniformity*

When the separation distance R between the antennas is much greater than the receiving antenna size, the length of the segment AB approximates very well the length l_{AB} of the arc on the circle centered in O and we could write that:

$$\theta_B \overset{\text{def}}{=} \frac{l_{AB}}{R} \approx \frac{AB}{R} = \frac{D}{2R} \tag{8.24}$$

Thus:

$$\frac{E_B}{E_A} = \frac{E(\theta_B)}{E_A} \approx 1 - \left(1 - \frac{1}{\sqrt{2}}\right)\left(\frac{2\frac{D}{2R}}{\alpha}\right)^2 = 1 - \left(1 - \frac{1}{\sqrt{2}}\right)\left(\frac{D}{R\alpha}\right)^2 \tag{8.25}$$

and

$$\Delta E = 1 - \frac{E_B}{E_A} \approx \left(1 - \frac{1}{\sqrt{2}}\right)\left(\frac{D}{R\alpha}\right)^2 \tag{8.26}$$

The above expression shows that big size receiving antenna (D) or narrow transmitting antenna main lobe (α) the field non-uniformity (ΔE) could be too great. The deviation from the constant amplitude distribution causes important gain measurement error. For instance, a value $\Delta E = 0.25\ dB$ yields a measured gain error of $0.1\ dB$.

b) *Non-uniformity of the field phase across the receiving antenna*. We showed earlier (eq. 8.4) that the maximum field phase variation across the receiving antenna depends on antenna size D and the separation distance R. For the maximum accepted value of $\pi/4$, the error in the measured gain is about $0.06\ dB$.

c) *Non-alignment of transmitting and receiving antennas*. When the directions of transmitting and receiving antennas' main lobes differ by an angle that represents a fraction k from the transmitting antenna main lobe beamwidth α, the error in the measured gain is $(24/R)(180/\pi)(k/\alpha)\ dB$. This value increases with $12(x/R\alpha)^2\ dB$, if the non-alignment between transmitting and probe antennas' main lobe directions represents a fraction x of α.

d) *Receiving of reflected waves (from ground or other obstacles)*. The received power of the reflected waves depends on transmitting antenna radiation pattern, the range configuration, neighboring obstacles distribution, and electrical properties (conductance, especially) of the ground and neighboring obstacles. Typical values for the reflection losses:

- Vegetation: $15 - 25\ dB$;
- Buildings: $5 - 15\ dB$;
- Absorbing materials: $20 - 40\ dB$;
- Ground (temperature < 5℃): $0 - 5\ dB$;
- Ground (temperature > 5℃): $1 - 15\ dB$.

e) *Measuring equipment errors*, such as:

- Transmitter frequency error (at frequency scanning, especially);
- Non-linearity of transmitter phase and amplitude characteristics;
- Calibration error of probe antenna.

In order to reduce the gain measurement total error, special measures should be taken. For instance:

- Determine the power of the reflected waves across the zone occupied by the antenna under test (but in its absence), by using as small as possible probe antenna. This power should be extracted from the received power measured at the output of antenna under test.
- Place obstacles on the ground between antennas that allow for diffuse reflection. Usually, a reflected waves' attenuation of about $10\ dB$ is obtained, but for an optimal obstacle configuration it could rise to $25\ dB$.
- Make corrections on measured data based on actual distribution of field amplitude and phase across the receiving antenna area.
- Use time domain measurement techniques that allows for separation of the reflected wave contribution from the total measured power. They use a low frequency modulation of the transmitted signal. The reflected wave is identified due to its phase delay, as it travels on a longer path.

Modern techniques use complex processing of measured data and significantly reduce the measurement error. They even allow for identification of antenna zone that contributes to the radiation pattern shape in a given solid angle.

8.5 – Near-field Measurements

Measurement Principles and Techniques

Near-field measurement techniques have all the advantages of the compact range measurements (moving on small distance, independence from weather conditions, and full security of equipment) plus the possibility to measure some details of the radiation pattern, not available otherwise. Moreover, because the antenna under test does not need to be moved during testing, large structures of antennas or fragile antennas could be precisely

tested. The main drawbacks are the great complexity of computing algorithms (hence, big computing time) and the impossibility to measure back lobes in some cases.

Near-field measurement process comprises two phases. Firstly, field amplitude and phase are measured in great number of points very close to antenna under test, in its radiative (near-field) zone, by moving an extremely small antenna (denoted as *probe*) on a surface with an appropriate chosen shape. In the second phase, the field created by the antenna in its far-field zone (radiation region) is computed based on the measured field distribution. Equations derived for aperture antenna radiation are used and they are integrated numerically.

Moving the probe on the surface of a sphere that includes the antenna under test is the logical implementation of the near-field measurement principle and maintaining the probe on spherical surface with minimum error is not a difficult task. But the numerical integration used to find out the radiated field involves complex mathematical algorithms and corrections by means of special functions (Bessel, Legendre polynomials) which lead to computing time increasing (tens of hours or even days). Keeping probe moving on plane surface near the antenna allows for using FFT algorithms for numerical integration, so the computing time is much shorter. But, the mechanical device used to keep the probe on a plane surface with small errors is very complex; moreover, the radiated field can be computed only inside a hemisphere (back lobes cannot be computed!). A trade-off is made when moving the probe on a cylindrical surface that includes the antenna under test. The complexity of the mechanism used to move the probe is not so high and the numerical algorithms of integration are not so time consuming.

The probe should not be closer than a wavelength to the antenna, in order to remain in the antenna radiative region. Also, its size should be much smaller than wavelength, in order for the field in the measuring point not to be modified in the presence of the probe. The separation between two successive measuring points should be much smaller than wavelength, such that measurement data be an accurate representation of actual field distribution on that surface.

For good measurement accuracy it is required for the level of reflected waves to be much smaller than the smallest side lobe level to be measured, while the probe itself should reflect the incident wave as little as possible. Knowing that the measurement process lasts for some hours, measures should be taken for the measured field distribution not to change during measuring time interval. This implies that the temperature inside the measurement room and the signal frequency should remain constant.

The near-field measurement principle was elaborated and applied since 1952, but accurate results appeared after 1963, due to the increased available computing power and the use of measured data correction based on actual probe parameters.

The measurement process consists in going through the following steps:

1. Measuring of field distribution (amplitude and phase) in all chosen points on two mutual perpendicular planes. Considering, for instance, that the probe moves in the xOy plane of a Cartesian coordinate system and that we choose to measure the field distribution in vertical and in horizontal planes, we obtain two functions $E_v(x, y)$ and $E_h(x, y)$ defined in discrete points in the xOy plane.

2. By means of bi-dimensional Fourier transformation we obtain their spatial frequency spectra: $A_v(k_x, k_y)$ and $A_h(k_x, k_y)$.

3. These spectra are modified according with the actual probe properties and the functions $A'_v(k_x, k_y)$ and $A'_h(k_x, k_y)$ are obtained.

4. Based on proved equations for radiation from arbitrary field distributions we compute the field in the radiation region $E_\theta(r,\theta,\phi)$ and $E_\phi(r,\theta,\phi)$ by means of numerical integration.

5. The above obtained functions represent partial radiation patterns. Total radiation pattern is given by $E=\sqrt{E_\theta^2+E_\phi^2}$. Its graphical representation allows for computing antenna under test parameters.

Optionally, functions E_θ and E_ϕ can be used to compute the actual current distribution through antenna. Thus, changes of this distribution could be made, if application asks for some modifications of the computed radiation pattern.

Mathematical models used by near-field measurement techniques are very accurate. Errors in the computed radiation pattern are caused by errors in probe positioning, field amplitude measuring, and to correction errors of measured data for field distribution distortion by reflections from walls, from antenna under test supporting structure, and from probe.

Requirements for Measuring Equipment

Gathering measuring data for field distribution determination needs probe moving and accurately positioning on a spherical, cylindrical or plane surface with size much greater than the probe size. Also, field must be precisely measured in each point and all the measured result must be recorded. Position of antenna under test must be known with an error smaller than one hundredth of wavelength at the operating frequency. If not imposed by other criteria, probe should be moved vertically, in order to avoid errors induced by gravitation. During probe moving care should be taken for the field phase not to be modified by continuously changing position of the probe connection cables. Transmitter and receiver must have very good stability of frequency and amplitude of the generated signals, appropriate availability of power, high limits for dynamic changes and linearity. Because data gathering lasts for some hours, room temperature stability and measuring equipment sensitivity to temperature become extremely important parameters.

Although the antennas are separated by much smaller distance than for the far-field measurements, the necessary transmitted power is about the same. This is due to the much lower gain of the small sized probe, to the greater losses in the probe connection cables, which are longer, and to the required small measuring error of side lobe levels.

Probe positioning error and its vibrations are reduced by reducing the scanning speed. In order to keep the total measuring time at reasonable values, in given point measuring is performed on all required frequencies. For each of the frequencies a waiting time is needed for frequency setting and stabilization and for receiver setting on the new frequency. Then, a time interval is needed for measuring field amplitude and phase and for storing measured data. Severe requirements result for the measuring equipment and for the associated computing system. Roughly speaking, measurement in a point lasts for $1-10\ ms$, so, for an usual scanning speed of $100\ mm/s$, there are $1-10$ measuring points on $1\ mm$.

Measurement Errors

Error types and their typical values are presented in Table no. 8.1. Error types 3 and 8 mainly influence main lobe direction measuring error, error types 4, 9, and 10 mainly influence gain measuring error, while the rest of error types mainly influence side lobes' direction and level measuring error.

Table no. 8.1 – *Main measuring errors for near-field measurements*

Crt. No.	Error type	Typical value	Minimum value
1.	Probe positioning	$0.5\ mm$	$0.1\ mm$
2.	Probe vibration	$0.1\ mm$	$0.01\ mm$
3.	Antenna under test alignment	$0.1°$	$0.01°$
4.	Probe gain	$0.5\ dB$	$0.1\ dB$
5.	Probe main lobe direction	$1°$	$0.25°$
6.	Probe radiation pattern	$1\ dB$	$0.25\ dB$
7.	Probe diffuse reflection	$-35\ dB$	$-50\ dB$
8.	Field phase	$5°$	$0.5°$
9.	Field amplitude	$1\ dB$	$0.2\ dB$
10.	Dynamic range	$60\ dB$	$40\ dB$

C h a p t e r I X

PROPAGATION OF ELECTROMAGNETIC WAVES

9.1 – Influence Factors

Electromagnetic wave propagation from the transmitting antenna to the receiving one is influenced by Earth surface and atmosphere. Their influence depends on the wave frequency, wave polarization, weather conditions, terrain configuration, building density, direction of propagation, antenna distance to the ground, etc. Large water surfaces or vegetation strongly modify propagation parameters.

Evaluation of communication links should take into account all the phenomena that influence wave propagation. When temperature or weather conditions exhibits large variations, statistical averages should be used to characterize propagation parameters. The statistical analysis of signal-to-noise ratio at the receiving point should provide reasonable margins, such that the fraction of time with low quality or impossible reception be extremely small. Limits depend on the type of communication system.

Communication systems' performance is evaluated based on the assumed or measured propagation characteristics. The weight of different factors influencing wave propagation strongly depends on wave frequency. From this point of view frequency domain comprises, roughly, four main bands:

- *extremely low frequency band (ELF).* It includes frequencies smaller than about $1\ kHz$. Their wavelength greater than $100\ km$ asks for large size antennas very close to the ground (some of them are underground!). The distance between the ground and the ionosphere is comparable with their wavelength and guided type propagation is possible, allowing for large distance propagation. This frequency band is mainly used for communication with submersed entities because the path loss through the salted water is still at reasonable values.

- *low frequency band (LF).* It includes frequencies smaller than a few tens of MHz. Wave propagation is strongly influenced by the Earth surface and is due mainly to the so-named surface wave. It covers propagation distance until a few hundreds of km. Typical communication systems are radio broadcasting on long and medium waves.

- *high frequency band (HF).* It comprises frequencies smaller than $50\ MHz$. Waves are reflected by the ionosphere and transmissions at large distance (thousands of km) are possible. Surface wave is still present. The greatest distance where surface wave could be received is much smaller than the distance from the transmitter where the ionosphere-reflected wave reaches the Earth surface. So, usually for these transmissions, a quiet zone appears, where none of these waves is received. Typical communication systems are short wave radio broadcastings that rely solely on the ionospheric waves (transmitter radiates directly to the ionosphere). The zone on the Earth surface where the ionospheric waves reach the ground is variable because it depends on the ionosphere electron density and on its height, both of these variables having great changes with time of the day and with seasons of the year. Ionospheric transmissions are associated with significant fading, that is slow-time

variations of received power density, in a large dynamic range. Thus, long time intervals appear when the reception is of bad quality or even impossible.

- *very/ultra high frequencies (VHF/UHF).* It comprises frequencies greater than $50\ MHz$. Antenna size is small enough to be placed at big distance from the ground. The ground parameters have small influence on wave propagation. Also, the ionospheric reflection is absent because waves are absorbed by the ionosphere. Ground reflected waves could appear and interference with direct wave is possible in some locations. For frequencies greater than $1GHz$ path loss is big and multiple reflections and diffractions (scattering) could appear. Typical communication systems: MF radio broadcasting, TV broadcasting, radio relay transmissions, etc. Path loss is very high for frequencies greater than $20\ GHz$ and communication is possible only over distances smaller than a few tens of meters. Typical communication systems: tele-control, inter-satellite links, etc.

9.2 – Propagation over a Plane Surface

Path Gain Factor

Figure no. 9. 1 displays the configuration of a transmission system with two antennas separated by distance d and placed at heights h_1 (transmitting antenna) and h_2 (receiving antenna), respectively, above a perfect plane surface.

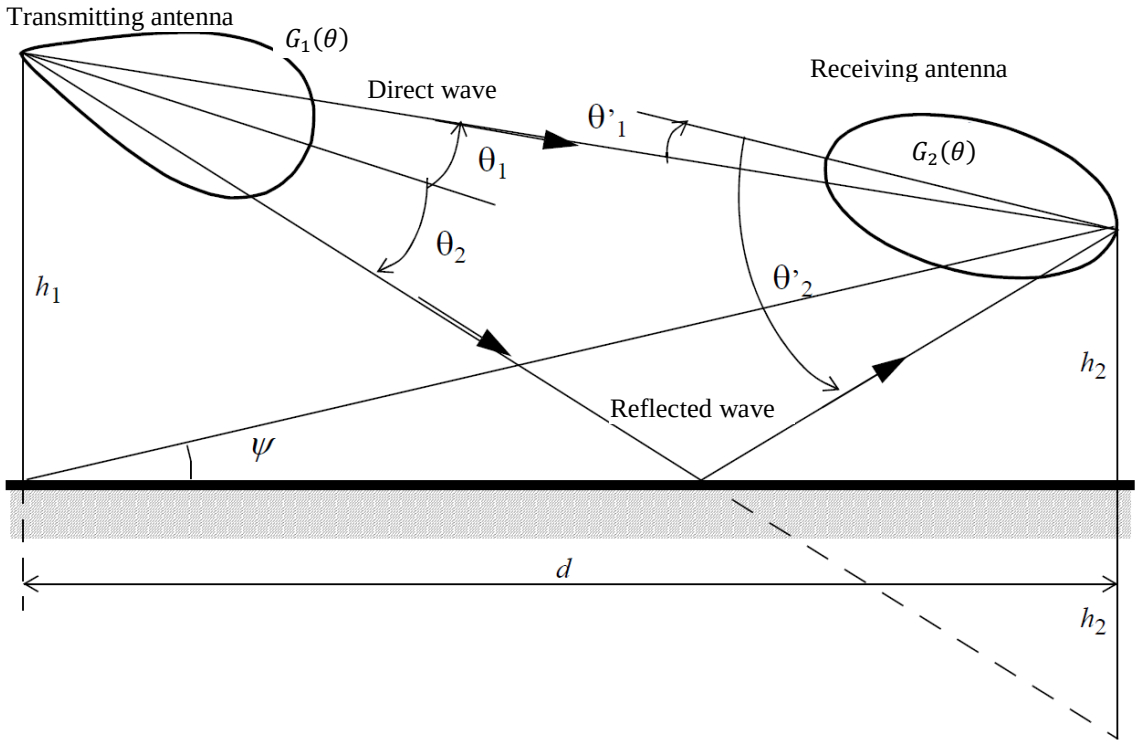


Figure no. 9.1 – *Transmission above a plane surface*

Direct wave path length is R_1, while the ground reflected wave path length is R_2. Due to different path lengths, the two waves arrive with different phases at the receiving antenna, so the vectorial sum of the electromagnetic field could be greater or smaller than the one of the direct wave, depending on the actual phase difference. Let's denote by $G(\theta)$ the gain function of an antenna, θ being measured from the antenna maximum gain direction (direction of the main lobe); in normalized representation $G(0) = 1$. In real systems, the two antennas' main lobes are not perfectly aligned to each other. Let's denote by θ_1 and θ_1',

respectively, the angles made by the direct path between antennas with transmitting antenna and receiving antenna main lobe directions. Also, we denote by θ_2 and θ_2', respectively, the angles made by the ground reflected wave path with the same directions.

Assuming that each antenna lies in the radiation region of the other, the complex amplitude of the field received on the direct path is:

$$E_d = G_1(\theta_1)\, G_2(\theta_1')\, \frac{e^{-jk_0R_1}}{4\pi R_1} \tag{9.1}$$

and the one of the field received on the ground reflected path is:

$$E_r = G_1(\theta_2)\, G_2(\theta_2')\, \rho e^{j\Phi}\, \frac{e^{-jk_0R_2}}{4\pi R_2} \tag{9.2}$$

where ρ is the ground reflection coefficient and Φ is field phase change at reflection.

In most of the cases the distance d between the antennas is much greater than their heights h_1 and h_2 above the ground and, so, angles θ_1, θ_1', θ_2, and θ_2' are small; as a result:

$$G_1(\theta_1) \approx G_1(\theta_2) \text{ and } G_2(\theta_1') \approx G_2(\theta_2') \tag{9.3}$$

For computing the total field modulus we assume that $R_1 \approx R_2$. Based on this assumption we can write the modulus of the total received field is:

$$|E_{rec}| = |E_d + E_r| = \left|G_1(\theta_1)G_2(\theta_1')\frac{e^{-jk_0R_1}}{4\pi R_1}\right| \left|1 + \rho e^{[\Phi - jk_0(R_1-R_2)]}\right| \tag{9.4}$$

The factor:

$$F = \left|1 + \rho e^{j[\Phi - k_0(R_1-R_2)]}\right| \tag{9.5}$$

is denoted as the *path gain factor* and it shows how the reflected wave modifies the direct one, generating the total received field.

By means of simple geometry and taking into account the above stated approximation $h_1, h_2 \ll d$ we get:

$$R_1 = \sqrt{d^2 + (h_1 - h_2)^2} = d\sqrt{1 + \left(\frac{h_1 - h_2}{d}\right)^2} \approx d\left[1 + \frac{1}{2}\left(\frac{h_1 - h_2}{d}\right)^2\right] \tag{9.6}$$

and

$$R_2 = \sqrt{d^2 + (h_1 + h_2)^2} = d\sqrt{1 + \left(\frac{h_1 + h_2}{d}\right)^2} \approx d\left[1 + \frac{1}{2}\left(\frac{h_1 + h_2}{d}\right)^2\right] \tag{9.7}$$

such that:

$$R_2 - R_1 \approx \frac{2h_1h_2}{d} \tag{9.8}$$

Using the above approximations, for lossless ($\rho = 1$) and phase reversal ($\Phi = \pi$) ground reflection we get the *ideal path gain factor*:

$$F \approx \left|1 + e^{\mathrm{j}\left(\pi - k_0\frac{2h_1h_2}{d}\right)}\right| = 2\left|\sin\frac{k_0h_1h_2}{d}\right| \tag{9.9}$$

In ideal conditions the path gain factor varies between 0 and 2, meaning that the direct wave could be completely compromised by the ground reflected wave or its modulus be doubled after combining with the latter. Of course, in actual situations the resultant wave modulus is strictly positive and less than the double of direct wave modulus.

Going back to ideal conditions, note from figure no. 9.1 that:

$$\frac{h_2}{d} = \tan\psi \tag{9.10}$$

Expression 9.9 can be rewritten as:

$$F \approx 2|\sin(k_0 h_1 \tan\psi)| \tag{9.11}$$

Path gain factor F reaches its maximum for:

$$\tan\psi = \frac{1}{k_0 h_1}\left(\frac{\pi}{2} + n\pi\right) = \frac{\lambda_0}{h_1}\left(\frac{1}{4} + \frac{n}{2}\right), \quad n = 0,1,2,\ldots\ldots \tag{9.12}$$

and its minimum for:

$$\tan\psi = \frac{1}{k_0 h_1} n\pi = \frac{\lambda_0}{h_1}\frac{n}{2}, \quad n = 0,1,2,\ldots\ldots \tag{9.13}$$

For given values of the transmitting antenna height h_1 and of the operating frequency (or, equivalently, wavelength λ_0) a diagram could be built to show path gain factor variation with the receiving antenna height h_2 and the separation distance of antennas d; this is denoted as the *coverage diagram*.

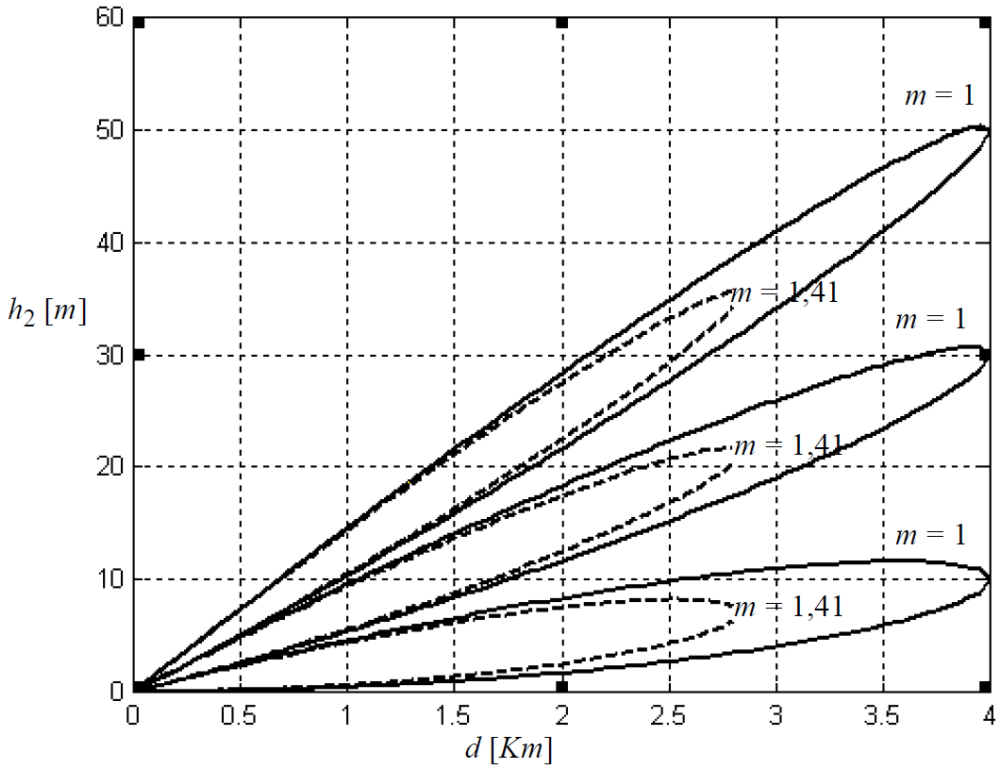


Figure no. 9.2 – *Coverage diagram*

Coverage diagram is usually presented as constant level curves $F = constant$ in (d, h_2) plane. For big values of d, the direct path length is approximated by d: $R_1 \approx d$. Values used for the constant level curves as multiples and submultiples of a reference value representing the field amplitude at an arbitrary distance R_0 from the transmitting antenna. When successive multiples are in a ratio $1/\sqrt{2}$ with each other, the constant level curves differ by 3 dB.

Figure no. 9.2 presents the coverage diagram for $h_1 = 100\lambda_0$ and $R_0 \approx d_0 = 2\ km$; for these values lobes are separated by angles of $\approx 0.3°$.

The *path gain factor in real cases* does not decrease to 0, nor does it increase to 2 because modulus of the reflection coefficient is smaller than 1, resulting in amplitude of the ground reflected wave being always smaller than the one of the direct wave. Moreover, phase of the reflection coefficient is π only for a limited interval of incident angle values.

For vertical polarized waves the reflection coefficient is:

$$\rho e^{j\Phi} = \frac{\left(\varepsilon_r - j\frac{\sigma}{\omega\varepsilon_0}\right)\sin\psi - \sqrt{\left(\varepsilon_r - j\frac{\sigma}{\omega\varepsilon_0}\right) - \cos^2\psi}}{\left(\varepsilon_r - j\frac{\sigma}{\omega\varepsilon_0}\right)\sin\psi + \sqrt{\left(\varepsilon_r - j\frac{\sigma}{\omega\varepsilon_0}\right) - \cos^2\psi}} \tag{9.14}$$

and for horizontal polarized waves it is:

$$\rho e^{j\Phi} = \frac{\sin\psi - \sqrt{\left(\varepsilon_r - j\frac{\sigma}{\omega\varepsilon_0}\right) - \cos^2\psi}}{\sin\psi + \sqrt{\left(\varepsilon_r - j\frac{\sigma}{\omega\varepsilon_0}\right) - \cos^2\psi}} \tag{9.15}$$

Here ε_r and σ are the ground electric relative permittivity and conductivity, respectively, ε_0 is the free-air electric permittivity, and ψ is the incidence angle in the reflection point (between the incident wave propagation direction and the ground plane).

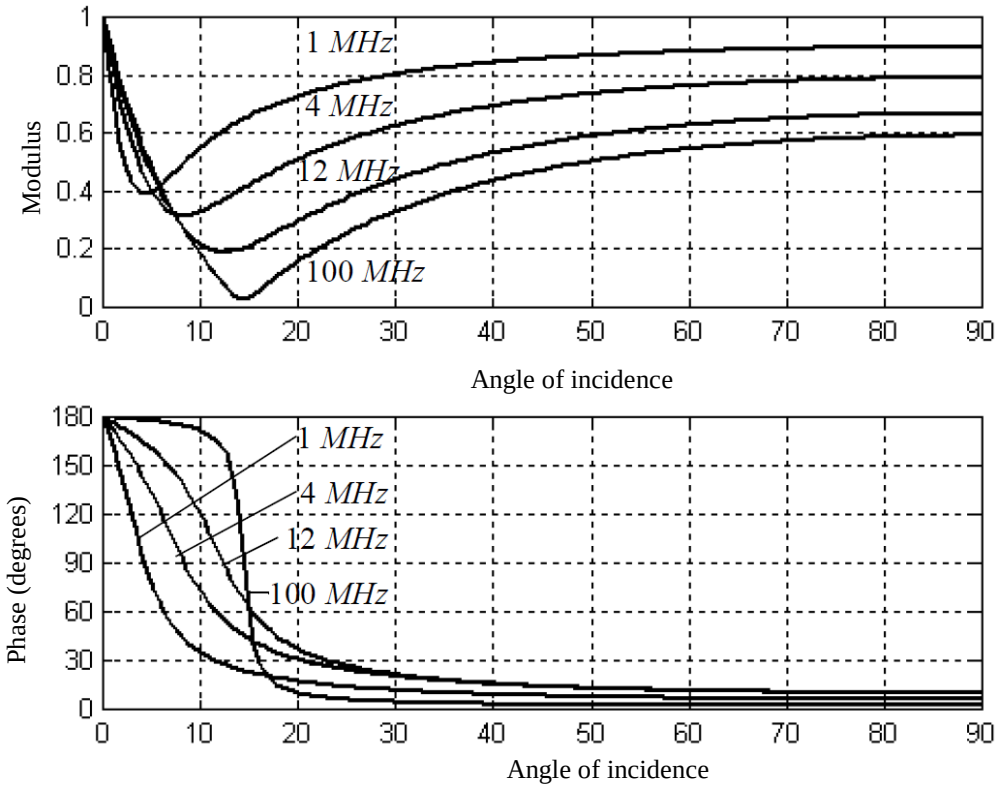


Figure no. 9.3 – *Reflection coefficient for vertical polarized waves*

Figures no. 9.3 and 9.4 display the dependence of the reflection coefficient modulus and phase on the incidence angle and frequency.

The curves for the vertical polarized waves have a minimum (maximum attenuation) of the modulus for an incidence angle of about $10° - 14°$; it is denoted as the *pseudo-Brewster angle*; minimum value is smaller at high frequency. Incident wave phase is reversed after reflection for incidence angles smaller than the pseudo-Brewster angle and it remains unchanged for greater incidence angles.

For horizontal polarized waves the modulus of the reflection coefficient monotonically decreases with the incidence angle, while the phase is reversed around π for all incidence angles and for all displayed frequencies.

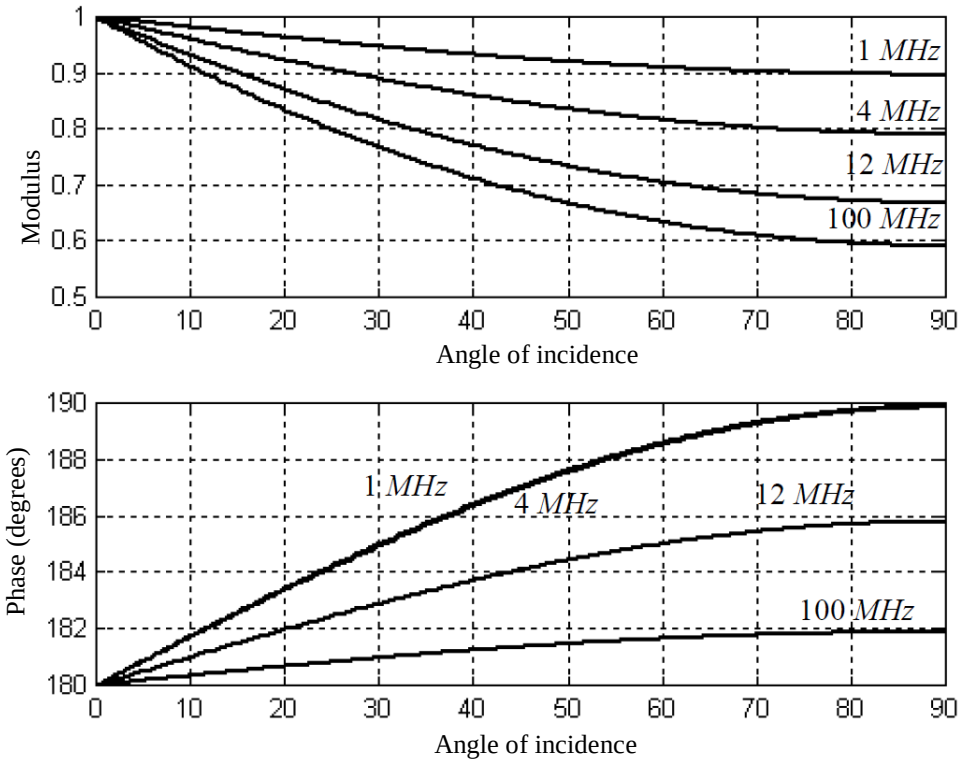


Figure no. 9.4 – *Reflection coefficient for horizontal polarized waves*

9.3 – Electromagnetic Wave Diffraction

Ground surface between communication antennas is far from being plane in real applications. There are numerous obstacles (hills, buildings, trees, etc.) high enough to appear close to the direct wave path, such that they influence the way this wave propagates along the path. They can have sharp vertices that allow for wave diffraction. Part of the diffracted wave reach the receiving point and combine with the direct wave, thus contributing to the total received power.

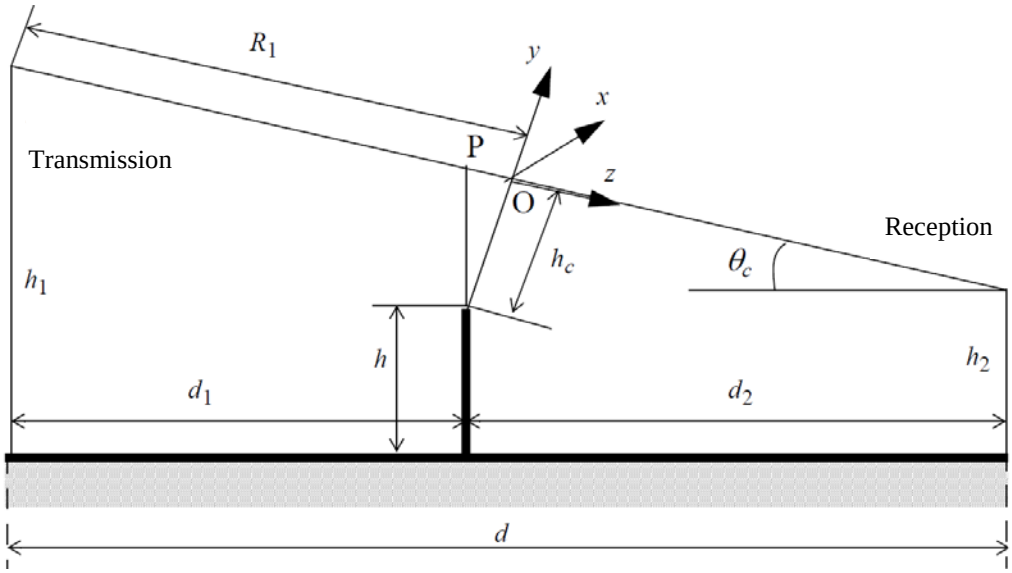


Figure no. 9.5 – Wave d*iffraction at a sharp vertex obstacle*

We consider for study the configuration presented in figure no. 9.5. A sharp obstacle peak is placed at distance h_c from the direct wave path. Based on notations in the figure, applying simple geometry, we can write:

$$h_c = \left(\frac{d_2 h_1 + d_1 h_2}{d} - h\right) \cos\theta_c \tag{9.16}$$

Usually, angle θ_c is very small, so $\cos\theta_c \approx 1$.

Assuming that point O is in the radiation region of transmitting antenna, the wave it creates in this point is a plane wave and, if it has a linear polarization with electric component oriented along Oy of a Cartesian coordinate system centered in O, we can write that:

$$\boldsymbol{E}_O = E_0 \frac{e^{-jk_0 R_1}}{R_1} \hat{\boldsymbol{y}} \tag{9.17}$$

Let's take a rectangular surface centered in O and included in xOy plane. Phase of the field in point Q of this surface differs from the one of the field in point O by a quantity that depends on the distance ρ between these points. As $\rho \ll R_1$, distance R_2 between point Q and transmitting antenna can be approximated by:

$$R_2 = \sqrt{R_1^2 + \rho^2} \approx R_1 + \frac{\rho^2}{2R_1} \tag{9.18}$$

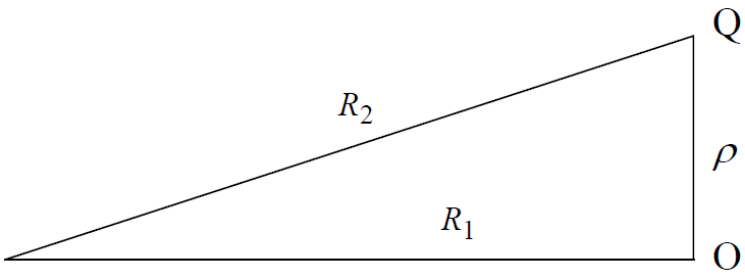


Figure no. 9.6 – *Approximation of distance R_2*

Let's assume that the main lobe of transmitting antenna is Gaussian, that is, antenna gain varies exponentially in the region around its maximum radiation direction:

$$G_\rho = G_O e^{-j\frac{\rho^2}{\alpha^2}} \tag{9.19}$$

where α is constant describing the gain change rate. Thus:

$$\boldsymbol{E}_Q = \frac{E_0}{R_1} e^{-j\frac{\rho^2}{\alpha^2}} e^{-jk_0\left(R_1 + \frac{\rho^2}{2R_1}\right)} \hat{\boldsymbol{y}} \tag{9.20}$$

Considering that the received field is the one radiated by the field distribution on the surface S, we can write it by using the results obtained when studying the radiation of apertures:

$$\boldsymbol{E}(r) = \frac{1}{4\pi^2} \int_S \boldsymbol{f}(k_x, k_y)\, e^{-j(k_x x + k_y y)} dk_x dk_y \tag{9.21}$$

where $\boldsymbol{f}$ is an integration constant whose tangential component in xOy plane is the bi-dimensional Fourier transform of the field distribution on the surface S:

$$\boldsymbol{f_t}(k_x, k_y) = \int_S \boldsymbol{E}_Q(x, y) e^{j(k_x x + k_y y)} dx dy \tag{9.22}$$

After some mathematical manipulations we obtain that:

$$\boldsymbol{E}(r) = \frac{jk_0E_0}{2\pi R_1(R-R_1)} e^{-jk_0R} \sqrt{\frac{\pi}{a}} \int_{-h_c}^{\infty} e^{-jay^2} dy \tag{9.23}$$

where R is the distance between antennas and:

$$a \stackrel{\text{def}}{=} jk_0 \left(\frac{1}{2z} + \frac{1}{2R_1}\right) + \frac{1}{\alpha^2} \tag{9.24}$$

For infinite value of h_c the expression 9.23 is the received field when the obstacle is absent. But:

$$\int_{-\infty}^{\infty} e^{-jay^2} dy = \sqrt{\frac{\pi}{a}} \tag{9.25}$$

so the ratio of the field received when the obstacle is present to the one when the obstacle is absent is:

$$F_d = \sqrt{\frac{a}{\pi}} \int_{-h_c}^{\infty} e^{-jay^2} dy \tag{9.26}$$

and it represents the *attenuation introduced by diffraction*.

The direct wave path remains unobstructed for values of $h_c > 0$; it becomes obstructed when $h_c = 0$ and it remains obstructed for all negative values of h_c (that is, vertex of the obstacle is *above* the line of sight between antennas). For $h_c = 0$, the diffraction attenuation factor is $1/2$, that is $-6\ dB$.

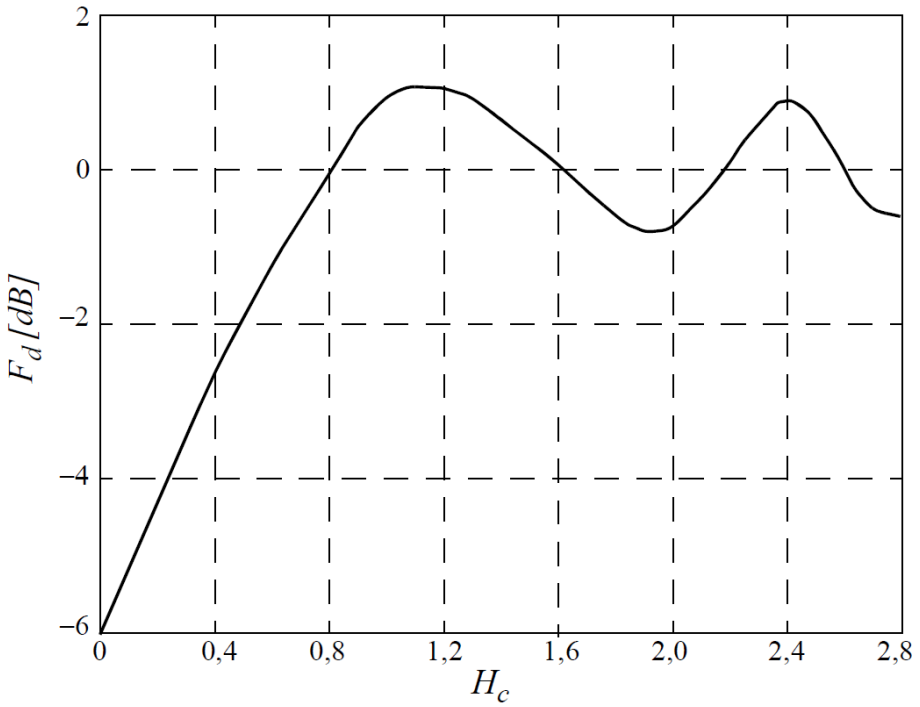


Figure no. 9.7 – *Diffraction attenuation vs. vertex height*

Figure no. 9.7 illustrates the diffraction attenuation factor as a function of an non-dimensional parameter:

$$H_c \stackrel{\text{def}}{=} \sqrt{\frac{2d}{\lambda_0} \frac{1}{d_1 d_2}}\ h_c \tag{9.27}$$

It reveals that the diffraction attenuation is smaller than $1\ dB$ for $H_c > 0.8$ or, equivalently, for:

$$\alpha^2 \gg \frac{\lambda_0 d_1 d_2}{2d} \tag{9.28}$$

which is fulfilled when $d \gg \lambda_0$.

When weather conditions allow for refraction index in the troposphere varying with height, the direct path is no longer a straight line, but a curved one below this line. Choosing antenna heights should take into account these situations, such that appropriate margin exists for H_c to remain greater than 0.8 in all situations.

9.4 – Surface Wave

When transmitting and receiving antennas are close to the ground the reflected wave has about the same level as the direct one, but with opposite phase in most of the cases. The combined wave has a very small level and the communication fails. Successful transmissions occur only at low frequency, where the wavelength is comparable with the distance between Earth surface and ionosphere and a guided-like propagation takes place. This wave is denoted as the *surface wave*.

The above mentioned conditions are met at frequency below 3 MHz, where long and medium wave radio broadcasting, terrestrial navigation systems (Omega, Loran, Decca), and other services are implemented. Wavelength is greater than $100\ m$ and antenna size is too great to allow for their positioning far for the ground. Usually, antennas for these applications are self-radiating masts placed at the ground level.

First study of this type of antennas was done by Sommerfeld in 1909. He took into consideration an electrical dipole (linear infinitesimal unity current element) placed on the Oz axis of a Cartesian coordinate system at distance $h > 0$ from the origin. It is considered that Earth surface is plane and placed in the xOy plane (figure no. 9.8).

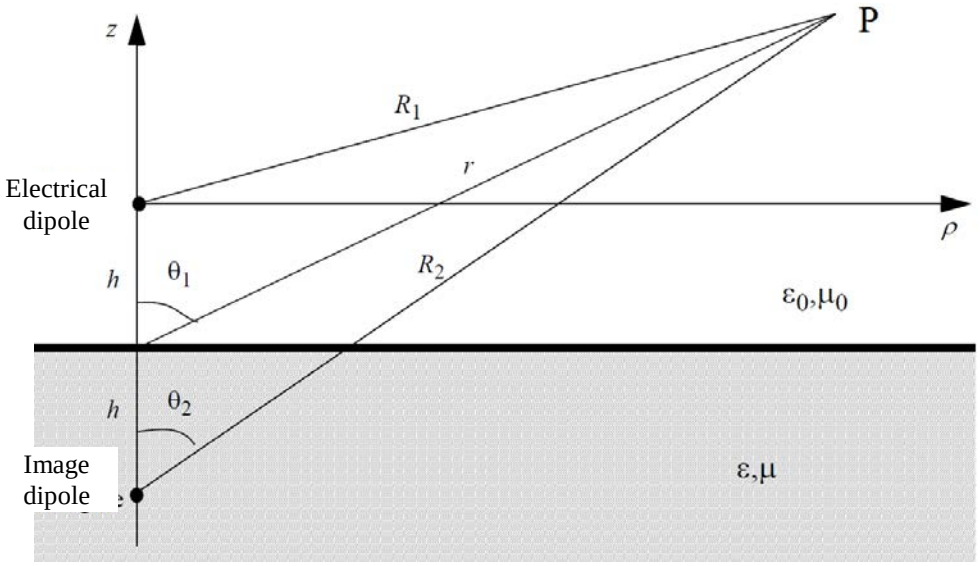


Figure no. 9.8 – *Electrical dipole radiation above a plane surface*

It was shown that vector potential created by the dipole in its radiation region has the modulus:

$$A = \frac{\mu_0}{4\pi}\left(\frac{e^{-jk_0R_1}}{R_1} - \frac{e^{-jk_0R_2}}{R_2} + 2\varepsilon_r\Im\right) \tag{9.29}$$

where ε_r is the ground relative permittivity, $\Im$ is special function in mathematics denoted as the *Zenneck integral*, and the other variables has the meaning showed in the figure.

Near to the ground level ($z = 0$), the first two terms cancel each other; at large distance ρ from the dipole:

$$A \approx \frac{\mu_0 \varepsilon_r}{2\pi} \Im = C \frac{e^{-jk_0 \rho \sqrt{\varepsilon_r + 1}}}{\sqrt{\rho}} \tag{9.30}$$

Zenneck integral explains analytically the presence of the surface wave and large distance wave propagation in such special conditions: the field modulus varies inverse proportionally only to the square root of the distance (not to the distance, as usually).

Accordingly, the electric component of the radiated wave is:

$$E = \frac{jk_0 \eta_0}{4\pi} \frac{e^{-jk_0 \rho}}{\rho} \frac{2(\varepsilon_r - 1)}{\varepsilon_r} a_s \tag{9.31}$$

where a_s is the surface wave attenuation factor; it has a complicated analytical expression that shows that it has a value of about 1 near the dipole and decreases to zero at a few wavelengths farther from the dipole.

9.5 – Ionospheric Propagation

Ionosphere

Ionosphere is the highest part of the atmosphere where a big concentration of ions and free electrons exists. Its height from the Earth surface varies between $90\ km$ and $1000\ km$ and free electron concentration varies between 10^{10} and 10^{12} electrons / m^3. It is generated by the cosmic radiation that ionizes the gas molecules. Ionosphere properties vary with the time of day and with seasons of the year, because the cosmic radiation varies like that. Also, a cycle of about 11 years is present in the ionosphere properties' variation and this is due to Sun activity variation.

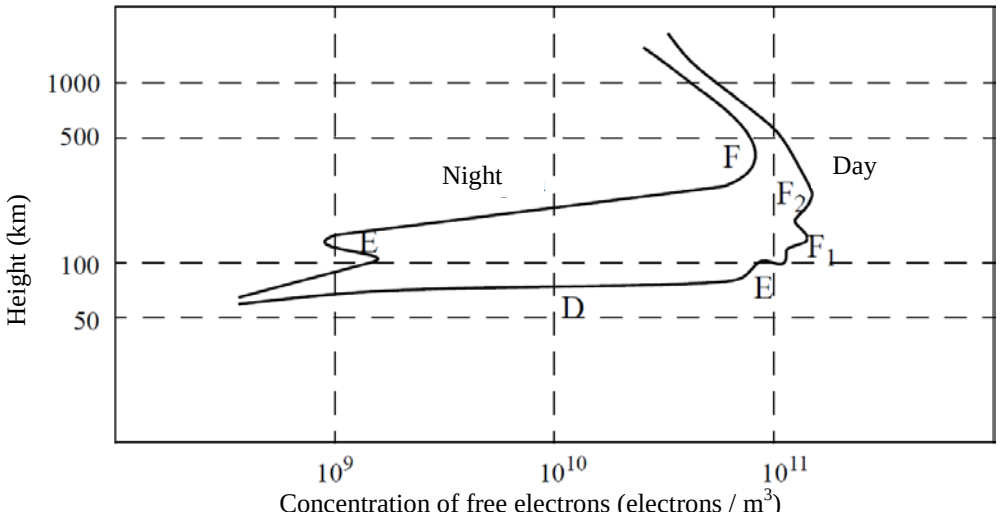


Figure no. 9.9 – *Ionosphere strata*

Ionosphere is not homogeneous. Strata with different free electron concentrations and with variable heights were identified. They are denoted with letters D, E and F. Stratum D could disappear temporarily. Stratum F could split in two sub-strata during the night. Figure no. 9.9 presents typical height and free electron concentration for the ionosphere strata. Also, figure no. 9.10 shows typical variations of ionosphere.

Ionosphere influences the propagation of electromagnetic waves with frequency below $50\ MHz$.

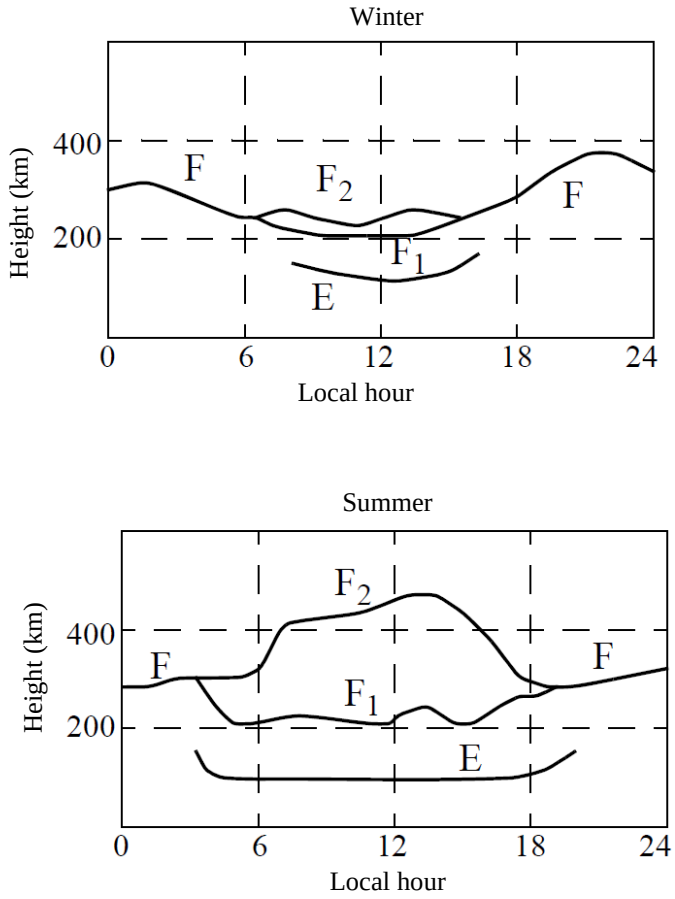


Figure no. 9.10 – *Hourly and seasonally variations of ionosphere*

Electrical Permittivity

Ionosphere contains both ions and free electrons, but its main properties are given by the free electrons as they have a greater mobility (free electron mass is 1800 times smaller than the mass of ions).

Movement of the free electron of mass m and electric charge $-q$ in an electrical field $\boldsymbol{E}$ is governed by the differential following equation:

$$m\frac{d\boldsymbol{v}}{dt} = -q\boldsymbol{E} \tag{9.32}$$

where $\boldsymbol{v}$ is the free electron speed. If $\boldsymbol{E}$ has a harmonic time variation, the free electron speed $\boldsymbol{v}$ varies harmonically in time, too, and the above equation could be rewritten in the Fourier transform domain as:

$$j\omega m\boldsymbol{v} = -q\boldsymbol{E} \tag{9.33}$$

Free electrons' movement generates an electrical current with density:

$$\boldsymbol{J} = -qN\boldsymbol{v} = \frac{q^2N}{j\omega m}\boldsymbol{E} \tag{9.34}$$

where N is the number of free electrons on unit volume (free electron density).

Maxwell's equation can be rewritten as:

$$\nabla \times \boldsymbol{H} = j\omega\varepsilon_0\boldsymbol{E} + \boldsymbol{J} = j\omega\varepsilon_0\boldsymbol{E} + \frac{q^2N}{j\omega m}\boldsymbol{E} = j\omega\varepsilon_0\left(1 - \frac{q^2N}{\omega^2 m\varepsilon_0}\right)\boldsymbol{E} \tag{9.35}$$

The quantity:

$$\omega_p \overset{\text{def}}{=} \sqrt{\frac{q^2 N}{m\varepsilon_0}} \tag{9.36}$$

is denoted as the eigen frequency of plasma oscillations; Maxwell's equation becomes:

$$\nabla \times \boldsymbol{H} = j\omega\varepsilon_0 \left(1 - \frac{\omega_p^2}{\omega^2}\right) \boldsymbol{E} \tag{9.37}$$

By comparing this equation with its standard expression, we note that the ionosphere behaves like a dielectric medium with the relative electrical permittivity:

$$\varepsilon_r = 1 - \frac{\omega_p^2}{\omega^2} = 1 - \frac{q^2 N}{\omega^2 m\varepsilon_0} \tag{9.38}$$

The ionosphere relative electrical permittivity is positive and smaller than 1 for $\omega > \omega_p$ and is negative for $\omega < \omega_p$; in this latter case the propagation constant $k = \omega\sqrt{\varepsilon\mu}$ becomes imaginary and the field modulus decays exponentially with distance: these frequencies are significantly attenuated by the ionosphere.

Virtual Height

Electromagnetic wave entering the ionosphere meets zones with increasing free electron density and, consequently, decreasing electrical permittivity, as it travels towards greater heights. Propagation path bends towards horizontal and, finally, changes direction towards the Earth surface. Total propagation path resembles the one of reflected wave.

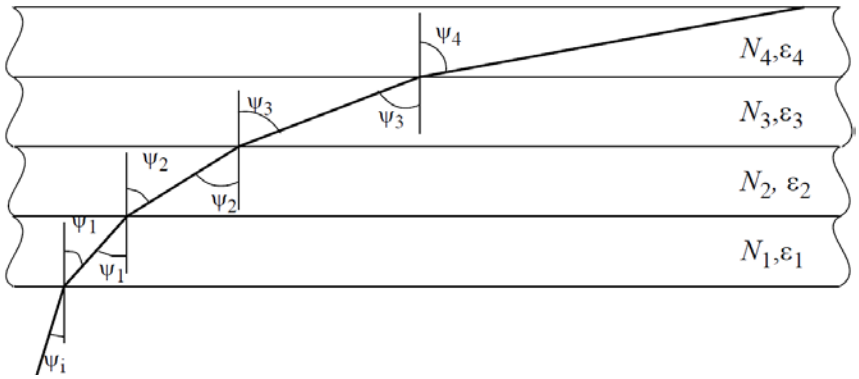


Figure no. 9.11 – *Wave diffraction by ionosphere*

In order to obtain an analytical tool for this phenomenon we assume that the ionosphere composed of successive strata of equal width and constant electrical permittivity (figure no. 9.11). As the free electron density increases with height, electrical permittivity decreases from a stratum to the other. After refraction at the separation surface between strata the wave path changes direction towards horizontal.

Analytically, because:

$$0 < N_1 < N_2 < N_3 < N_4 < \cdots \tag{9.39}$$

then:

$$\varepsilon_0 > \varepsilon_1 > \varepsilon_2 > \varepsilon_3 > \varepsilon_4 > \cdots \tag{9.40}$$

Refraction laws ask for:

$$\sin\psi_i = \sqrt{\varepsilon_{r1}}\sin\psi_1 = \sqrt{\varepsilon_{r2}}\sin\psi_2 = \sqrt{\varepsilon_{r3}}\sin\psi_3 = \sqrt{\varepsilon_{r4}}\sin\psi_4 = \cdots \tag{9.41}$$

so that:

$$\psi_i < \psi_1 < \psi_2 < \psi_3 < \psi_4 < \cdots \tag{9.42}$$

In every refraction point the wave bends farther from vertical, it becomes horizontal in some point (total reflection), and, finally, changes direction towards the Earth surface. The refraction phenomenon is identical for the new direction and the wave leaves the ionosphere under the same angle as it entered.

Regarding the ionosphere as a homogeneous stratum with an equivalent constant free electron concentration equal to the average value of the actual variable one (and, consequently, behaving like a homogeneous medium with average constant electrical permittivity), in order for the total reflection to appear it is mandatory that:

$$1 \cdot \sin\psi_i = \sqrt{\varepsilon_{r,avg}} \sin\frac{\pi}{2} \Longrightarrow \sin\psi_i = \sqrt{\varepsilon_{r,avg}} \qquad (9.43)$$

Using this result in equation 9.38 we obtain the critical value of the free electron density that allows for wave reflection:

$$N_{critic} = \frac{\omega^2 m\varepsilon_0}{q^2}\cos^2\psi_i \qquad (9.44)$$

The above expression could be viewed, also, as putting a limit for the maximum frequency that can be reflected by ionosphere when maximum free electron density is N_{max}:

$$f_{max} = \frac{q}{2\pi\cos\psi_i}\sqrt{\frac{N_{max}}{m\varepsilon_0}} \qquad (9.45)$$

Total wave path resembles the one of the wave reflected by a horizontal surface situated at height h' above the Earth (figure no. 9.12): this is the *ionosphere virtual height*. Virtual height increases with frequency because electrical permittivity increases, too.

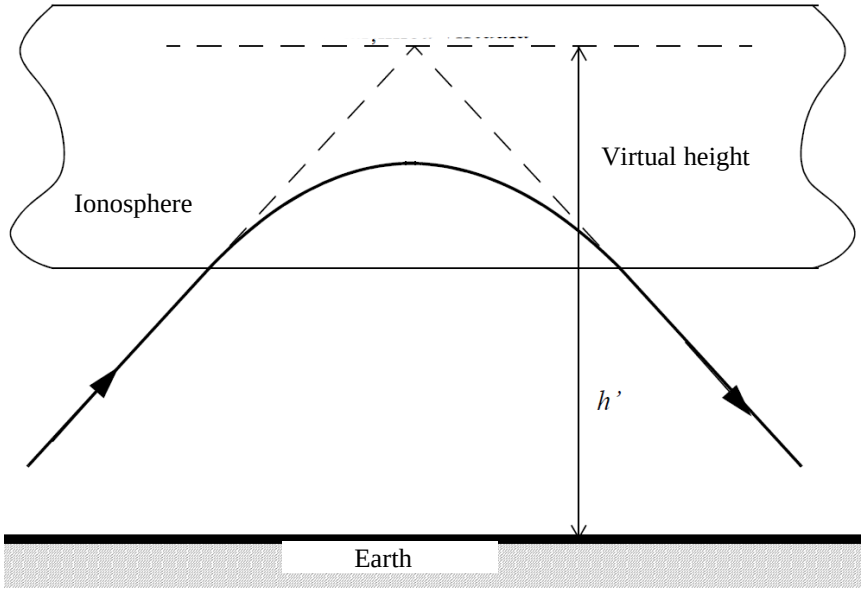


Figure no. 9.12 – *Ionosphere vertical height*

Ionospheric Transmissions

Figure no. 9.13 illustrates the geometry of transmission from point A to point B, both of them situated on the Earth surface. There are denoted in the figure: Earth center O, Earth radius R_p, reflection point in ionosphere P, and ionosphere virtual height h'. For given positions of points A and B and operating frequency, one should compute the needed incidence angle ψ_i and, finally, the transmitting antenna main lobe direction (measured with respect to local horizontal plane).

Applying sinus theorem in triangle OAP we get:

$$\frac{OP}{\sin\Phi} = \frac{OA}{\sin\psi_i} \tag{9.46}$$

Based on notations in the figure and observing that $\Phi = \pi - \theta - \psi_i$, the above equation can be rewritten as:

$$\frac{R_p + h'}{\sin(\theta+\psi_i)} = \frac{R_p}{\sin\psi_i} \tag{9.47}$$

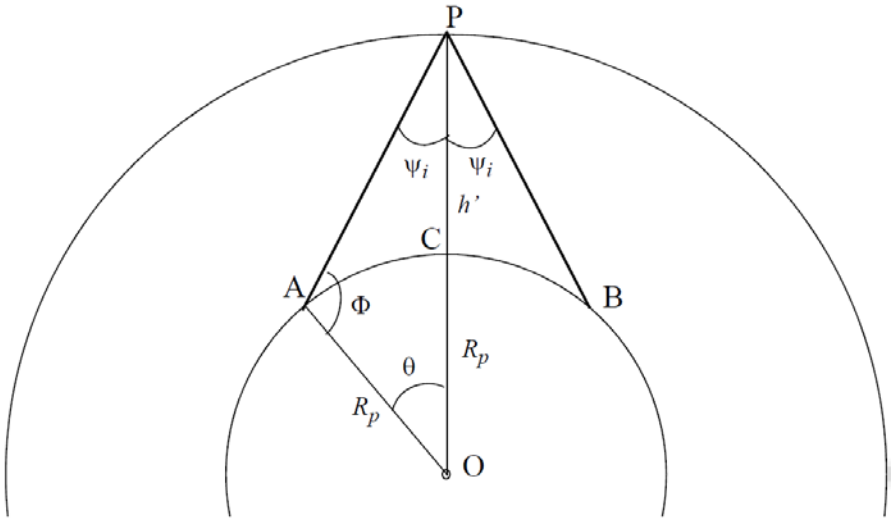


Figure no. 9.13 – *Ionospheric transmission*

But, if d is the length of the circle arc AB, then:

$$\theta = \frac{d/2}{R_p} \tag{9.48}$$

and equation 9.47 can be rewritten as:

$$1 + \frac{h'}{R_p} = \sin\left(\frac{d}{2R_p}\right)\cot\psi_i + \cos\psi_i \tag{9.49}$$

The above equation allows for computing the incidence angle ψ_i.

The transmitting antenna main lobe direction with respect to local horizontal plane is:

$$\Phi' = \Phi - \frac{\pi}{2} = \frac{\pi}{2} - \psi_i - \frac{d}{2R_p} \tag{9.50}$$

Knowing the strong diurnal and seasonal variability of the ionosphere parameters, its virtual height varies, too and, in order to maintain transmission between the same points A and B, transmitting antenna orientation or operating frequency should be modified accordingly.

Faraday Rotation

Free electron movement in the ionosphere is influenced by the Earth magnetic field. If $\boldsymbol{v_b}$ is the free electron speed component perpendicular to the Earth magnetic field $\boldsymbol{B_0}$, then the force applied to electrons is:

$$\boldsymbol{F_B} = \boldsymbol{v_b} \times \boldsymbol{B_0} \tag{9.51}$$

This force induces a rotation component on the electron movement modifying the polarization state of the electromagnetic wave; thus, the reflected wave polarization state

differs from the one of the incident wave. This phenomenon is denoted as the *Faraday rotation* and it becomes significant at frequencies around the so denoted *cyclotron frequency*:

$$f_c = \frac{qB_0}{2\pi m} \tag{9.52}$$

One important consequence of the Faraday rotation is supplementary loss in received power due to polarization mismatching between receiving wave and antenna.

9.6 – Microwave Propagation

Electromagnetic waves at frequency greater than $1\ GHz$ are denoted as microwaves. Their frequency is greater than the plasma eigen frequency ω_p and are not reflected by the ionosphere. Also, their ground reflection is diffuse (not specular), because their wavelength is comparable with the ground irregularities, and the reflected wave is practically zero. So, power transmission is due solely to the direct wave. Total path loss is sum of partial path losses due the following phenomena:

- *rain*. Water drops are almost spherical and have a good conductibility, so they strongly reflect incident electromagnetic waves and contribute significantly at diminishing received power. Path loss is characterized by the specific path loss on the unity length of the path across the rainy area: $L = aR^b$, R – rain intensity in mm/h, a, b – constants depending on frequency and rain temperature. For a moderate rain intensity (about $5\ mm/h$), path loss is about $0.074\ dB/km$ at $1\ GHz$, $0.85\ dB/km$ at $10\ GHz$, and $3.42\ dB/km$ at $100\ GHz$.
- *fog*. Water drops are smaller and denser than those associated with rain and path loss is greater: about $1\ dB/km$ at $1\ GHz$, for instance.
- *ice and snow*. Physical phenomenon is similar to the ones in the previous case, but path loss is smaller, because the imaginary part of the electrical permittivity is smaller for dry ice and snow. For wet snow or a mixture of ice and water (hail) the path loss is equal or even greater than the one associated with rain.
- *gases*. Different gases with dipolar molecules (oxygen, water vapors, etc.) that are present in the atmosphere attenuate selectively electromagnetic waves (gas absorption), because the wave aligns the dipoles with its magnetic component and it consumes some power for this. Absorption is greater at resonance frequency (gas spectral line): $13.5\ mm$ and $1.67\ mm$ wavelength for water vapors, $5\ mm$ and $2.5\ mm$ wavelength for oxygen, etc. Specific path loss at the corresponding frequencies can be as big as $10\ dB/km$.

9.7 – Fading

Fading slow-rate big changes of received power density value. It is a phenomenon specific for ionospheric transmissions, but it could appear for microwave communications when weather conditions fluctuate. Large changes of received power mean that received power is below nominal input power of the receiving equipment or signal-to-noise ratio smaller than the threshold for some time intervals, causing bad quality or impossible reception. Fading intensity is appreciated through the percent of non-reception time. Design of communication systems should take into account the worst case; operating frequency, transmitting power, transmitting and receiving antenna placement should be chosen such that the percent of non-receiving time be smaller than a given threshold (0.01 %, for instance).

One of the most used protective measures against fading is *diversity*. This means creating more than one independent *versions* of the desired signal at the receiving point. By means of combining them, the percent of non-receiving time can be significantly reduced.

Based on the technique used to create these versions, diversity includes:

- *frequency diversity*. Useful information is simultaneously transmitted by using two or more frequencies. Because path loss is dependent on frequency, the probability that all the versions be simultaneously smaller than the receiver nominal power decreases. The method is costly because we need supplementary transmitters and antennas; moreover, spectrum efficiency is reduced.
- *spatial diversity*. Useful signal is simultaneously received in two or more points separated by more than a wavelength. Received signals differ because they travel on different paths from transmitter to receiving points.
- *time diversity*. Useful information transmission is repeated periodically. Signals received at different moments differ because propagation conditions change in time.
- *polarization diversity*. Useful signal is transmitted with different polarization waves; antennas with matched polarization properties are used for reception. Signals differ because path loss depends on wave polarization.
- *angle diversity*. Signals are transmitted by much different paths by using directive antennas. Received signals differ significantly from each other because they travel on completely different paths, but the technique involves great costs.

Based on the method used to obtain the receiver input signal from different received versions, diversity can be with:

- *switched combining*. Receiver scans versions until one with signal-to-ratio greater than a given threshold is found and it keeps receiving it as long as this condition is fulfilled. When the condition does not hold anymore, receiver switches to another acceptable version.
- *selection combining*. Receiver scans versions periodically and it selects each time the version with maximum signal-to-noise ratio.
- *equal-gain combining.* Receiver uses the sum of all versions. Improvement in the signal-to-noise ratio is possible because useful signal components of different versions are somehow correlated as they come from the same transmitter, while the noise components are completely uncorrelated. Thus, the useful signal component is enhanced by summation, while the noise component is not.
- *maximal-ratio (optimal) combining*. Versions are summed after weighting based on their measured sign-to-noise ratios. Sum signal has signal-to-noise ratio equal to sum of sign-to-noise ratios of all the versions (the maximum possible value !).