

OBJECTIVES:

- To enable the student to understand the basic principles in antenna and microwave system design
- To enhance the student knowledge in the area of various antenna designs.
- To enhance the student knowledge in the area of microwave components and antenna for practical applications.

UNIT I INTRODUCTION TO MICROWAVE SYSTEMS AND ANTENNAS**9**

Microwave frequency bands, Physical concept of radiation, Near- and far-field regions, Fields and Power Radiated by an Antenna, Antenna Pattern Characteristics, Antenna Gain and Efficiency, Aperture Efficiency and Effective Area, Antenna Noise Temperature and G/T, Impedance matching, Friis transmission equation, Link budget and link margin, Noise Characterization of a microwave receiver.

UNIT II RADIATION MECHANISMS AND DESIGN ASPECTS**9**

Radiation Mechanisms of Linear Wire and Loop antennas, Aperture antennas, Reflector antennas, Microstrip antennas and Frequency independent antennas, Design considerations and applications.

UNIT III ANTENNA ARRAYS AND APPLICATIONS**9**

Two-element array, Array factor, Pattern multiplication, Uniformly spaced arrays with uniform and non-uniform excitation amplitudes, Smart antennas.

UNIT IV PASSIVE AND ACTIVE MICROWAVE DEVICES**9**

Microwave Passive components: Directional Coupler, Power Divider, Magic Tee, attenuator, resonator, Principles of Microwave Semiconductor Devices: Gunn Diodes, IMPATT diodes, Schottky Barrier diodes, PIN diodes, Microwave tubes: Klystron, TWT, Magnetron.

UNIT V MICROWAVE DESIGN PRINCIPLES**9**

Impedance transformation, Impedance Matching, Microwave Filter Design, RF and Microwave Amplifier Design, Microwave Power amplifier Design, Low Noise Amplifier Design, Microwave Mixer Design, Microwave Oscillator Design

TOTAL: 45 PERIODS**OUTCOMES:**

The student should be able to:

- Apply the basic principles and evaluate antenna parameters and link power budgets
- Design and assess the performance of various antennas
- Design a microwave system given the application specifications

TEXTBOOKS:

1. John D Krauss, Ronald J Marhefka and Ahmad S. Khan, "Antennas and Wave Propagation: Fourth Edition, Tata McGraw-Hill, 2006. (UNIT I, II, III)
2. David M. Pozar, "Microwave Engineering", Fourth Edition, Wiley India, 2012. (UNIT I, IV, V)

REFERENCES:

1. Constantine A. Balanis, —Antenna Theory Analysis and Design||, Third edition, John Wiley India Pvt Ltd., 2005.
2. R.E. Collin, "Foundations for Microwave Engineering", Second edition, IEEE Press, 2001

UNIT I INTRODUCTION TO MICROWAVE SYSTEMS AND ANTENNAS

Electromagnetic spectrum:

Electromagnetic spectrum

Microwaves occupy a place in the electromagnetic spectrum with frequency above ordinary radio waves, and below infrared light:

Electromagnetic spectrum			
Name	Wavelength	Frequency (Hz)	Photon energy (eV)
Gamma ray	< 0.02 nm	> 15 EHz	> 62.1 keV
X-ray	0.01 nm – 10 nm	30 EHz – 30 PHz	124 keV – 124 eV
Ultraviolet	10 nm – 400 nm	30 PHz – 750 THz	124 eV – 3 eV
Visible light	390 nm – 750 nm	770 THz – 400 THz	3.2 eV – 1.7 eV
Infrared	750 nm – 1 mm	400 THz – 300 GHz	1.7 eV – 1.24 meV
Microwave	1 mm – 1 m	300 GHz – 300 MHz	1.24 meV – 1.24 μ eV
Radio	1 m – 100 km	300 MHz – 3 kHz	1.24 μ eV – 12.4 feV

Microwave Frequency Bands:

Microwaves Frequency Bands	
Band	Frequency range
HF Band	3 to 30 MHz
VHF Band	30 to 300 MHz
UHF Band	300 to 1000 MHz
L Band	1 to 2 GHz
S Band	2 to 4 GHz
C Band	4 to 8 GHz
X Band	8 to 12 GHz
Ku Band	12 to 18 GHz
K Band	18 to 27 GHz
Ka Band	27 to 40 GHz
V Band	40 to 75 GHz
W Band	75 to 110 GHz
mm Band	110 to 300 GHz

Physical Concept of Radiation (Radiation Mechanism)

One of the first questions that may be asked concerning antennas would be “how is radiation accomplished?”

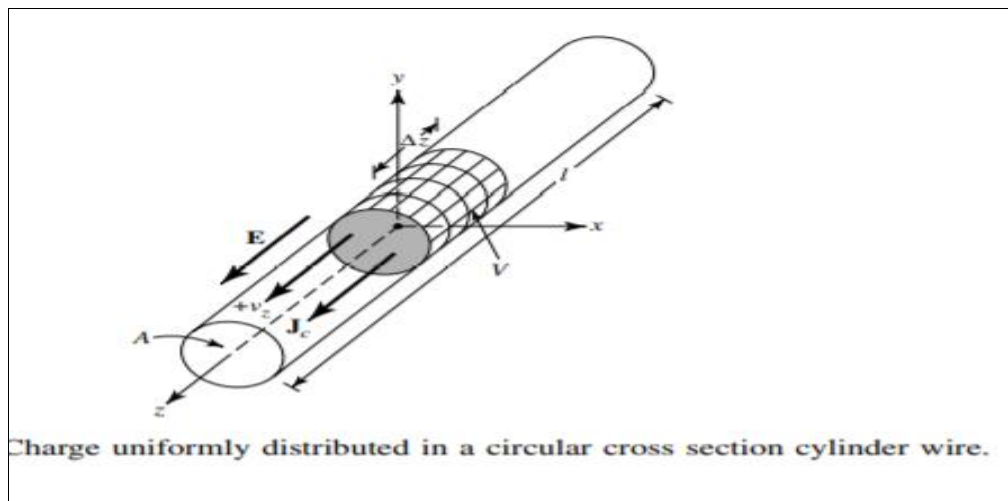
In other words, how are the electromagnetic fields generated by the source, contained and guided within the transmission line and antenna, and finally “detached” from the antenna to form a free-space wave?

Let us first examine some basic sources of radiation.

Radiation from Single Wire

Conducting wires are material whose prominent characteristic is the motion of electric charges and the creation of current flow.

Let us assume that an electric volume charge density, represented by q_v (coulombs/m³), is distributed uniformly in a circular wire of cross-sectional area A and volume V , as shown in Figure.



The total charge Q within volume V is moving in the z direction with a uniform velocity v_z (meters/sec). It can be shown that the current density J_z (amperes/m²) over the cross section of the wire is given by

$$J_z = q_v v_z \quad (1a)$$

If the wire is made of an ideal electric conductor, the current density J_s (amperes/m) resides on the surface of the wire and it is given by

$$J_s = q_s v_z \quad (1b)$$

where q_s (coulombs/m²) is the surface charge density.

If the wire is very thin (ideally zero radius), then the current in the wire can be represented by

$$I_z = q_l v_z \quad (1c)$$

where q_l (coulombs/m) is the charge per unit length. Instead of examining all three current densities, we will primarily concentrate on the very thin wire. The conclusions apply to all three.

If the current is time varying, then the derivative of the current of (1c) can be written as

$$dI_z / dt = q_l dv_z / dt = q_l a_z \quad (2)$$

where $dv_z / dt = a_z$ (meters/sec²) is the acceleration. If the wire is of length l , then (2) can be written as

$$l dI_z / dt = l q_l dv_z / dt = l q_l a_z \quad (3)$$

Equation (3) is the basic relation between current and charge, and it also serves as the fundamental relation of electromagnetic radiation.

It simply states that to create radiation, there must be a time-varying current or an acceleration (or deceleration) of charge. We usually refer to currents in time-harmonic applications while charge is most often mentioned in transients. To create charge acceleration (or deceleration) the wire must be curved, bent, discontinuous, or terminated. Periodic charge acceleration (or deceleration) or time-varying current is also created when charge is oscillating in a time-harmonic motion.

Important Conclusions:

(i) If a charge is not moving, current is not created and there is no radiation.

(ii) If charge is moving with a uniform velocity:

(a) There is no radiation if the wire is straight, and infinite in extent.

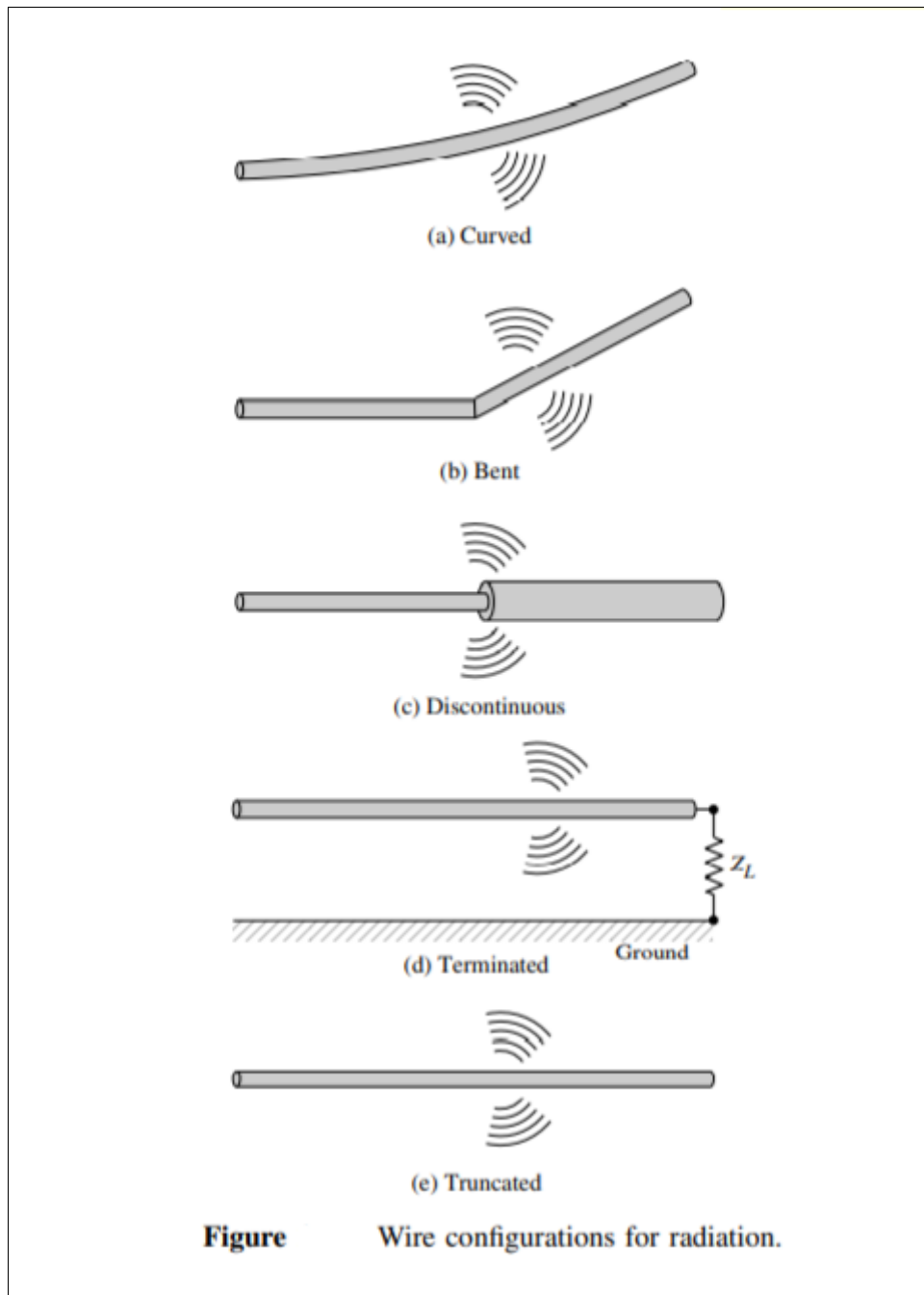
(b) There is radiation if the wire is curved, bent, discontinuous, terminated, or truncated, as shown in Figure

(iii) If charge is oscillating in a time-motion, it radiates even if the wire is straight.

A qualitative understanding of the radiation mechanism may be obtained by considering a pulse source attached to an open-ended conducting wire, which may be connected to the ground through a discrete load at its open end, as shown in Figure (d).

When the wire is initially energized, the charges (free electrons) in the wire are set in motion by the electrical lines of force created by the source. When charges are accelerated in the source-end of the wire and decelerated (negative acceleration with respect to original motion) during reflection from its end, it is suggested that radiated fields are produced at each end and along the remaining part of the wire.

Stronger radiation with a more broad frequency spectrum occurs if the pulses are of shorter or more compact duration while continuous time-harmonic oscillating charge produces, ideally, radiation of single frequency determined by the frequency of oscillation.



The acceleration of the charges is accomplished by the external source in which forces set the charges in motion and produce the associated field radiated. The deceleration of the charges at the end of the wire is accomplished by the internal (self) forces associated with the induced field due to the build up of charge concentration at the ends of the wire. The internal forces receive energy from the charge build up as its velocity is reduced to zero at the ends of the wire. **Therefore, charge acceleration due to an exciting electric field and deceleration due to impedance discontinuities or smooth curves of the wire are mechanisms responsible for electromagnetic radiation.**

Radiation from Two-Wires

Let us consider a voltage source connected to a two-conductor transmission line which is connected to an antenna. This is shown in Figure. Applying a voltage across the two-conductor transmission line creates an electric field between the conductors. The electric field has associated with it electric lines of force which are tangent to the electric field at each point and their strength is proportional to the electric field intensity. The electric lines of force have a tendency to act on the free electrons (easily detachable from the atoms) associated with each conductor and force them to be displaced. The movement of the charges creates a current that in turn creates a magnetic field intensity. Associated with the magnetic field intensity are magnetic lines of force which are tangent to the magnetic field.

We have accepted that electric field lines start on positive charges and end on negative charges. They also can start on a positive charge and end at infinity, start at infinity and end on a negative charge, or form closed loops neither starting or ending on any charge. Magnetic field lines always form closed loops encircling current-carrying conductors because physically there are no magnetic charges.

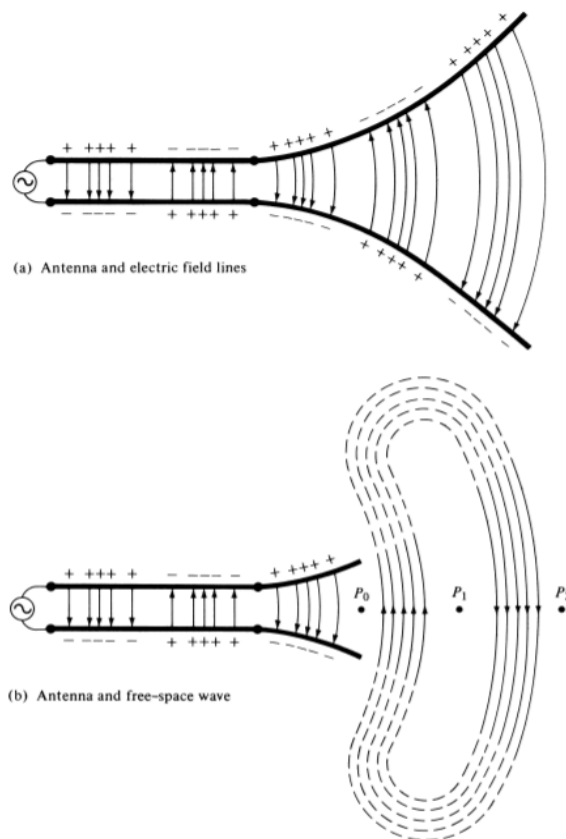


Figure Source, transmission line, antenna, and detachment of electric field lines.

The electric field lines drawn between the two conductors help to exhibit the distribution of charge. If we assume that the voltage source is sinusoidal, we expect the electric field between the conductors to also be sinusoidal with a period equal to that of the applied source. The relative magnitude of the electric field intensity is indicated by the density (bunching) of

the lines of force with the arrows showing the relative direction (positive or negative). The creation of time-varying electric and magnetic fields between the conductors forms electromagnetic waves which travel along the transmission line, as shown in Figure (a).

The electromagnetic waves enter the antenna and have associated with them electric charges and corresponding currents. If we remove part of the antenna structure, as shown in Figure (b), free-space waves can be formed by “connecting” the open ends of the electric lines (shown dashed).

The free-space waves are also periodic but a constant phase point P_0 moves outwardly with the speed of light and travels a distance of $\lambda/2$ (to P_1) in the time of one-half of a period. It has been shown that close to the antenna the constant phase point P_0 moves faster than the speed of light but approaches the speed of light at points far away from the antenna (analogous to phase velocity inside a rectangular waveguide).

free-space waves and water waves -analogy

The question still unanswered is how the guided waves are detached from the antenna to create the free-space waves that are indicated as closed loops. Before we attempt to explain that, let us draw a parallel between the guided and free-space waves, and water waves created by the dropping of a pebble in a calm body of water or initiated in some other manner.

Once the disturbance in the water has been initiated, water waves are created which begin to travel outwardly. If the disturbance has been removed, the waves do not stop or extinguish themselves but continue their course of travel. If the disturbance persists, new waves are continuously created which lag in their travel behind the others. The same is true with the electromagnetic waves created by an electric disturbance.

If the initial electric disturbance by the source is of a short duration, the created electromagnetic waves travel inside the transmission line, then into the antenna, and finally are radiated as free-space waves, even if the electric source has ceased to exist (as was with the water waves and their generating disturbance). If the electric disturbance is of a continuous nature, electromagnetic waves exist continuously and follow in their travel behind the others. This is shown in Figure for a biconical antenna. When the electromagnetic waves are within the transmission line and antenna, their existence is associated with the presence of the charges inside the conductors. However, when the waves are radiated, they form closed loops and there are no charges to sustain their existence. This leads us to conclude that electric charges are required to excite the fields but are not needed to sustain them and may exist in their absence. This is in direct analogy the water waves.

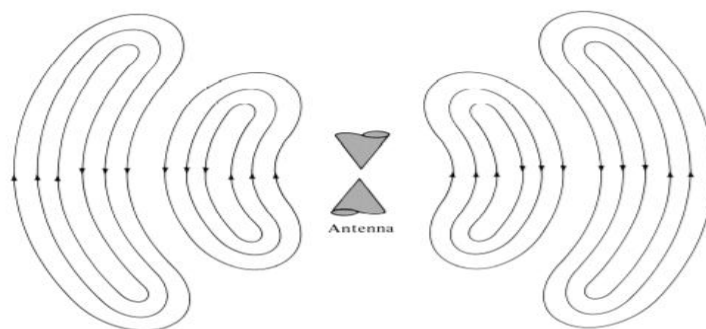


Figure 1.13 Electric field lines of free-space wave for biconical antenna.

Isotropic, Directional, and Omnidirectional Patterns

An isotropic radiator is defined as “a hypothetical lossless antenna having equal radiation in all directions.” Although it is ideal and not physically realizable, it is often taken as a reference for expressing the directive properties of actual antennas.

A directional antenna is one “having the property of radiating or receiving electromagnetic waves more effectively in some directions than in others”. Example of antenna with directional radiation patterns is shown in Figure.

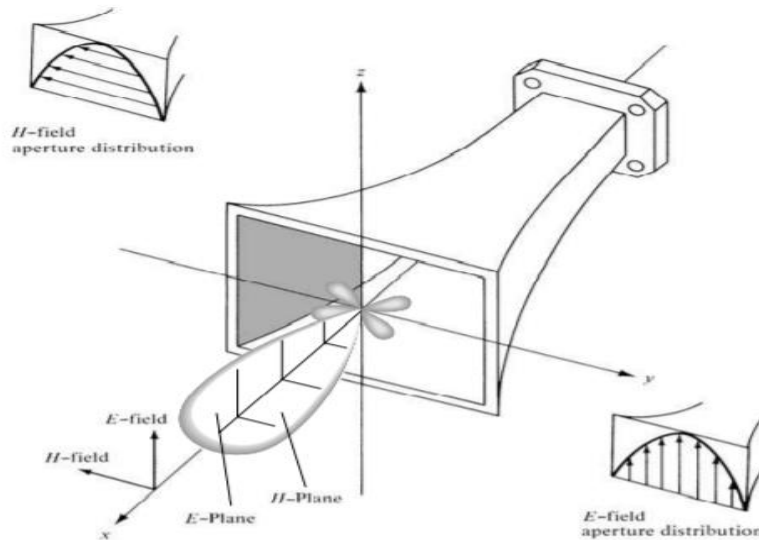


Figure . Principal *E*- and *H*-plane patterns for a pyramidal horn antenna.

It is seen that the pattern in Figure below is nondirectional in the azimuth plane [$f(\varphi)$, $\theta = \pi/2$] and directional in the elevation plane [$g(\theta)$, $\varphi = \text{constant}$]. This type of a pattern is designated as omnidirectional, and it is defined as one “having an essentially nondirectional pattern in a given plane (in this case in azimuth) and a directional pattern in any orthogonal plane (in this case in elevation).” An omnidirectional pattern is then a special type of a directional pattern.

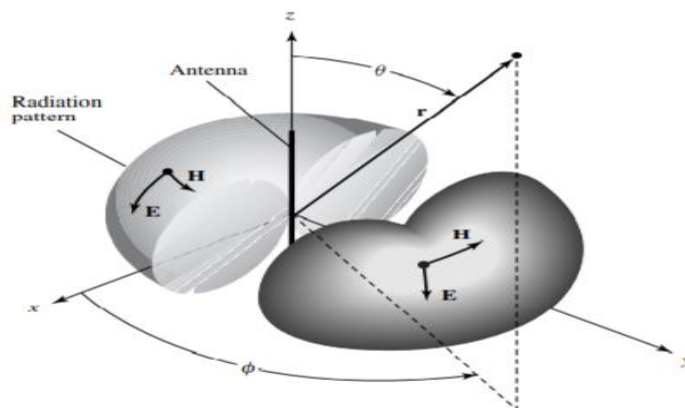


Figure . Omnidirectional antenna pattern.

Antenna near and far field(Field Regions of Antenna)

The space surrounding an antenna is usually subdivided into three regions:

- (a) reactive near-field,
- (b) radiating near-field (Fresnel) and
- (c) far-field (Fraunhofer) regions as shown in Figure .

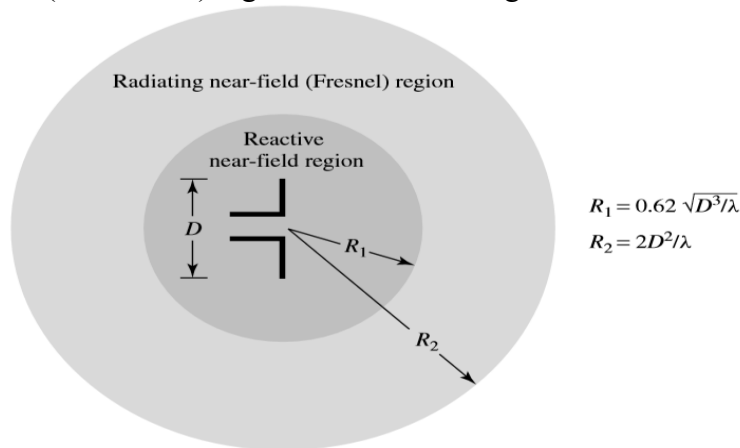


Figure 1.1 Field regions of an antenna.

These regions are so designated to identify the field structure in each. Although no abrupt changes in the field configurations are noted as the boundaries are crossed, there are distinct differences among them.

Reactive near-field region

Reactive near-field region is defined as “that portion of the near-field region immediately surrounding the antenna wherein the reactive field predominates.” For most antennas, the outer boundary of this region is commonly taken to exist at a distance $R < 0.62 \sqrt{D^3/\lambda}$ from the antenna surface, where λ is the wavelength and D is the largest dimension of the antenna. “For a very short dipole, or equivalent radiator, the outer boundary is commonly taken to exist at a distance $\lambda/2\pi$ from the antenna surface.”

Radiating near-field (Fresnel) region

Radiating near-field (Fresnel) region is defined as “that region of the field of an antenna between the reactive near-field region and the far-field region wherein radiation fields predominate and wherein the angular field distribution is dependent upon the distance from the antenna.

If the antenna has a maximum dimension that is not large compared to the wavelength, this region may not exist. For an antenna focused at infinity, the radiating near-field region is sometimes referred to as the Fresnel region on the basis of analogy to optical terminology. If the antenna has a maximum overall dimension which is very small compared to the wavelength, this field region may not exist.” The inner boundary is taken to be the distance $R \geq 0.62 \sqrt{D^3/\lambda}$ and the outer boundary the distance $R < 2D^2/\lambda$ where D is the largest* dimension of the antenna. This criterion is based on a maximum phase error of $\pi/8$. In this region the field pattern is, in general, a function of the radial distance and the radial field component may be appreciable.

*To be valid, D must also be large compared to the wavelength ($D > \lambda$)

Far-field (Fraunhofer) region

Far-field (Fraunhofer) region is defined as “that region of the field of an antenna where the angular field distribution is essentially independent of the distance from the antenna. If the antenna has a maximum* overall dimension D , the far-field region is commonly taken to exist at distances greater than $2D^2/\lambda$ from the antenna, λ being the wavelength.

The far-field patterns of certain antennas, such as multibeam reflector antennas, are sensitive to variations in phase over their apertures. For these antennas $2D^2/\lambda$ may be inadequate. In physical media, if the antenna has a maximum overall dimension, D , which is large compared to $\pi/|\gamma|$, the far-field region can be taken to begin approximately at a distance equal to $|\gamma|D^2/\pi$ from the antenna, γ being the propagation constant in the medium.

For an antenna focused at infinity, the far-field region is sometimes referred to as the Fraunhofer region on the basis of analogy to optical terminology.” In this region, the field components are essentially transverse and the angular distribution is independent of the radial distance where the measurements are made. The inner boundary is taken to be the radial distance $R = 2D^2/\lambda$ and the outer one at infinity.

The amplitude pattern of an antenna in different regions

The amplitude pattern of an antenna, as the observation distance is varied from the reactive near field to the far field, changes in shape because of variations of the fields, both magnitude and phase.

A typical progression of the shape of an antenna, with the largest dimension D , is shown in Figure. It is apparent that in the reactive near field region the pattern is more spread out and nearly uniform, with slight variations. As the observation is moved to the radiating near-field region(Fresnel), the pattern begins to smooth and form lobes. In the far-field region (Fraunhofer), the pattern is well formed, usually consisting of few minor lobes and one, or more, major lobes.

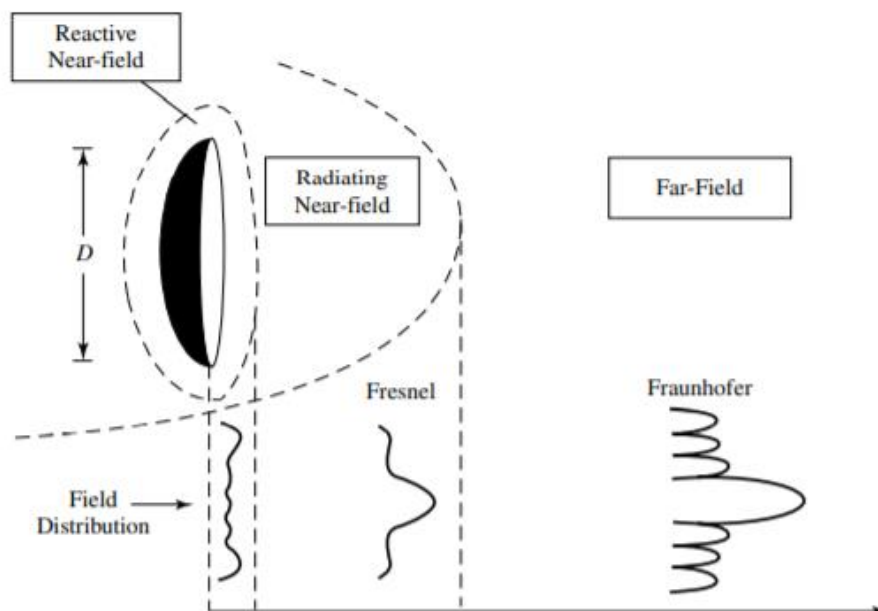


Figure: Typical changes of antenna amplitude pattern shape from reactive near field toward the far field

Antenna Parameters

(a) Radiation pattern.

The radiation pattern of an antenna is a plot of the magnitude of the far-zone field strength versus position around the antenna, at a fixed distance from the antenna.

Thus the radiation pattern can be plotted from the pattern function $F_\theta(\theta, \varphi)$ or $F_\varphi(\theta, \varphi)$, versus either the angle θ (for an elevation plane pattern) or the angle φ (for an azimuthal plane pattern). The choice of plotting either F_θ or F_φ is dependent on the polarization of the antenna.

(b) main lobe, side lobe, minor lobe and back lobe with reference to antenna radiation pattern.

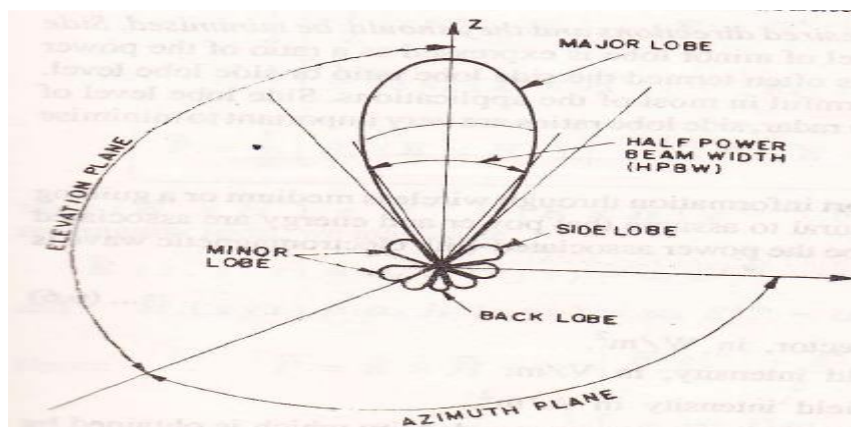
Major Lobe: Major lobe is also called as main beam and is defined as “the radiation lobe containing the direction of maximum radiation”. In some antennas, there may be more than one major lobe.

Minor lobe: All the lobes except the major lobes are called minor lobe.

Side lobe: A side lobe is adjacent to the main lobe.

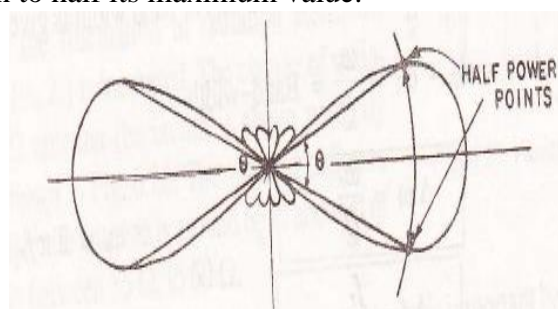
Back lobe: Normally refers to a minor lobe that occupies the hemisphere in a direction opposite to that of the major(main) lobe .

- Minor lobes normally represents radiation in undesired directions and they should be minimized.



(c) Half Power Beam Width (HPBW) of an antenna.

Half Power Beam Width is a measure of directivity of an antenna. It is an angular width in degrees, measured on the radiation pattern (main lobe) between points where the radiated power has fallen to half its maximum value.



(d) beam solid angle

The beam area or beam solid angle Ω_A for antenna is given by integral of the normalized power pattern over a sphere.

$$\Omega_A = \int_0^{2\pi} \int_0^{\pi} P_n(\theta, \phi) d\Omega \quad \text{steradian}$$

$P_n(\theta, \phi) = \text{Normalized power pattern}$

Beam solid angle is also given approximately by

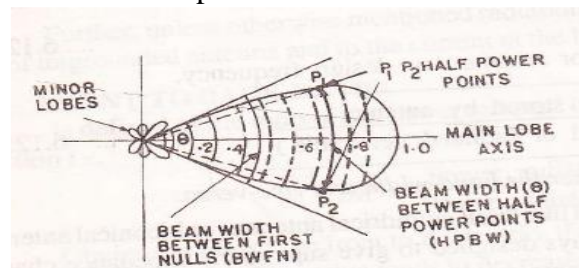
$$\Omega_A = \theta_{HP} \phi_{HP} \quad \text{steradian}$$

$\theta_{HP} = \text{HPBW in } E\text{-plane or } \theta \text{ plane}$

$\phi_{HP} = \text{HPBW in } H\text{-plane or } \phi \text{ plane}$

(e) Beam Width between First Null

Beam width between first null (BWFN) is the angular width in degrees, measured on the radiation pattern between first null points on either side of the main lobe.



(f) Radiation Intensity

Radiation Intensity $U(\theta, \phi)$ in given direction is defined as the power per unit solid angle in that direction.

- The power radiated per unit area in any direction is given by pointing vector P .
- For distant field for which E and H are orthogonal in a plane normal to the radius vector,

$$\text{The power flow per unit area is given by } P = \frac{E^2}{\eta_v} \text{ watts / sqm}$$

- There are r^2 square meters of surface area per unit solid angle (or steradian).
- $U(\theta, \phi) = r^2 P = \frac{r^2 E^2}{\eta_v} \text{ watts/unit solid angle}$

The radiation intensity gives the variation in radiated power versus position around the antenna. We can find the total power radiated by the antenna by integrating the Poynting vector over the surface of a sphere that encloses the antenna. This is equivalent to integrating the radiation intensity over a unit sphere.

$$P_{rad} = \text{Power radiated} = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} U(\theta, \phi) \sin\theta d\theta d\phi$$

(g) Directivity of an antenna

The directivity (D) of an antenna is defined as the ratio of the maximum value of the power radiated per unit solid angle to the average power radiated per unit solid angle.

That is, directivity is ratio of the maximum radiation intensity in the main beam to the average radiation intensity over all space.

$$D = \frac{U_{max}}{U_{avg}} = \frac{U_{max}}{P_{rad}/4\pi} = \frac{4\pi U_{max}}{\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} U(\theta, \phi) \sin\theta d\theta d\phi}$$

Thus, the directivity measures how intensely the antenna radiates in its preferred direction than an isotropic radiator would when fed with the same total power.

Directivity is a dimensionless ratio of power, and is usually expressed in dB as $D(\text{dB}) = 10 \log(D)$

directivity of isotropic radiator:

An isotropic radiator is a hypothetical loss less radiator having equal radiation in all directions.

$U(\theta, \phi) = 1$ for isotropic antenna. Applying the integral identity, $\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \sin\theta d\theta d\phi = 4\pi$, we have,

$$D = \frac{4\pi U_{max}}{\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} U(\theta, \phi) \sin\theta d\theta d\phi} = 1$$

The directivity of an isotropic antenna is $D = 1$, or 0 dB.

Relationship between Directivity and beamwidth

Beamwidth and directivity are both measures of the focusing ability of an antenna: an antenna pattern with a narrow main beam will have a high directivity, while a pattern with a wide beam will have a lower directivity.

Approximate relation between beam width and directivity that apply with reasonable accuracy for antennas with pencil beam patterns is the following:

$$D \cong \frac{32,000}{\theta_1 \theta_2} \quad \text{where } \theta_1 \text{ and } \theta_2 \text{ are the beam widths in two orthogonal planes of the main}$$

beam, in degrees. This approximation does not work well for omnidirectional patterns because there is a well-defined main beam in only one plane for such patterns.

(h) radiation efficiency of antenna

Radiation efficiency of an antenna is defined as the ratio of the radiated output power to the supplied input power.

$$\eta_{rad} = \frac{P_{rad}}{P_{in}} = \frac{P_{in} - P_{loss}}{P_{in}} = 1 - \frac{P_{loss}}{P_{in}}$$

where P_{rad} is the power radiated by the antenna, P_{in} is the power supplied to the input of the antenna, and P_{loss} is the power lost in the antenna(dissipative losses) due to metal conductivity or dielectric loss with in the antenna.

(i)Gain of an antenna

The gain of the antenna is closely related to the directivity, it is a measure that takes into account the efficiency of the antenna as well as its directional capabilities.

Antenna gain is defined as the product of directivity and efficiency:

$$Gain = G = \eta_{rad} \times D.$$

Thus, gain is always less than or equal to directivity.

(j) Aperture efficiency

Aperture efficiency is defined as the ratio of the actual directivity of an aperture antenna to the maximum directivity of aperture antenna.

The maximum directivity that can be obtained from an electrically large aperture of area A is given as, $D_{max} = \frac{4\pi A}{\lambda^2}$

$$\eta_{ap} = \text{aperture efficiency} = \frac{D}{D_{max}}$$

(k) Effective aperture area

Received power is proportional to the power density, or Poynting vector, of the incident wave. Since the Poynting vector has dimensions of W/m², and the received power, P_r, has dimensions of W, the proportionality constant must have units of area.

$$\text{We have, } P_r = A_e \times S_{avg}$$

where A_e is defined as the effective aperture area of the receive antenna. The effective aperture area has dimensions of m², and can be interpreted as the “capture area” of a receive antenna, intercepting part of the incident power density radiated toward the receive antenna.

relation between effective aperture area and Directivity(gain)

The maximum effective aperture area of an antenna is related to the directivity of the antenna as,

$$A_e = \frac{D\lambda^2}{4\pi}$$

The maximum effective aperture area as defined above does not include the effect of losses in the antenna, which can be accounted for by replacing D with G, the gain, of the antenna.

$$A_e = \frac{G\lambda^2}{4\pi}$$

(l)Antenna Brightness temperature

When the antenna beam width is broad enough that different parts of the antenna pattern see different background temperatures, the effective brightness temperature seen by the antenna can be found by weighting the spatial distribution of background temperature by the pattern function of the antenna.

Mathematically we can write the brightness temperature T_b seen by the antenna as

$$T_b = \frac{\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} T_B(\theta, \phi) D(\theta, \phi) \sin\theta d\theta d\phi}{\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} D(\theta, \phi) \sin\theta d\theta d\phi}$$

Where T_B(θ, φ) is the distribution of the background temperature, and D(θ, φ) is the directivity (or the power pattern function) of the antenna. Antenna brightness temperature is referenced at the terminals of the antenna. Observe that when T_B is a constant, T_b = T_B

(m) Antenna Noise Temperature

If a receiving antenna has dissipative loss, so that its radiation efficiency η_{rad} is less than unity, the power available at the terminals of the antenna is reduced by the factor η_{rad} from that intercepted by the antenna (the definition of radiation efficiency is the ratio of output to input power).

This reduction applies to received noise power, as well as received signal power, so the noise temperature of the antenna will be reduced from the brightness temperature by the factor η_{rad}.

In addition, thermal noise will be generated internally by resistive losses in the antenna, and this will increase the noise temperature of the antenna. We can find the resulting noise temperature seen at the antenna terminals as,

$$T_A = \eta_{rad} T_b + (1 - \eta_{rad}) T_p.$$

The equivalent temperature T_A is called the antenna noise temperature, and is a combination of the external brightness temperature seen by the antenna and the thermal noise generated by the antenna.

Note: This temperature is referenced at the output terminals of the antenna.

$T_A = T_b$ for a lossless antenna with $\eta_{\text{rad}} = 1$.

If the radiation efficiency is zero, meaning that the antenna appears as a matched load and does not see any external background noise, then $T_A = T_p$, due to the thermal noise generated by the losses.

(n)G/T ratio

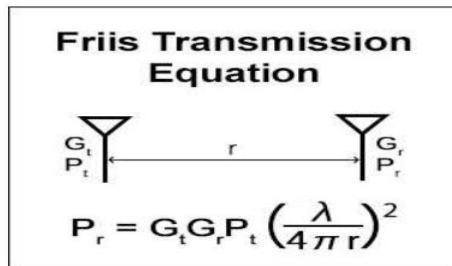
Useful figure of merit for receive antennas is the G/T ratio, defined as $10 \log(G/T_A)$ dB/K, where G is the gain of the antenna, and T_A is the antenna noise temperature.

This quantity is important because, the signal-to-noise ratio (SNR) at the input to a receiver is proportional to G/T_A . The ratio G/T can often be maximized by increasing the gain of the antenna, since this increases the numerator and usually minimizes reception of noise from hot sources at low elevation angles. Of course, higher gain requires a larger and more expensive antenna, and high gain may not be desirable for applications requiring omnidirectional coverage (e.g., cellular telephones or mobile data networks), so often a compromise must be made.

Note: that the dimensions given for $10 \log(G/T)$ are not actually decibels per degree kelvin, but this is the nomenclature that is commonly used for this quantity.

Friis Transmission Formula

This formula gives the power received over a radio communication link.



Let the transmitter feed a power P_t to a transmitting antenna of effective aperture A_{et} . At a distance r a receiving antenna of effective aperture A_{er} , intercepts some of the power radiated by the transmitting antenna and delivers it to the receiver.

Assuming that the transmitting antenna is isotropic, the power per unit area at the receiving antenna is

$$S_r = \frac{P_t}{4\pi r^2} \quad (\text{W}) \quad \text{----- 1}$$

If the transmitting antenna has gain G_t , the power per unit area at the receiving antenna will be increased in proportion as given by,

$$S_r = \frac{P_t G_t}{4\pi r^2} \quad (\text{w}) \quad \text{----- 2}$$

Now, the power collected by the receiving antenna of effective aperture A_{er} is,

$$P_r = A_{er} S_r = \frac{A_{er} P_t G_t}{4\pi r^2} \quad (\text{w}) \quad \text{----- 3}$$

The gain of the transmitting antenna can be expressed as,

$$G_t = \frac{4\pi A_{et}}{\lambda^2} \quad \text{----- 4a}$$

$$G_r = \frac{4\pi A_{er}}{\lambda^2} \quad \text{----- 4b}$$

Substituting for gain in equation 3, we have,

$$P_r = \frac{A_{er} P_t A_{et} 4\pi}{4\pi r^2 \lambda^2} = \frac{A_{er} P_t A_{et}}{r^2 \lambda^2} \quad \text{----- 5a}$$

In terms of antenna gain, received power can be expressed as,

$$P_r = \frac{G_r P_t G_t \lambda^2}{4\pi \times 4\pi r^2} = \frac{G_r P_t G_t \lambda^2}{(4\pi r)^2} \quad \text{----- 5b}$$

Equation 5 is Friis transmission formula

$$P_r = \frac{G_r P_t G_t \lambda^2}{(4\pi r)^2}$$

P_r = Received power (antenna matched) in W

P_t = power in to transmitting antenna in W

A_{et} = Effective aperture of transmitting antenna, m^2

A_{er} = Effective aperture of Receiving antenna, m^2

r = distance between transmitting and receiving antenna, m

λ = wave length, m

EIRP and it's significance

The product $P_t G_t$ is defined as the Effective Isotropic Radiated Power (EIRP).

$$EIRP = P_t G_t \quad W$$

For a given frequency, range, and receiver antenna gain, the received power is proportional to the EIRP of the transmitter and received power can only be increased by increasing the EIRP.

This can be done by increasing the transmit power, or the transmit antenna gain, or both.

Path Loss

Path loss is the quantity that account for the free-space reduction in signal strength with distance between the transmitter and receiver.

$$\text{Path loss} = \text{Transmitted power} - \text{Received power} = P_t - P_r$$

Assuming unity gain antennas, path loss is given as(using Friis formula)

$$\text{path loss (dB)} = 20 \log \left(\frac{4\pi r}{\lambda} \right)$$

Link Budget and Link Margin

The various terms in the Friis formula are often tabulated separately in a link budget, where each of the factors can be individually considered in terms of its net effect on the received power.

Additional loss factors, such as line losses or impedance mismatch at the antennas, atmospheric attenuation, and polarization mismatch can also be added to the link budget.

One of the terms in a link budget is the path loss, accounting for the free-space reduction in signal strength with distance between the transmitter and receiver.

$$\text{Path loss} = \text{Transmitted power} - \text{Received power} = P_t - P_r$$

Assuming unity gain antennas, path loss is given as(using Friis formula)

$$\text{path loss (dB)} = 20 \log \left(\frac{4\pi r}{\lambda} \right)$$

We can write the budget as shown in the following link budget:

Transmit power	P_t
Transmit antenna line loss	$(-)L_t$
Transmit antenna gain	G_t
Path loss (free-space)	$(-)L_0$
Atmospheric attenuation	$(-)L_A$
Receive antenna gain	G_r
Receive antenna line loss	$(-)L_r$
<hr/>	
Receive power	P_r

We have also included loss terms for atmospheric attenuation and line attenuation. Assuming that all of the above quantities are expressed in dB (or dBm, in the case of P_t), we can write the receive power as

$$P_r(\text{dBm}) = P_t - L_t + G_t - L_0 - L_A + G_r - L_r$$

If the transmit and/or receive antenna is not impedance matched to the transmitter/receiver (or to their connecting lines), impedance mismatch will reduce the received power by the factor $(1 - |\Gamma|^2)$ where Γ is the appropriate reflection coefficient.

The resulting impedance mismatch loss,

$$L_{\text{imp}}(\text{dB}) = -10 \log(1 - |\Gamma|^2) \geq 0,$$

can be included in the link budget to account for the reduction in received power.

Another possible entry in the link budget relates to the polarization matching of the transmit and receive antennas, as maximum power transmission between transmitter and receiver requires both antennas to be polarized in the same manner.

If a transmit antenna is vertically polarized, for example, maximum power will only be delivered to a vertically polarized receiving antenna, while zero power would be delivered to a horizontally polarized receive antenna, and half the available power would be delivered to a circularly polarized antenna.

Link Margin

In practical communications systems it is usually desired to have the received power level greater than the threshold level required for the minimum acceptable quality of service (usually expressed as the minimum carrier-to-noise ratio (CNR), or minimum SNR).

This design allowance for received power is referred to as the link margin, and can be expressed as the difference between the design value of received power and the minimum threshold value of receive power:

$$\text{Link margin (dB)} = \text{LM} = P_r - P_r(\text{min}) > 0, \text{ where all quantities are in dB.}$$

Link margin should be a positive number; typical values may range from 3 to 20 dB. Having a reasonable link margin provides a level of robustness to the system to account for variables such as signal fading due to weather, movement of a mobile user, multipath propagation problems, and other unpredictable effects that can degrade system performance.

Link margin for a given communication system can be improved by increasing the received power (by increasing transmit power or antenna gains), or by reducing the minimum threshold power (by improving the design of the receiver, changing the modulation method, or by other means)

Fade margin.

Signal fading occur due to weather, movement of a mobile user, multipath propagation problems, and other unpredictable effects that can degrade system performance and quality of service. Link margin that is used to account for fading effects is sometimes referred to as fade margin.

Noise Characterization of a Microwave Receiver

(i) NOISE FIGURE and EQUIVALENT NOISE TEMPERATURE of a SYSTEM

(General concepts)

The signal-to-noise ratio is the ratio of desired signal power to undesired noise power, and so is dependent on the signal power.

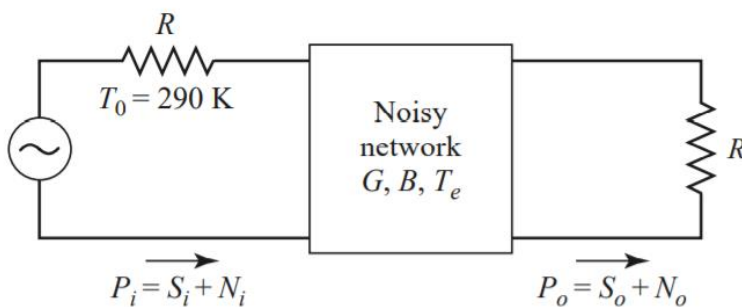
When noise and a desired signal are applied to the input of a noiseless network, both noise and signal will be attenuated or amplified by the same factor, so that the signal-to-noise ratio will be unchanged.

However, if the network is noisy, the output noise power will be increased more than the output signal power, so that the output signal-to-noise ratio will be reduced.

The noise figure, F , is a measure of this reduction in signal-to-noise ratio, and is defined as,

$$F = \frac{S_i/N_i}{S_o/N_o} \geq 1 \quad \text{----- (1)}$$

where S_i , N_i are the input signal and noise powers, and S_o , N_o are the output signal and noise powers. By definition, the input noise power is assumed to be the noise power resulting from a matched resistor at $T_0 = 290$ K; that is, $N_i = kT_0B$.



Determining the noise figure of a noisy network.

Consider Figure shown above, which shows noise power N_i and signal power S_i being fed into a noisy two-port network.

The network is characterized by a gain, G , a bandwidth, B , and an equivalent noise temperature, T_e .

The input noise power is $N_i = kT_0B$, and the output noise power is a sum of the amplified input noise and the internally generated noise: $N_o = kGB(T_0 + T_e)$.

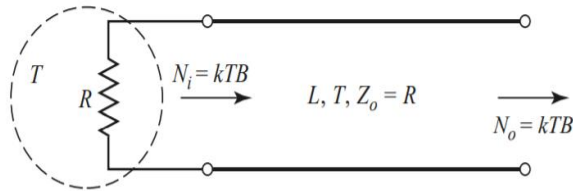
The output signal power is $S_o = GS_i$. Using these results in (1) gives the noise figure as,

$$F = \frac{S_i}{kT_0B} \times \frac{kGB(T_0+T_e)}{GS_i} = 1 + \frac{T_e}{T_0} \geq 1 \quad \text{-----(2)}$$

$$T_e = (F - 1)T_0 \quad \text{-----(3)}$$

It is important to keep in mind two things concerning the definition of noise figure: noise figure is defined for a matched input source, and for a noise source equivalent to a matched load at temperature $T_0 = 290$ K. Noise figure and equivalent noise temperatures are interchangeable characterizations of the noise properties of a component.

An important special case occurs in practice for a two-port network consisting of a passive, lossy component, such as an attenuator or lossy transmission line, held at a physical temperature T . Consider such a network with a matched source resistor that is also at temperature T , as shown in Figure.



Determining the noise figure of a lossy line or attenuator with loss L and temperature T .

The power gain, G , of a lossy network is less than unity; the loss factor, L , can be defined as $L = 1/G > 1$. Because the entire system is in thermal equilibrium at the temperature T , and has a driving point impedance of R , the output noise power must be $N_o = kTB$. However, we can also think of this power as coming from the source resistor (attenuated by the lossy line), and from the noise generated by the line itself. Thus we also have that

$$N_o = kTB = GkTB + G N_{\text{added}} \text{ -----(4)}$$

Where N_{added} is the noise generated by the line, as if it appeared at the input terminals of the line. Solving (4) for this power gives

$$N_{\text{added}} = \frac{(1-G)}{G} kTB = (L - 1)kTB \text{ ----- (5)}$$

Then (5) shows that the lossy line has an equivalent noise temperature (referred to the input) given by,

$$T_e = (L - 1)T \text{ ----- (6)}$$

Noise figure is,

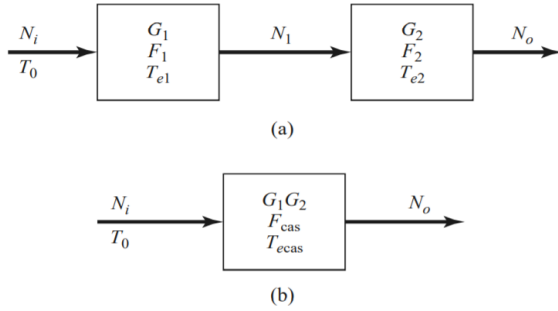
$$F = 1 + \frac{(L-1)T}{T_o} \geq 1 \text{ -----(7)}$$

Noise Figure of a Cascaded System

In a typical microwave system the input signal travels through a cascade of many different components, each of which may degrade the signal-to-noise ratio to some degree. If we know the noise figure (or noise temperature) of the individual stages, we can determine the noise figure (or noise temperature) of the cascade connection of stages.

We will see that the noise performance of the first stage is usually the most critical, an interesting result that is very important in practice.

Consider the cascade of two components, having gains G_1, G_2 , noise figures F_1, F_2 , and equivalent noise temperatures T_{e1}, T_{e2} , as shown in Figure.



Noise figure and equivalent noise temperature of a cascaded system. (a) Two cascaded networks. (b) Equivalent network.

We wish to find the overall noise figure and equivalent noise temperature of the cascade, as if it were a single component. The overall gain of the cascade is $G_1 G_2$.

Using noise temperatures, we can write the noise power at the output of the first stage as

$$N_1 = G_1 k T_0 B + G_1 k T_{e1} B \quad \text{----- (8)}$$

since $N_i = k T_0 B$ for noise figure calculations. The noise power at the output of the second stage is

$$N_o = G_2 N_1 + G_2 k T_{e2} B$$

$$N_o = G_1 G_2 k T_0 B + G_1 G_2 k T_{e1} B + G_2 k T_{e2} B$$

$$N_o = G_1 G_2 k B \left(T_0 + T_{e1} + \frac{T_{e2}}{G_1} \right) \quad \text{----- (9)}$$

For the equivalent system we have,

$$N_o = G_1 G_2 k B (T_0 + T_{cas}) \quad \text{----- (10)}$$

Where,

$$T_{cas} = T_{e1} + \frac{T_{e2}}{G_1} \quad \text{----- (11)}$$

Using (3) to convert the temperatures in (11) to noise figures yields the noise figure of the cascade system as,

$$F_{cas} = F_1 + \frac{(F_2 - 1)}{G_1} \quad \text{----- (12)}$$

Equations (11) and (12) show that the noise characteristics of a cascaded system are dominated by the characteristics of the first stage since the effect of the second stage is reduced by the gain of the first (assuming $G_1 > 1$).

Thus, for the best overall system noise performance, the first stage should have a low noise figure and at least moderate gain. Expense and effort should be devoted primarily to the first stage, as opposed to later stages, since later stages have a diminished impact on the overall noise performance.

Equations (11) and (12) can be generalized to an arbitrary number of stages, as

$$T_{cas} = T_{e1} + \frac{T_{e2}}{G_1} + \frac{T_{e3}}{G_1 G_2} + \dots \dots \quad \text{-----(13)}$$

$$F_{cas} = F_1 + \frac{(F_2-1)}{G_1} + \frac{(F_3-1)}{G_1 G_2} + \dots \dots \quad \text{-----(14)}$$

(ii) **Noise Characterization of Receiver**

We can now analyze the noise characteristics of a complete antenna–transmission line–receiver front end, as shown in Figure. In this system the total noise power at the output of the receiver, N_o , will be due to contributions from the antenna pattern, the loss in the antenna, the loss in the transmission line, and the receiver components.

This noise power will determine the minimum detectable signal level for the receiver and, for a given transmitter power, the maximum range of the communication link.

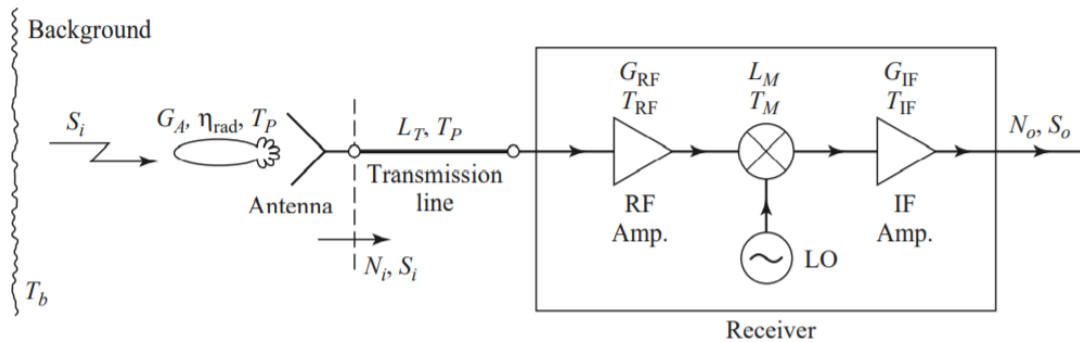


FIGURE 14.14 Noise analysis of a microwave receiver front end, including antenna and transmission line contributions.

The receiver components in Figure consist of an RF amplifier with gain G_{RF} and noise temperature T_{RF} , a mixer with an RF-to-IF conversion loss factor L_M and noise temperature T_M , and an IF amplifier with gain G_{IF} and noise temperature T_{IF} .

The noise effects of later stages can usually be ignored since the overall noise figure is dominated by the characteristics of the first few stages.

The component noise temperatures can be related to noise figures as $T = (F - 1)T_0$.

The equivalent noise temperature of the receiver can be found as

$$T_{REC} = T_{RF} + \frac{T_M}{G_{RF}} + \frac{T_{IF} L_M}{G_{RF}} \quad \text{----- (1)}$$

The transmission line connecting the antenna to the receiver has a loss L_T , and is at a physical temperature T_p . So, its equivalent noise temperature is

$$T_{TL} = (L_T - 1)T_p \quad \text{----- (2)}$$

We can find that the noise temperature of the transmission line (TL) and receiver (REC) cascade is

$$T_{TL+REC} = T_{TL} + L_T T_{REC} = (L_T - 1)T_p + L_T T_{REC} \quad \text{----- (3)}$$

This noise temperature is defined at the antenna terminals (the input to the transmission line). The entire antenna pattern can collect noise power. If the antenna has a reasonably high gain with relatively low sidelobes, we can assume that all noise power comes via the main beam, so that the noise temperature of the antenna is given by,

$$T_A = \eta_{rad}T_b + (1 - \eta_{rad})T_p \text{ ----- (4)}$$

where η_{rad} is the efficiency of the antenna, T_p is its physical temperature, and T_b is the equivalent brightness temperature of the background seen by the main beam.

The noise power at the antenna terminals, which is also the noise power delivered to the transmission line, is

$$N_i = kBT_A = kB[\eta_{rad}T_b + (1 - \eta_{rad})T_p] \text{ -----(5)}$$

where B is the system bandwidth. If S_i is the received power at the antenna terminals, then the input SNR at the antenna terminals is S_i / N_i .

The output signal power is,

$$S_o = \frac{S_i G_{RF} G_{IF}}{L_T L_M} = S_i G_{SYS} \text{ ----- (6)}$$

where G_{SYS} has been defined as a system power gain.

The output noise power is,

$$N_o = (N_i + kBT_{TL+REC})G_{SYS}$$

$$N_o = (kBT_A + kBT_{TL+REC})G_{SYS}$$

$$N_o = kB(T_A + T_{TL+REC})G_{SYS} = kBT_{SYS}G_{SYS} \text{ ----- (7)}$$

where T_{SYS} has been defined as the overall system noise temperature.

The output SNR is,

$$\frac{S_o}{N_o} = \frac{S_i}{kBT_{SYS}} \text{ ----- (8)}$$

$$\frac{S_o}{N_o} = \frac{S_i}{kB[\eta_{rad}T_b + (1 - \eta_{rad})T_p + (L_T - 1)T_p + L_T T_{REC}]}$$

It may be possible to improve this output SNR by various signal processing techniques.

Impedance Matching

- Impedance matching, is often an important part of a larger design process for a microwave component or system.
- The basic idea of impedance matching is illustrated in Figure, which shows an impedance **matching network** placed between a load impedance and a transmission line.
- The **matching network** is ideally lossless, to avoid unnecessary loss of power, and is usually designed so that the impedance seen looking into the matching network is Z_0 .
- Then reflections will be eliminated on the transmission line to the left of the matching network, although there will usually be multiple reflections between the matching network and the load.
- This procedure is sometimes referred to as tuning. Impedance matching or tuning is important for the following reasons:

Impedance Matching



FIGURE 5.1 A lossless network matching an arbitrary load impedance to a transmission line.

Impedance matching or tuning is important for the following reasons:

- 1) Maximum power is delivered when the load is matched to the line (assuming the generator is matched), and power loss in the feed line is minimized.
- 2) Impedance matching sensitive receiver components (antenna, low-noise amplifier, etc.) may improve the signal-to-noise ratio of the system.
- 3) Impedance matching in a power distribution network (such as an antenna array feed network) may reduce amplitude and phase errors.

- As long as the load impedance, Z_L , has a positive real part, a matching network can always be found.
- Many choices are available.

Factors that may be important in the selection of a particular matching network include the following:

- Complexity
- Bandwidth
- Implementation
- Adjustability

Complexity

- The simplest design that satisfies the required specifications is generally preferable.
- A simpler matching network is usually cheaper, smaller, more reliable, and less lossy than a more complex design.

Bandwidth

- Any type of matching network can ideally give a perfect match (zero reflection) at a single frequency.
- In many applications, however, it is desirable to match a load over a band of frequencies.
- There are several ways of doing this, with, of course, a corresponding increase in complexity.

Implementation

- Depending on the type of transmission line or waveguide being used, one type of matching network may be preferable to another.
- For example, tuning stubs are much easier to implement in waveguide than are multi section quarter-wave transformers.

Adjustability

- In some applications the matching network may require adjustment to match a variable load impedance.
- It is possible in some types of matching networks.

Impedance matching techniques:

- **MATCHING WITH LUMPED ELEMENTS (L NETWORKS)**
- **SINGLE-STUB TUNING**
- **DOUBLE-STUB TUNING**
- **THE QUARTER-WAVE TRANSFORMER**
- **TAPERED LINES**

MATCHING WITH LUMPED ELEMENTS (L NETWORKS)

- The simplest type of matching network is the L-section, which uses two reactive elements to match an arbitrary load impedance to a transmission line.
- There are two possible configurations for this network, as shown in Figure:

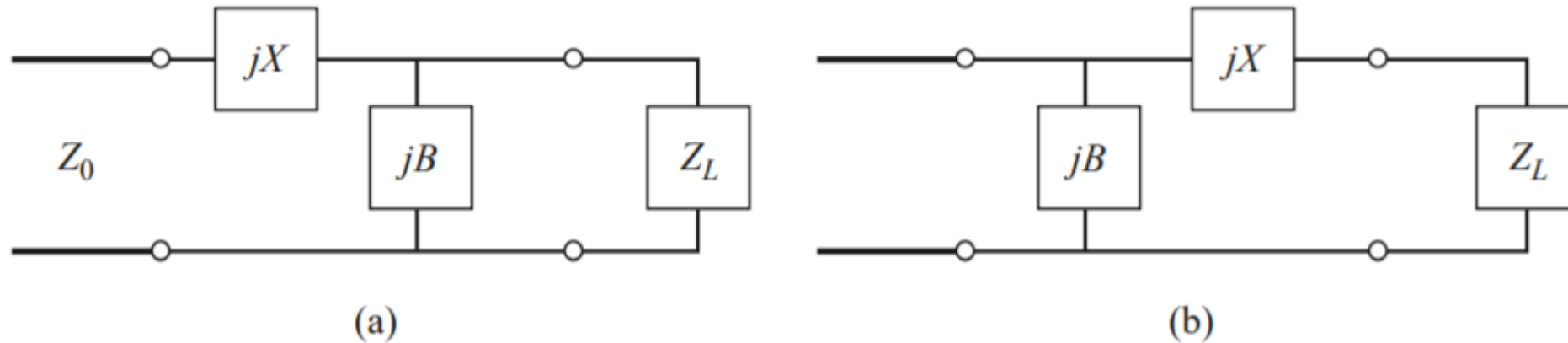


FIGURE 5.2 L-section matching networks. (a) Network for z_L inside the $1 + jx$ circle. (b) Network for z_L outside the $1 + jx$ circle.

- If the normalized load impedance, $z_L = Z_L / Z_0$, is inside the $1 + jx$ circle on the Smith chart, then the circuit of Figure (a) should be used.
- If the normalized load impedance is outside the $1 + jx$ circle on the Smith chart, the circuit of Figure (b) should be used.
- The $1 + jx$ circle is the resistance circle on the impedance Smith chart for which $r = 1$.
- In either of the configurations of Figure, the reactive elements may be either inductors or capacitors, depending on the load impedance.

- Thus, there are eight distinct possibilities for the matching circuit for various load impedances.
- If the frequency is low enough and/or the circuit size is small enough, actual lumped-element capacitors and inductors can be used. This may be feasible for frequencies up to about 1 GHz or so.
- Modern microwave integrated circuits may be small enough such that lumped elements can be used at higher frequencies as well.

Determination of component values of matching network:

consider the circuit of Figure (b)

This circuit is used when z_L is outside the $1 + jx$ circle on the Smith chart, which implies that $R_L < Z_0$.

The admittance seen looking into the matching network, followed by the load impedance, must be equal to $1/Z_0$ for an impedance-matched condition:

$$\frac{1}{Z_0} = jB + \frac{1}{R_L + j(X + X_L)}$$

Rearranging and separating into real and imaginary parts gives two equations for the two unknowns, X and B :

$$BZ_0(X + X_L) = Z_0 - R_L,$$

$$(X + X_L) = BZ_0R_L.$$

Solving for X and B gives

$$X = \pm\sqrt{R_L(Z_0 - R_L)} - X_L,$$

$$B = \pm\frac{\sqrt{(Z_0 - R_L)/R_L}}{Z_0}.$$

Because $R_L < Z_0$, the arguments of the square roots are always positive. Again, note that two solutions are possible.

For circuit in fig (a), component values can be determined as follows:

$$Z_0 = jX + \frac{1}{jB + 1/(R_L + jX_L)}.$$

Rearranging and separating into real and imaginary parts gives two equations for the two unknowns, X and B :

$$B(XR_L - X_L Z_0) = R_L - Z_0,$$

$$X(1 - BX_L) = BZ_0 R_L - X_L.$$

Solving (5.2a) for X and substituting into (5.2b) gives a quadratic equation for B . The solution is

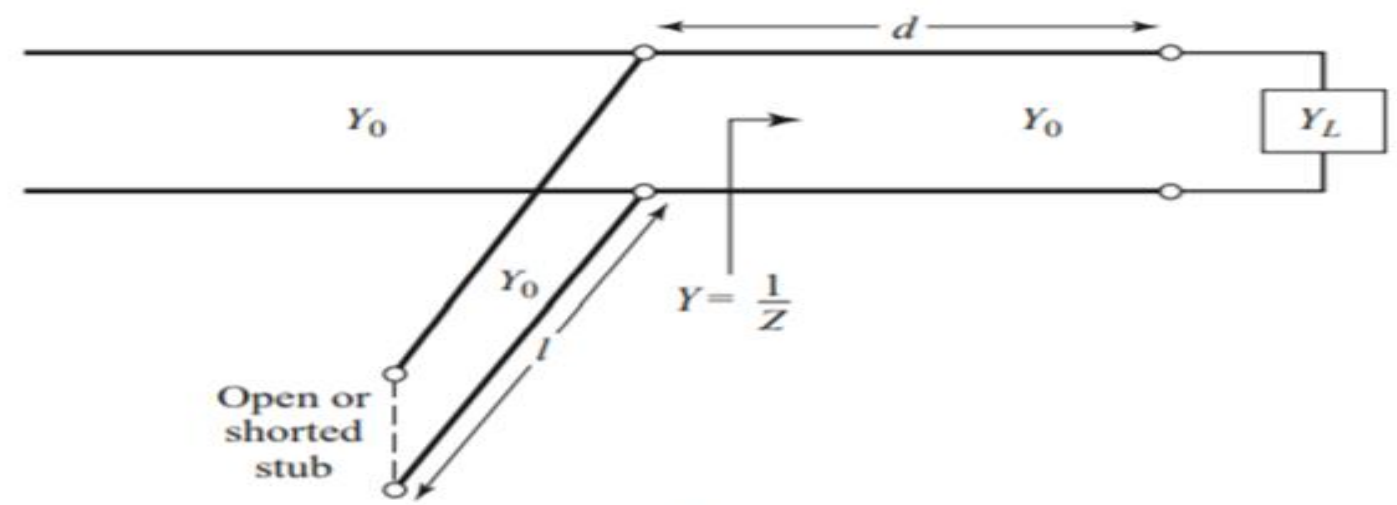
$$B = \frac{X_L \pm \sqrt{R_L/Z_0} \sqrt{R_L^2 + X_L^2 - Z_0 R_L}}{R_L^2 + X_L^2}.$$

Note that since $R_L > Z_0$, the argument of the second square root is always positive. Then the series reactance can be found as

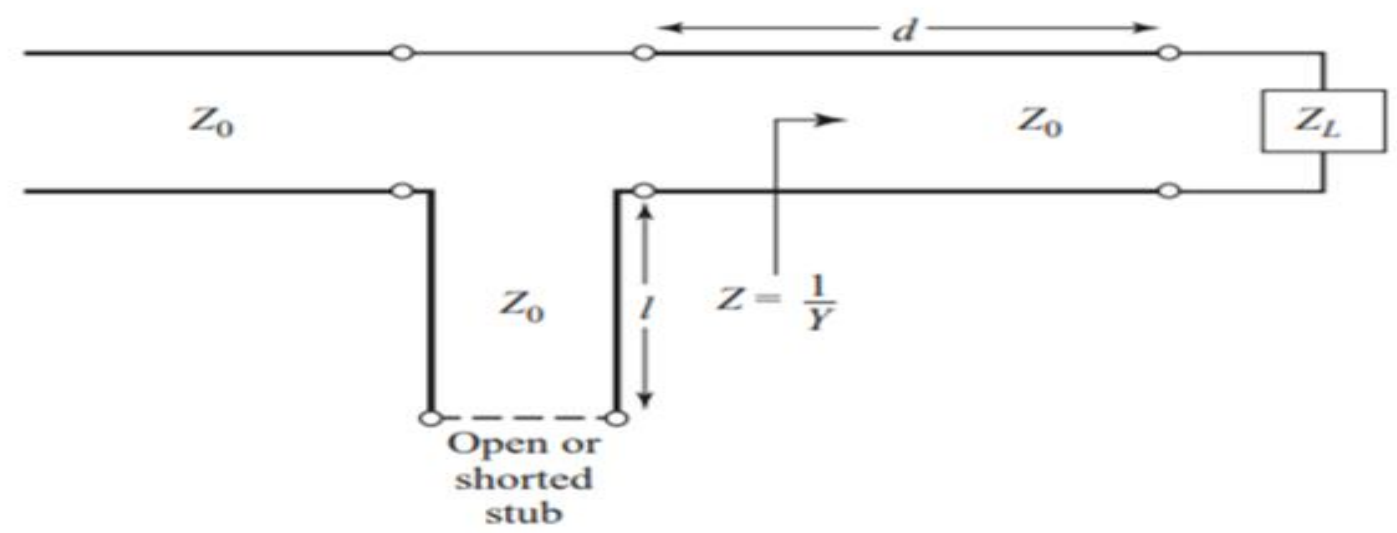
$$X = \frac{1}{B} + \frac{X_L Z_0}{R_L} - \frac{Z_0}{B R_L}.$$

SINGLE-STUB TUNING

- Another popular matching technique uses a single open-circuited or short-circuited length of transmission line (a stub) connected either in parallel or in series with the transmission feed line at a certain distance from the load, as shown in Figure.
- Such a single-stub tuning circuit is often very convenient because the stub can be fabricated as part of the transmission line media of the circuit, and lumped elements are avoided.
- Shunt stubs are preferred for microstrip line or stripline, while series stubs are preferred for slotline or coplanar waveguide.



(a)



(b)

FIGURE 5.4 Single-stub tuning circuits. (a) Shunt stub. (b) Series stub.

- In single-stub tuning the two adjustable parameters are the distance, d , from the load to the stub position, and the value of susceptance or reactance provided by the stub.
- For the shunt-stub case, the basic idea is to select d so that the admittance, Y , seen looking into the line at distance d from the load is of the form $Y_0 + j B$. Then the stub susceptance is chosen as $-j B$, resulting in a matched condition.
- For the series-stub case, the distance d is selected so that the impedance, Z , seen looking into the line at a distance d from the load is of the form $Z_0 + j X$. Then the stub reactance is chosen as $-j X$, resulting in a matched condition.

DOUBLE-STUB TUNING

- The single-stub tuner is able to match any load impedance (having a positive real part) to a transmission line, but suffers from the disadvantage of requiring a variable length of line between the load and the stub.
- This may not be a problem for a fixed matching circuit, but would probably pose some difficulty if an adjustable tuner is desired.
- In this case, the double-stub tuner, which uses two tuning stubs in fixed positions, can be used. Such tuners are often fabricated in coaxial line with adjustable stubs connected in shunt to the main coaxial line.

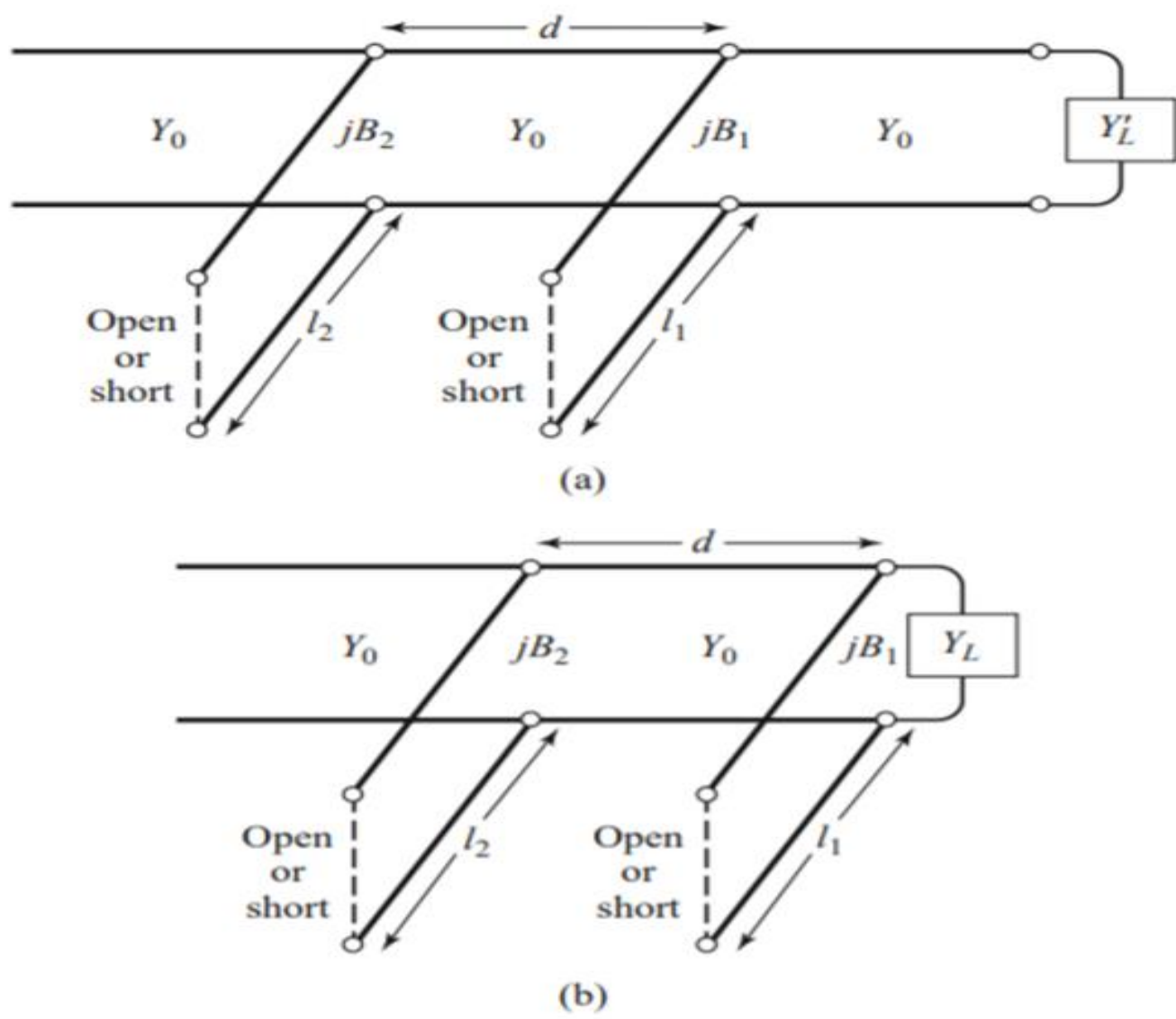


FIGURE 5.7 Double-stub tuning. (a) Original circuit with the load an arbitrary distance from the first stub. (b) Equivalent circuit with the load transformed to the first stub.

THE QUARTER-WAVE TRANSFORMER

- The quarter-wave transformer is a simple and useful circuit for matching a real load impedance to a transmission line.
- An additional feature of the quarter-wave transformer is that it can be extended to multi section designs in a methodical manner to provide broader bandwidth.
- If only a narrow band impedance match is required, a single-section transformer may suffice.
- One drawback of the quarter-wave transformer is that it can only match a real load impedance.

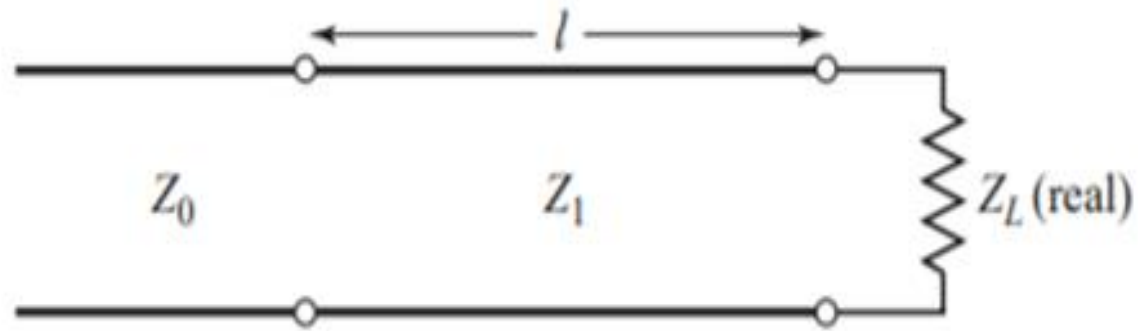


FIGURE 5.10 A single-section quarter-wave matching transformer. $\ell = \lambda_0/4$ at the design frequency f_0 .

The single-section quarter-wave matching transformer circuit is shown in Fig with the characteristic impedance of the matching section given as

$$Z_1 = \sqrt{Z_0 Z_L}.$$

UNIT II
RADIATION MECHANISMS AND DESIGN ASPECTS

- **RADIATION MECHANISMS OF LINEAR WIRE ANTENNAS**
 - Alternating Current Element (Oscillating Dipole/ Hertzian dipole)
 - Half-wave Dipole Antenna*
- **LOOP ANTENNAS***
- **APERTURE ANTENNAS**
 - Wire Antennas Vs Aperture Antennas
 - Field Equivalence Principle#
 - Horn Antennas*
 - Design Principle
 - Rectangular Horn Antennas and Solved Problem
 - Conical Horn Antennas
 - Ridge Horns
 - Septum Horns
 - Corrugated Horns
 - Aperture-Matched Horn
 - Slot Antennas
 - Methods of Feeding*
 1. Coaxial feed
 2. Offset feed
 3. Boxed-in Slot Antenna
 4. Waveguide-fed Slot
 5. Broadside Array of Slots in a Waveguide
 - Babinet's Principle#
 - Booker's Extension of Babinet's Principle#
 - Impedance of Slot Antenna and Solved Problem#
- **REFLECTOR ANTENNAS**
 - Reflectors of various shapes
 - Parabolic Reflector*
 - f/d ratio
 - Feed systems for Parabolic Reflectors
 1. Axial/Front Feed
 2. Offset Feed
 3. Cassegrain Feed
 4. Gregorian Feed

- **MICROSTRIP ANTENNAS**

- Basic Characteristics of Microstrip antennas*
- Feeding Methods*
 - Microstrip Line Feed
 - Coaxial Probe Feed
 - Aperture Coupled Feed
 - Proximity Coupled Feed
- Methods of Analysis
 - Transmission-Line Model
 - Cavity Model

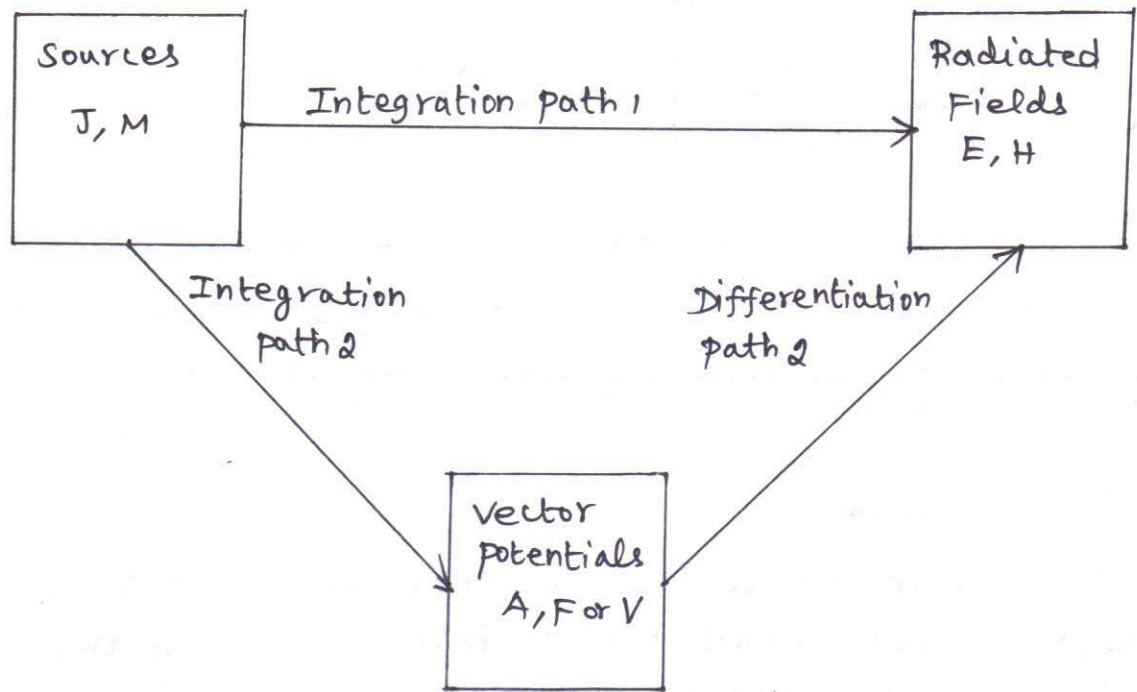
- **FREQUENCY INDEPENDENT ANTENNAS**

- Rumsey's principle#
- Frequency-Independent Planar Log Spiral Antenna*
- Frequency Independent Conical Spiral Antenna
- Log-Periodic Antenna*
 - Basic concept
 - Regions of LPDA
 - Radiation pattern
 - Log-periodic behaviour
 - Design equations for LPDA
 - Solved Problem

Sections requiring additional attention in the perspective of Part-A.

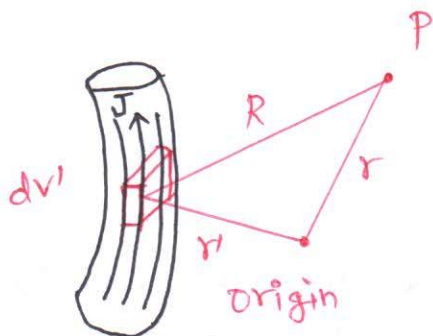
* Sections requiring additional attention in the perspective of Part-B.

Block Diagram for computing fields radiated by Electric and magnetic sources:-



Retarded potential:-

In dealing with antennas or radiating systems, the propagation time is a matter of great importance. Thus, if an alternating current is flowing in the short element, the effect of the current is not felt instantaneously at the point P but only after a time interval equal to the time required for the disturbance to propagate over the distance R.



Time required for the disturbance to propagate is nothing but the propagation time of EM waves. Thus, a current carrying conductor has an elemental volume dv' which is r' away from origin. point P is at a distance R away from dv' . v is the velocity which the field progresses or wave travels.

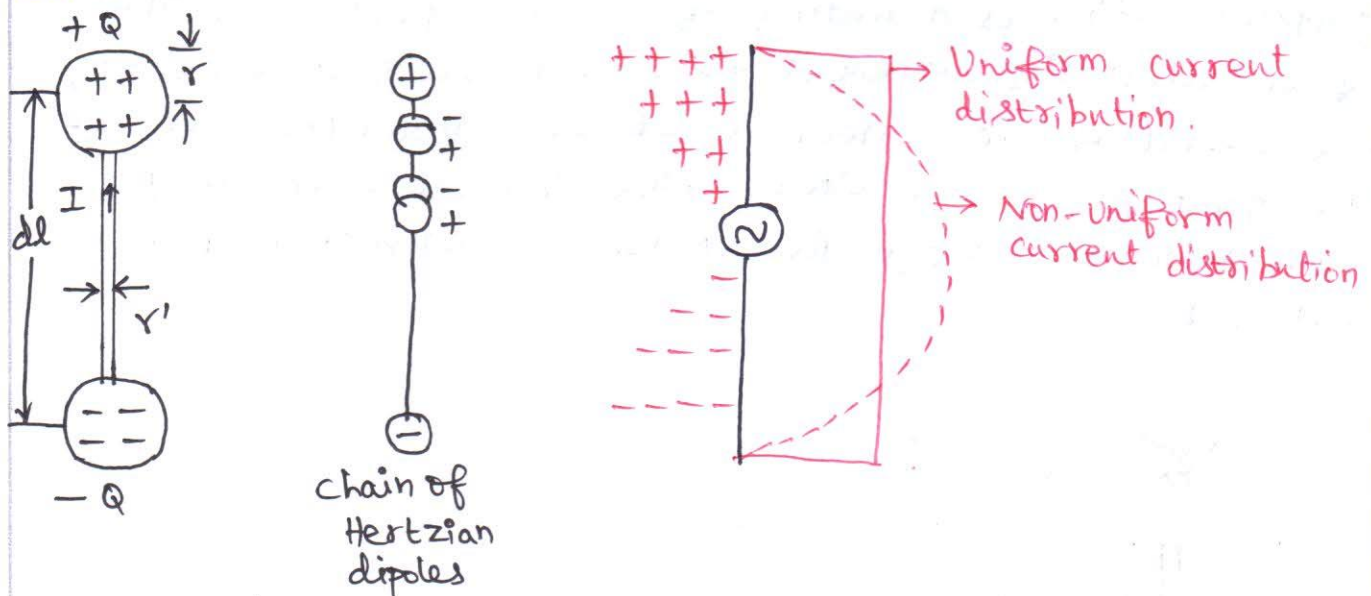
$$v(r, t) = \frac{1}{4\pi\epsilon} \int \frac{\rho(r', t - R/v)}{R} dv'$$

$$A(r, t) = \frac{\mu}{4\pi} \int \frac{J(r', t - R/v)}{R} dv' \Rightarrow \text{Vector Magnetic Potential.}$$

In these expressions, a time delay of $\frac{R}{v}$ seconds has been introduced and potentials are delayed or retarded by a factor of $\frac{R}{v}$ and hence referred to as **retarded potential**.

Hertzian Dipole :-

A Hertzian dipole can be conceived as a very short current element terminated at both the ends in two very small spheres or discs. The wire joining the spheres is very thin making the distributed capacitance between the spheres negligible. In view of the short length of the wire, the current can be assumed to be uniform throughout.



In above figure, r' is the radius of the wire connecting the two spheres (of radius r) on which the charges are residing, dl is the length of the wire and λ is the wavelength. The relative values of these parameters have to be such that

$$r' \ll r, \quad r \ll dl \quad \text{and} \quad dl \ll \lambda.$$

If the aforementioned conditions are met, the very short, very thin wire can be considered to be composed of a chain of spheres of diminishing radii. The adjacent spheres shall carry charges of opposite polarity. If the charges carried by alternate spheres are equal in magnitudes, these will get canceled everywhere except at the top and the bottom tips of the wire and the current distribution along the wire will be uniform. In case of incomplete cancellation, the charge and the current distribution along the wire will be non-uniform. Thus, an oscillating dipole can be considered as a Hertzian dipole.

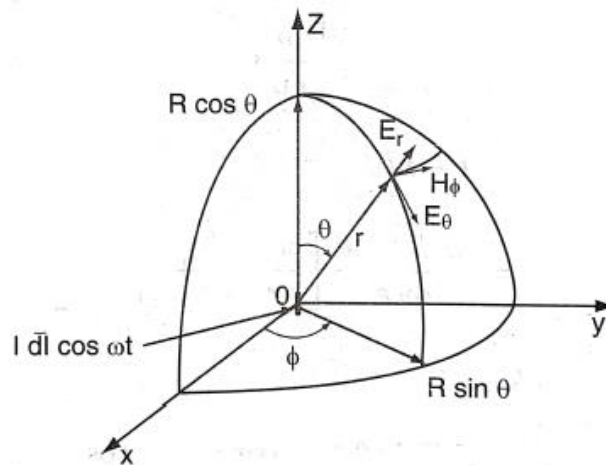
Far-field due to an Alternating Current Element (Oscillating Dipole/ Hertzian dipole)

Consider that a time varying current I is flowing in a very short and very thin wire of length dl in the z -direction. This current is given by $I dl \cos \omega t$. Since the current is in the z -direction, the current density J will have only a z -component.

$$J = a_z J_z \quad (1)$$

The vector magnetic potential A will also have only a z -component.

$$A = a_z A_z \quad (2)$$



Configuration of filamentary current carrying conductor

For a line charge density,

$$A_z = \int \frac{\mu I dl}{4\pi R} = \frac{\mu I dl \cos \omega(t - r/v)}{4\pi r} \quad (3)$$

We know that $A_x = 0, A_y = 0$ and $A_z \neq 0$. Since the three-dimensional radiation problem needs to be tackled in a spherical co-ordinate system, A_z needs to be transformed into the spherical co-ordinate system.

$$\begin{bmatrix} A_r \\ A_\theta \\ A_\phi \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{bmatrix} \begin{bmatrix} A_x \\ A_y \\ A_z \end{bmatrix}$$

$$\begin{bmatrix} A_r \\ A_\theta \\ A_\phi \end{bmatrix} = \begin{bmatrix} \sin\theta\cos\phi & \sin\theta\sin\phi & \cos\theta \\ \cos\theta\cos\phi & \cos\theta\sin\phi & -\sin\theta \\ -\sin\phi & \cos\phi & 0 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ A_z \end{bmatrix}$$

$$A_r = A_z \cos\theta ; A_\theta = -A_z \sin\theta ; A_\phi = 0 \quad (4)$$

$$A_r = \frac{\mu I dl \cos \omega(t - r/v)}{4\pi r} \cos\theta \quad (5)$$

$$A_\theta = -\frac{\mu I dl \cos \omega(t - r/v)}{4\pi r} \sin\theta \quad (6)$$

Further from the relation $B = \nabla \times A$, the components of $\nabla \times A$ are obtained as below.

$$B = \nabla \times A = \frac{1}{r^2 \sin \theta} \begin{vmatrix} a_r & ra_\theta & r \sin \theta a_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ A_r & rA_\theta & r \sin \theta A_\phi \end{vmatrix}$$

We know that $A_\phi = 0$ and A_r & A_θ are independent of ϕ . So $\frac{\partial}{\partial \phi} = 0$.

$$B = \nabla \times A = \frac{1}{r^2 \sin \theta} \begin{vmatrix} a_r & ra_\theta & r \sin \theta a_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & 0 \\ A_r & rA_\theta & 0 \end{vmatrix}$$

$$B = \nabla \times A = \frac{1}{r^2 \sin \theta} \left[a_r(0) - ra_\theta(0) + r \sin \theta a_\phi \left(\frac{\partial(rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) \right]$$

$$(\nabla \times A)_r = B_r = \mu H_r = 0 \Rightarrow H_r = 0 \quad (7)$$

$$(\nabla \times A)_\theta = B_\theta = \mu H_\theta = 0 \Rightarrow H_\theta = 0 \quad (8)$$

$$(\nabla \times A)_\phi = B_\phi = \mu H_\phi = \frac{1}{r^2 \sin \theta} \left\{ r \sin \theta \left(\frac{\partial(rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right) \right\}$$

$$\Rightarrow H_\phi = \frac{1}{\mu r} \left[\frac{\partial(rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right] \quad (9)$$

To find H_ϕ

Using eq(6),

$$\begin{aligned} \frac{\partial(rA_\theta)}{\partial r} &= \frac{\partial}{\partial r} \left(-r \frac{\mu I dl \cos \omega(t - r/v)}{4\pi r} \sin \theta \right) = -\frac{\mu I dl \sin \theta}{4\pi} \frac{\partial}{\partial r} [\cos \omega(t - r/v)] \\ &= -\frac{\mu I dl \sin \theta}{4\pi} \left(\frac{-\omega}{v} \right) [-\sin \omega(t - r/v)] \end{aligned}$$

$$\frac{\partial(rA_\theta)}{\partial r} = -\frac{\mu I dl \sin \theta}{4\pi} \left(\frac{\omega}{v} \right) \sin \omega(t - r/v) \quad (10)$$

Using eq(5)

$$\frac{\partial A_r}{\partial \theta} = \frac{\partial}{\partial \theta} \left(\frac{\mu I dl \cos \omega(t - r/v)}{4\pi r} \cos \theta \right) = \frac{\mu I dl \cos \omega(t - r/v)}{4\pi r} \frac{\partial}{\partial \theta} [\cos \theta]$$

$$\frac{\partial A_r}{\partial \theta} = -\frac{\mu I dl \cos \omega(t - r/v)}{4\pi r} \sin \theta \quad (11)$$

Using eq(10) and eq(11) in eq(9),

$$H_\phi = \frac{1}{\mu r} \left[\frac{\partial(rA_\theta)}{\partial r} - \frac{\partial A_r}{\partial \theta} \right]$$

$$\begin{aligned}
H_\phi &= \frac{1}{\mu r} \left[\left(-\frac{\mu I dl \sin\theta}{4\pi} \left(\frac{\omega}{v} \right) \sin\omega(t-r/v) \right) - \left(-\frac{\mu I dl \cos\omega(t-r/v)}{4\pi r} \sin\theta \right) \right] \\
H_\phi &= \frac{1}{\mu r} \left[-\frac{\mu I dl \sin\theta}{4\pi} \left(\frac{\omega}{v} \right) \sin\omega(t-r/v) + \frac{\mu I dl \cos\omega(t-r/v)}{4\pi r} \sin\theta \right] \\
H_\phi &= \frac{I dl \sin\theta}{4\pi} \left[-\frac{\mu}{\mu r} \left(\frac{\omega}{v} \right) \sin\omega(t-r/v) + \frac{\mu \cos\omega(t-r/v)}{\mu r^2} \right] \\
\mathbf{H}_\phi &= \frac{I dl \sin\theta}{4\pi} \left[-\left(\frac{\omega}{rv} \right) \sin\omega(t-r/v) + \frac{\cos\omega(t-r/v)}{r^2} \right] \tag{12}
\end{aligned}$$

Now our objective is to find the electric field components from the magnetic field strength relations.

From Maxwell's Equation,

$$\nabla \times H = J + \partial D / \partial t = \sigma E + \varepsilon \partial E / \partial t$$

The observation point P lies at a distance r away from the antenna. Moreover, the medium surrounding the antenna element is air. Hence $\sigma = 0$.

$$\nabla \times H = \varepsilon \partial E / \partial t$$

$$\Rightarrow E = \frac{1}{\varepsilon} \int (\nabla \times H) dt$$

$$\nabla \times H = \frac{1}{r^2 \sin\theta} \begin{vmatrix} a_r & ra_\theta & r \sin\theta a_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & \frac{\partial}{\partial \phi} \\ H_r & rH_\theta & r \sin\theta H_\phi \end{vmatrix}$$

Using eq(7), eq(8) and eq(9), $H_r = 0, H_\theta = 0$ & $H_\phi \neq 0$

From eq(12), we can observe that H_ϕ is independent of $\phi \Rightarrow \frac{\partial H_\phi}{\partial \phi} = 0$

$$\nabla \times H = \frac{1}{r^2 \sin\theta} \begin{vmatrix} a_r & ra_\theta & r \sin\theta a_\phi \\ \frac{\partial}{\partial r} & \frac{\partial}{\partial \theta} & 0 \\ 0 & 0 & r \sin\theta H_\phi \end{vmatrix}$$

$$\nabla \times H = \frac{1}{r^2 \sin\theta} \left[a_r \frac{\partial}{\partial \theta} (r \sin\theta H_\phi) - ra_\theta \frac{\partial}{\partial r} (r \sin\theta H_\phi) \right]$$

$$(\nabla \times H)_r = \frac{1}{r^2 \sin\theta} \left[\frac{\partial}{\partial \theta} (r \sin\theta H_\phi) \right] \tag{13}$$

$$(\nabla \times H)_\theta = \frac{1}{r^2 \sin\theta} \left[-r \frac{\partial}{\partial r} (r \sin\theta H_\phi) \right] = \frac{-1}{r \sin\theta} \left[\frac{\partial}{\partial r} (r \sin\theta H_\phi) \right] \tag{14}$$

$$(\nabla \times H)_\phi = 0 \Rightarrow \mathbf{E}_\phi = \mathbf{0} \tag{15}$$

Now we have to find E_r and E_θ

To find E_r

$$E_r = \frac{1}{\varepsilon} \int (\nabla \times H)_r dt$$

$$(\nabla \times H)_r = \frac{1}{r^2 \sin\theta} \left[\frac{\partial}{\partial \theta} (r \sin\theta H_\phi) \right]$$

Using eq(12),

$$(\nabla \times H)_r = \frac{1}{r^2 \sin\theta} \left[\frac{\partial}{\partial \theta} \left\{ r \sin\theta \frac{I d l \sin\theta}{4\pi} \left(- \left(\frac{\omega}{r v} \right) \sin\omega(t - r/v) + \frac{\cos\omega(t - r/v)}{r^2} \right) \right\} \right]$$

$$= \frac{1}{r^2 \sin\theta} \frac{r I dl}{4\pi} \left(- \left(\frac{\omega}{r v} \right) \sin\omega(t - r/v) + \frac{\cos\omega(t - r/v)}{r^2} \right) \left[\frac{\partial}{\partial \theta} (\sin^2\theta) \right]$$

$$= \frac{1}{r \sin\theta} \frac{I dl}{4\pi} \left(- \left(\frac{\omega}{r v} \right) \sin\omega(t - r/v) + \frac{\cos\omega(t - r/v)}{r^2} \right) [2 \sin\theta \cos\theta]$$

$$(\nabla \times H)_r = \frac{1}{r} \frac{I dl}{4\pi} \left(- \left(\frac{\omega}{r v} \right) \sin\omega(t - r/v) + \frac{\cos\omega(t - r/v)}{r^2} \right) (2 \cos\theta)$$

$$(\nabla \times H)_r = \frac{2 I dl \cos\theta}{4\pi} \left(- \left(\frac{\omega}{r^2 v} \right) \sin\omega(t - r/v) + \frac{\cos\omega(t - r/v)}{r^3} \right)$$

Using above equation,

$$E_r = \frac{1}{\varepsilon} \int (\nabla \times H)_r dt$$

$$E_r = \frac{2 I dl \cos\theta}{4\pi \varepsilon} \int \left[- \left(\frac{\omega}{r^2 v} \right) \sin\omega(t - r/v) + \frac{\cos\omega(t - r/v)}{r^3} \right] dt$$

$$= \frac{2 I dl \cos\theta}{4\pi \varepsilon} \left[- \left(\frac{\omega}{r^2 v} \right) \frac{-\cos\omega(t - r/v)}{\omega} + \frac{\sin\omega(t - r/v)}{\omega r^3} \right]$$

$$E_r = \frac{2 I dl \cos\theta}{4\pi \varepsilon} \left[\frac{\cos\omega(t - r/v)}{r^2 v} + \frac{\sin\omega(t - r/v)}{\omega r^3} \right] \quad (16)$$

To find E_θ

$$E_\theta = \frac{1}{\varepsilon} \int (\nabla \times H)_\theta dt$$

$$(\nabla \times H)_\theta = \frac{-1}{r \sin\theta} \left[\frac{\partial}{\partial r} (r \sin\theta H_\phi) \right]$$

Using eq(12),

$$(\nabla \times H)_\theta = \frac{-1}{r \sin\theta} \frac{\partial}{\partial r} \left\{ r \sin\theta \frac{I d l \sin\theta}{4\pi} \left[- \left(\frac{\omega}{r v} \right) \sin\omega(t - r/v) + \frac{\cos\omega(t - r/v)}{r^2} \right] \right\}$$

$$\begin{aligned}
(\nabla \times H)_\theta &= \frac{-1}{r \sin \theta} \sin \theta \frac{I d \sin \theta}{4\pi} \frac{\partial}{\partial r} \left[- \left(\frac{r\omega}{rv} \right) \sin \omega(t - r/v) + \frac{r \cos \omega(t - r/v)}{r^2} \right] \\
&= \frac{-I d \sin \theta}{4\pi r} \frac{\partial}{\partial r} \left[- \left(\frac{\omega}{v} \right) \sin \omega(t - r/v) + \frac{\cos \omega(t - r/v)}{r} \right] \\
&= \frac{-I d \sin \theta}{4\pi r} \left[\left(\frac{\omega}{v} \right)^2 \cos \omega(t - r/v) + \frac{-r \left(\frac{-\omega}{v} \right) \sin \omega(t - r/v) - \cos \omega(t - r/v)}{r^2} \right] \\
&= \frac{-I d \sin \theta}{4\pi r} \left[\left(\frac{\omega}{v} \right)^2 \cos \omega(t - r/v) + \frac{\left(\frac{r\omega}{v} \right) \sin \omega(t - r/v)}{r^2} - \frac{\cos \omega(t - r/v)}{r^2} \right] \\
(\nabla \times H)_\theta &= \frac{-I d \sin \theta}{4\pi r} \left[\frac{\omega^2 \cos \omega(t - r/v)}{v^2} + \frac{\omega \sin \omega(t - r/v)}{rv} - \frac{\cos \omega(t - r/v)}{r^2} \right]
\end{aligned}$$

$$E_\theta = \frac{1}{\varepsilon} \int (\nabla \times H)_\theta dt$$

$$\begin{aligned}
E_\theta &= \frac{1}{\varepsilon} \int \frac{-I d \sin \theta}{4\pi r} \left[\frac{\omega^2 \cos \omega(t - r/v)}{v^2} + \frac{\omega \sin \omega(t - r/v)}{rv} - \frac{\cos \omega(t - r/v)}{r^2} \right] dt \\
&= \frac{-I d \sin \theta}{4\pi \varepsilon r} \left[\frac{\omega^2 \sin \omega(t - r/v)}{v^2 \omega} - \frac{\omega \cos \omega(t - r/v)}{rv \omega} - \frac{\sin \omega(t - r/v)}{\omega r^2} \right] \\
&= \frac{-I d \sin \theta}{4\pi \varepsilon r} \left[\frac{\omega \sin \omega(t - r/v)}{v^2} - \frac{\cos \omega(t - r/v)}{rv} - \frac{\sin \omega(t - r/v)}{\omega r^2} \right] \\
E_\theta &= \frac{I d \sin \theta}{4\pi \varepsilon} \left[- \frac{\omega \sin \omega(t - r/v)}{rv^2} + \frac{\cos \omega(t - r/v)}{r^2 v} + \frac{\sin \omega(t - r/v)}{\omega r^3} \right]
\end{aligned}$$

Putting $t' = t - r/v$,

$$E_\theta = \frac{I d \sin \theta}{4\pi \varepsilon} \left[- \frac{\omega \sin \omega t'}{rv^2} + \frac{\cos \omega t'}{r^2 v} + \frac{\sin \omega t'}{\omega r^3} \right] \quad (17)$$

Putting $t' = t - r/v$, eq(16) becomes

$$E_r = \frac{2 I d \cos \theta}{4\pi \varepsilon} \left[\frac{\cos \omega t'}{r^2 v} + \frac{\sin \omega t'}{\omega r^3} \right] \quad (18)$$

Putting $t' = t - r/v$, eq(12) becomes

$$H_\phi = \frac{I d \sin \theta}{4\pi} \left[- \left(\frac{\omega}{rv} \right) \sin \omega t' + \frac{\cos \omega t'}{r^2} \right] \quad (19)$$

Inference

The expressions of E_r , E_θ and H_ϕ involve three types of terms:

1. The terms inversely proportional to r^3 represent *electrostatic field*.
2. The terms inversely proportional to r^2 represent *induction or near-field*.
3. The terms which are inversely proportional to r represent *radiation field (distant or far-field)*

It can be noted the magnitudes of the two bracketed terms in eq(19) will become equal if the following relation is satisfied:

$$\frac{\omega}{rv} = \frac{1}{r^2} \Rightarrow r = \frac{v}{\omega} = \frac{f\lambda}{2\pi f} = \frac{\lambda}{2\pi} \Rightarrow r \approx \frac{\lambda}{6}$$

Far-field Region or Radiation Zone

When the distance between the antenna and observation point (r) is very large, the terms $1/r^2$ and $1/r^3$ in eq(17), eq(18) and eq(19) can be neglected in favour of terms of $1/r$.

$$E_\theta \approx \frac{I d\sin\theta}{4\pi\epsilon} \left[-\frac{\omega\sin\omega t'}{rv^2} \right] \quad (20)$$

$$E_r \approx 0 \quad (21)$$

$$H_\phi \approx \frac{I d\sin\theta}{4\pi} \left[-\left(\frac{\omega}{rv}\right) \sin\omega t' \right] \quad (22)$$

The amplitudes in the far-field will be

$$|E_\theta| = \frac{I d\sin\theta}{4\pi\epsilon} \left(\frac{\omega}{rv^2}\right) \quad (23)$$

$$|H_\phi| = \frac{I d\sin\theta}{4\pi} \left(\frac{\omega}{rv}\right) = \frac{I d\sin\theta}{4\pi} \left(\frac{2\pi f}{rv}\right) = \frac{I d\sin\theta}{2r} \left(\frac{f}{v}\right) = \frac{I d\sin\theta}{2\lambda r} \quad (24)$$

Using above two equations,

$$\frac{|E_\theta|}{|H_\phi|} = \frac{\frac{I d\sin\theta}{4\pi\epsilon} \left(\frac{\omega}{rv^2}\right)}{\frac{I d\sin\theta}{4\pi} \left(\frac{\omega}{rv}\right)} = \frac{1}{v\epsilon}$$

We know that velocity of a wave in a medium is given by

$$v = \frac{1}{\sqrt{\mu\epsilon}}$$

$$\frac{|E_\theta|}{|H_\phi|} = \frac{1}{v\epsilon} = \frac{1}{\left(\frac{1}{\sqrt{\mu\epsilon}}\right)\epsilon} = \frac{\sqrt{\mu\epsilon}}{\epsilon} = \sqrt{\frac{\mu}{\epsilon}} = \eta = 120\pi \text{ (or) } 377\Omega \quad (25)$$

Where η is the intrinsic impedance of the medium.

Power Radiated and Radiation Resistance

It can be noted that E has no ϕ component and H contains only a ϕ component.

$$E = E_r a_r + E_\theta a_\theta$$

$$H = H_\phi a_\phi$$

The power flow can be given by the Poynting vector

$$S = \frac{1}{2} \text{Re} (E \times H^*) = \frac{1}{2} \text{Re} \{ (E_r a_r + E_\theta a_\theta) \times a_\phi H_\phi^* \} = \frac{1}{2} \text{Re} \{ -a_\theta E_r H_\phi^* + a_r E_\theta H_\phi^* \}$$

The total radiated power is given by

$$P_{rad} = \iiint W \cdot ds = \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} S \cdot (a_r r^2 \sin\theta \, d\theta \, d\phi)$$

$$\begin{aligned} P_{rad} &= \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} \frac{1}{2} \text{Re} \{ -a_\theta E_r H_\phi^* + a_r E_\theta H_\phi^* \} \cdot (a_r r^2 \sin\theta \, d\theta \, d\phi) \\ &= \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} \frac{1}{2} \text{Re} \{ E_\theta H_\phi^* \} r^2 \sin\theta \, d\theta \, d\phi = \frac{1}{2} \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} |E_\theta| |H_\phi^*| r^2 \sin\theta \, d\theta \, d\phi \end{aligned}$$

Since $|E_\theta| / |H_\phi| = \eta \Rightarrow |E_\theta| = \eta |H_\phi|$

$$\begin{aligned} P_{rad} &= \frac{1}{2} \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} \eta |H_\phi| |H_\phi^*| r^2 \sin\theta \, d\theta \, d\phi \\ &= \frac{1}{2} \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} \eta |H_\phi|^2 r^2 \sin\theta \, d\theta \, d\phi \\ &= \frac{1}{2} \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} \eta \left(\frac{I \, dl \sin\theta}{2\lambda r} \right)^2 r^2 \sin\theta \, d\theta \, d\phi \\ &= \frac{1}{2} \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} \eta \frac{I^2 dl^2 \sin^2\theta}{4\lambda^2 r^2} r^2 \sin\theta \, d\theta \, d\phi \\ &= \frac{\eta}{2} \frac{I^2 dl^2}{4\lambda^2} \int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} \sin^3\theta \, d\theta \, d\phi \\ &= \frac{\eta}{2} \frac{I^2}{4} \left(\frac{dl}{\lambda} \right)^2 \int_{\phi=0}^{\phi=2\pi} d\phi \int_{\theta=0}^{\theta=\pi} \sin^3\theta \, d\theta \\ &= \frac{\eta}{2} \frac{I^2}{4} \left(\frac{dl}{\lambda} \right)^2 (2\pi) \int_{\theta=0}^{\theta=\pi} \sin^3\theta \, d\theta \end{aligned}$$

$$P_{rad} = \frac{\eta}{2} \frac{I^2}{4} \left(\frac{dl}{\lambda}\right)^2 (2\pi) \int_{\theta=0}^{\theta=\pi} \sin^3\theta \, d\theta$$

Using the following identity, above equation can be re-written as

$$\begin{aligned} \sin^3\theta &= \frac{3\sin\theta - \sin 3\theta}{4} \\ P_{rad} &= \frac{\eta}{2} \frac{I^2}{4} \left(\frac{dl}{\lambda}\right)^2 (2\pi) \int_{\theta=0}^{\theta=\pi} \left(\frac{3\sin\theta - \sin 3\theta}{4}\right) d\theta \\ &= \frac{\eta}{2} \frac{I^2}{4} \left(\frac{dl}{\lambda}\right)^2 \frac{(2\pi)}{4} \left[-3\cos\theta + \frac{\cos 3\theta}{3}\right]_0^\pi \\ &= \frac{\eta}{2} \frac{I^2}{4} \left(\frac{dl}{\lambda}\right)^2 \frac{(2\pi)}{4} \left[3 - \frac{1}{3} + 3 - \frac{1}{3}\right] \\ &= \frac{\eta}{2} \frac{I^2}{4} \left(\frac{dl}{\lambda}\right)^2 \frac{(2\pi)}{4} \left[\frac{16}{3}\right] \\ P_{rad} &= \eta \frac{\pi}{3} I^2 \left(\frac{dl}{\lambda}\right)^2 \end{aligned}$$

Assuming the antenna is lossless, the power radiated by the dipole will be equal to the power delivered to the dipole.

$$P_{rad} = P_{delivered}$$

$$\eta \frac{\pi}{3} I^2 \left(\frac{dl}{\lambda}\right)^2 = I_{rms}^2 R_r$$

We know that

$$I_{rms} = \frac{I}{\sqrt{2}}$$

$$\eta \frac{\pi}{3} I^2 \left(\frac{dl}{\lambda}\right)^2 = \left(\frac{I}{\sqrt{2}}\right)^2 R_r$$

$$120\pi \frac{\pi}{3} I^2 \left(\frac{dl}{\lambda}\right)^2 = \frac{I^2}{2} R_r$$

$$R_r = 80 \pi^2 \left(\frac{dl}{\lambda}\right)^2 = 790 \left(\frac{dl}{\lambda}\right)^2$$

Directivity

$$D = \frac{4\pi}{\int \int P_n(\theta, \phi) d\Omega}$$

$$D = \frac{4\pi}{\int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} P_n(\theta, \phi) \sin\theta d\theta d\phi}$$

$$D = \frac{4\pi}{\int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} (\sin^2\theta) \sin\theta d\theta d\phi}$$

$$= \frac{4\pi}{\int_{\phi=0}^{\phi=2\pi} \int_{\theta=0}^{\theta=\pi} \sin^3\theta d\theta d\phi}$$

$$= \frac{4\pi}{\int_{\phi=0}^{\phi=2\pi} d\phi \int_{\theta=0}^{\theta=\pi} \sin^3\theta d\theta}$$

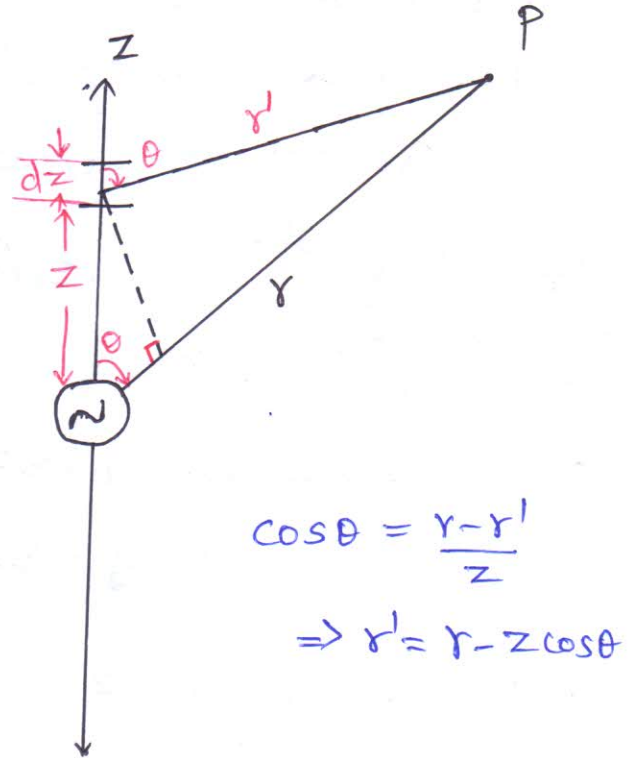
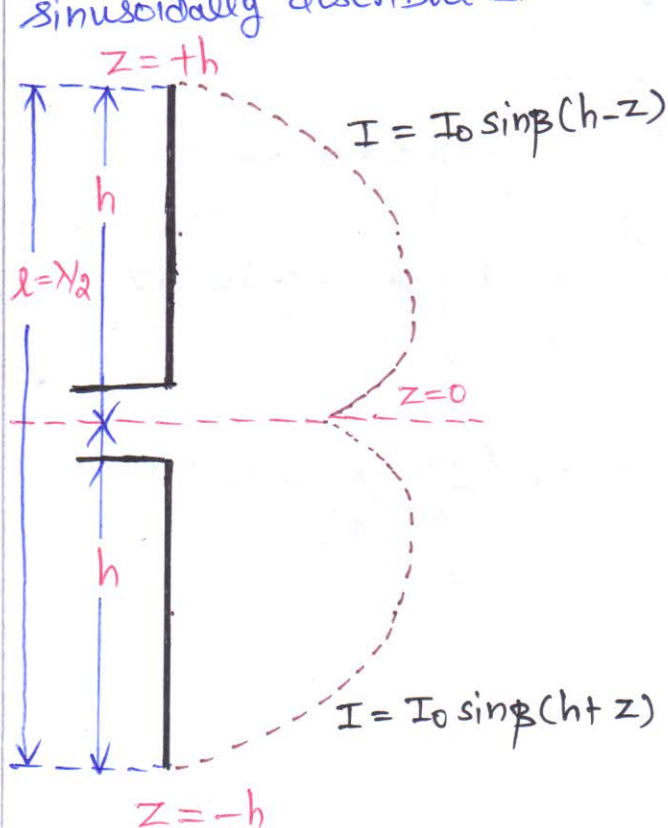
$$= \frac{4\pi}{\int_{\phi=0}^{\phi=2\pi} d\phi \int_{\theta=0}^{\theta=\pi} \sin^3\theta d\theta}$$

$$= \frac{4\pi}{(2\pi) \int_{\theta=0}^{\theta=\pi} \sin^3\theta d\theta}$$

$$D = \frac{4\pi}{(2\pi) \left(\frac{4}{3}\right)} = \frac{3}{2} = 1.5$$

Radiation from a half-wave dipole antenna:-

A $\lambda/2$ antenna is a metal rod or thin wire which has a physical length of half-wavelength in free space at the frequency of operation. It is assumed that current is sinusoidally distributed as shown in the figure.



The current distribution on the N_2 dipole is

$$I = I_0 \sin \beta (h - z) \quad , \quad 0 < z < h$$

$$I = I_0 \sin \beta (h + z) \quad , \quad -h < z < 0$$

The expression for the vector potential at a point P due to the current element $I dz$ will be

$$dA_z = \frac{\mu I e^{-j\beta r'}}{4\pi r'} dz$$

the total vector potential at p due to all the current elements will be

$$A_z = \frac{\mu}{4\pi} \int_{-h}^0 \frac{I_0 \sin \beta (h + z)}{r'} e^{-j\beta r'} dz + \frac{\mu}{4\pi} \int_0^h \frac{I_0 \sin \beta (h - z)}{r'} e^{-j\beta r'} dz$$

①

Since the distant or radiation fields are only required in this problem, it is possible to make some simplifying approximations.

For distance terms, $r' \approx r$

For phase terms, $r' \approx r - z \cos \theta$

Using above approximations, Eq ① is modified as

$$A_z = \frac{\mu I_0 e^{-j\beta r}}{4\pi r} \left[\int_{-h}^0 \sin \beta (h + z) e^{j\beta z \cos \theta} dz + \int_0^h \sin \beta (h - z) e^{j\beta z \cos \theta} dz \right]$$

$$\begin{aligned} \sin \beta (h + z) &= \sin (\beta h + \beta z) = \sin \left(\frac{\pi}{2} + \beta z \right) \\ &= \cos \beta z \end{aligned}$$

$$\sin \beta (h - z) = \cos \beta z$$

For $h = \lambda/4$,

$$A_z = \frac{\mu I_0 e^{-j\beta r}}{4\pi r} \left[\int_{-\lambda/4}^{\lambda/4} e^{j\beta z \cos\theta} \cos\beta z \cdot dz \right] \quad \text{--- (2)}$$

We know that eq (2) resembles

$$\int e^{az} \cdot \cos bz \, dz = \frac{e^{az}}{a^2 + b^2} (a \cos bz + b \sin bz)$$

where

$$a = j\beta \cos\theta \quad \text{and} \quad b = \beta$$

$$A_z = \frac{\mu I_0 e^{-j\beta r}}{4\pi r} \left[\frac{e^{j\beta z \cos\theta} (j\beta \cos\theta \cos\beta z + \beta \sin\beta z)}{-\beta^2 \cos^2\theta + \beta^2} \right]_{-\lambda/4}^{\lambda/4}$$

$$= \frac{\mu I_0 e^{-j\beta r}}{4\pi r \beta^2} \left[\frac{e^{j\beta z \cos\theta} (j \cos\theta \cos\beta z + \sin\beta z)}{\sin^2\theta} \right]_{-\lambda/4}^{\lambda/4}$$

$$A_z = \frac{\mu I_0 e^{-j\beta r}}{4\pi r \beta} \left[\frac{e^{\frac{j\pi}{2} \cos\theta} (0+1) - e^{-\frac{j\pi}{2} \cos\theta} (0-1)}{\sin^2\theta} \right]$$

$$A_z = \frac{\mu I_0 e^{-j\beta r}}{4\pi r \sin^2\theta \beta} \left[e^{\frac{j\pi}{2} \cos\theta} + e^{-\frac{j\pi}{2} \cos\theta} \right]$$

Using the identity $e^{jx} + e^{-jx} = 2 \cos x$,

$$A_z = \frac{\mu I_0 e^{-j\beta r}}{2\pi r \beta \sin^2\theta} \left[2 \cos\left(\frac{\pi}{2} \cos\theta\right) \right]$$

$$A_z = \frac{\mu I_0 e^{-j\beta r}}{2\pi r \beta} \left[\frac{\cos\left(\frac{\pi}{2} \cos\theta\right)}{\sin^2\theta} \right] \quad \text{--- (3)}$$

Maxwell Equation: $B = \nabla \times A \Rightarrow H = \frac{1}{\mu} (\nabla \times A)$

$$H_{\phi} = \frac{1}{\mu} (\nabla \times A)_{\phi} = \frac{1}{\mu} \left[\frac{1}{r} \left\{ \frac{\partial (rA_{\theta})}{\partial r} - \frac{\partial A_r}{\partial \theta} \right\} \right]$$

Far-field term alone is considered,

$$H_{\phi} = \frac{1}{\mu} \left[\frac{1}{r} \left\{ \frac{\partial (rA_{\theta})}{\partial r} \right\} \right] \text{ where } A_{\theta} = -A_z \sin \theta$$

$$= \frac{1}{\mu} \left[\frac{1}{r} \frac{\partial (-rA_z \sin \theta)}{\partial r} \right] = \frac{1}{\mu} \left[-\frac{\sin \theta}{r} \frac{\partial (rA_z)}{\partial r} \right]$$

The derivative in above eq. is found as

$$\frac{\partial (rA_z)}{\partial r} = \frac{\mu I_0}{2\pi \beta} \frac{\cos(\frac{\pi}{2} \cos \theta)}{\sin^2 \theta} \left\{ \frac{\partial}{\partial r} \left(\frac{r e^{-j\beta r}}{r} \right) \right\}$$

$$= \frac{\mu I_0}{2\pi \beta} \frac{\cos(\pi/2 \cos \theta)}{\sin^2 \theta} (-j\beta) e^{-j\beta r}$$

$$\Rightarrow H_{\phi} = \frac{1}{\mu} \left[+\frac{\sin \theta}{r} \cdot \frac{\mu I_0}{2\pi} \frac{\cos(\pi/2 \cos \theta)}{\sin^2 \theta} (+j) e^{-j\beta r} \right]$$

$$H_{\phi} = j \frac{I_0}{2\pi r} \frac{\cos(\pi/2 \cos \theta)}{\sin \theta} e^{-j\beta r} \quad \text{--- (4)}$$

we know that $H = \frac{E_{\theta}}{H_{\phi}} \Rightarrow E_{\theta} = H H_{\phi} = 120\pi \times H_{\phi}$

$$E_{\theta} = 120\pi \left[\frac{j I_0}{2\pi r} \frac{\cos(\pi/2 \cos \theta)}{\sin \theta} e^{-j\beta r} \right]$$

$$E_{\theta} = j \frac{60 I_0}{r} e^{-j\beta r} \frac{\cos(\pi/2 \cos \theta)}{\sin \theta} \quad \text{--- (5)}$$

Time averaged power density is

$$P_{av} = \frac{1}{2} \mathbf{E} \times \mathbf{H}^* \Rightarrow P_{av} = \frac{1}{2} |E_\theta| |H_\phi|^* \hat{a}_r \sin 90^\circ$$

(E_θ and H_ϕ are in phase and orthogonal)

$$P_{av} = \frac{1}{2} \eta |H_\phi|^2 \hat{a}_r = \frac{\eta}{2} \left[\frac{I_0}{2\pi r} \frac{\cos(\pi/2 \cos\theta)}{\sin\theta} \right]^2 \hat{a}_r$$

(since $|e^{j\beta r} e^{-j\beta r}| = 1$)

$$P_{av} = \eta \frac{I_0^2}{8\pi^2 r^2} \frac{\cos^2(\pi/2 \cos\theta)}{\sin^2\theta} \hat{a}_r \quad \text{--- (6)}$$

The total power radiated through a hemispherical surface of radius r will equal

$$P_{rad} = \oint P_{av} \cdot d\mathbf{s} = \int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \frac{\eta I_0^2}{8\pi^2 r^2} \frac{\cos^2(\pi/2 \cos\theta)}{\sin^2\theta} \sin\theta d\theta d\phi$$

$$= \eta \frac{I_0^2}{8\pi^2} \int_{\phi=0}^{2\pi} d\phi \int_{\theta=0}^{\pi} \frac{\cos^2(\pi/2 \cos\theta)}{\sin\theta} d\theta$$

$$= \eta \frac{I_0^2}{2 \cdot 8\pi^2} (2\pi) \times 2 \times \int_{\theta=0}^{\pi/2} \frac{\cos^2(\pi/2 \cos\theta)}{\sin\theta} d\theta$$

$$P_{rad} = \frac{\eta I_0^2}{2\pi} \times 0.609 = I_{rms}^2 R_r = \left(\frac{I_0}{\sqrt{2}} \right)^2 R_r = \frac{I_0^2 R_r}{2}$$

$$\Rightarrow R_r = 2 \times \frac{\eta}{2\pi} \times 0.609 = \frac{120\pi}{\pi} \times 0.609$$

$$R_r = 73 \Omega$$

for a half-wave dipole, $Z = 73 + j42.5 \Omega$

The Directivity of $\lambda/2$ dipole antenna is

$$\begin{aligned} D &= \frac{4\pi}{\int_{\phi=0}^{2\pi} \int_{\theta=0}^{\pi} \frac{P(\theta, \phi)}{P(\theta, \phi)_{\max}} d\Omega} = \frac{4\pi}{\int_{\phi=0}^{2\pi} d\phi \int_{\theta=0}^{\pi} \frac{\cos^2(\pi/2 \cos\theta)}{\sin\theta} d\theta} \\ &= \frac{4\pi}{2\pi \times 2 \times \int_{\theta=0}^{\pi/2} \frac{\cos^2(\pi/2 \cos\theta)}{\sin\theta} d\theta} = \frac{4\pi}{2\pi \times 2 \times 0.609} \end{aligned}$$

$$D = 1.642$$

EC8701

**ANTENNAS AND MICROWAVE
ENGINEERING**

Course Instructor: M.Lingeshwaran M.E.,(Ph.D.)

UNIT II

RADIATION MECHANISMS AND DESIGN ASPECTS

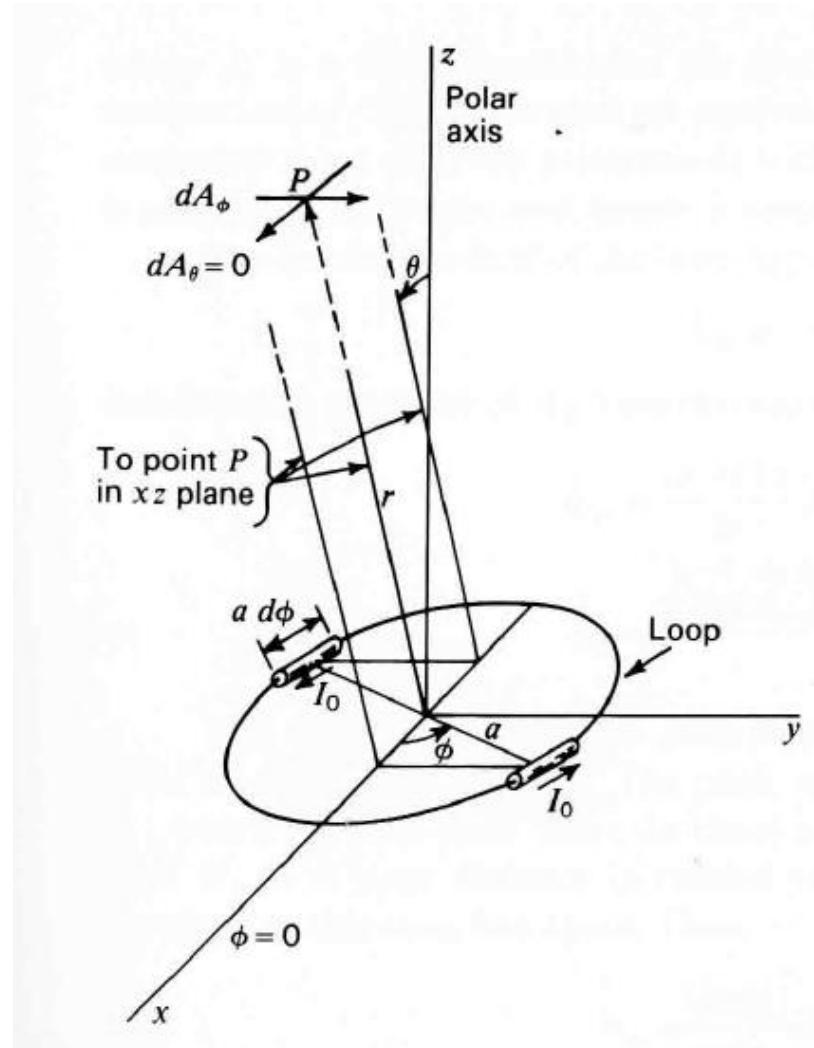
Design considerations and applications

- Radiation Mechanisms of Linear Wire and Loop antennas
- Aperture antennas
- Reflector antennas
- Microstrip antennas
- Frequency independent antennas

Loop Antenna

- Let us assume a loop antenna of radius a which is large enough i.e., the circumference is almost equal to one wavelength.
- The center of the large loop is at the origin of the spherical co-ordinate system.
- It is assumed that the current in the loop is *uniform* and *in-phase*.
- The far-field components can be derived from vector magnetic potential which is a function of current distribution over the loop.
- As shown in the following figure, we assume two small current carrying elements (or short dipoles) located diametrically opposite on the loop.
- If these dipoles are too short and makes an angle of $d\phi$ at the center of the loop then the length of each short dipole is given as $a d\phi$.
- As the loop is lying in the xy -plane only (current is confined to the loop) , hence, the vector magnetic potential will have only ϕ component (A_ϕ) and other components are zero ($A_r = A_\theta = 0$).

Loop Antenna



Loop in spherical co-ordinate system

Loop Antenna

- The observation point P is at distance r from origin, hence, the infinitesimal component of the vector magnetic potential in ϕ direction due to diametrically opposite short dipoles is given as

$$dA_{\phi} = \frac{\mu dM}{4\pi r}$$

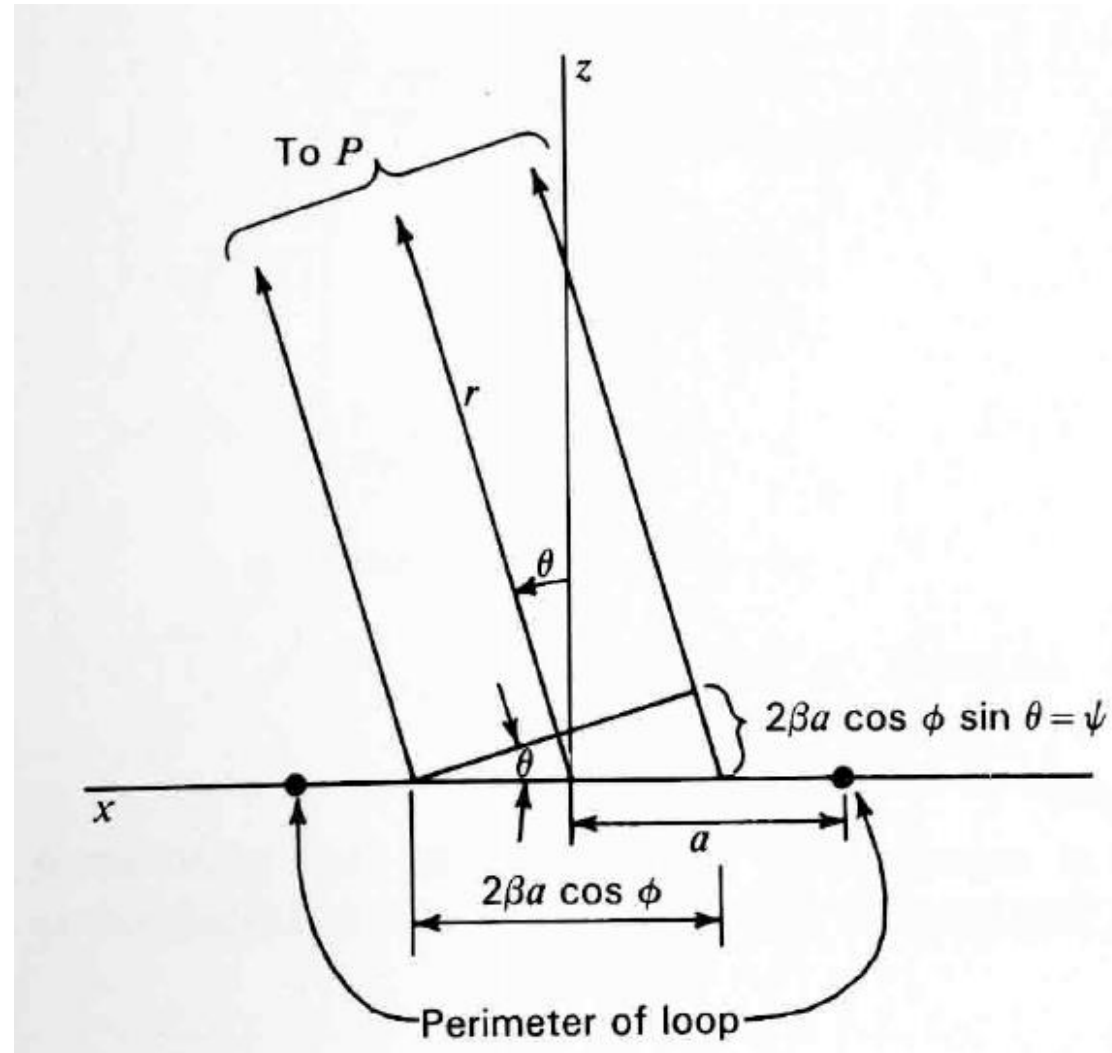
where dM is the current moment produced by short dipoles of length $a d\phi$

- In xz -plane or ($\phi = 0^{\circ}$ plane), the ϕ component of retarded current moment due to one short dipole is defined as

$$dM_1 = [I](a d\phi) \cos\phi$$

where $[I] = I_0 e^{j\omega[t-(r/c)]}$ and I_0 is the peak current on the loop.

Loop Antenna



Cross section of the loop in xz -plane

Loop Antenna

- The resultant current moment due to two short dipoles is

$$dM = 2 j [I] a d\phi \cos\phi \sin\frac{\psi}{2}$$

where $\psi = 2\beta a \cos\phi \sin\theta$ and $\beta = \frac{2\pi}{\lambda}$

- By substituting the value of ψ in above expression, the resultant current moment comes out as

$$dM = 2 j [I] a \cos\phi [\sin(\beta a \cos\phi \sin\theta)]d\phi$$

- This expression can be substituted in dA_ϕ ,

$$dA_\phi = \frac{\mu dM}{4\pi r} = \frac{\mu \{2 j [I] a \cos\phi [\sin(\beta a \cos\phi \sin\theta)]d\phi\}}{4\pi r}$$

$$dA_\phi = \frac{j\mu [I] a}{2\pi r} \cos\phi [\sin(\beta a \cos\phi \sin\theta)]d\phi$$

Loop Antenna

- Integrating dA_ϕ yields,

$$A_\phi = \frac{j\mu [I] a}{2\pi r} \int_{\phi=0}^{\phi=\pi} \sin(\beta a \sin\theta \cos\phi) \cos\phi d\phi$$

$$\Rightarrow A_\phi = \frac{j\mu [I] a}{2r} J_1(\beta a \sin\theta)$$

where J_1 is a Bessel function of the first order and of argument $(\beta a \sin\theta)$.

- Note

$$J_1(\beta a \sin\theta) = \frac{1}{\pi} \int_{\phi=0}^{\phi=\pi} \sin(\beta a \sin\theta \cos\phi) \cos\phi d\phi$$

Loop Antenna

- The far electric field of the loop has only ϕ component given by

$$E_{\phi} = -j\omega A_{\phi}$$

- Substituting A_{ϕ} in above expression yields

$$E_{\phi} = -j\omega A_{\phi} = -j\omega \left\{ \frac{j\mu [I] a}{2r} J_1(\beta a \sin\theta) \right\} \Rightarrow \boxed{E_{\phi} = \frac{\mu\omega [I] a}{2r} J_1(\beta a \sin\theta)}$$

Since $\omega = 2\pi f$, $c = f\lambda$ and $\beta = \frac{2\pi}{\lambda}$

$$E_{\phi} = \frac{\mu (2\pi f) [I] a}{2r} \times \left(\frac{\lambda}{\lambda} \right) \times J_1(\beta a \sin\theta) = \frac{2\pi}{\lambda} \times f\lambda \times \mu \times \frac{[I] a}{2r} J_1(\beta a \sin\theta)$$

$$E_{\phi} = \beta c \mu \frac{[I] a}{2r} J_1(\beta a \sin\theta)$$

Loop Antenna

$$E_{\phi} = \beta c \mu \frac{[I]a}{2r} J_1(\beta a \sin\theta)$$

- Substituting the values of c and μ in above expression yields

$$E_{\phi} = (3 \times 10^8) \times (4\pi \times 10^{-7}) \frac{\beta a [I]}{2r} J_1(\beta a \sin\theta)$$

$$E_{\phi} = \frac{60\pi\beta a [I]}{r} J_1(\beta a \sin\theta)$$

- This expression gives the instantaneous electric field at a distance r from a loop of any radius a .
- The magnetic field H_{θ} at a large distance is related to E_{ϕ} by intrinsic impedance of the free space η . Thus,

$$H_{\theta} = \frac{E_{\phi}}{\eta} = \frac{E_{\phi}}{120\pi} = \frac{\beta a [I]}{2r} J_1(\beta a \sin\theta)$$

Far-field Patterns of Circular Loop Antenna

- The far-field patterns for a loop of any size is given by

$$E_{\phi} = \frac{60\pi\beta a [I]}{r} J_1(\beta a \sin\theta) \text{ and } H_{\theta} = \frac{\beta a [I]}{2r} J_1(\beta a \sin\theta)$$

- For a loop of a given size, βa is constant and shape of the far-field pattern is given as a function of θ by

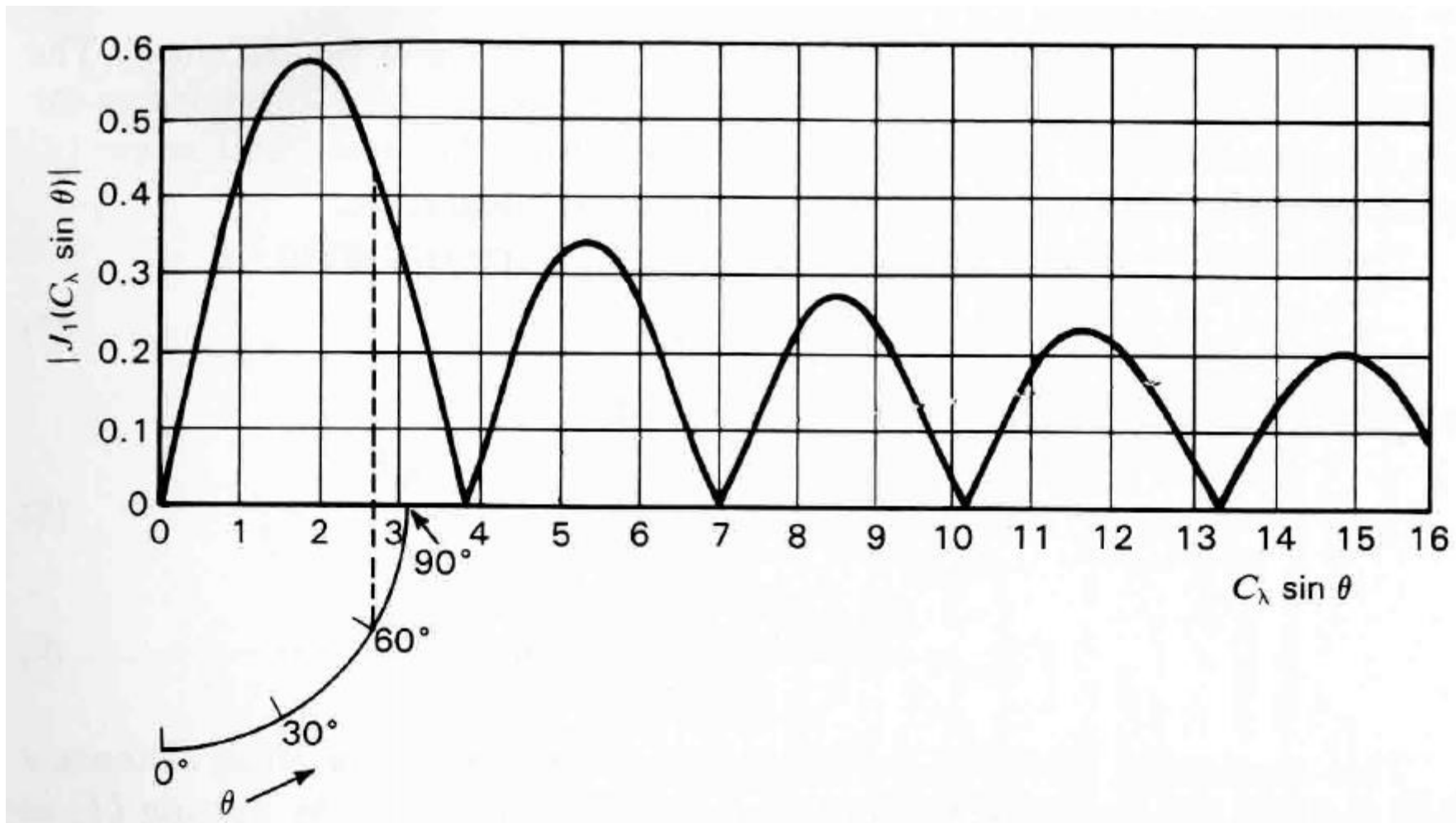
$$J_1(C_{\lambda} \sin \theta)$$

where C_{λ} is the circumference of the loop in wavelengths. That is,

$$C_{\lambda} = \frac{2\pi}{\lambda} a = \beta a$$

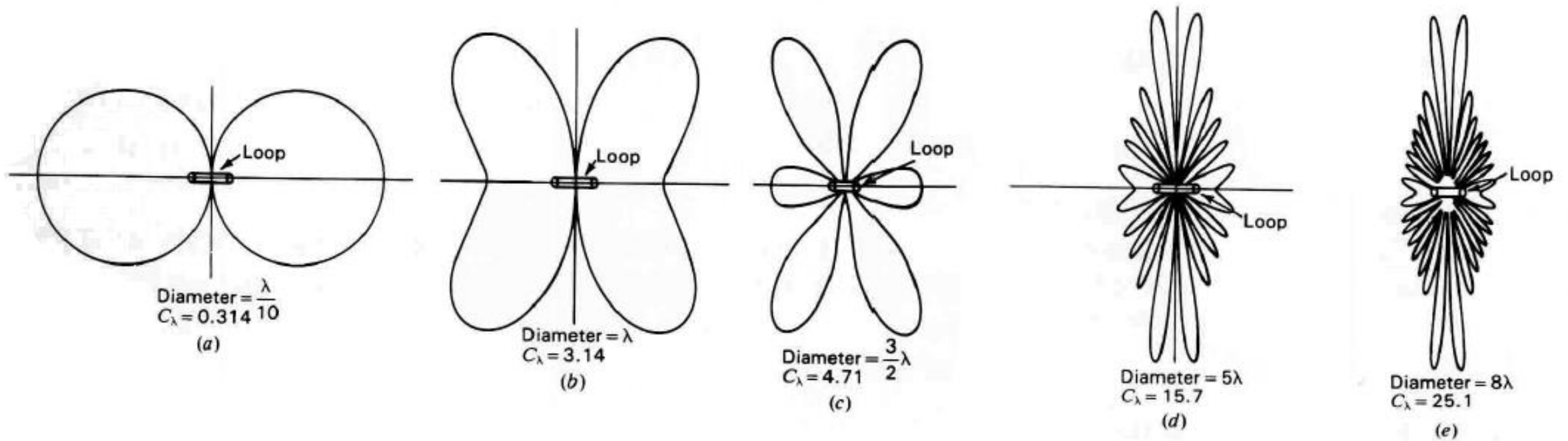
- $0 \leq |\sin \theta| \leq 1$
- When $\theta = 90^{\circ}$, the relative field is $J_1(C_{\lambda})$
- As θ decreases to zero, the values of the relative field vary in accordance with the J_1 curve from $J_1(C_{\lambda})$ to zero.

Far-field Patterns of Circular Loop Antenna



Rectified first-order Bessel curve for patterns of loops

Far-field Patterns of Circular Loop Antenna



Far-field patterns of loops of $0.1, 1, 1.5, 5$ and 8λ diameter.

Uniform in-phase current is assumed on the loops

Small Loop as a Special Case

- The far-field patterns for a loop of any size is given by

$$E_{\phi} = \frac{60\pi\beta a [I]}{r} J_1(\beta a \sin\theta)$$
$$H_{\theta} = \frac{\beta a [I]}{2r} J_1(\beta a \sin\theta)$$

- For small arguments of the first-order Bessel function, the following approximate relation can be used

$$J_1(x) = \frac{x}{2}$$

- When $x = \frac{1}{3}$, the approximation is about 1% in error.
- The relation becomes exact as x approaches to zero.

Small Loop as a Special Case

- Thus, if the perimeter of the loop is $\lambda/3$ or less ($C_\lambda < \frac{\lambda}{3}$), far-field equations for a small loop is

$$E_\phi = \frac{60\pi\beta a [I]}{r} \left(\frac{\beta a \sin\theta}{2} \right) = \frac{60\pi\beta^2 a^2 [I]}{2r} (\sin\theta) = \frac{60 \left(\frac{2\pi}{\lambda} \right)^2 \pi a^2 [I]}{2r} (\sin\theta)$$

$$E_\phi = \frac{120 \pi^2 [I] \sin \theta}{r} \frac{A}{\lambda^2}$$

$$H_\theta = \frac{\beta a [I]}{2r} \left(\frac{\beta a \sin\theta}{2} \right) = \frac{\beta^2 a^2 [I] \sin\theta}{4r} = \frac{\left(\frac{2\pi}{\lambda} \right)^2 a^2 [I] \sin\theta}{4r}$$

$$H_\theta = \frac{\pi [I] \sin\theta}{r} \frac{A}{\lambda^2}$$

Aperture Antennas

- The term *aperture* refers to an opening in an otherwise closed surface.
- As applied to antennas, aperture antennas represent a class of antennas that are generally analyzed considering the antenna as an opening in a surface.
- Typical antennas that fall in this category are the *slot, horn, reflector, and lens antennas*.
- Aperture antennas are most common at *microwave frequencies*.
- Aperture antennas are very practical for *space applications*, because they can be flush mounted on the surface of the spacecraft or aircraft.
- Their opening can be covered with a dielectric material to protect them from environmental conditions. This type of mounting does not disturb the aerodynamic profile of the craft, which in high-speed applications is critical.

Wire Antennas Vs Aperture Antennas

- The radiation characteristics of wire antennas can be determined once the current distribution on the wire is known.
- For many configurations, however, *the current distribution is not known exactly* and only physical intuition or experimental measurements can provide a reasonable approximation to it. This is even more evident in aperture antennas (slits, slots, waveguides, horns, reflectors, lenses).
- The central idea used in the analysis of aperture type antennas is the conversion of the original antenna geometry into an equivalent geometry which can be looked at as radiation through an aperture in a closed surface.
- In case of wire antennas, the fields of an antenna are expressed in terms of the vector potential which, in turn, is an integral over the current distribution on the antenna surface. The major difficulty in computing the fields of an aperture type antenna is the *integration over a complex surface of the antenna*.

Wire Antennas Vs Aperture Antennas

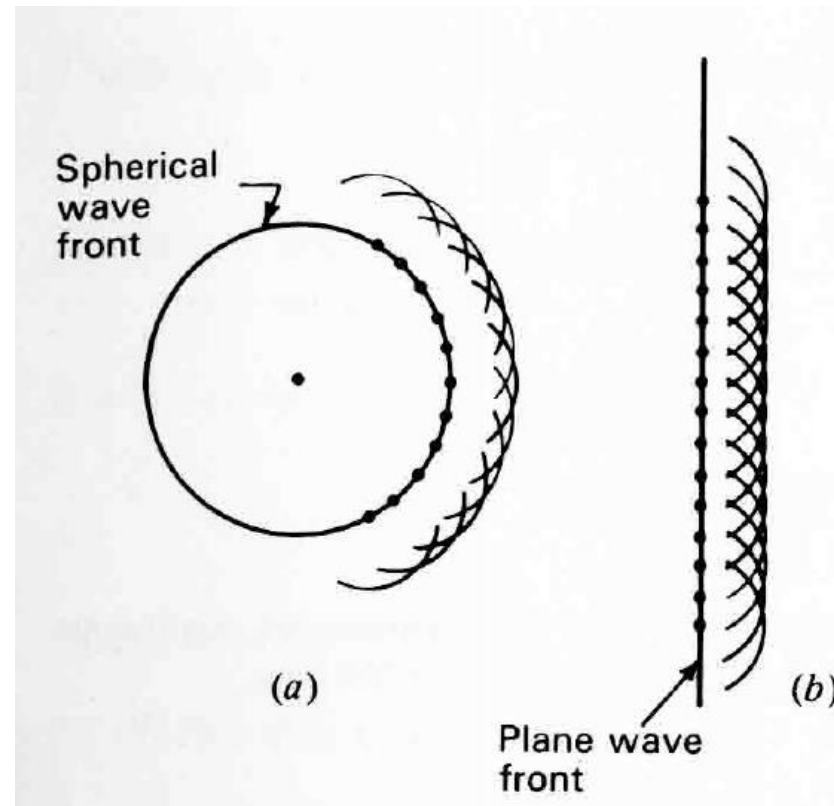
- This issue is somewhat simplified by converting the antenna into an aperture problem, using the *field equivalence principle*.
- The aperture geometry is conveniently chosen as some regular surface so that the integration can be carried out with much less effort.
- All that is needed is *the knowledge of the tangential E or H fields in the aperture to compute the far-fields of the antenna*
- Obviously, some approximation is involved in determining the tangential fields in the aperture, but in general, the computed far-fields are fairly accurate for all practical purposes, if sufficient care is taken in arriving at this approximation.

Field Equivalence Principle

- The field equivalence is a principle by which actual sources, such as an antenna and transmitter, are replaced by equivalent sources. The fictitious sources are said to be *equivalent within a region* because they *produce the same fields within that region*.
- FEP is based on
 - Huygens' Principle
 - Uniqueness Theorem

Huygens' Principle

- *“Each point on a primary wavefront can be considered to be a new source of a secondary spherical wave and that a secondary wave front can be constructed as the envelope of these secondary waves”.*



Uniqueness Theorem

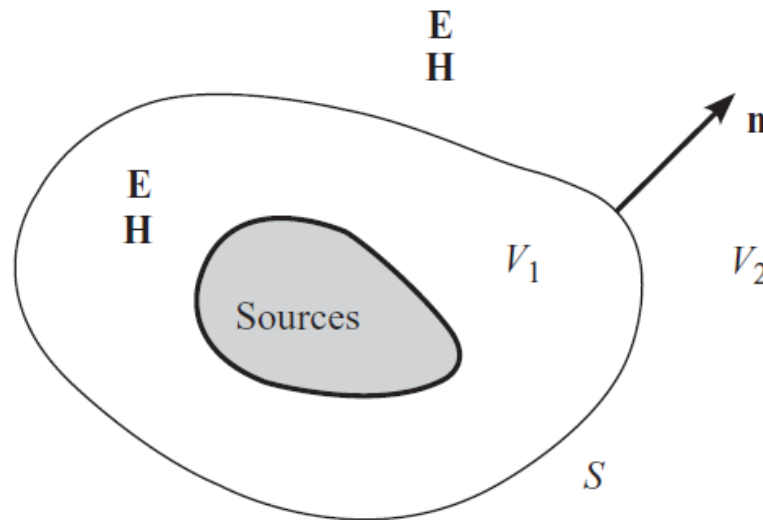
- *“A field in a lossy region is uniquely specified by the sources within the region plus the tangential components of the electric field over the boundary, or the tangential components of the magnetic field over the boundary, or the former over part of the boundary and the latter over the rest of the boundary”.*
- *“For a given set of sources and boundary conditions in a lossy medium, the solution to Maxwell’s equations is unique”.*
- Consider a source-free volume \mathbf{V} in an isotropic homogeneous medium bounded by a surface \mathbf{S} , and let $(\mathbf{E}_1, \mathbf{H}_1)$ be the fields inside it produced by a set of sources external to the volume.
- Now, let $(\mathbf{E}_2, \mathbf{H}_2)$ be another possible set of fields in the volume \mathbf{V} .
- It can be shown that if either the tangential \mathbf{E} or the tangential \mathbf{H} is the same on the surface \mathbf{S} for the two sets of solutions, the fields are identical everywhere in the volume. This is known as the uniqueness theorem.

Uniqueness Theorem

- It is important to note that it is sufficient to equate either the tangential E or the tangential H on S for the solution to be unique.
- In other words, in a source-free region the fields are completely determined by the tangential E or the tangential H on the bounding surface.
- Although the uniqueness theorem is derived for a dissipative medium, one can prove the theorem for a lossless medium by a limiting process as loss tends to zero.

Love's Field Equivalence Principle

- Consider a set of current sources in a homogeneous isotropic medium producing electromagnetic fields \mathbf{E} and \mathbf{H} everywhere.
- Enclose all the sources by a closed surface S , separating the entire space into two parts, volume V_1 containing the sources and the volume V_2 being source-free.
- Let the surface S be chosen such that it is also source-free.
- Let $\hat{\mathbf{n}}$ be a unit normal to the surface drawn from V_1 into V_2 .



Fields and Sources

Love's Field Equivalence Principle

- According to the field equivalence principle, the fields in V_2 due to the sources in volume V_1 can also be generated by an equivalent set of virtual sources on surface S , given by

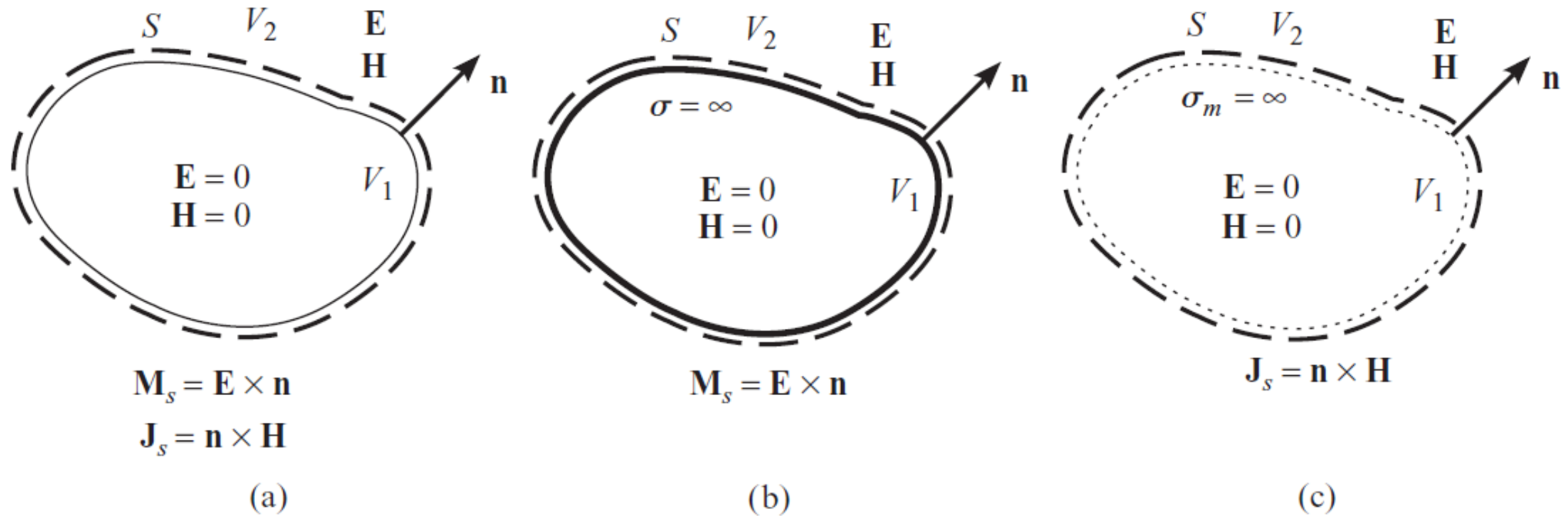
$$\mathbf{J}_S = \hat{\mathbf{n}} \times \mathbf{H}$$

$$\mathbf{M}_S = -\hat{\mathbf{n}} \times \mathbf{E} = \mathbf{E} \times \hat{\mathbf{n}}$$

where \mathbf{E} and \mathbf{H} are the fields on the surface S produced by the original set of sources in volume V_1 .

- Further the set of virtual sources produce null fields everywhere in V_1 .
- Here \mathbf{M}_S represents the magnetic surface current density and \mathbf{J}_S the electric surface current density.

Love's Field Equivalence Principle



Three forms of the field equivalence principle: (a) surface current densities \mathbf{J}_s and \mathbf{M}_s on the surface S , (b) surface current density \mathbf{M}_s alone on the surface S , which is a conducting surface, and (c) surface current density \mathbf{J}_s alone on the surface S , which is a magnetic conductor surface

Love's Field Equivalence Principle

- The proof of this principle makes use of the uniqueness theorem.
- Consider a situation where the fields in volume V_2 are the same as before, (\mathbf{E}, \mathbf{H}) , but we delete all the sources in V_1 and assume the fields to be identically zero everywhere in V_1 .
- At the boundary surface, S , the fields are discontinuous and, hence, cannot be supported unless we introduce sources on the discontinuity surface.
- Specifically, we introduce surface current sheets on S , such that $\mathbf{J}_S = \hat{\mathbf{n}} \times \mathbf{H}$ and $\mathbf{M}_S = \mathbf{E} \times \hat{\mathbf{n}}$, so that the boundary conditions are satisfied.
- Since the tangential \mathbf{E} and \mathbf{H} satisfy the boundary conditions, it is a solution of Maxwell's equations, and from the uniqueness theorem, it is the only solution.
- Thus, the original sources in V_1 and the new set of surface current sources produce the same fields (\mathbf{E}, \mathbf{H}) in the volume V_2 . Therefore, these are equivalent problems as far as the fields in the volume V_2 are concerned.

Love's Field Equivalence Principle

- Let us assume that the medium inside S is replaced by a perfect electric conductor ($\sigma = \infty$). The *introduction of the perfect electric conductor shorts out electric current density ($J_S = 0$)* and there exists only a magnetic current density M_S as shown in Fig.(b).
- Let us assume that instead of placing a perfect electric conductor within S , we *introduce a perfect magnetic conductor which will short out the magnetic current density* and reduce the equivalent problem to that shown in Fig.(c).

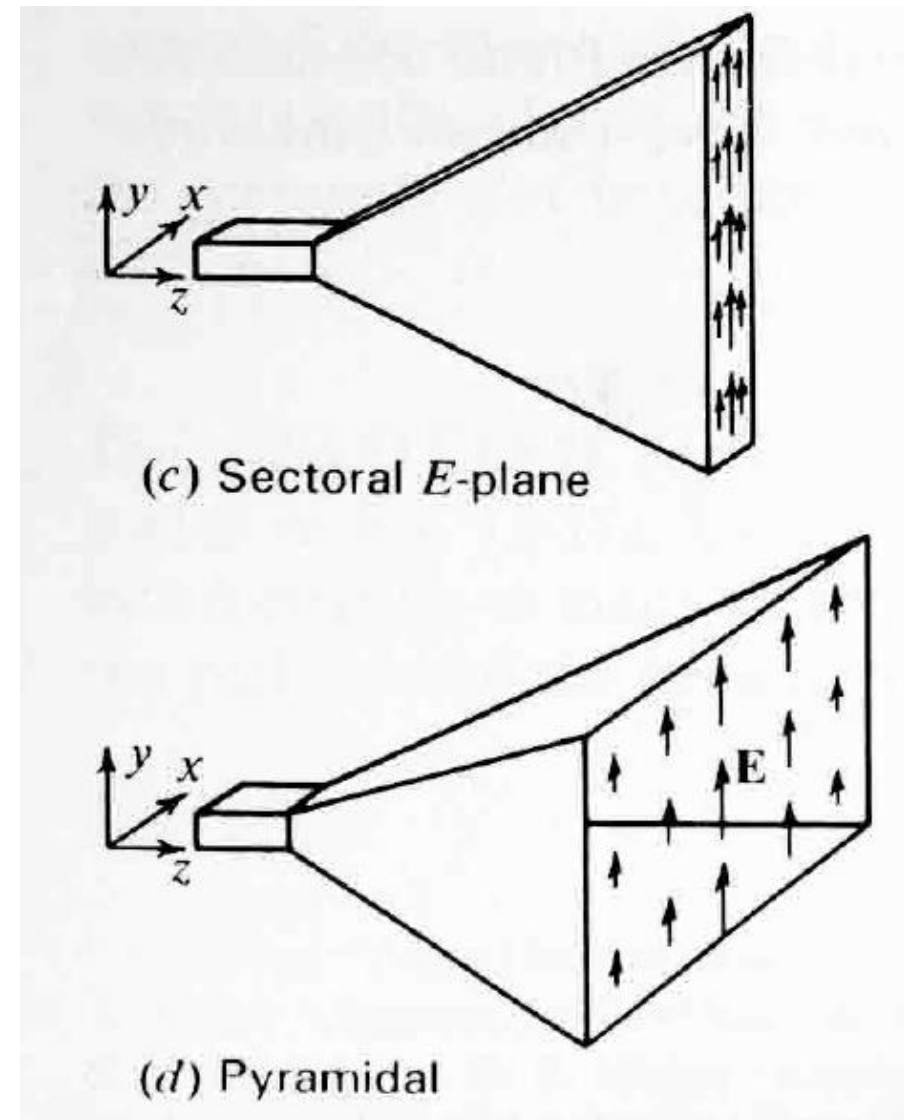
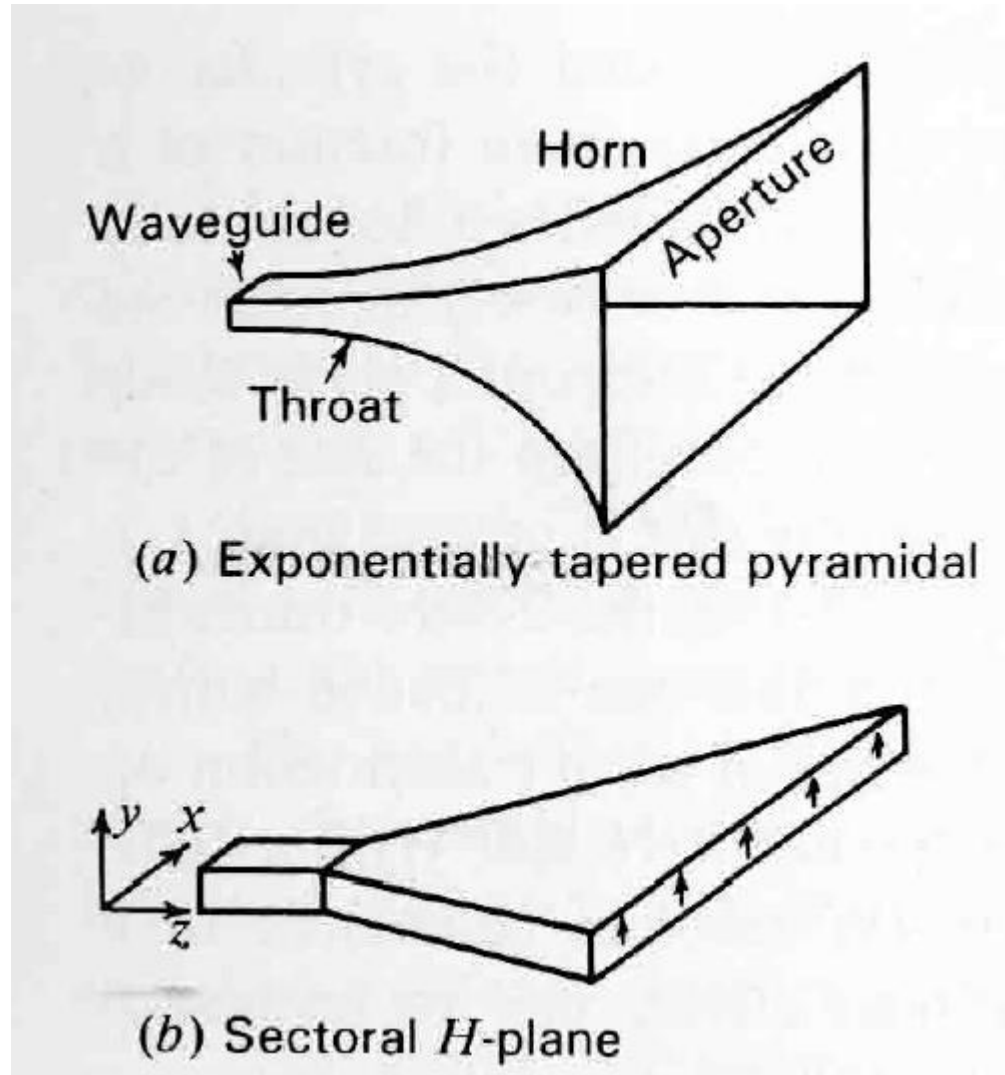
Horn Antennas

- Microwave Antenna
- Used as a feed element for large radio astronomy, satellite tracking and communication dishes.
- Used as a feed for reflectors and lenses
- It is a common element of phased arrays and serves as a universal standard for calibration and gain measurements of other high-gain antennas.
- Its widespread applicability stems from its
 - simplicity in construction
 - ease of excitation
 - versatility
 - large gain and preferred overall performance.

Horn Antennas

- A horn antenna may be regarded as a *flared-out (or opened-out) waveguide*.
- Function of the horn is to produce a uniform phase front with a large aperture than that of the waveguide and hence *greater directivity*.
- Jagadis Chandra Bose constructed a pyramidal horn in 1897.
- Horn antennas (Rectangular/Circular) are energized from waveguides.
- To minimize the reflections of the guided wave, the transition region or horn between the waveguide at the throat and free space at the aperture could be given a *gradual exponential taper* as in **fig(a)** or **fig(e)**.
- Assuming that the rectangular waveguide is energized with a TE_{10} mode wave electric field (\mathbf{E} in the y direction), the horn in **fig(b)** is flared out in a plane perpendicular to \mathbf{E} . This is the plane of the magnetic field \mathbf{H} . Hence, this type of horn is called a sectoral horn flared in H plane or simply an ***H-plane sectoral horn***.

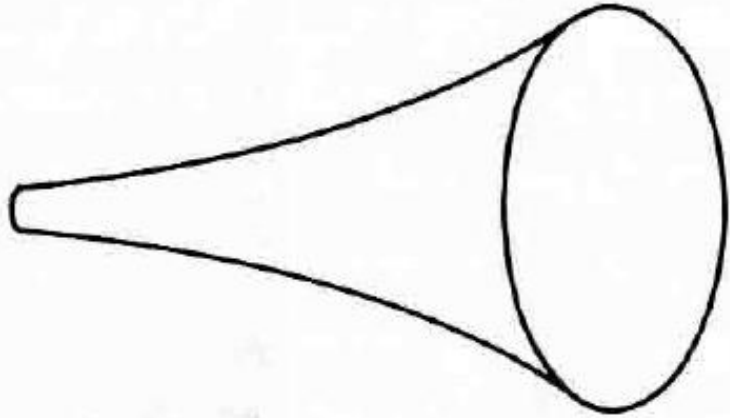
Rectangular Horn Antennas



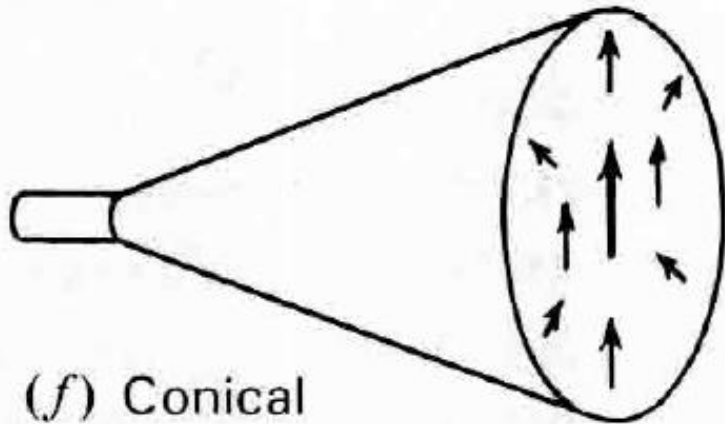
Horn Antennas

- The horn in **fig(c)** is flared out in the plane of the electric field \mathbf{E} , and hence is called an *E-plane sectoral horn*.
- A rectangular horn with flare in both planes, as in **fig(d)**, is called a *pyramidal horn*.
- Horn shown in **fig(f)** is a *conical type*.
- When excited with a circular waveguide carrying a TE_{11} mode wave, the electric field distribution at the aperture is as shown by the arrows.
- Horn in **fig(g)** and **fig(h)** are *biconical* types.
- TEM Biconical antenna is excited in the TEM mode by a vertical radiator.
- TE_{01} Biconical antenna is excited in the TE_{01} mode by a small horizontal loop antenna
- Biconical horns are non-directional in the horizontal plane.

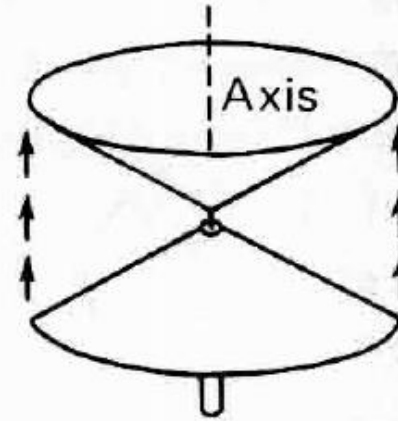
Circular Horn Antennas



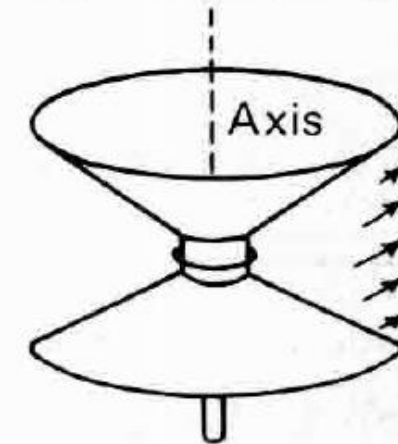
(e) Exponentially tapered



(f) Conical



(g) TEM biconical

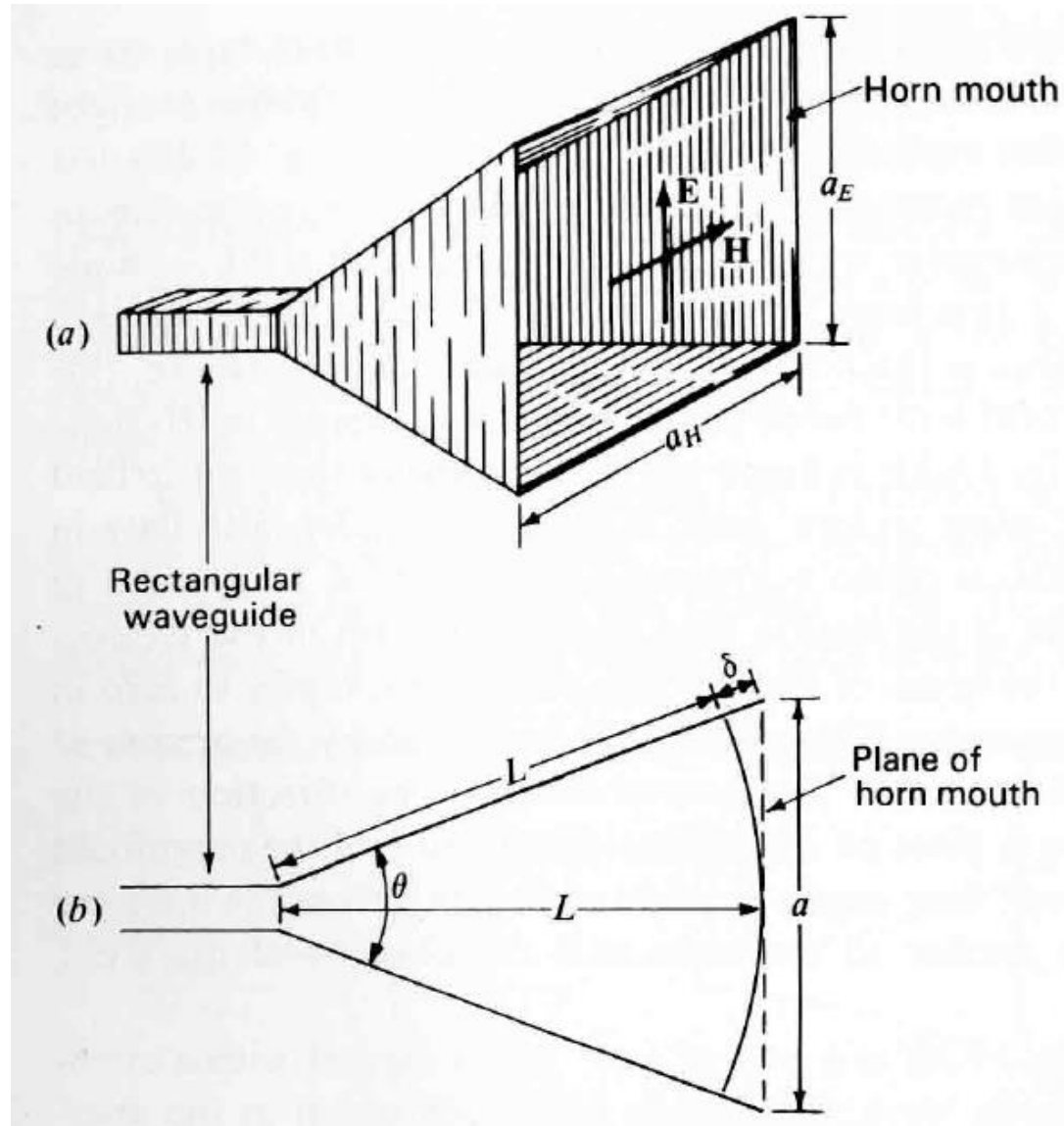


(h) TE_{01} biconical

Horn Antennas – Design Principle

- The *principle of equality of path length* (*Fermat's principle*) is applicable to the horn design but with a different emphasis.
- *Instead of requiring a constant phase across the horn mouth, the requirement is relaxed to a one where the phase may deviate, but by less than a specified amount δ , equal to the path length difference between a ray traveling along the side and along the axis of the horn.*
- Design parameters:
 - E-plane : θ_E is the flare angle (deg) and a_E is the aperture dimension (m)
 - H-pane: θ_H is the flare angle (deg) and a_H is the aperture dimension (m)
 - L = Horn Length (m)
 - δ = path length difference (m)

Horn Antennas – Design Principle



(a) Pyramidal horn antenna

(b) Cross section with dimensions used in analysis

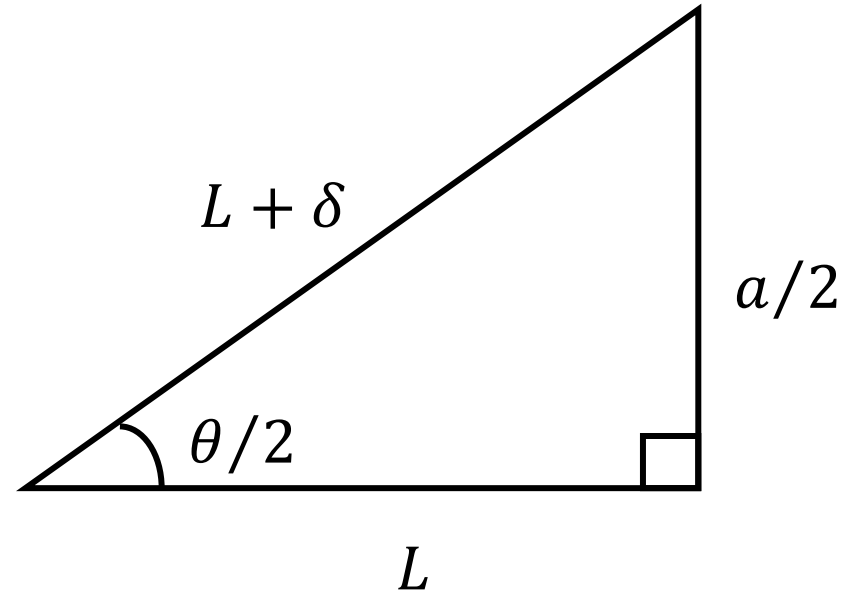
Horn Antennas – Design Equations

- From the figure,

$$\cos \frac{\theta}{2} = \frac{L}{L + \delta}$$

$$\sin \frac{\theta}{2} = \frac{a/2}{L + \delta} = \frac{a}{2(L + \delta)}$$

$$\tan \frac{\theta}{2} = \frac{\frac{a}{2(L + \delta)}}{\frac{L}{L + \delta}} = \frac{a}{2L}$$



- From the geometry, we have also that

$$(L + \delta)^2 = L^2 + \left(\frac{a}{2}\right)^2 \Rightarrow L^2 + \delta^2 + 2L\delta = L^2 + \frac{a^2}{4} \Rightarrow \delta^2 + 2L\delta = \frac{a^2}{4}$$

$$\Rightarrow L = \frac{a^2}{8\delta} \quad (\delta \ll L) \text{ and } \theta = 2 \tan^{-1} \frac{a}{2L} = 2 \cos^{-1} \frac{L}{L + \delta}$$

Horn Antennas – Design Constraints

- In E-plane of horn, $\delta \leq 0.25\lambda$
- In H-plane, $\delta \approx 0.4\lambda$ (Since \mathbf{E} goes to zero at the horn edges)
- To obtain as uniform an aperture distribution as possible, a very long horn with a small flare angle is required. Practically, the horn should be as short as possible.
- If δ is a sufficiently small fraction of a wavelength, the field has nearly uniform phase over the entire aperture.
- For a constant length L , the directivity of the horn increases (beamwidth decreases) as the aperture a and flare angle θ are increased.
- However, if the aperture and flare angle become so large that $\delta \cong 180^\circ$, the field at the edge of the aperture is in phase opposition to the field on the axis.
- For all but very large flare angles, the ratio $L/(L + \delta)$ is so nearly unity that the effect of the additional path length δ on the distribution of the field magnitude can be neglected.

Horn Antennas – Design Constraints

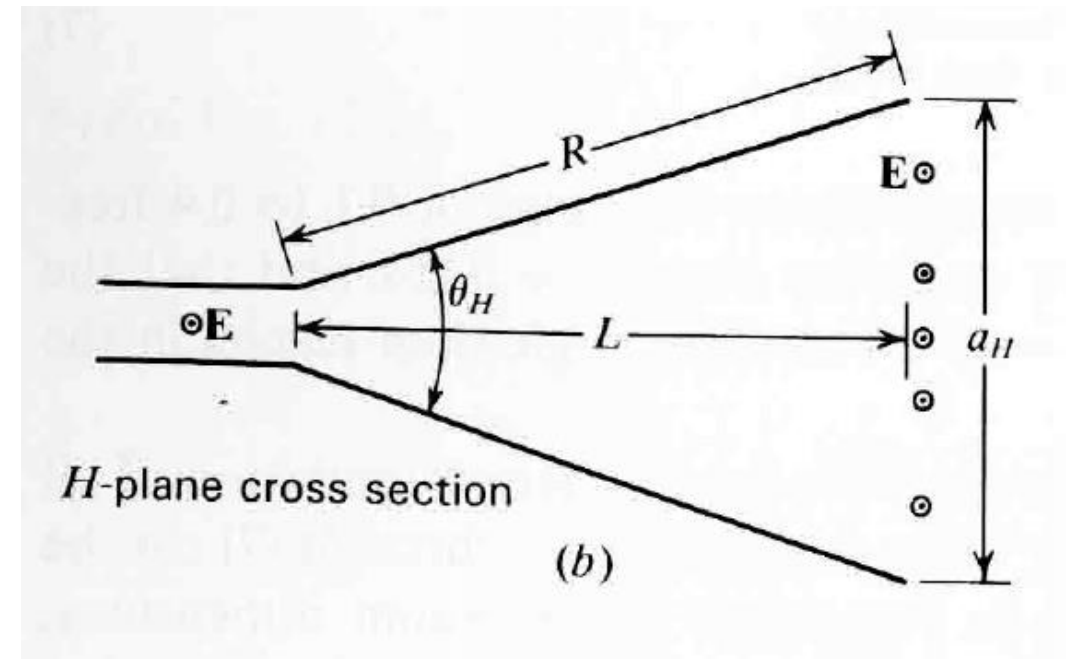
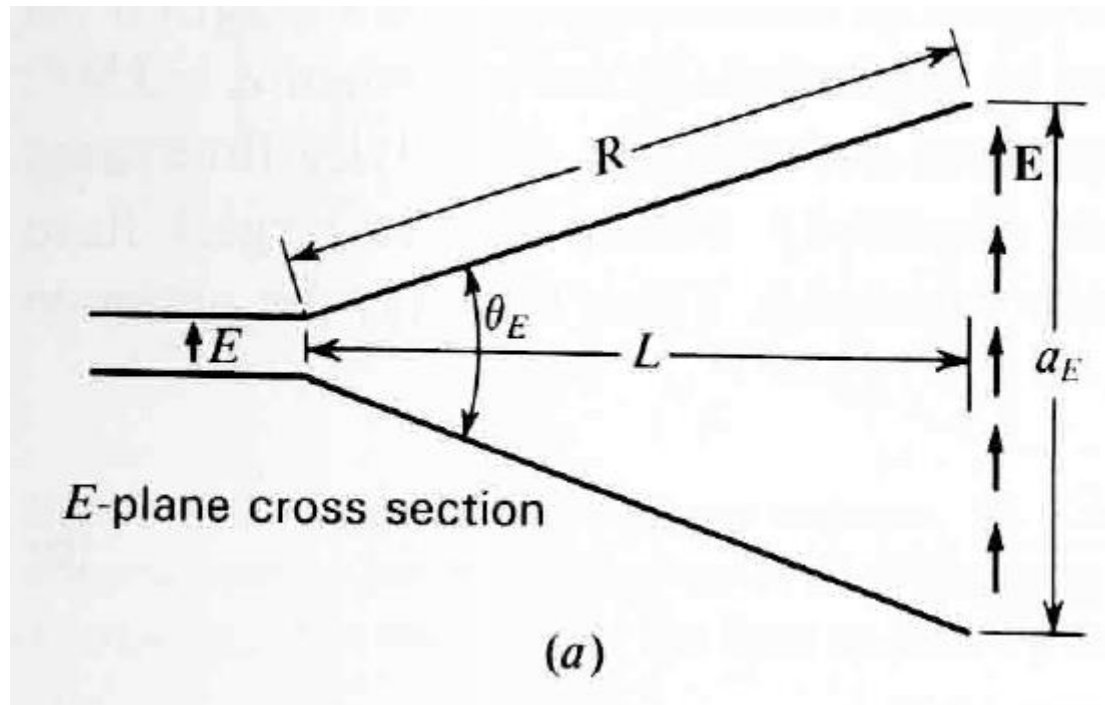
- However, when $\delta = 180^\circ$, the phase reversal at the edges of the aperture reduces the directivity (increases sidelobes).
- It follows that the maximum directivity occurs at the largest flare angle for which δ does not exceed a certain value (δ_0).
- Thus, the *optimum horn dimensions* can be related by

$$\delta_0 = \frac{L}{\cos(\theta/2)} \quad - \quad L = \text{optimum } \delta$$

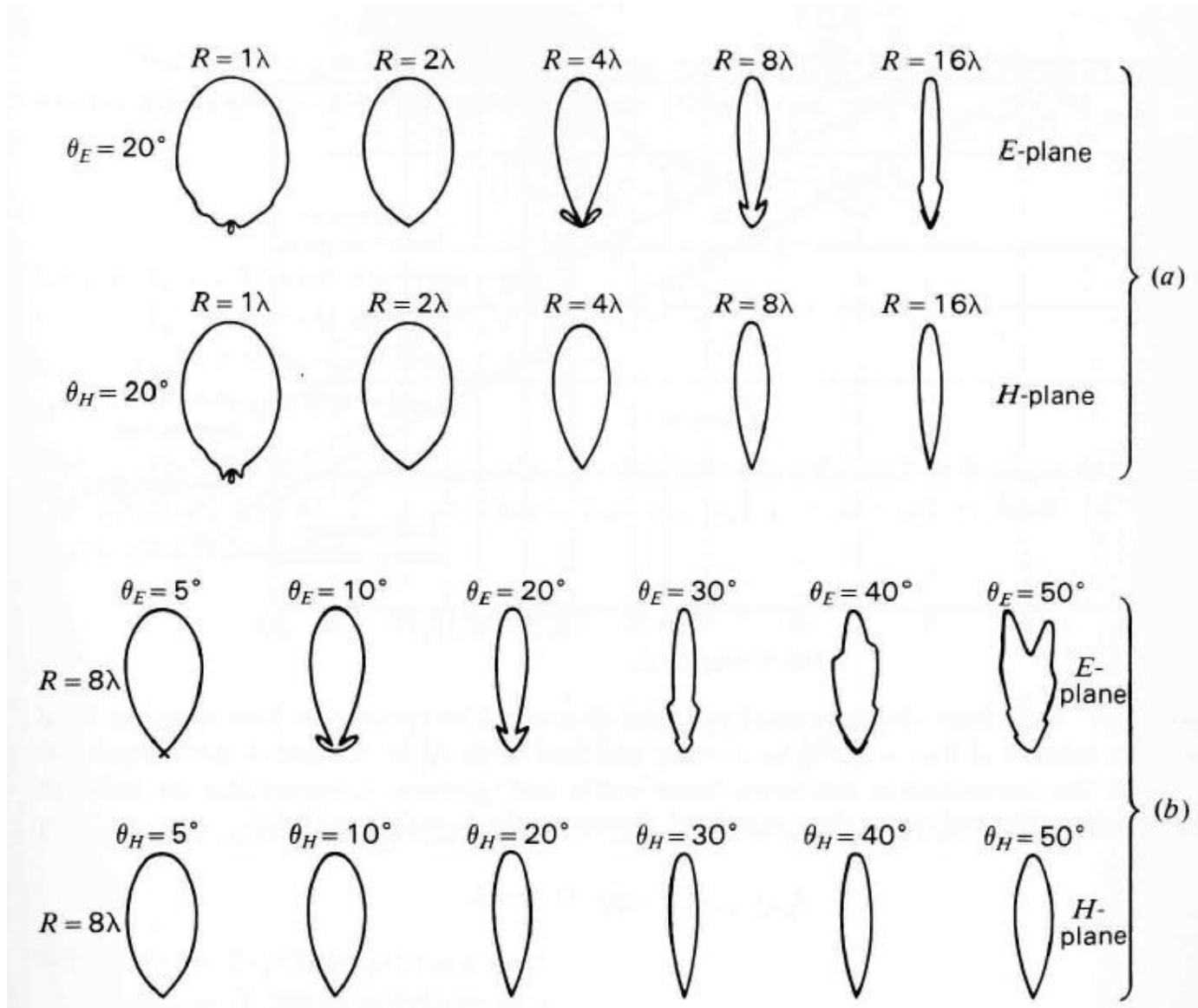
$$L = \frac{\delta_0 \cos(\theta/2)}{1 - \cos(\theta/2)} = \text{optimum length}$$

Rectangular Horn Antennas

- Referring to the following figure, the total flare angle in the E plane is θ_E and the total flare angle in the H plane is θ_H . The axial length of the horn from throat to aperture is L and the radial length is R .



Rectangular Horn Antennas



Measured E-plane and H-plane field patterns of rectangular horns as a function of flare angle and horn length

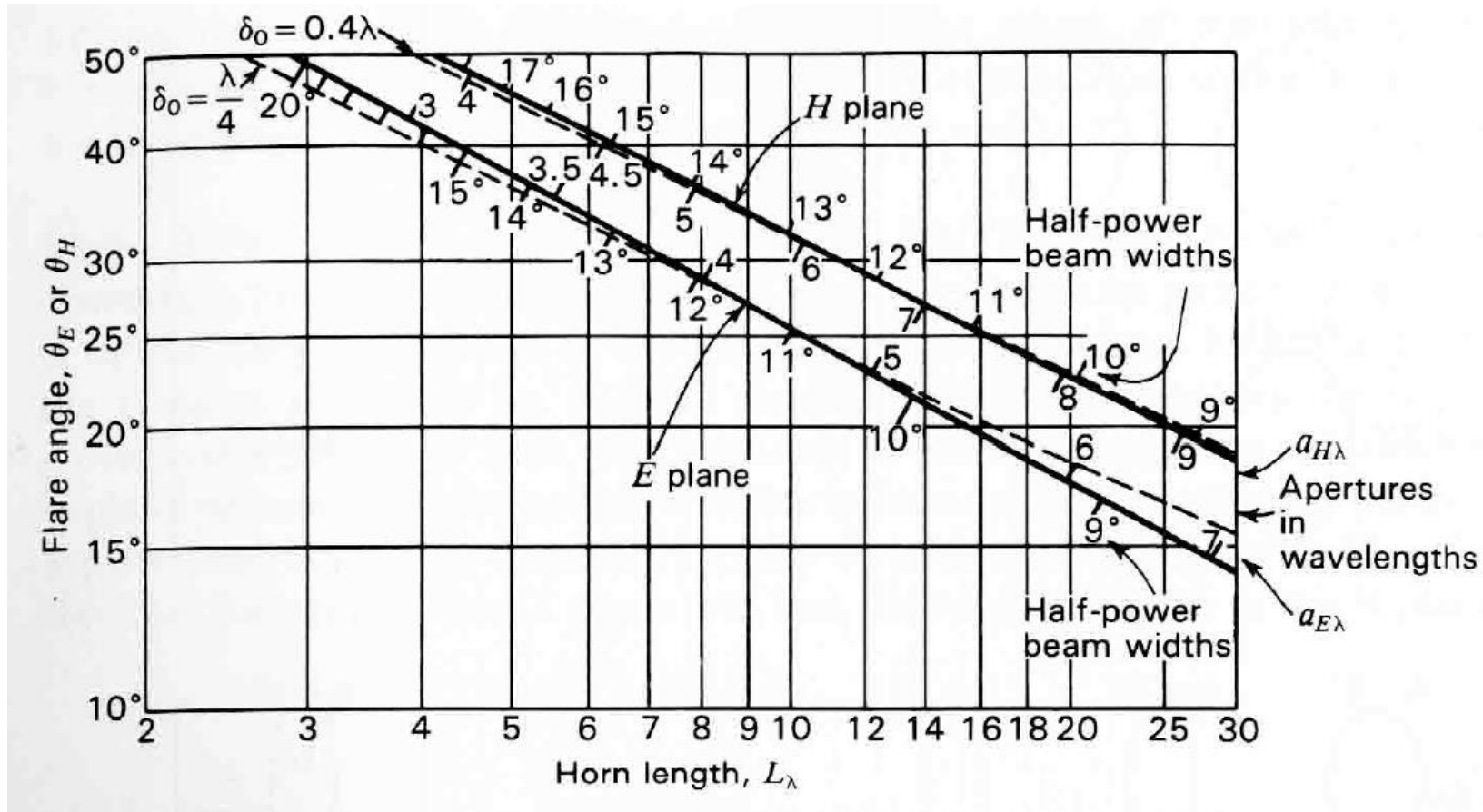
Rectangular Horn Antennas

- Patterns measured by Donald Rhodes are shown in above figure.
- In (a), the patterns in the E plane and H plane are compared as a function of R . Both sets are for a flare angle of 20° . The E-plane patterns have minor lobes whereas the H-plane patterns have practically none.
- In (b), measured patterns for horns with $R = 8\lambda$ are compared as a function of flare angle. In the upper row E-plane patterns are given as a function of the E-plane flare angle θ_E and in the lower row H-plane patterns are shown as a function of the H-plane flare angle θ_H .

Rectangular Horn Antennas

- For a flare angle $\theta_E = 50^\circ$, the E-plane pattern is split, whereas for $\theta_H = 50^\circ$, the H-plane pattern is not.
- This is because a given phase shift at the aperture in the E-plane horn has more effect on the pattern than the same phase shift in the H-plane horn.
- In the H-plane horn, the field goes to zero at the edges of the aperture, so the phase near the edge is relatively less important.
- Accordingly, we should expect the value of δ_0 for the H plane to be larger than for the E plane.
- From Rhodes's experimental patterns, optimum dimensions were selected for both E- and H-plane flare as a function of flare angle and horn length L.
- These optimum dimensions are indicated by the solid lines in the following figure. The corresponding half-power beam widths and apertures in wavelengths are also indicated.

Rectangular Horn Antennas



Experimentally determined optimum dimensions for rectangular horn antennas. Solid curves give relation of flare angle θ_E in E plane and flare angle θ_H in H plane to horn length. The corresponding half-power beamwidths and apertures in wavelengths are indicated along the curves. Dashed curves show calculated dimensions for $\delta_0 = 0.25\lambda$ and $\delta_0 = 0.4\lambda$.

Rectangular Horn Antennas

- The dashed curves show the calculated dimensions for a path length $\delta_0 = 0.25\lambda$ and $\delta_0 = 0.4\lambda$.
- The value of 0.25λ gives a curve close to the experimental curve for E-plane flare, while the value of 0.4λ gives a curve close to the experimental one for H-plane flare over a considerable range of horn length.
- Thus, the tolerance in path length is greater for H-plane flare than for E-plane flare, as indicated above.

Rectangular Horn Antennas

- The directivity (or gain, assuming no loss) of a horn antenna can be expressed in terms of its effective aperture. Thus,

$$D = \frac{4\pi A_e}{\lambda^2} = \frac{4\pi \varepsilon_{ap} A_p}{\lambda^2}$$

where A_e = effective aperture, m^2 ; A_p = physical aperture, m^2 ; ε_{ap} = aperture efficiency; λ = wavelength, m

- For a rectangular horn $A_p = a_E a_H$ and for a conical horn $A_p = \pi r^2$, where r = aperture radius. It is assumed that a_E , a_H or r are all at least 1λ .
- Taking $\varepsilon_{ap} \simeq 0.6$,

$$D \simeq \frac{7.5 A_p}{\lambda^2} \Rightarrow D \simeq 10 \log \left(\frac{7.5 A_p}{\lambda^2} \right) \quad (dBi) \Rightarrow \boxed{D \simeq 10 \log(7.5 a_{E\lambda} a_{H\lambda})}$$

where $a_{E\lambda}$ = E-plane aperture in λ ; $a_{H\lambda}$ = H-plane aperture in λ

Rectangular Horn Antennas - Example

- (a) Determine the length L , H -plane aperture and flare angles θ_E and θ_H (in the E and H planes respectively) of a pyramidal horn for which the E-plane aperture $a_E = 10\lambda$. The horn is fed by a rectangular waveguide with TE_{10} mode. Let $\delta = 0.2\lambda$ in the E-plane and 0.375λ in the H-plane. (b) What are the beamwidths? (c) What is the directivity?

Solution:

Taking $\delta = 0.2\lambda$ in the E-plane, the required horn length

$$L = \frac{a^2}{8\delta} = \frac{(10\lambda)^2}{8 \times 0.2\lambda} = \frac{100\lambda}{1.6} = 62.5\lambda$$

Flare angle in E-plane is

$$\theta_E = 2 \tan^{-1} \frac{a_E}{2L} = 2 \tan^{-1} \left(\frac{10\lambda}{2 \times 62.5\lambda} \right) = 2 \tan^{-1} \left(\frac{10}{125} \right) = 9.1^\circ$$

Rectangular Horn Antennas - Example

Taking $\delta = 0.375 \lambda$ in the H-plane, the flare angle in the H-plane

$$\theta_H = 2 \cos^{-1} \frac{L}{L + \delta} = 2 \cos^{-1} \left(\frac{62.5\lambda}{62.5\lambda + 0.375\lambda} \right) = 2 \cos^{-1} \left(\frac{62.5}{62.875} \right) = 12.52^\circ$$

H-plane aperture is

$$a_H = 2L \tan \frac{\theta_H}{2} = 2 \times 62.5\lambda \tan \left(\frac{12.52^\circ}{2} \right) = 125\lambda \times 0.109 = 13.7\lambda$$

$$HPBW (E - Plane) = \frac{56^\circ}{a_{E\lambda}} = \frac{56^\circ}{10} = 5.6^\circ$$

$$HPBW (H - Plane) = \frac{67^\circ}{a_{H\lambda}} = \frac{67^\circ}{13.7} = 4.9^\circ$$

$$D \simeq 10 \log(7.5 a_{E\lambda} a_{H\lambda}) = 10 \log(7.5 \times 10 \times 13.7) = 30.1 \text{ dBi}$$

Conical Horn Antennas

- Conical Horn can be directly excited from a circular waveguide.
- Dimensions can be determined from the following expressions by taking $\delta_0 = 0.32\lambda$,

$$\delta_0 = \frac{L}{\cos(\theta/2)} - L = \text{optimum } \delta$$
$$L = \frac{\delta_0 \cos(\theta/2)}{1 - \cos(\theta/2)} = \text{optimum length}$$

- For optimum conical horns,

$$HPBW (E - plane) = 60/a_{E\lambda}$$

$$HPBW (H - plane) = 70/a_{H\lambda}$$

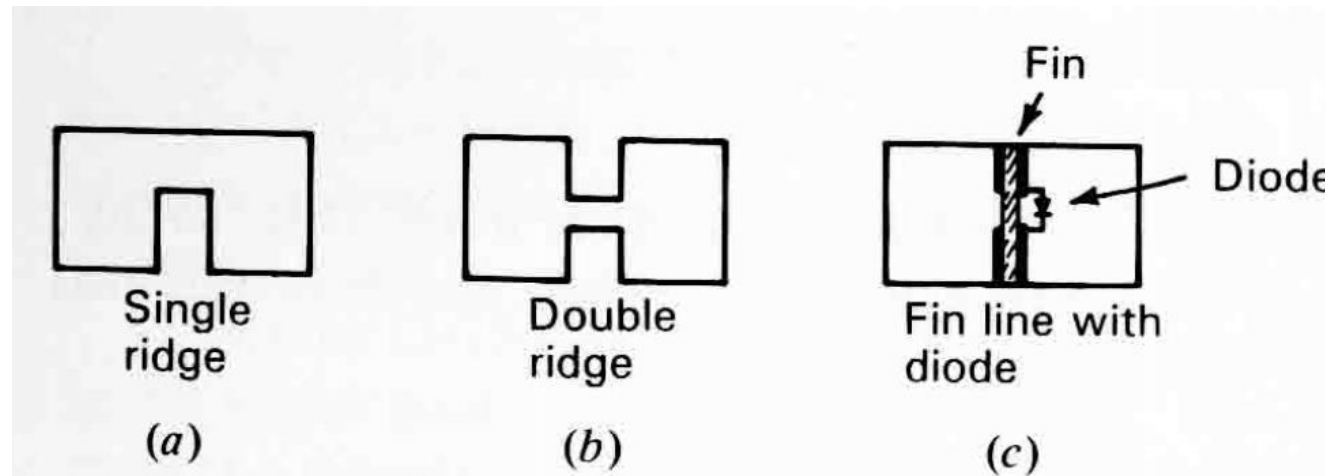
These values are about 6% more than the values for a rectangular horn.

Conical Horn Antennas

- Biconical horns have patterns that are non-directional in the horizontal plane (axis of horns vertical).
- These horns may be regarded as modified pyramidal horns with a 360° flare angle in the horizontal plane.
- The optimum vertical-plane flare angle is about the same as for a sectoral horn of the same cross section excited in the same mode.

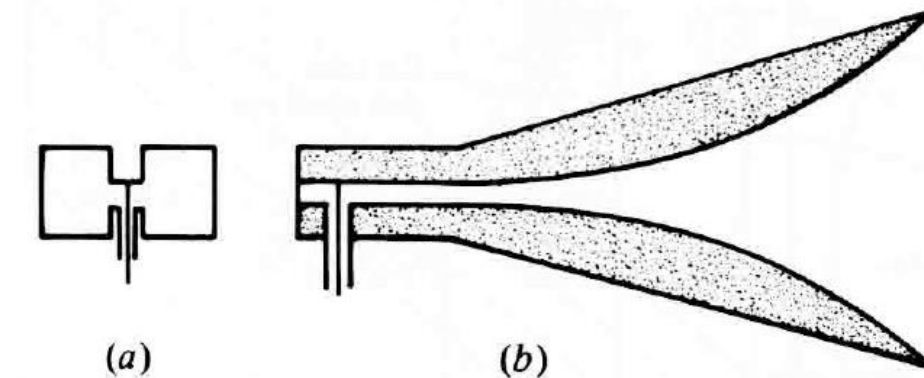
Ridge Horns

- A central ridge loads a waveguide and increases its useful bandwidth by lowering the cutoff frequency of the dominant mode.
- Rectangular guides with single and double ridges are shown in the following figures(a) and (b)
- A very thin ridge or fin is also effective in producing the loading of a central ridge. It may consist of a metal-clad ceramic sheet which facilitates the installation of shunt circuit elements as suggested in figure(c).



Ridge Horns

- Cutoff frequency can be lowered by placing dielectric material in the waveguide, but this does not increase the bandwidth and it may increase losses.
- By continuing a double-ridge structure from a waveguide into a pyramidal horn as suggested in the following figure, the useful bandwidth of the horn can be increased manifold.

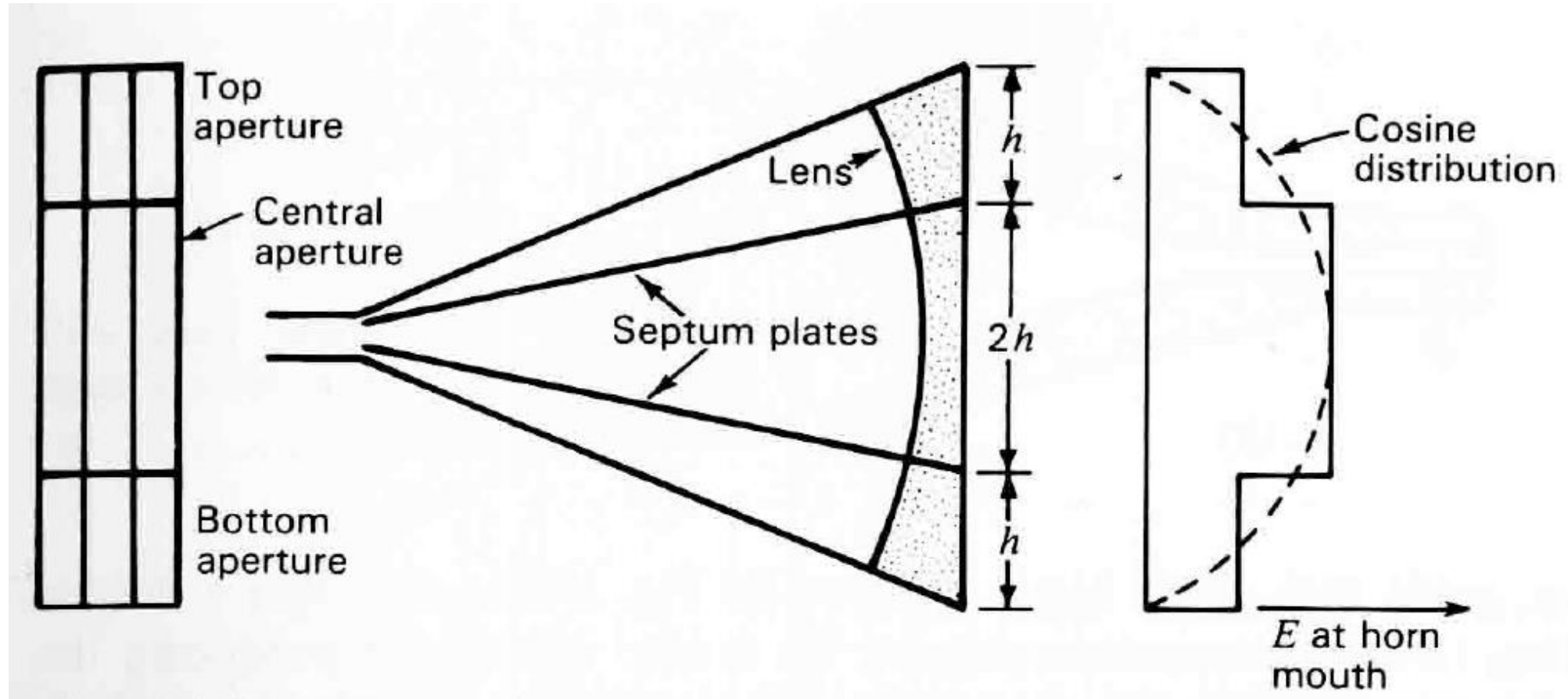


Double-ridge or Vivaldi horn with coaxial feed.
The view at (a) is a cross section at the feed point.

Septum Horns

- Although the electric field in the H plane of a pyramidal horn tends to zero at the edges, resulting in a tapered distribution and reduced sidelobes, the electric field in the E plane may be close to uniform in amplitude to the edges, resulting in significant sidelobes.
- By introducing septum plates bonded to the horn walls, a stepped-amplitude distribution can be achieved in the E plane with a reduction in E-plane sidelobes.
- Typically, the first sidelobes of a uniform amplitude distribution are down about 13 dB.
- A cosine field distribution is approximated with a 1:2:1 stepped amplitude distribution with apertures also in the ratio 1:2:1 as suggested in the following figure. To achieve this distribution, the septums must be appropriately spaced at the throat of the horn.

Septum Horns



Two-septum horn with 1:2:1 stepped amplitude distribution in field intensity at mouth of horn

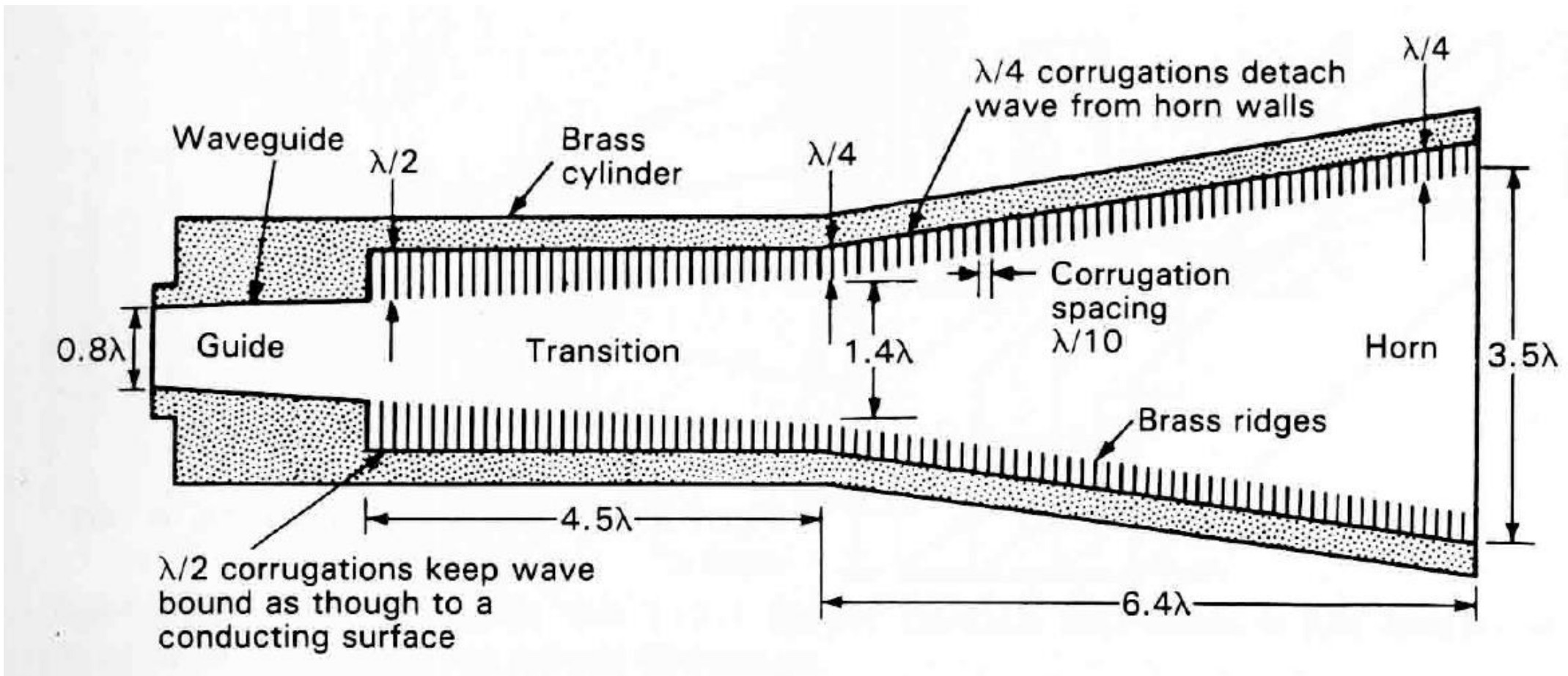
Corrugated Horns

- Corrugated horns can provide reduced edge diffraction, improved pattern symmetry and reduced cross-polarization (less E field in the H plane).
- Corrugations on the horn walls acting as $\lambda/4$ chokes are used to reduce E to very low values at all horn edges for all polarizations. These prevent waves from diffracting around the edges of the horn (or surface currents flowing around the edge and over the outside).
- The reactance at open end of corrugation is

$$X \simeq 377 \tan\left(\frac{2\pi d}{\lambda}\right) \quad (\Omega)$$

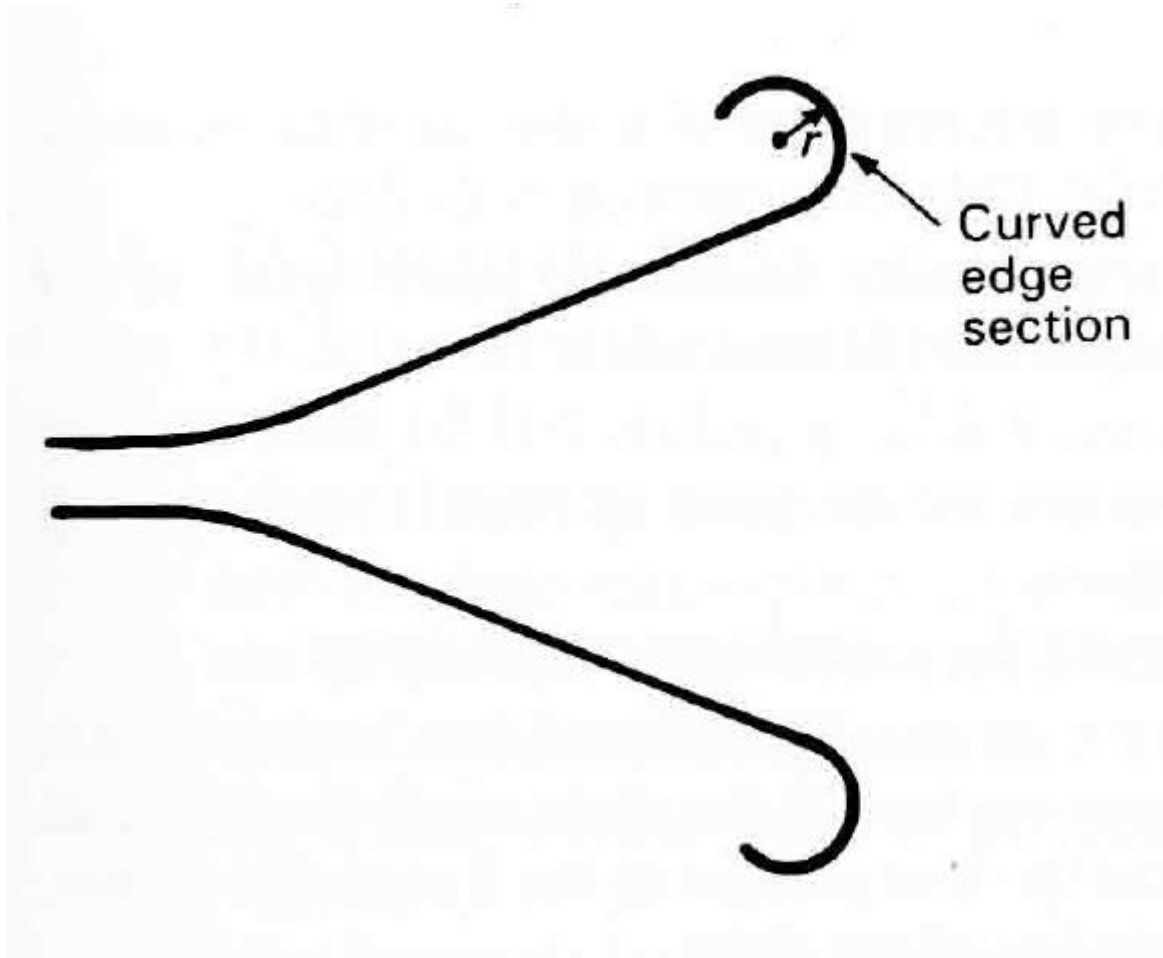
- Corrugations with depth $d = \lambda/2$ act as conducting surface ($X = 0$)
- Corrugations with depth $d = \lambda/4$ present a high impedance ($X = \infty$)

Corrugated Horn



Cross section of circular waveguide-fed corrugated horn with corrugated transition

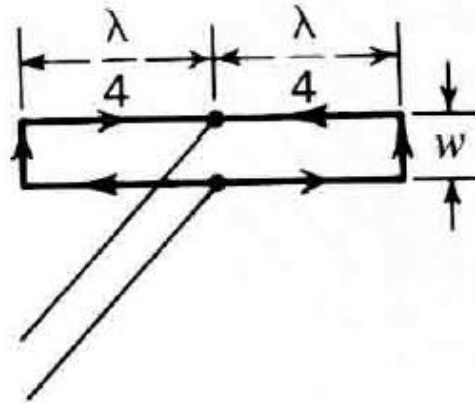
Aperture-Matched Horn



- By attaching a smooth curved (or rolled) surface section to the outside of the aperture edge of a horn, a significant improvement in the pattern, impedance and bandwidth characteristics can be achieved.
- This arrangement is an attractive alternative to a corrugated horn.
- The shape of the rolled edge is not critical but its radius of curvature should be at least $\lambda/4$.

Slot Antennas

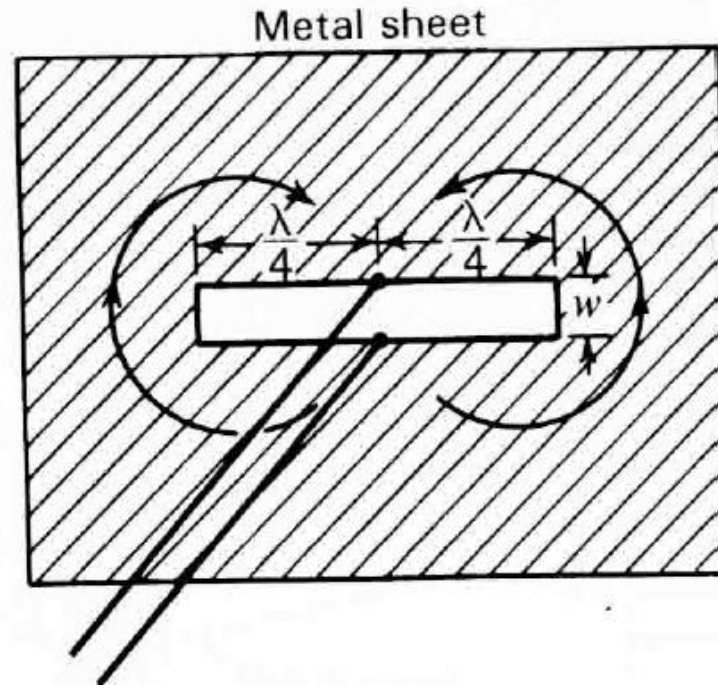
- Slot antennas are useful in many applications, especially where low-profile or flush installations are required as, for example, on high-speed aircraft.



- The antenna shown in above figure, consisting of two resonant $\lambda/4$ stubs connected to a 2-wire transmission line, forms an *inefficient radiator*. The long wires are closely spaced ($w \ll \lambda$) and carry currents of opposite phase so that their fields tend to cancel. The end wires carry currents in the same phase, but they are too short to radiate efficiently. Hence, enormous currents may be required to radiate appreciable amounts of power.

Slot Antennas

- The antenna shown in the following figure is a *very efficient radiator*. In this arrangement a $\lambda/2$ slot is cut in a flat metal sheet. Although the width of the slot is small ($w \ll \lambda$), the currents are not confined to the edges of the slot but spread out over the sheet. This is a simple type of slot antenna. Radiation occurs equally from both sides of the sheet. If the slot is horizontal, as shown, the radiation normal to the sheet is vertically polarized.

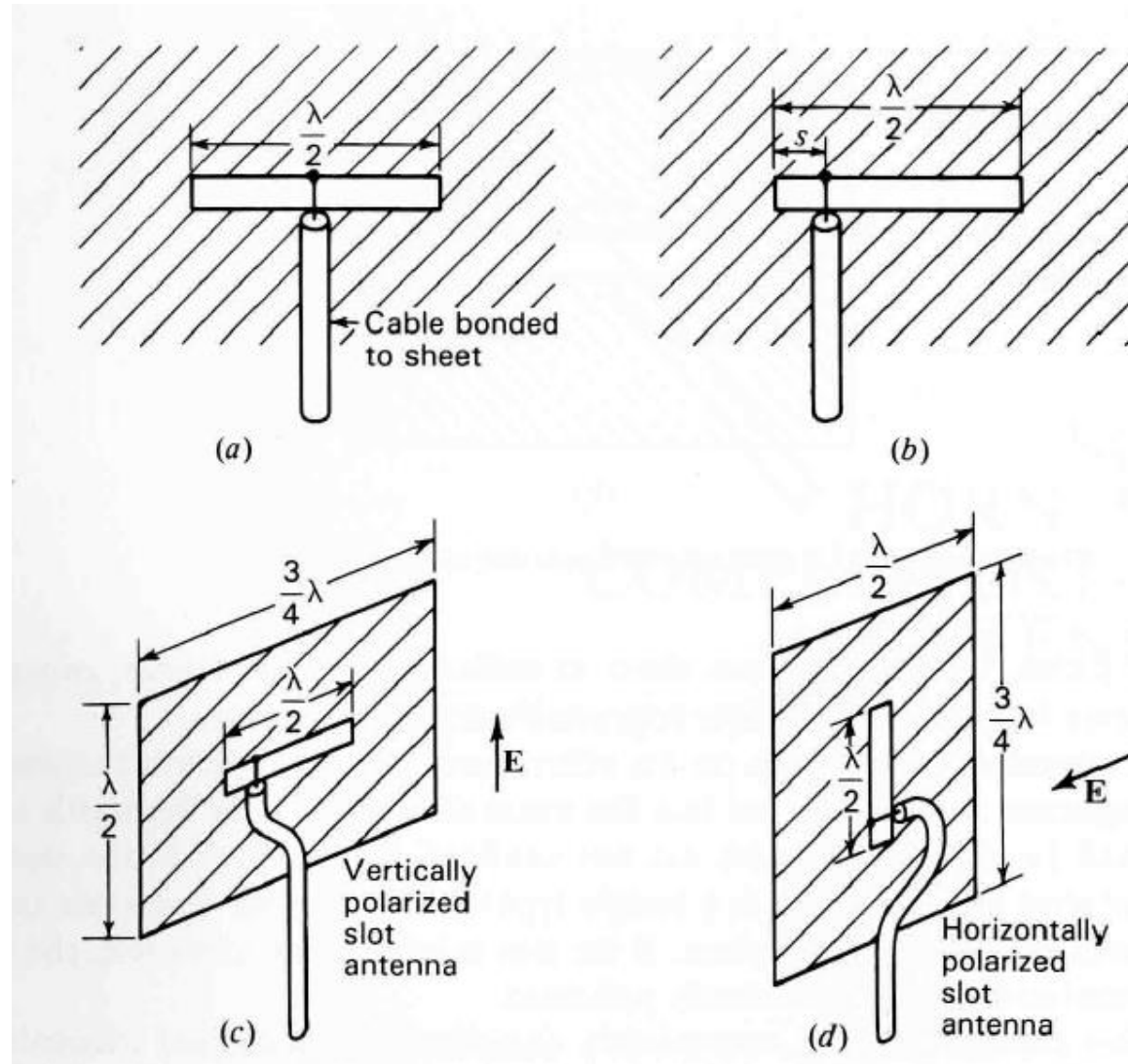


Slot Antennas - Methods of Feeding

- A slot antenna may be conveniently energized with a *coaxial transmission line* as in the following figure(a).
- The outer conductor of the cable is bonded to the metal sheet. Since the terminal resistance at the center of a resonant $\lambda/2$ slot in a large sheet is about 500Ω and the characteristic impedance of coaxial transmission lines is usually much less, an *offset feed* as shown in figure(b) may be used to provide a *better impedance match*.
- For a 50Ω coaxial cable, the distance $s \approx \lambda/20$.

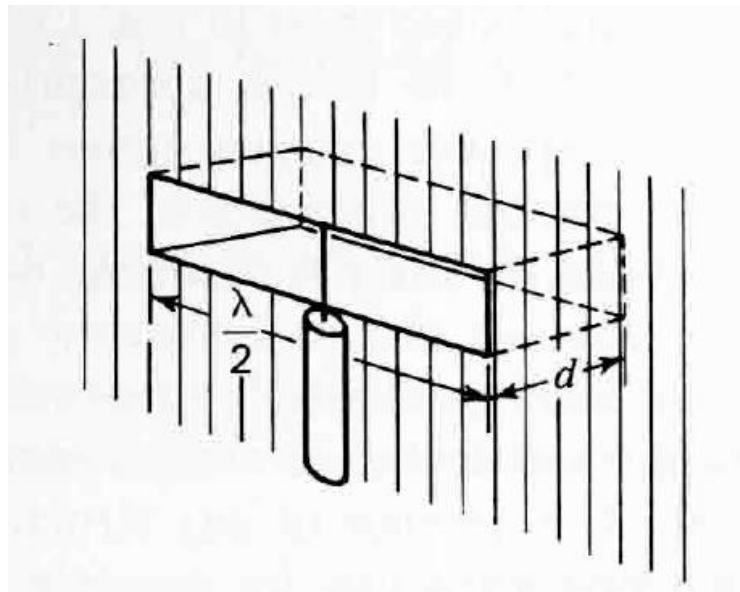
Slot Orientation	Horizontal Slot as shown in figure (c)	Vertical Slot as shown in figure (d)
Wave Polarization	Vertical Polarization	Horizontal Polarization

Slot antennas fed by coaxial transmission lines



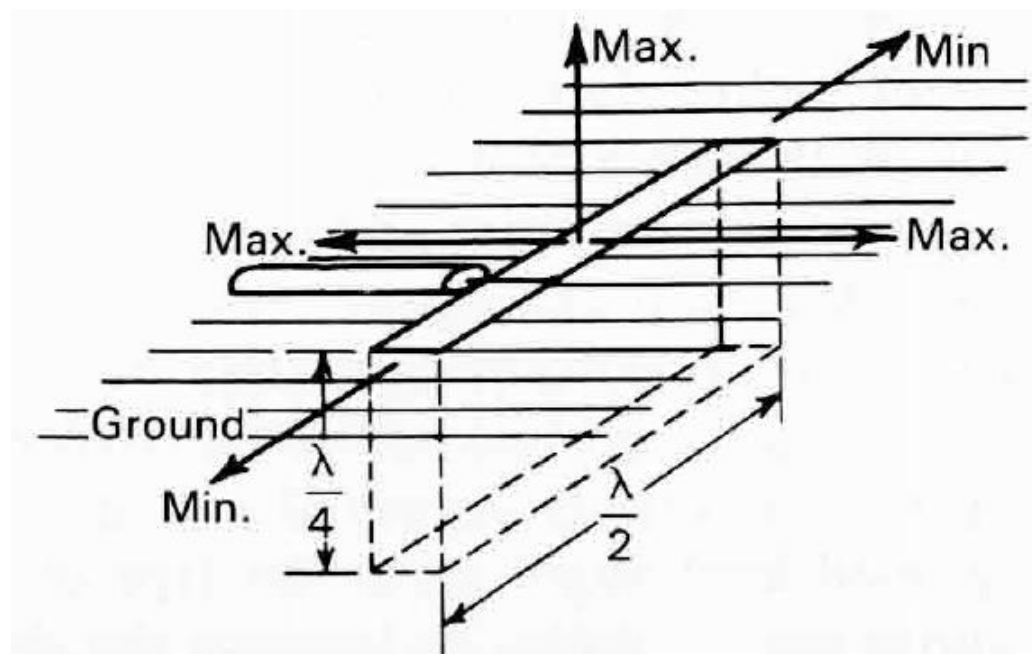
Boxed-in Slot Antenna

- A flat sheet with a $\lambda/2$ slot radiates equally on both sides of the sheet. However, if the sheet is very large (ideally infinite) and boxed in as shown in the following figure, radiation occurs only from one side.
- If the depth d of the box is appropriate ($d \sim \lambda/4$ for a thin slot), no appreciable shunt susceptance appears across the terminals. With such a zero susceptance box, the terminal impedance of the resonant $\lambda/2$ slot is nonreactive and approximately twice its value without the box, or about 1000Ω .



Boxed-in Slot Antenna as Flush Radiator

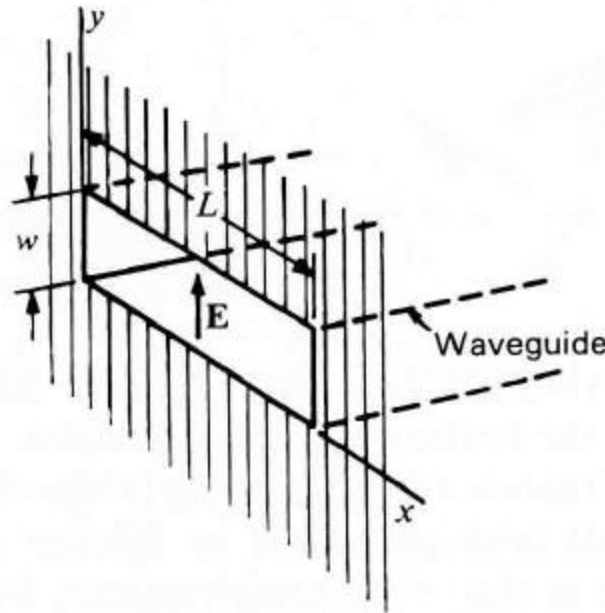
- The boxed-in slot antenna might be applied even at relatively long wavelengths by using the ground as the flat conducting sheet and excavating a trench $\lambda/2$ long by $\lambda/4$ deep as shown in the following figure.



- Radiation is maximum in all directions at right angles to the slot and is zero along the ground in the directions of the ends of the slot. The radiation along the ground is vertically polarized.

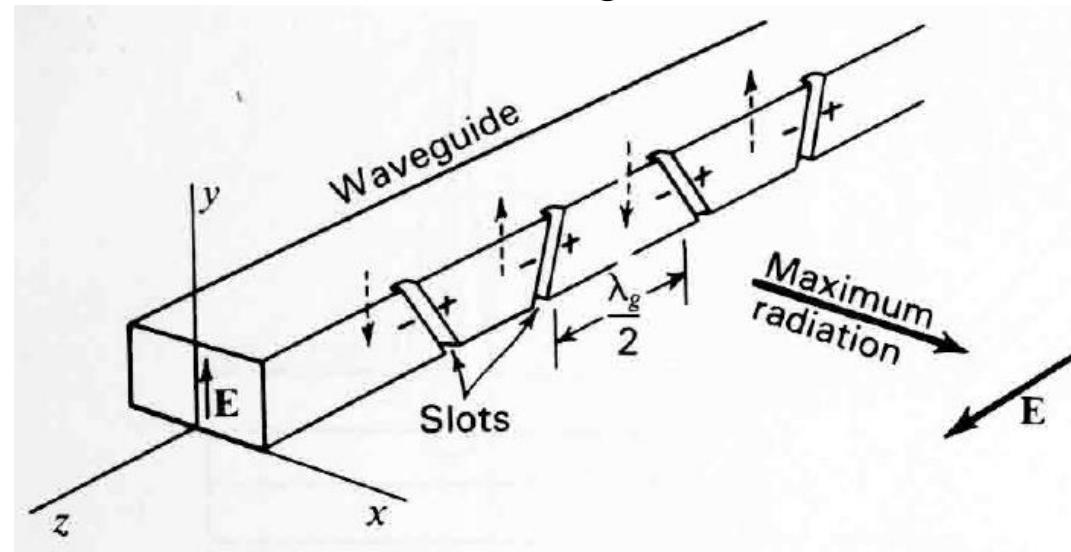
Waveguide-fed Slot

- Radiation from only one side of a large flat sheet may also be achieved by a slot fed with a waveguide as shown in the figure. With the transmission in the guide in the TE_{10} mode, the direction of the electric field \mathbf{E} is as shown.
- The width L of the guide must be more than $\lambda/2$ to transmit energy, but it should be less than 1λ to suppress higher-order transmission modes. With the horizontal slot, the radiation normal to sheet is vertically polarized.



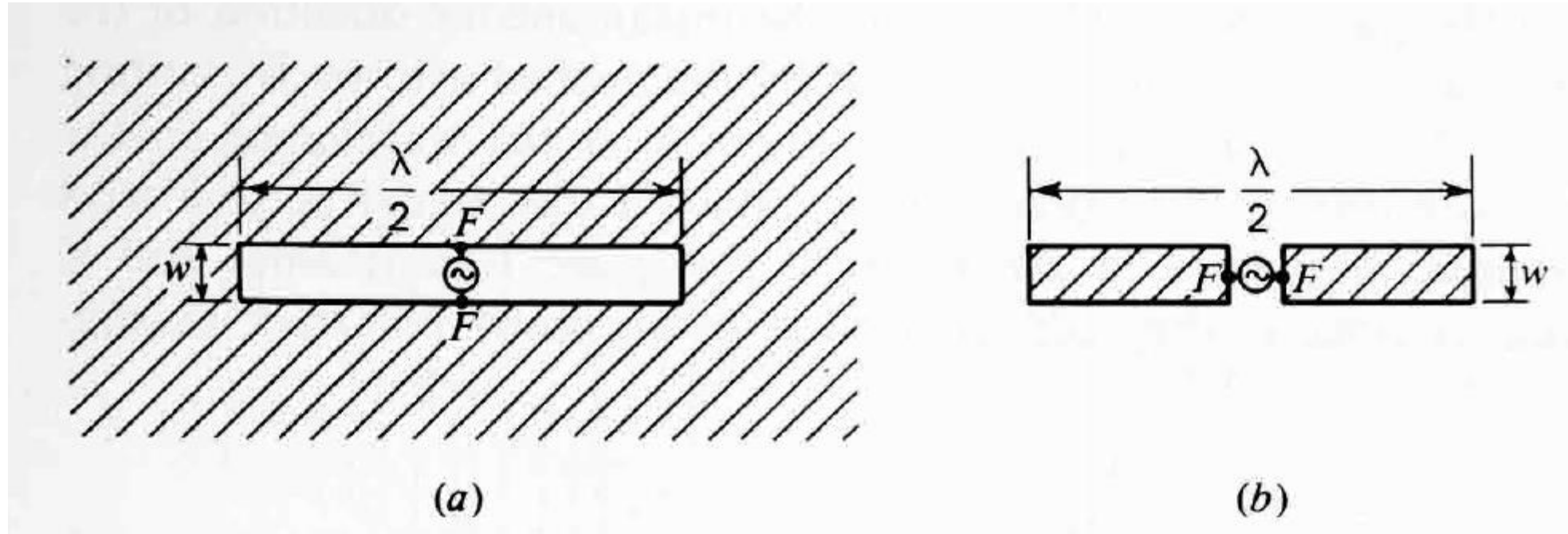
Broadside Array of Slots in a Waveguide

- An array of slots may be cut in the waveguide as shown in the figure so as to produce a directional radiation pattern.
- By cutting inclined slots as shown at intervals of $\lambda_g/2$ (where λ_g is the guide wavelength), the slots are energized in phase and produce a directional pattern with maximum radiation broadside to the guide.



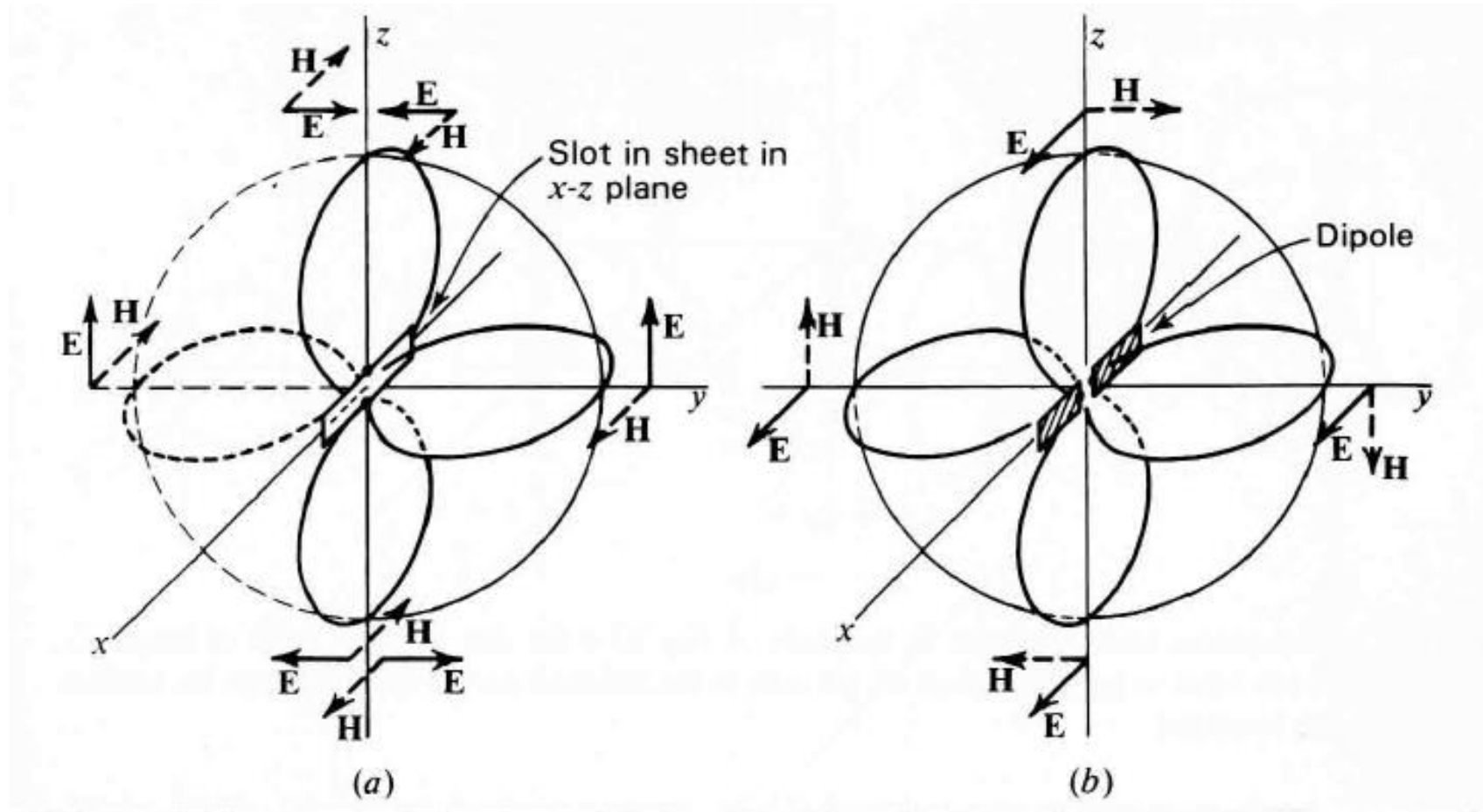
- If the guide is horizontal and \mathbf{E} inside the guide is vertical, the radiated field is horizontally polarized.

Slot Antenna Vs Complementary Dipole



A $\lambda/2$ slot in an infinite sheet (a) and a complementary $\lambda/2$ dipole antenna (b)

Slot Antenna Vs Complementary Dipole



Radiation-field patterns of slot in an infinite sheet (a) and of complementary dipole antenna (b). The patterns have the same shape but with \mathbf{E} and \mathbf{H} interchanged.

Babinet's Principle

- Babinet's principle in optics states that *when the field behind a screen with an opening is added to the field of a complementary structure, the sum is equal to the field when there is no screen*.
- Let a source and 2 imaginary planes, plane of screens A and plane of observation B, be arranged as in the following figure.
- **Case 1:** Let a perfectly absorbing screen be placed in plane A. Then in plane B there is a region of shadow as indicated. Let the field behind this screen be some function f_1 of x, y and z . Thus,

$$F_S = f_1(x, y, z)$$

- **Case 2:** Let the first screen be replaced by its complementary screen and the field behind it be given by

$$F_{CS} = f_2(x, y, z)$$

Babinet's Principle

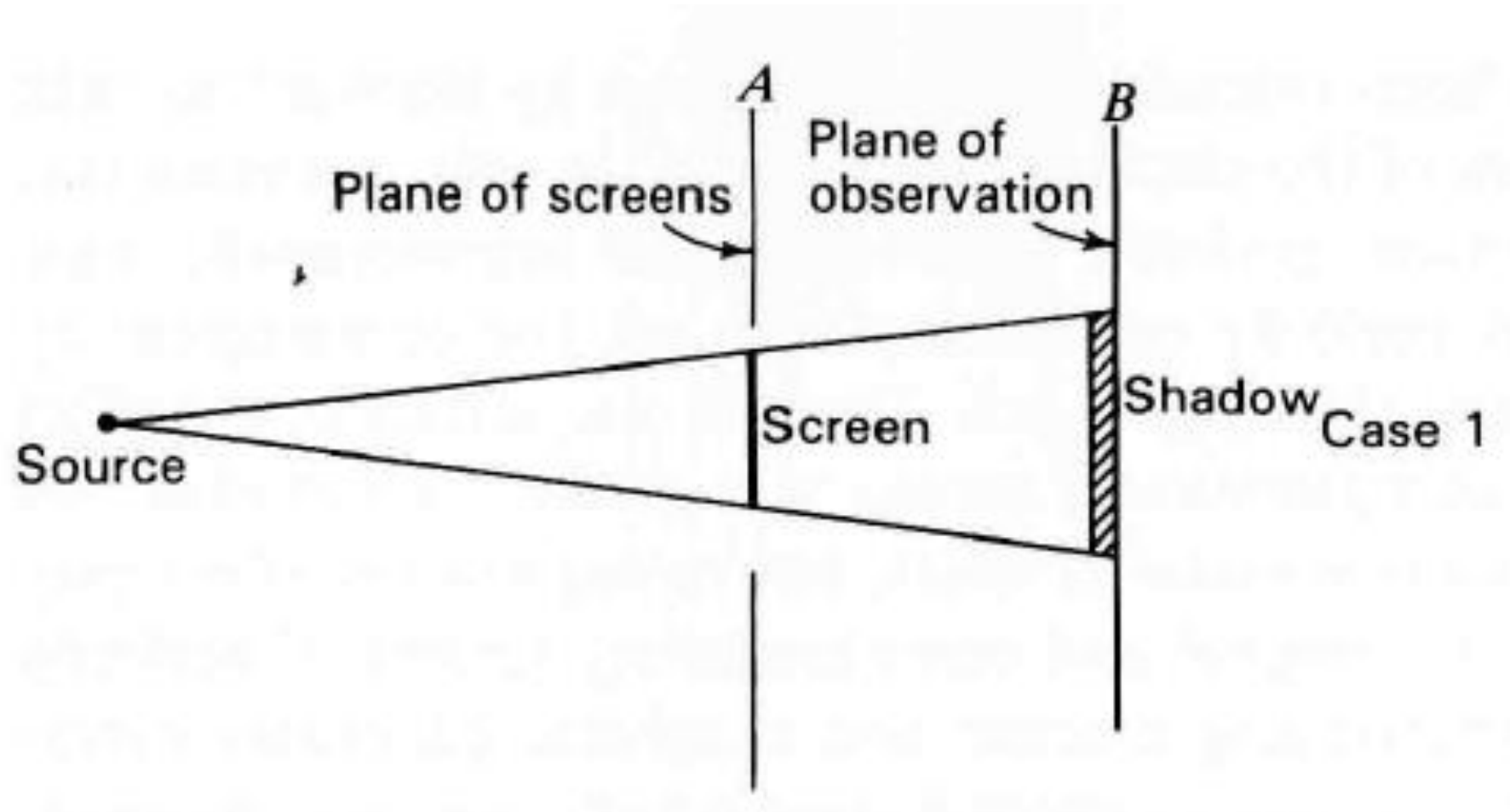
- **Case 3:** With no screen, the field is

$$F_0 = f_3(x, y, z)$$

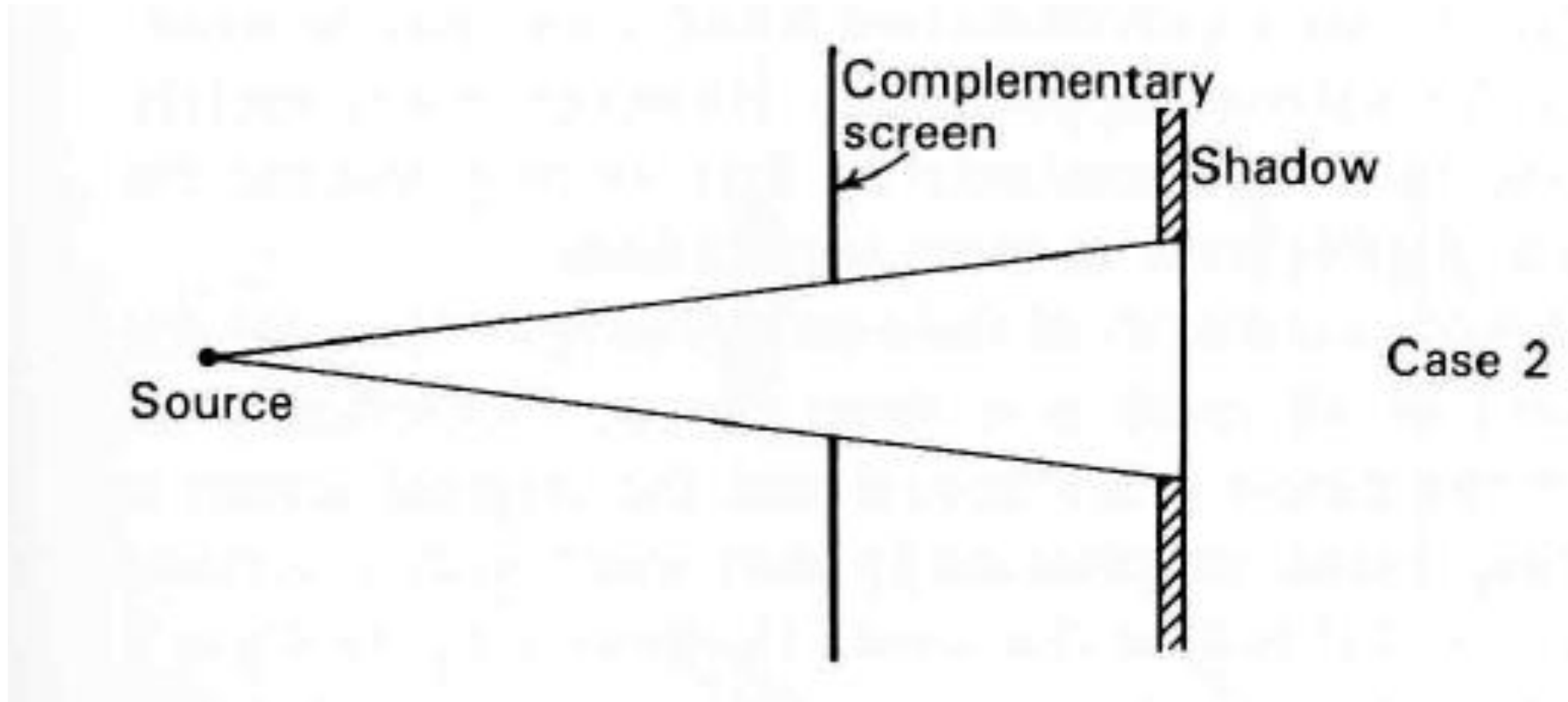
- Babinet's principle asserts that at the same point x, y, z ,

$$F_S + F_{CS} = F_0$$

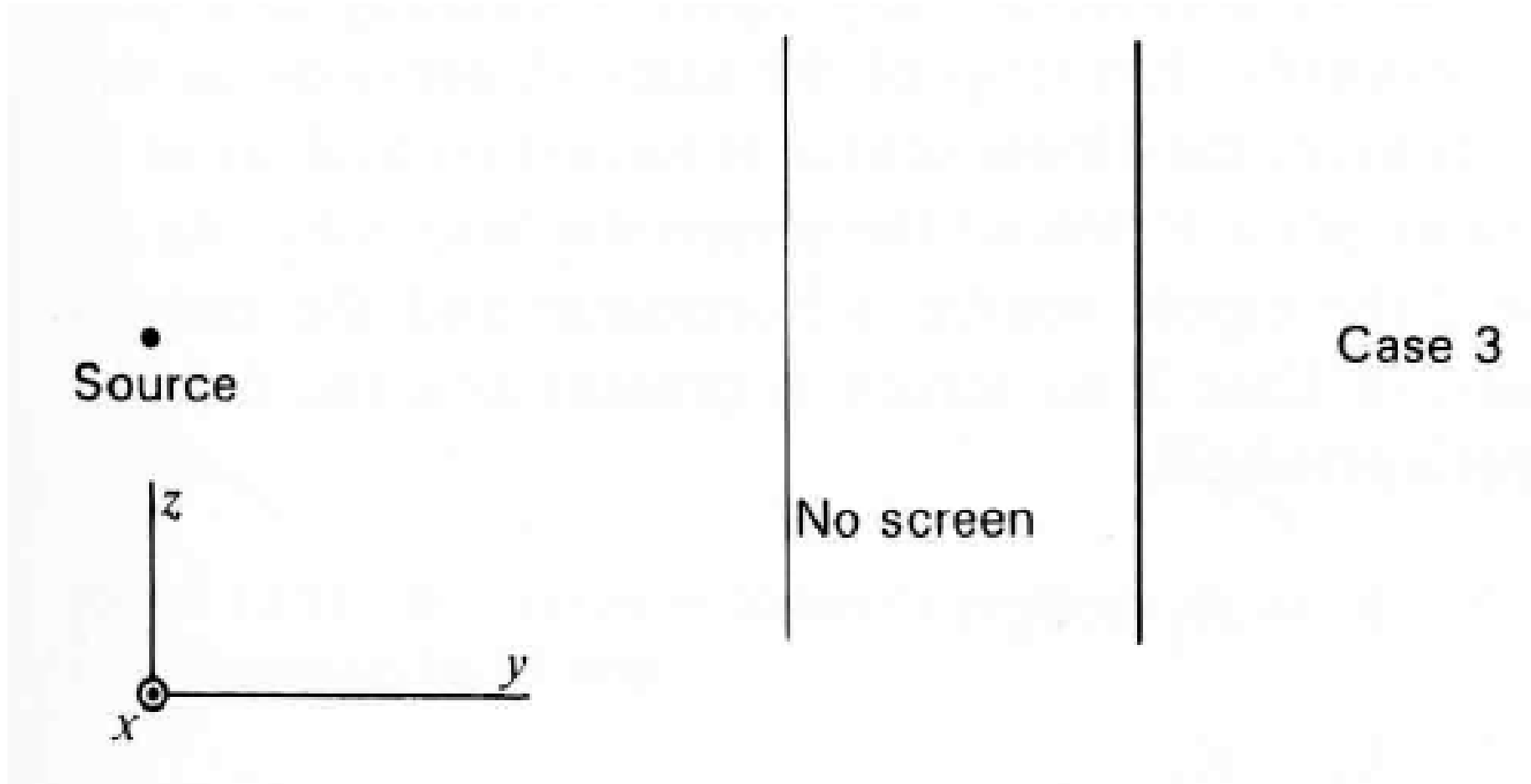
Babinet's Principle



Babinet's Principle



Babinet's Principle



Booker's Extension of Babinet's Principle

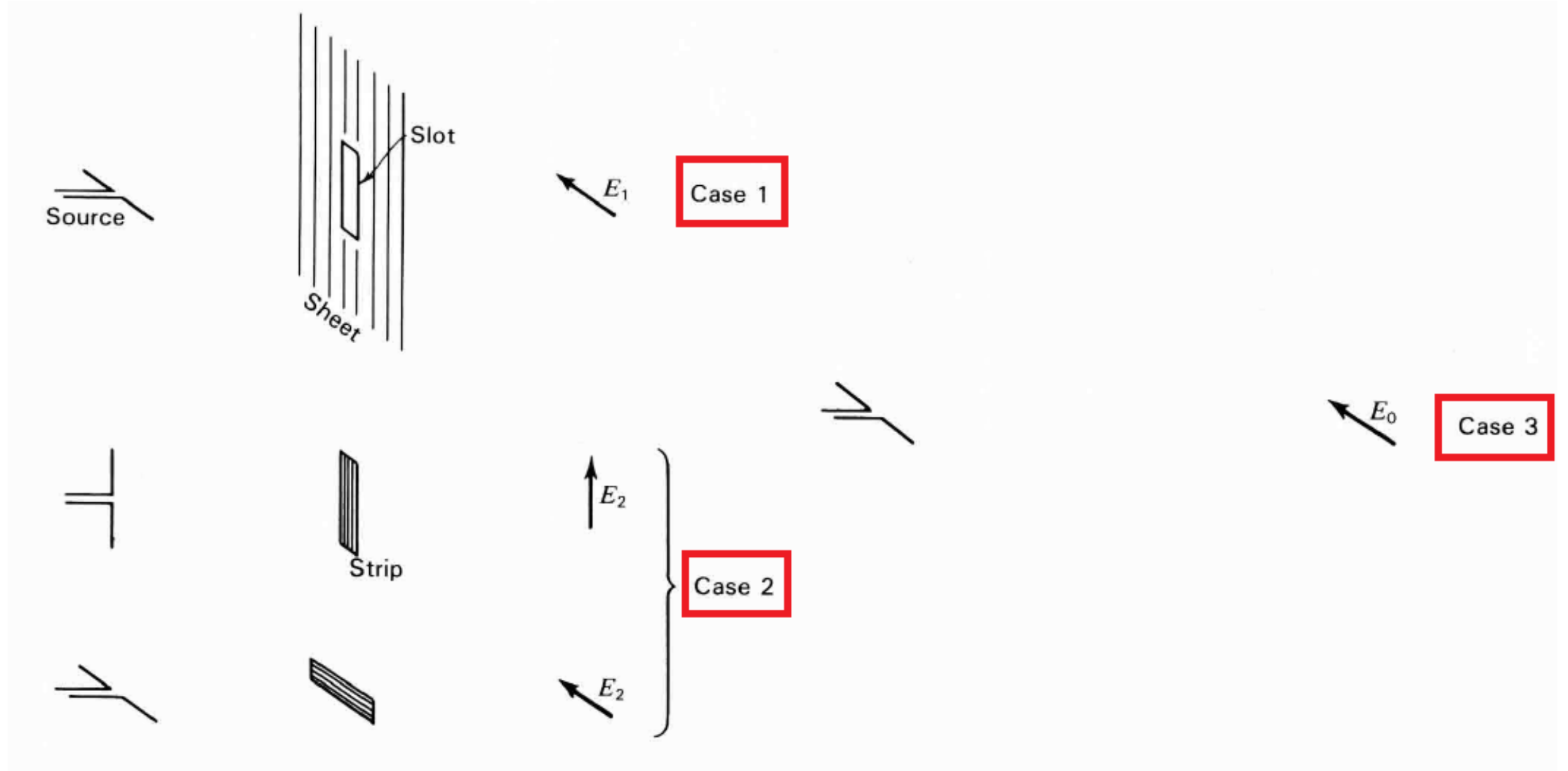
- Babinet's principle has been extended and generalized by Booker to take into account the vector nature of the electromagnetic field.
- In this extension, it is assumed that the screen is plane, perfectly conducting and infinitesimally thin.
- Furthermore, if one screen is perfectly conducting ($\sigma = \infty$), the complementary screen must have infinite permeability ($\mu = \infty$).
- Thus, if one screen is a perfect conductor of electricity, the complementary screen is a perfect conductor of magnetism.
- No infinitely permeable material exists, but the equivalent effect may be obtained by making both the original and complementary screens of perfectly conducting material and interchanging electric and magnetic quantities everywhere.

Booker's Extension of Babinet's Principle

- **Case 1:** The dipole is horizontal and the original screen is an infinite, perfectly conducting, plane, infinitesimally thin sheet with a vertically cut out slot as indicated. At a point P behind the screen the field is E_1 .
- **Case 2:** The original screen is replaced by the complementary screen consisting of a perfectly conducting, plane, infinitesimally thin strip of the same dimensions as the slot in the original screen. In addition, the dipole source is turned vertical so as to interchange **E** and **H**. At the same point P behind the screen the field is E_2 .
- **Case 2 Alternative:** The dipole source is horizontal and the strip is also turned horizontal.
- **Case 3:** No screen is present and the field at point P is E_0 . Then, by Babinet's principle,

$$E_1 + E_2 = E_0 \implies \frac{E_1}{E_0} + \frac{E_2}{E_0} = 1$$

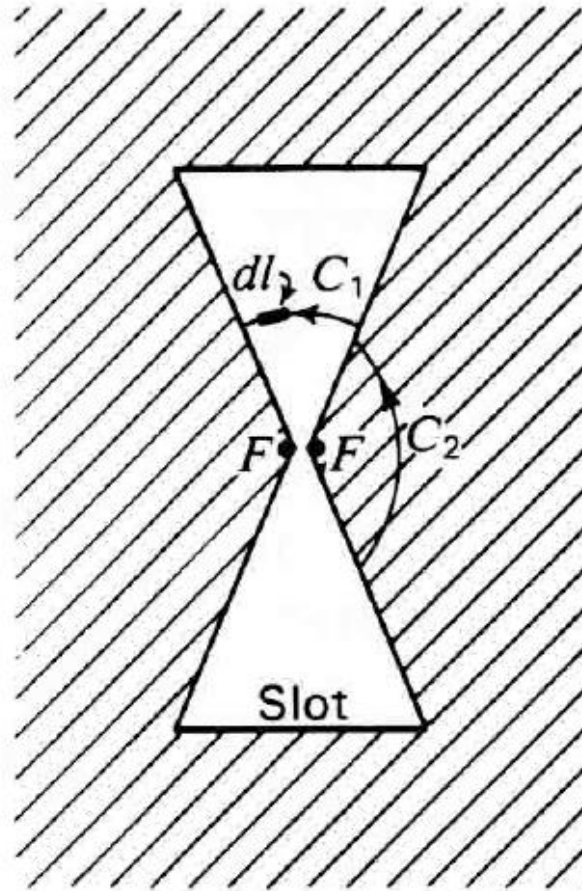
Booker's Extension of Babinet's Principle



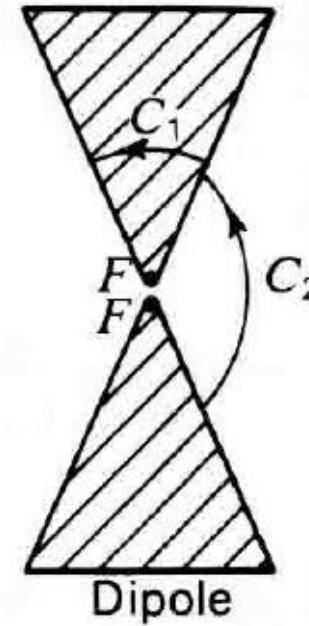
Impedance of Slot Antenna

- Knowing the impedance Z_d of the complementary dipole antenna, the impedance Z_s of the slot antenna can be determined.
- Consider the slot antenna and the complementary dipole antenna as shown in the following figure. The terminals of each antenna are indicated by FF, and it is assumed that they are separated by an infinitesimal distance. It is assumed that the dipole and slot are cut from an infinitesimally thin, plane, perfectly conducting sheet.
- Let a generator be connected to the terminals of the slot antenna. The driving point impedance Z_s at the terminals is the ratio of the terminal voltage V_s to the terminal current I_s .
- Let E_s and H_s be the electric and magnetic fields of the slot at any point P.

Impedance of Slot Antenna



(a)



(b)

Slot antenna (a) and complementary dipole antenna (b)

Impedance of Slot Antenna

- The voltage V_S at the terminals FF of the slot is given by the line integral of E_S over the path C_1 as C_1 approaches zero. Thus,

$$V_S = \lim_{C_1 \rightarrow 0} \int_{C_1} E_S \cdot dl \quad (1)$$

where dl = infinitesimal vector element of length dl along the contour or path C_1 .

- The current I_S at the terminals of the slot is

$$I_S = 2 \lim_{C_2 \rightarrow 0} \int_{C_2} H_S \cdot dl \quad (2)$$

- The path C_2 is just outside the metal sheet and parallel to its surface. The factor 2 enters because only $\frac{1}{2}$ of the closed line integral is taken, the line integral over the other side of the sheet being equal by symmetry.

Impedance of Slot Antenna

- Let a generator be connected to the terminals of the dipole. The driving-point impedance Z_d at the terminals is the ratio of the terminal voltage V_d to the terminal current I_d . Let E_d and H_d be the electric and magnetic fields of the dipole at any point P. Then the voltage at the dipole terminals is

$$V_d = \lim_{C_2 \rightarrow 0} \int_{C_2} E_d \cdot dl \quad (3)$$

- Current at the dipole terminals is

$$I_d = 2 \lim_{C_1 \rightarrow 0} \int_{C_1} H_d \cdot dl \quad (4)$$

Impedance of Slot Antenna

- However,

$$\lim_{C_2 \rightarrow 0} \int_{C_2} E_d \cdot dl = Z_0 \lim_{C_2 \rightarrow 0} \int_{C_2} H_s \cdot dl \quad (5)$$

$$\lim_{C_1 \rightarrow 0} \int_{C_1} H_d \cdot dl = \frac{1}{Z_0} \lim_{C_1 \rightarrow 0} \int_{C_1} E_s \cdot dl \quad (6)$$

where Z_0 is the intrinsic impedance of the surrounding medium.

- Substituting (3) and (2) in (5) yields,

$$V_d = \frac{Z_0}{2} I_s \quad (7)$$

- Substituting (4) and (1) in (6) gives,

$$V_s = \frac{Z_0}{2} I_d \quad (8)$$

Impedance of Slot Antenna

- Multiplying (7) and (8) we have

$$\frac{V_s}{I_s} \frac{V_d}{I_d} = \frac{Z_0^2}{4} \tag{9}$$
$$\Rightarrow Z_s Z_d = \frac{Z_0^2}{4} \Rightarrow \boxed{Z_s = \frac{Z_0^2}{4Z_d}}$$

- For free space $Z_0 = 376.7 \Omega$,

$$\boxed{Z_s = \frac{Z_0^2}{4Z_d} = \frac{35476}{Z_d} (\Omega)}$$

Impedance of Slot Antenna-Example Calculation

The impedance of an infinitesimally thin $\lambda/2$ antenna ($L = 0.5\lambda, L/D = \infty$) is $73 + j42.5 \Omega$. Calculate the terminal impedance of an infinitesimally thin $\lambda/2$ slot antenna.

Given:

$$Z_d = 73 + j 42.5 \Omega$$

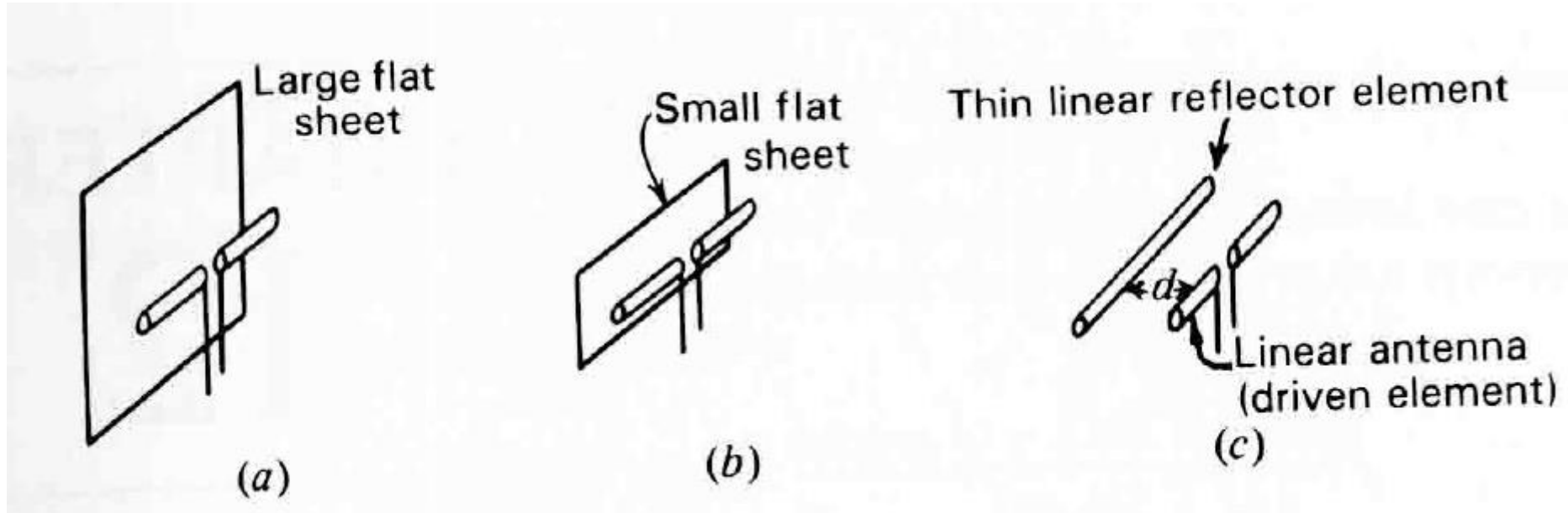
Solution:

$$Z_s = \frac{Z_0^2}{4Z_d} = \frac{35476}{Z_d} = \frac{35476}{73 + j 42.5} = 363 - j 211 \Omega$$

Reflector Antennas

- Reflectors are widely used to modify the radiation pattern of a radiating element.
- For example, the backward radiation from an antenna may be eliminated with a plane sheet reflector of large enough dimensions.
- In the more general case, a beam of pre-determined characteristics may be produced by means of a large, suitably shaped, and illuminated reflector surface.
- Reflector antenna is basically a combination of feed and reflecting surface to achieve high directivity at microwave frequencies.
- The feed is known as *primary antenna* and it can be slot, dipole or horn antenna.
- The reflector is also known as *secondary antenna* and it is a metallic plate which may be curved or flat and used to direct the incident energy in a specific direction to achieve high directivity
- **Applications:** Radars, radio astronomy, microwave communication, and satellite tracking

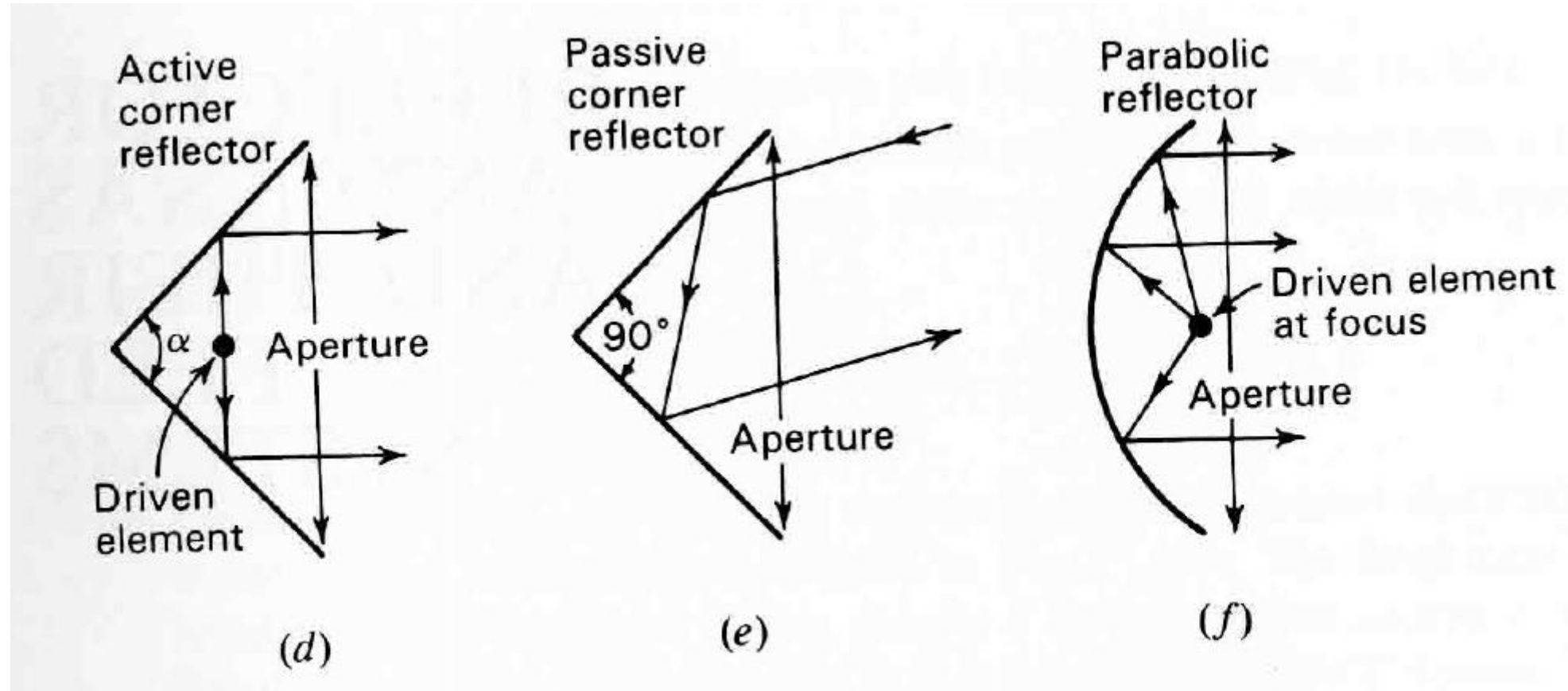
Reflectors of various shapes



Reflectors of various shapes

- Figure(a) has a large, flat sheet reflector near a linear dipole antenna to reduce the backward radiation.
- With small spacings between the antenna and sheet, this arrangement also yields a substantial gain in the forward radiation.
- The desirable properties of the sheet reflector may be largely preserved with the reflector reduced in size as in figure(b) and even in the limiting case of figure(c).
- Here the sheet has degenerated into a thin reflector element.
- Whereas the properties of the large sheet are relatively insensitive to small frequency changes, the thin reflector element is highly sensitive to frequency changes.

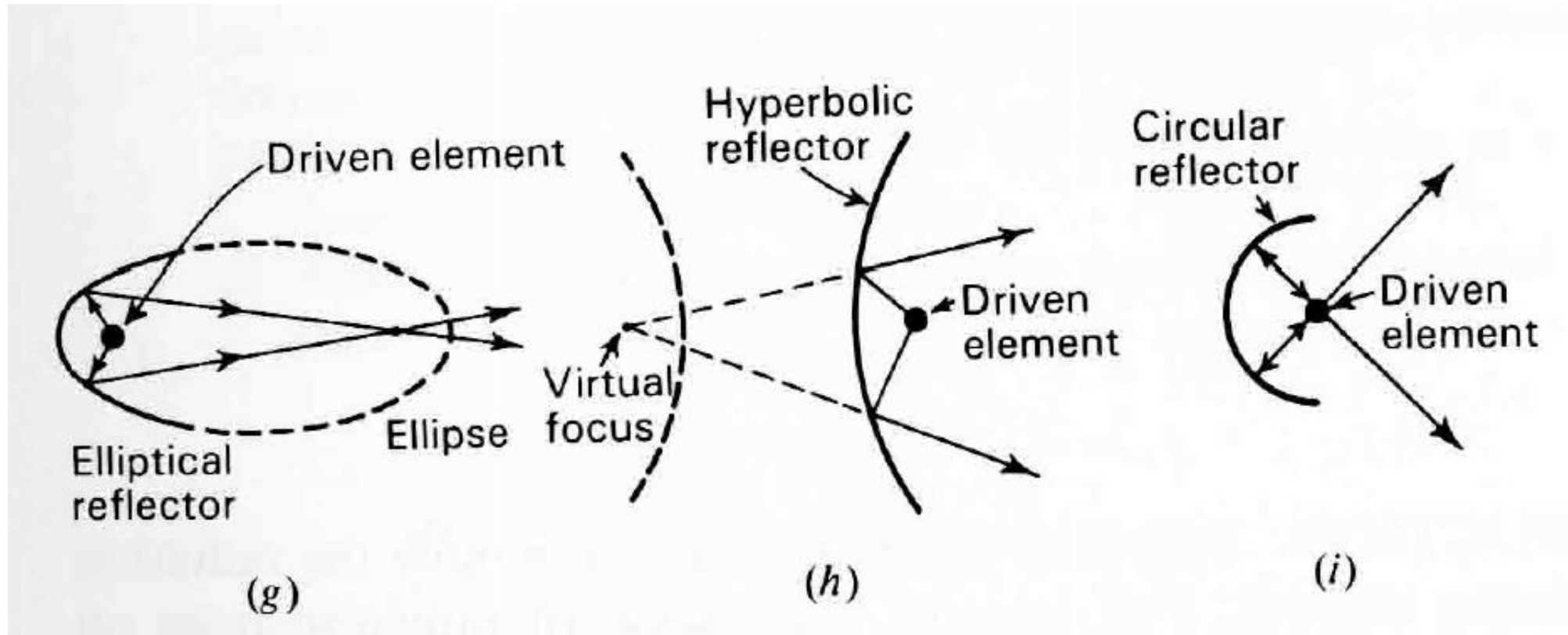
Reflectors of various shapes



Reflectors of various shapes

- With two flat sheets intersecting at an angle α ($< 180^\circ$) as in figure(d), a sharper radiation pattern than from a flat sheet reflector ($\alpha = 180^\circ$) can be obtained. This arrangement, called an *active corner reflector antenna*, is most practical where apertures of 1 or 2λ are of convenient size.
- A corner reflector without an exciting antenna can be used as a *passive reflector* or target or radar waves. In this application the aperture may be many wavelengths, and the corner angle is always 90° . Reflectors with this angle have the property that an incident wave is reflected back toward its source as in figure(e), the corner acting as a *retroreflector*.
- Parabolic reflectors can be used to provide highly directional antennas as shown in figure(f). The parabola reflects the waves originating from a source at the focus into a parallel beam, the parabola transforming the curved wave front from the feed antenna at the focus into a plane wave front.

Reflectors of various shapes



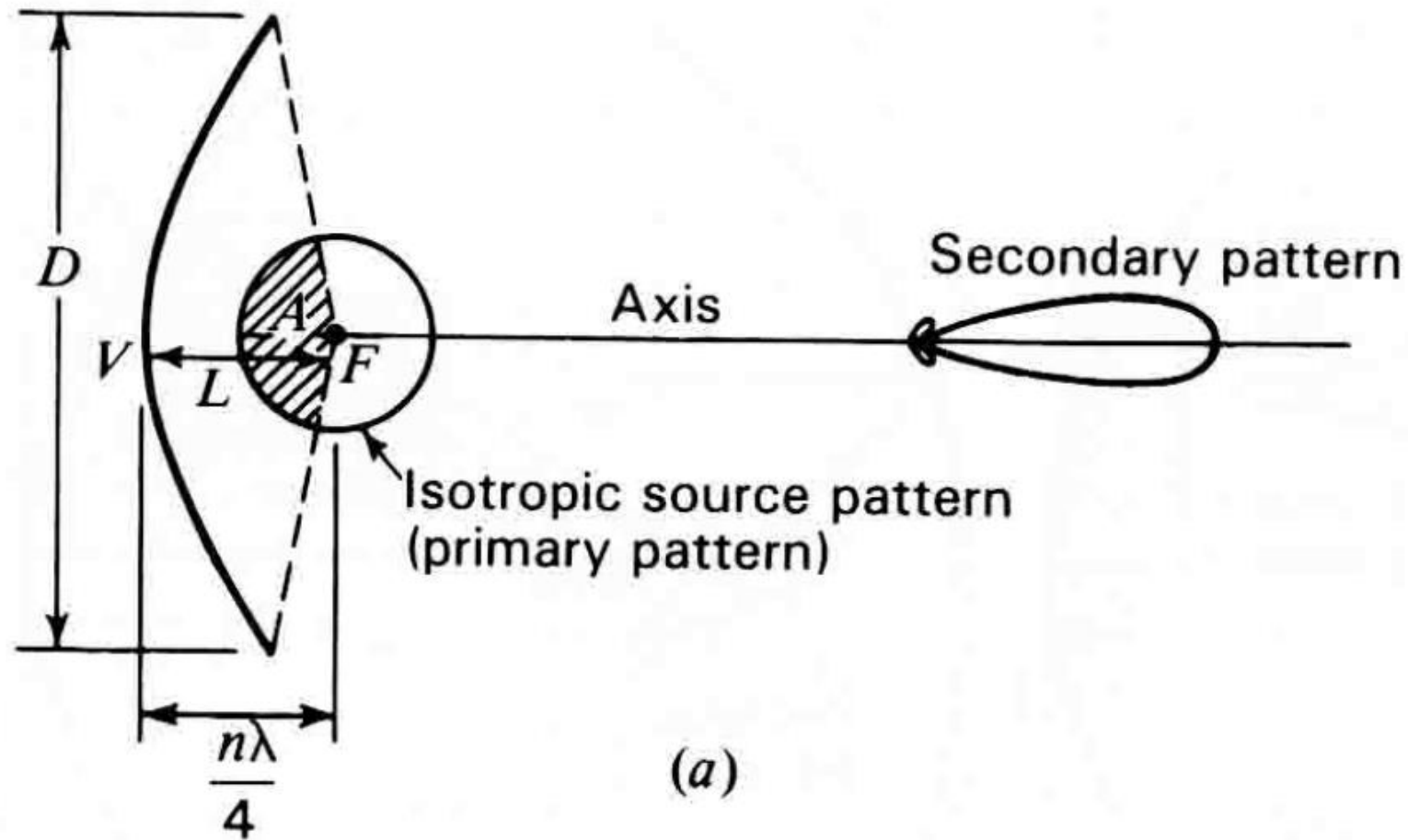
Reflectors of various shapes

- Many other shapes of reflectors can be employed for special applications.
- For instance, with an antenna at one focus, the elliptical reflector as shown in figure (g) produces a diverging beam with all reflected waves passing through the second focus of the ellipse.
- Examples of reflectors of other shapes are the hyperbolic and the circular reflectors as shown in figure(h) and (i).

Parabolic Reflector

- The surface generated by the revolution of a parabola around its axis is called a *paraboloid* or a *parabola of revolution*.
- If an isotropic source is placed at the focus of a paraboloidal reflector as in figure(a), the portion A of the source radiation that is intercepted by the paraboloid is reflected as a plane wave of circular cross section provided that the reflector surface deviates from a true parabolic surface by no more than a small fraction of a wavelength.
- If the distance L between the focus and vertex of the paraboloid is an even number of $\lambda/4$, the direct radiation in the axial direction from the source will be in opposite phase and will tend to cancel the central region of the reflected wave.
- However, if $L = n\lambda/4$ where $n = 1,3,5, \dots$, the direct radiation in the axial direction from the source will be in the same phase and will tend to reinforce the central region of the reflected wave.

Parabolic Reflector

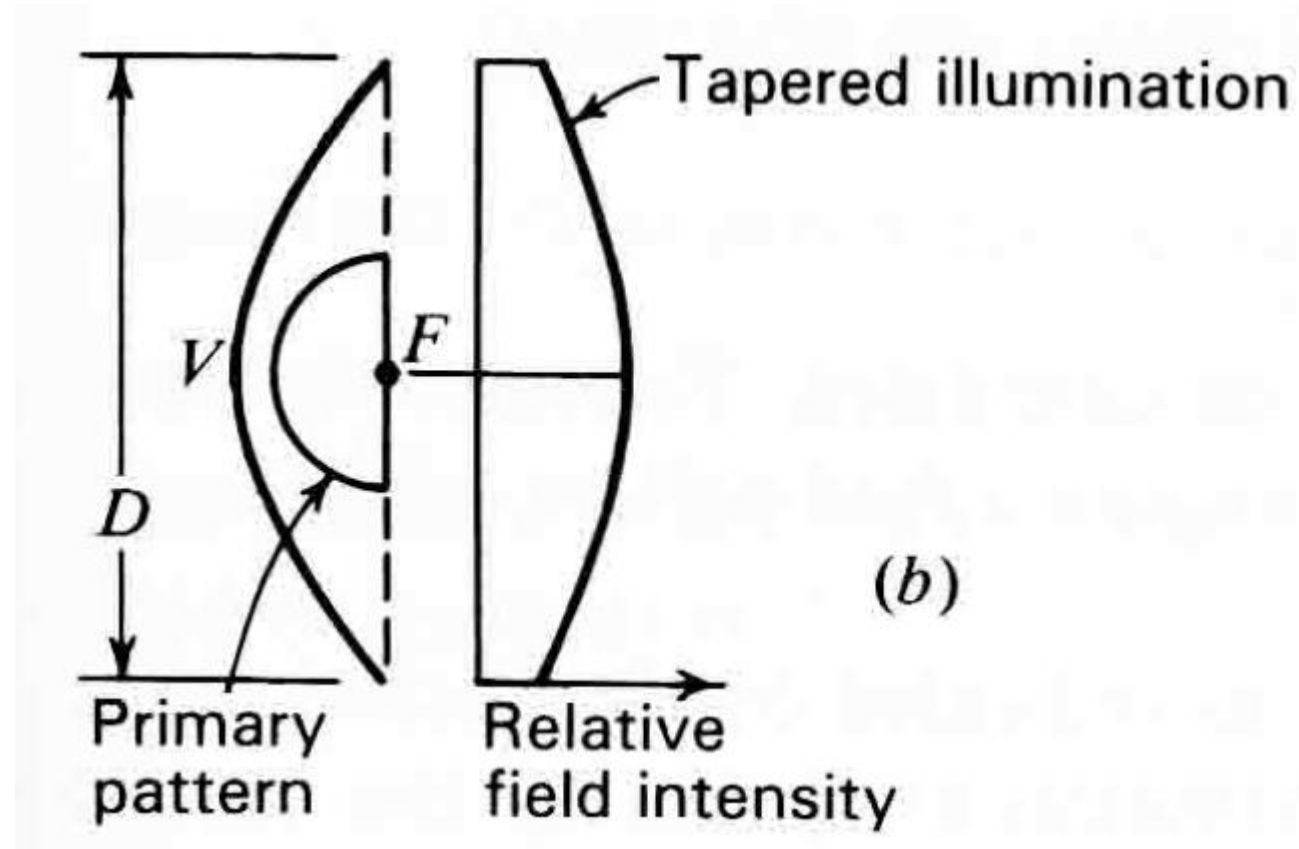


Parabolic reflectors of different focal lengths (L) with sources of different patterns

Parabolic Reflector

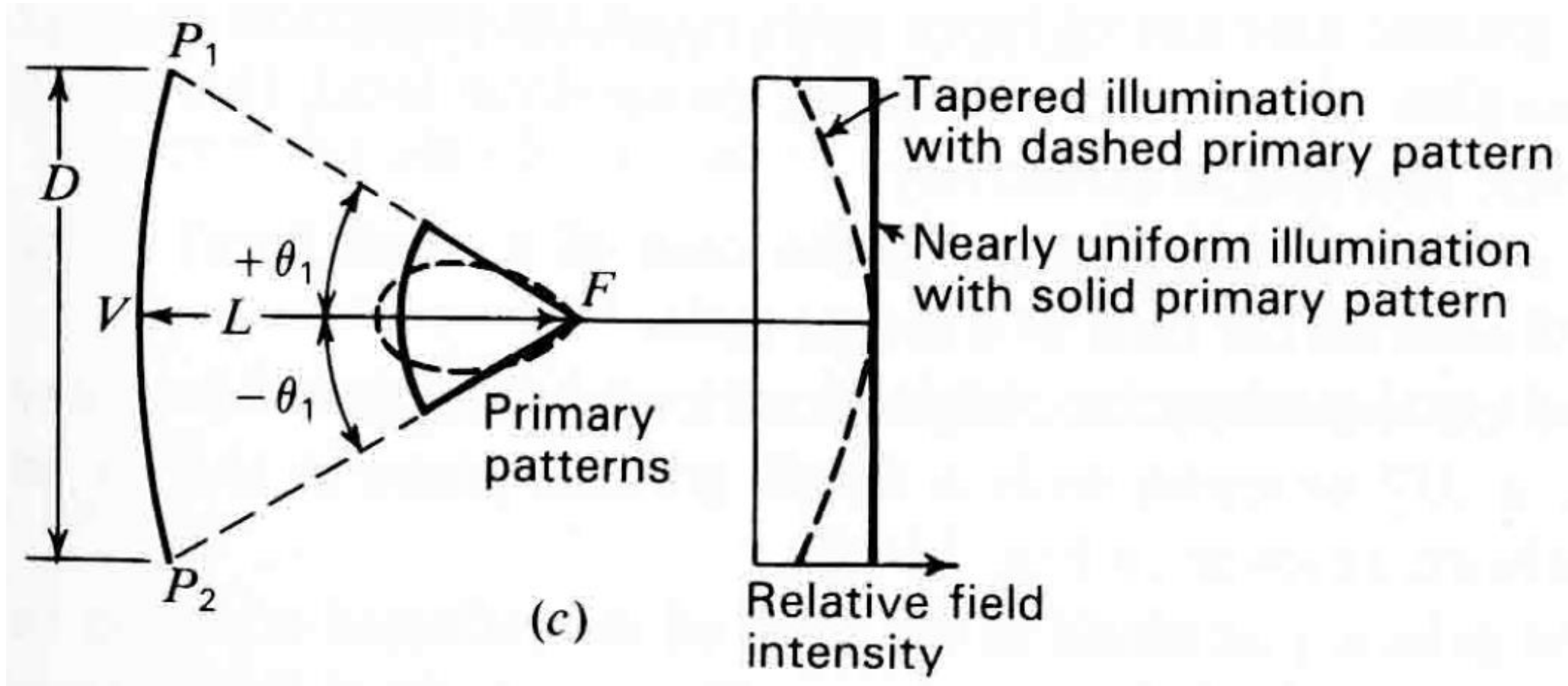
- Direct radiation from the source can be eliminated by means of a directional source or primary antenna as in figure(b) and (c).
- A primary antenna with the idealized hemispherical pattern is shown in figure(b) results in a wave of uniform phase over the reflector aperture. However, the amplitude is tapered as indicated.
- To obtain a more uniform aperture field distribution or illumination, it is necessary to make θ_1 small, as suggested in figure(c), by increasing the focal length L while keeping the reflector diameter D constant.
- A typical pattern for a directional source as indicated by the dashed curve at (c) (left) gives a more tapered aperture distribution as shown by the dashed curve at (c) (right). The greater amount of taper with resultant reduction in edge illumination may be desirable in order to reduce the minor-lobe level, this being achieved, however, at some sacrifice in directivity.

Parabolic Reflector



Parabolic reflectors of different focal lengths (L) with sources of different patterns

Parabolic Reflector



Parabolic reflectors of different focal lengths (L) with sources of different patterns

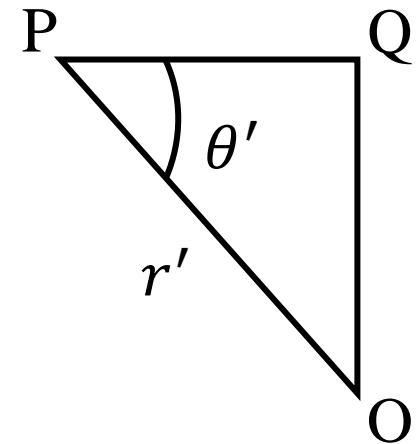
Parabolic Reflector-(f/d) ratio

- The surface of a paraboloidal reflector is formed by rotating a parabola about its axis. Its surface must be a paraboloid of revolution so that rays emanating from the focus of the reflector are transformed into plane waves. The design is based on optical techniques, and it does not take into account any deformations (diffractions) from the rim of the reflector.
- Referring to the following figure and choosing a plane perpendicular to the axis of the reflector through the focus, it follows that

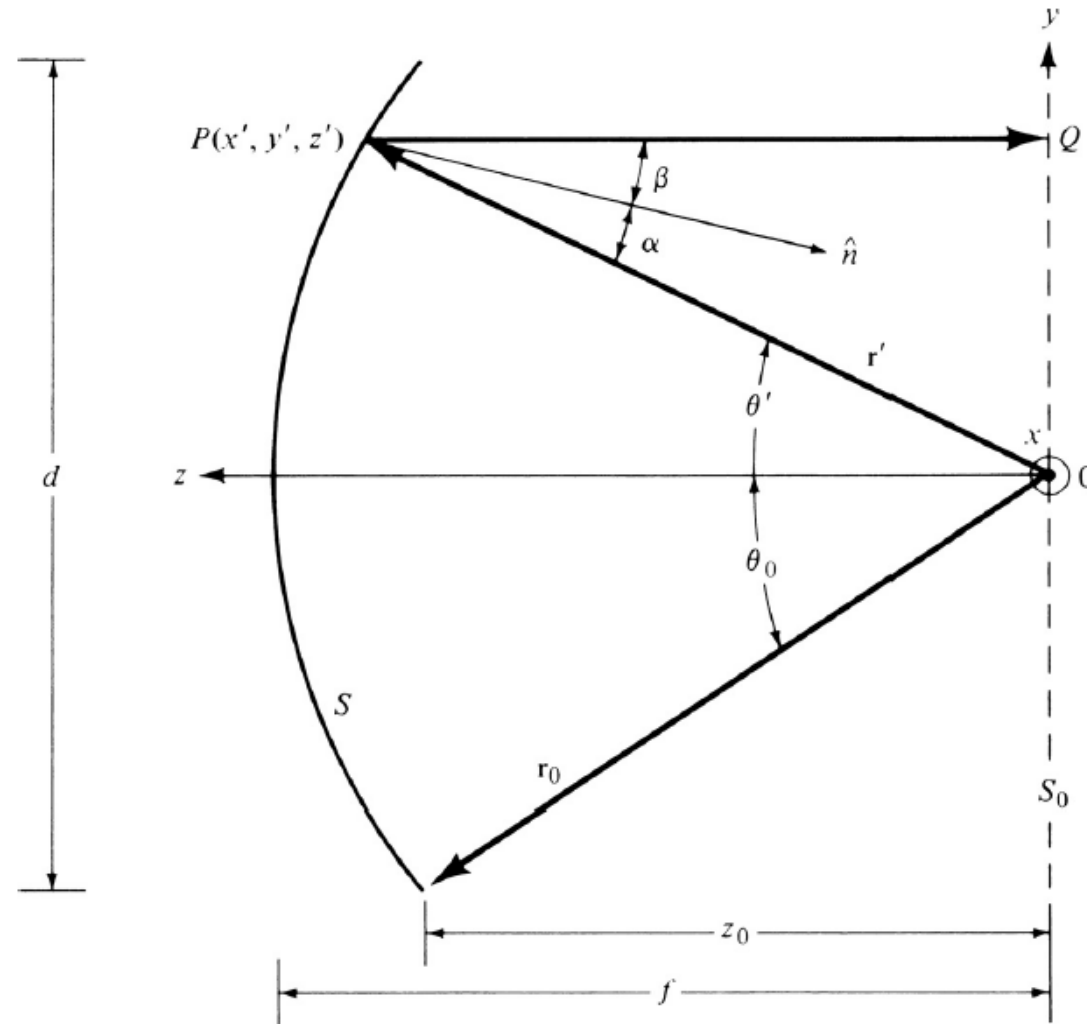
$$OP + PQ = \text{constant} = 2f \quad (1)$$

- From the figure,

$$\cos \theta' = \frac{PQ}{OP} = \frac{PQ}{r'} \Rightarrow PQ = r' \cos \theta'$$



Parabolic Reflector-(f/d) ratio



Two-dimensional configuration of a paraboloidal reflector

Parabolic Reflector-(f/d) ratio

- Eq.(1) can be written as

$$r' (1 + \cos \theta') = 2f \Rightarrow r' = \frac{2f}{1 + \cos \theta'} = f \sec^2 \left(\frac{\theta'}{2} \right) \quad \theta \leq \theta_0$$

- Since a paraboloid is a parabola of revolution (about its axis), above equation is also the equation of a paraboloid in terms of the spherical coordinates r', θ', ϕ' . Because of its rotational symmetry, there are no variations with respect to ϕ' . Above equation can also be written in terms of the rectangular coordinates x', y', z' . That is,

$$r' + r' \cos \theta' = \sqrt{(x')^2 + (y')^2 + (z')^2} + z' = 2f$$

$$\Rightarrow \sqrt{(x')^2 + (y')^2 + (z')^2} = 2f - z'$$

Parabolic Reflector-(f/d) ratio

$$\Rightarrow \sqrt{(x')^2 + (y')^2 + (z')^2} = 2f - z'$$

- Squaring on both sides,

$$(x')^2 + (y')^2 + (z')^2 = (2f - z')^2$$

$$\Rightarrow (x')^2 + (y')^2 + (z')^2 = 4f^2 + (z')^2 - 4fz'$$

$$(x')^2 + (y')^2 = 4f(f - z') \text{ with } (x')^2 + (y')^2 \leq (d/2)^2$$

- Another expression that is usually very prominent in the analysis of reflectors is that relating the subtended angle θ_0 to the f/d ratio. From the geometry of previous figure,

$$\theta_0 = \tan^{-1} \left(\frac{d/2}{z_0} \right)$$

where z_0 is the distance along the axis of the reflector from the focal point to the edge of the rim.

Parabolic Reflector-(f/d) ratio

- We can similarly have for point R (x_0, y_0, z_0) as

$$(x_0)^2 + (y_0)^2 = 4f (f - z_0) \Rightarrow \frac{(x_0)^2 + (y_0)^2}{4f} = f - z_0$$

$$z_0 = f - \frac{(x_0)^2 + (y_0)^2}{4f} \Rightarrow z_0 = f - \frac{(d/2)^2}{4f} \Rightarrow z_0 = f - \frac{d^2}{16f}$$

- Substituting above expression in $\theta_0 = \tan^{-1} \left(\frac{d/2}{z_0} \right)$ reduces it to

$$\theta_0 = \tan^{-1} \left(\frac{d/2}{z_0} \right) = \tan^{-1} \left| \frac{\frac{d}{2}}{f - \frac{d^2}{16f}} \right|$$

$$\Rightarrow \theta_0 = \tan^{-1} \left| \frac{\frac{d}{2}}{\frac{d^2}{f} \left[\left(\frac{f}{d} \right)^2 - \frac{1}{16} \right]} \right|$$

Parabolic Reflector-(f/d) ratio

$$\theta_0 = \tan^{-1} \left| \frac{\frac{d}{2}}{\frac{d^2}{f} \left[\left(\frac{f}{d}\right)^2 - \frac{1}{16} \right]} \right| \Rightarrow \theta_0 = \tan^{-1} \left| \frac{\frac{1}{2} \left(\frac{f}{d}\right)}{\left(\frac{f}{d}\right)^2 - \frac{1}{16}} \right|$$

- It can also be shown that another form of above expression is

$$f = \frac{d}{4} \cot \left(\frac{\theta_0}{2} \right)$$

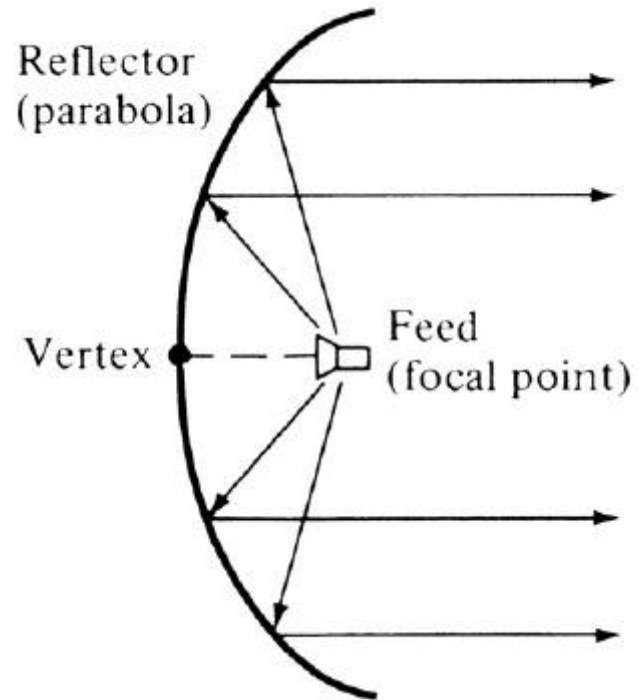
Feed systems for Parabolic Reflectors

Axial/Front Feed:

- The overall radiation characteristics (antenna pattern, antenna efficiency, polarization discrimination, etc.) of a reflector can be improved if the structural configuration of its surface is upgraded.
- It has been shown by geometrical optics that if a beam of parallel rays is incident upon a reflector whose geometrical shape is a parabola, the radiation will converge (focus) at a spot which is known as the *focal point*.
- In the same manner, if a point source is placed at the focal point, the rays reflected by a parabolic reflector will emerge as a parallel beam.
- This is one form of the principle of reciprocity, and it is demonstrated geometrically in the following figure.
- The symmetrical point on the parabolic surface is known as the *vertex*. Rays that emerge in a parallel formation are usually said to be collimated.

Feed systems for Parabolic Reflectors

Axial/Front Feed:



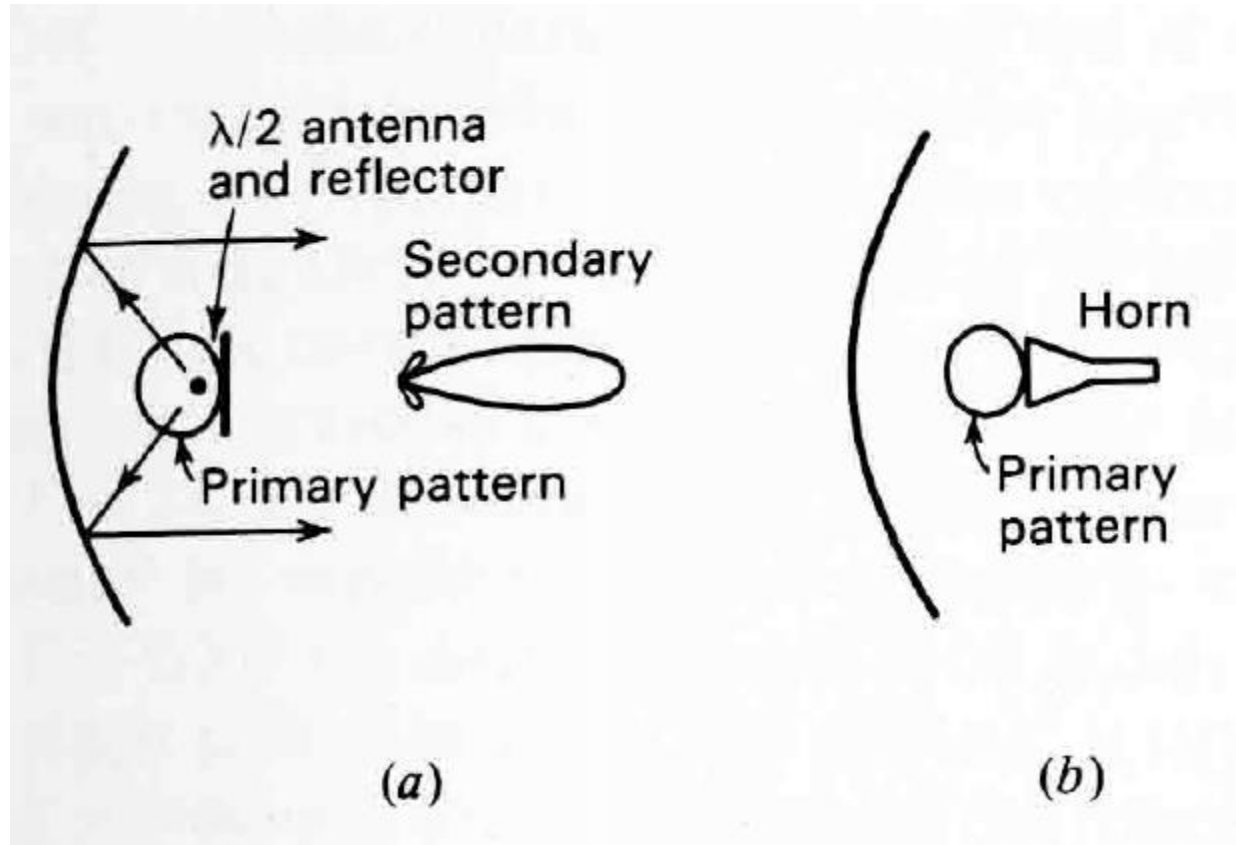
Feed systems for Parabolic Reflectors

Axial/Front Feed:

- In practice, collimation is often used to describe the highly directional characteristics of an antenna even though the emanating rays are not exactly parallel. Since the transmitter (receiver) is placed at the focal point of the parabola, the configuration is usually known as front fed.
- A $\lambda/2$ antenna with small ground plane is shown in figure(a) and a small horn antenna in figure(b).
- **Drawbacks:** The Presence of primary antenna in the path of reflected wave has 2 principal disadvantages:
 - Waves reflected from the parabola back to the primary antenna produce interaction and mismatching.
 - Primary antenna acts as an obstruction, blocking out the central portion of the aperture and increasing the minor lobes. (**Aperture Blockage**).

Feed systems for Parabolic Reflectors

Axial/Front Feed:



Feed systems for Parabolic Reflectors

Axial/Front Feed:

- The disadvantage of the front-fed arrangement is that the transmission line from the feed must usually be long enough to reach the transmitting or the receiving equipment, which is usually placed behind or below the reflector.
- This may necessitate the use of long transmission lines whose losses may not be tolerable in many applications, especially in low-noise receiving systems.
- In some applications, the transmitting or receiving equipment is placed at the focal point to avoid the need for long transmission lines.
- However, in some of these applications, especially for transmission that may require large amplifiers and for low-noise receiving systems where cooling and weatherproofing may be necessary, the equipment may be too heavy and bulky and will provide undesirable blockage.

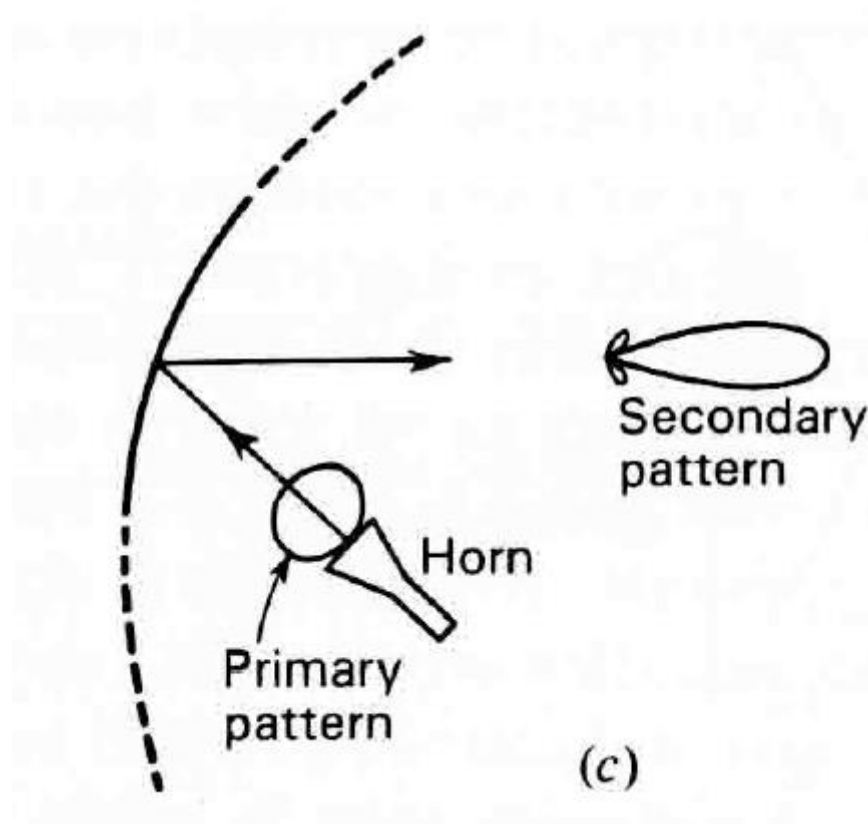
Feed systems for Parabolic Reflectors

Offset Feed:

- To eliminate some of the deficiencies of the symmetric configurations, offset-parabolic reflector designs have been developed for single- and dual-reflector systems.
- Offset-reflector designs reduce aperture blocking and VSWR. In addition, they lead to the use of larger f/d ratios while maintaining acceptable structural rigidity, which provide an opportunity for improved feed pattern shaping and better suppression of cross-polarized radiation emanating from the feed.
- However, offset-reflector configurations generate cross-polarized antenna radiation when illuminated by a linearly polarized primary feed. Circularly polarized feeds eliminate depolarization, but they lead to squinting of the main beam from boresight.
- In addition, the structural asymmetry of the system is usually considered a major drawback

Feed systems for Parabolic Reflectors

Offset Feed:



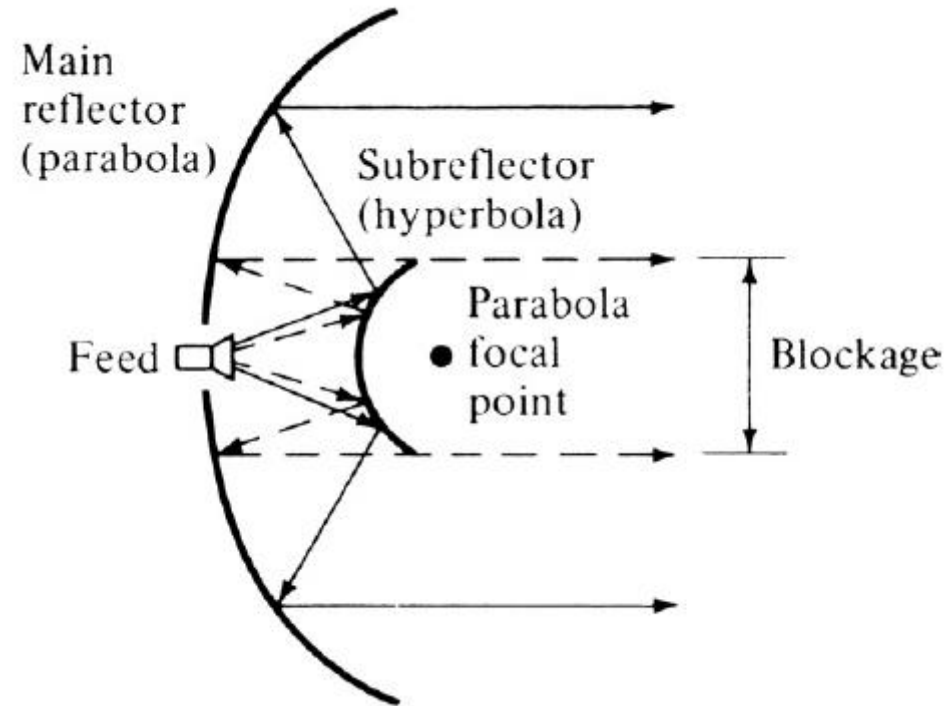
Feed systems for Parabolic Reflectors

Cassegrain Feed:

- Another arrangement that avoids placing the feed (transmitter and/or receiver) at the focal point is that shown in the following figure, and it is known as the Cassegrain feed.
- Through geometrical optics, Cassegrain, a famous astronomer (hence its name), showed that incident parallel rays can be focused to a point by utilizing two reflectors.
- To accomplish this, the main (primary) reflector must be a parabola, the secondary reflector (subreflector) a hyperbola, and the feed placed along the axis of the parabola usually at or near the vertex.
- Cassegrain used this scheme to construct optical telescopes, and then its design was copied for use in radio frequency systems.

Feed systems for Parabolic Reflectors

Cassegrain Feed:



Feed systems for Parabolic Reflectors

Cassegrain Feed:

- For this arrangement, the rays that emanate from the feed illuminate the subreflector and are reflected by it in the direction of the primary reflector, as if they originated at the focal point of the parabola (primary reflector).
- The rays are then reflected by the primary reflector and are converted to parallel rays, provided the primary reflector is a parabola and the subreflector is a hyperbola.
- Diffractions occur at the edges of the subreflector and primary reflector, and they must be taken into account to accurately predict the overall system pattern, especially in regions of low intensity.
- Even in regions of high intensity, diffractions must be included if an accurate formation of the fine ripple structure of the pattern is desired.

Feed systems for Parabolic Reflectors

Cassegrain Feed:

- With the Cassegrain-feed arrangement, the transmitting and/or receiving equipment can be placed behind the primary reflector. This scheme makes the system relatively more accessible for servicing and adjustments.
- In general, the Cassegrain arrangement provides a variety of benefits, such as the
 - ability to place the feed in a convenient location
 - reduction of spillover and minor lobe radiation
 - ability to obtain an equivalent focal length much greater than the physical length
 - capability for scanning and/or broadening of the beam by moving one of the reflecting surfaces

Feed systems for Parabolic Reflectors

Cassegrain Feed:

- To achieve good radiation characteristics, the subreflector or subdish must be several, at least a few, wavelengths in diameter.
- However, its presence introduces shadowing which is the principal limitation of its use as a microwave antenna.
- The shadowing can significantly degrade the gain of the system, unless the main reflector is several wavelengths in diameter.
- Therefore the Cassegrain is usually attractive for applications that require gains of 40 dB or greater.

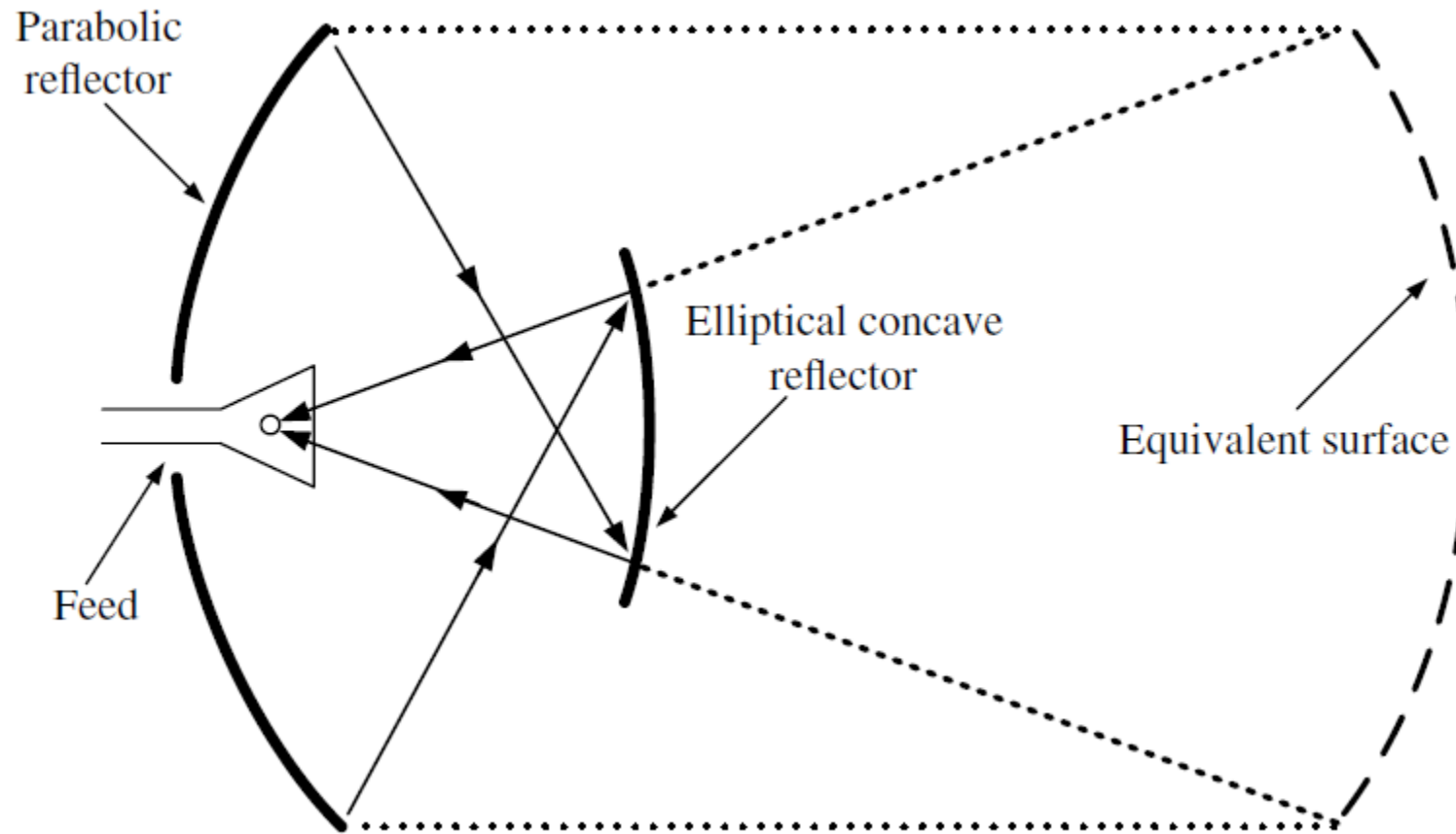
Feed systems for Parabolic Reflectors

Gregorian Feed:

- One of the other reflector arrangements is the classical Gregorian design as shown in the following figure, where the main reflector is a parabola while the subreflector is a concave ellipse.
- The focal point is between the main reflector and subreflector. Its equivalent parabola is shown dashed in the figure.
- When the overall size and the feed beamwidth of the classical Gregorian are identical to those of the classical Cassegrain, the Gregorian requires a shorter focal length for the main dish.

Feed systems for Parabolic Reflectors

Gregorian Feed:



Microstrip antennas

- In high-performance aircraft, spacecraft, satellite, and missile applications, where size, weight, cost, performance, ease of installation, and aerodynamic profile are constraints, low-profile antennas may be required.
- Microstrip antennas are
 - low profile
 - conformable to planar and nonplanar surfaces
 - simple and inexpensive to manufacture using modern printed-circuit technology
 - mechanically robust when mounted on rigid surfaces
 - compatible with MMIC designs and
 - when the particular patch shape and mode are selected, they are very versatile in terms of resonant frequency, polarization, pattern, and impedance.

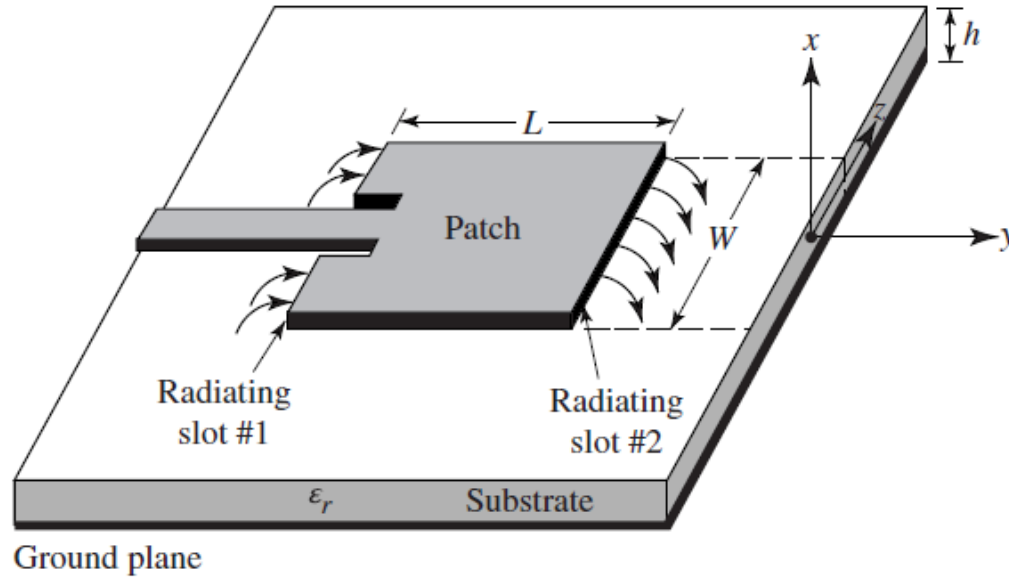
Microstrip antennas

- Major operational **disadvantages of microstrip antennas** are their
 - low efficiency
 - low power
 - High Q (sometimes in excess of 100)
 - poor polarization purity
 - poor scan performance
 - spurious feed radiation and
 - very narrow frequency bandwidth, which is typically only a fraction of a percent or at most a few percent.

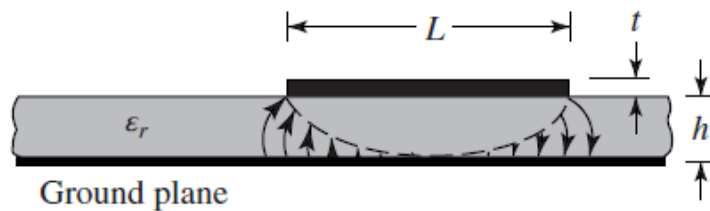
Basic Characteristics of Microstrip antennas

- Microstrip antennas received considerable attention starting in the 1970s, although the idea of a microstrip antenna can be traced to 1953 and a patent in 1955.
- Microstrip antennas, as shown in Figure (a), consist of a *very thin* ($t \ll \lambda_0$, where λ_0 is the free-space wavelength) *metallic strip (patch)* placed a small fraction of a wavelength ($h \ll \lambda_0$, usually $0.003\lambda_0 \leq h \leq 0.05\lambda_0$) above a *ground plane*.
- The microstrip patch is designed so its pattern maximum is normal to the patch (*broadside radiator*). This is accomplished by properly choosing the mode (field configuration) of excitation beneath the patch.
- *End-fire radiation* can also be accomplished by judicious mode selection.
- For a rectangular patch, the length L of the element is usually $\lambda_0/3 < L < \lambda_0/2$.
- The strip(patch) and the ground plane are separated by a *dielectric sheet* (referred to as the substrate), as shown in Figure (a).

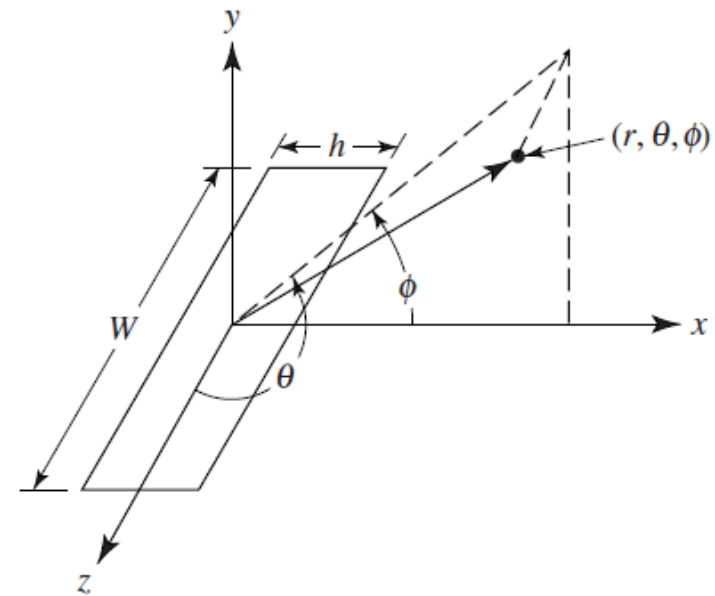
Basic Characteristics of Microstrip antennas



(a) Microstrip antenna



(b) Side view



(c) Coordinate system for each radiating slot

Basic Characteristics of Microstrip antennas

- There are numerous substrates that can be used for the design of microstrip antennas, and their *dielectric constants* are usually in the range of $2.2 \leq \epsilon_r \leq 12$.
- The ones that are most desirable for good antenna performance are *thick substrates* whose dielectric constant is in the lower end of the range because they provide better efficiency, larger bandwidth, loosely bound fields for radiation into space, but at the expense of larger element size.
- *Thin substrates* with higher dielectric constants are desirable for microwave circuitry because they require tightly bound fields to minimize undesired radiation and coupling, and lead to smaller element sizes; however, because of their greater losses, they are less efficient and have relatively smaller bandwidths.
- Since microstrip antennas are often integrated with other microwave circuitry, a compromise has to be reached between good antenna performance and circuit design.

Basic Characteristics of Microstrip antennas

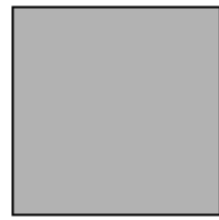
Company	Substrate	Thickness (mm)	Frequency (GHz)	ϵ_r	$\tan\delta$
Rogers Corporation	Duroid [®] 5880	0.127	0 – 40	2.20	0.0009
	RO 3003	1.575	0 – 40	3.00	0.0010
	RO 3010	3.175	0 – 10	10.2	0.0022
	RO 4350	0.168	0 – 10	3.48	0.0037
—	—	0.508	—	—	—
—	FR4	1.524	—	—	—
—	FR4	0.05 – 100	0.001	4.70	—
DuPont	HK 04J	0.025	0.001	3.50	0.005
Isola	IS 410	0.05 – 3.2	0.1	5.40	0.035
Arlon	DiClad 870	0.091	0 – 10	2.33	0.0013
Polyflon	Polyguide	0.102	0 – 10	2.32	0.0005
Neltec	NH 9320	3.175	0 – 10	3.20	0.0024
Taconic	RF-60A	0.102	0 – 10	6.15	0.0038

Typical Substrates and Their Parameters

Basic Characteristics of Microstrip antennas

- Often microstrip antennas are also referred to as *patch antennas*.
- The radiating elements and the feed lines are usually photoetched on the dielectric substrate.
- The radiating patch may be square, rectangular, thin strip (dipole), circular, elliptical, triangular, or any other configuration.
- *Square, rectangular, dipole (strip), and circular* are the most common because of ease of analysis and fabrication, and their attractive radiation characteristics, especially low cross-polarization radiation.
- *Microstrip dipoles* are attractive because they inherently possess a large bandwidth and occupy less space, which makes them attractive for arrays.
- *Linear and circular polarizations* can be achieved with either single elements or arrays of microstrip antennas.

Basic Characteristics of Microstrip antennas



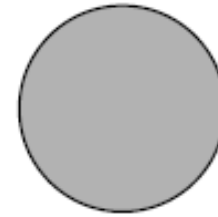
(a) Square



(b) Rectangular



(c) Dipole



(d) Circular



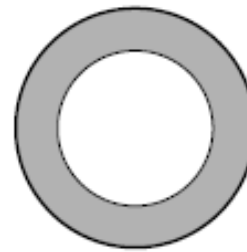
(e) Elliptical



(f) Triangular



(g) Disc sector



(h) Circular ring



(i) Ring sector

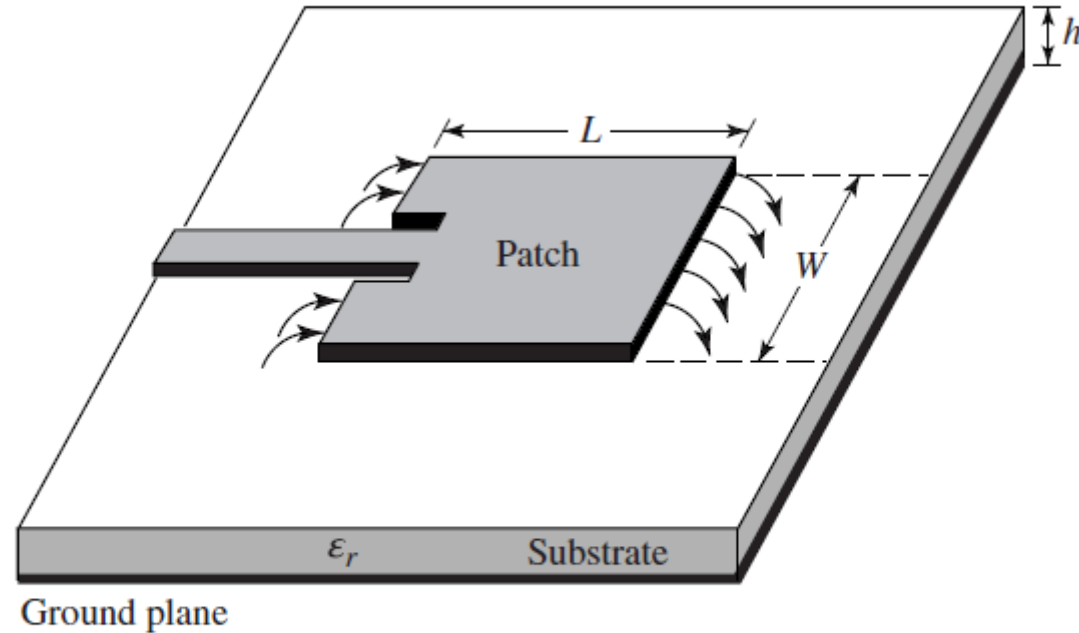
Representative shapes of microstrip patch elements

Microstrip antennas - Feeding Methods

- The four most popular configurations that can be used to feed microstrip antennas are
 - Microstrip Line Feed
 - Coaxial Probe Feed
 - Aperture Coupled Feed
 - Proximity Coupled Feed

Microstrip antennas - Feeding Methods

1. Microstrip Line Feed:



(a) Microstrip line feed

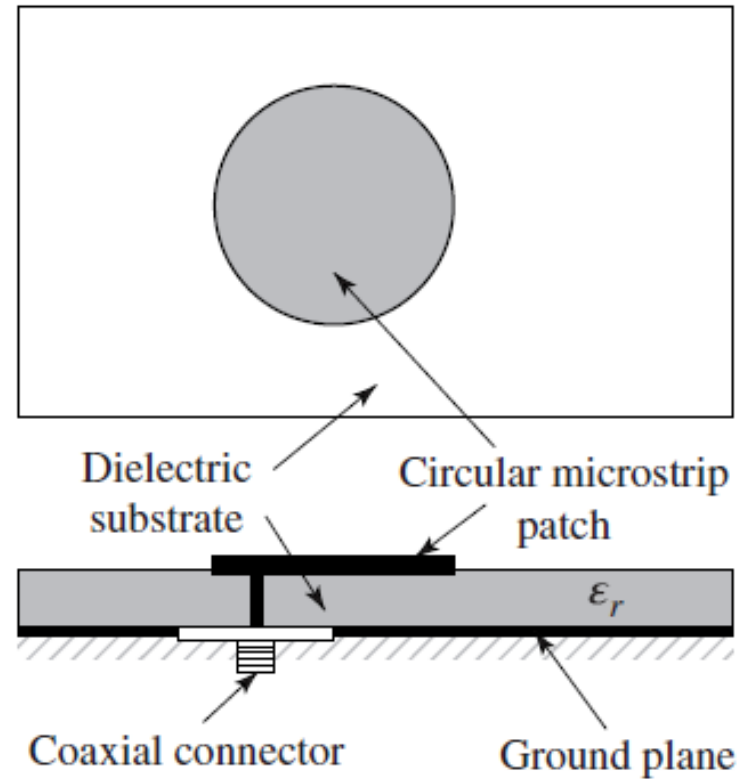
Microstrip antennas - Feeding Methods

1. Microstrip Line Feed:

- Easy to fabricate
- Simple to match by controlling inset position
- Simple to model
- As substrate thickness increases, surface waves and spurious feed radiation increase
 - This limits the bandwidth

Microstrip antennas - Feeding Methods

2. Coaxial line feed:



(b) Probe feed

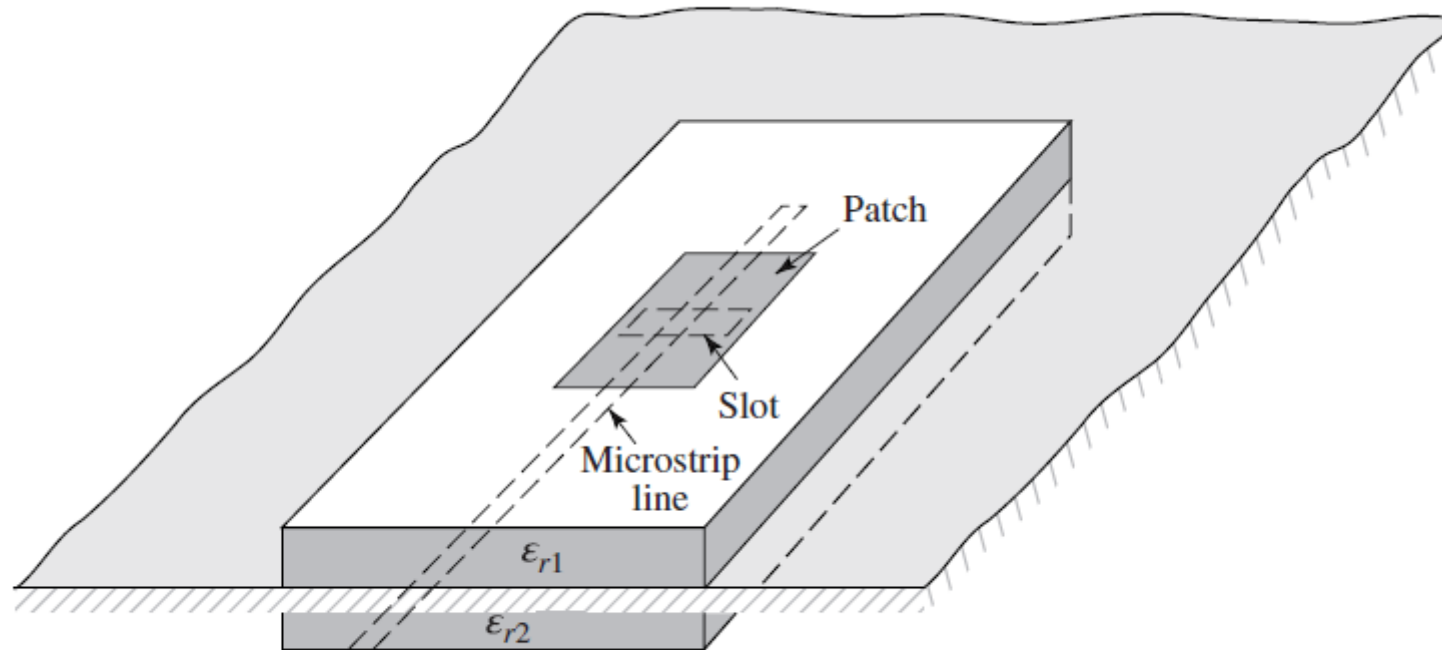
Microstrip antennas - Feeding Methods

2. Coaxial line feed:

- Inner conductor attached to the patch and outer conductor to ground plane
- Easy to fabricate and to match
- Low spurious radiation
- Narrow bandwidth
- Difficult to model, especially for thick substrates

Microstrip antennas - Feeding Methods

3. Aperture-coupled feed:



(c) Aperture-coupled feed

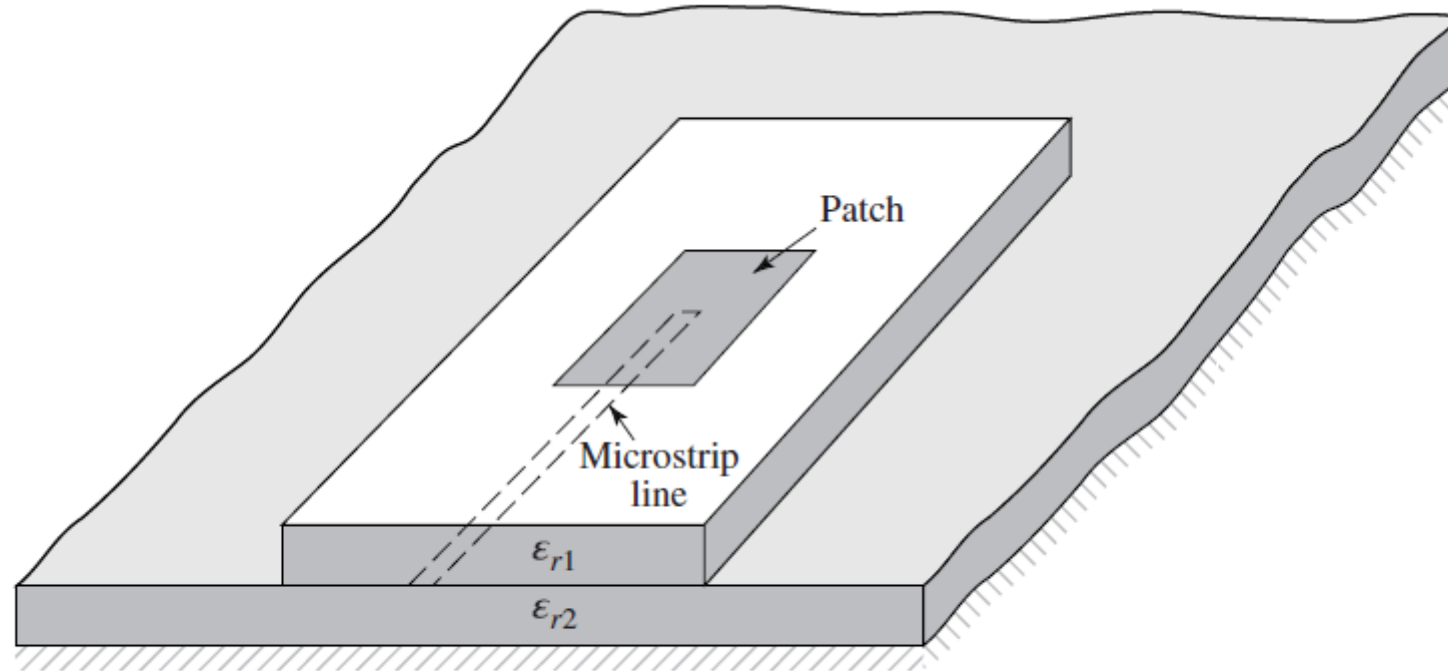
Microstrip antennas - Feeding Methods

3. Aperture-coupled feed:

- Non-contacting feed
- Most difficult to fabricate
- Narrow bandwidth
- Somewhat easier to model
- Moderate spurious radiation
- Independent optimization of feed mechanism and radiating element possible
- Feed line width and slot length used to control matching

Microstrip antennas - Feeding Methods

4. Proximity-coupled feed:



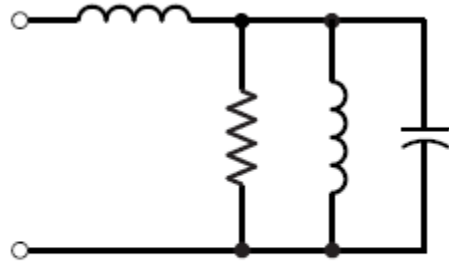
(d) Proximity-coupled feed

Microstrip antennas - Feeding Methods

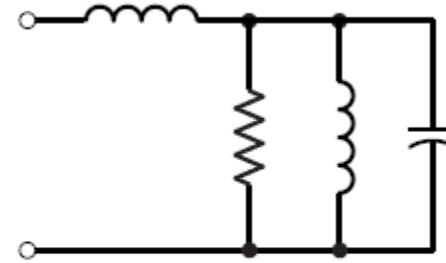
4. Proximity-coupled feed:

- Non-contacting feed
- Largest bandwidth
- Somewhat easy to model
- Low spurious radiation
- Fabrication somewhat more difficult
- Length of feeding stub and width-to-line ratio of patch can be used to control matching

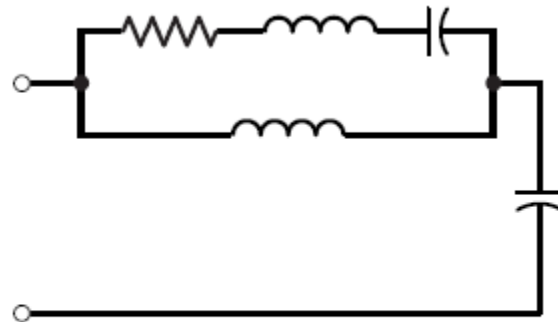
Microstrip antennas - Feeding Methods



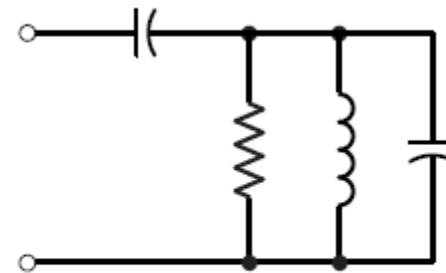
(a) Microstrip line



(b) Probe



(c) Aperture-coupled



(d) Proximity-coupled

Equivalent circuits for typical feeds

Microstrip antennas - Methods of Analysis

- There are many methods of analysis for microstrip antennas. The most popular models are the *transmission-line, cavity and full wave* (which include primarily integral equations/Moment Method).
- The *transmission-line model* is the easiest of all, it gives good physical insight, but is less accurate and it is more difficult to model coupling.
- Compared to the transmission-line model, the *cavity model* is more accurate but at the same time more complex. However, it also gives good physical insight and is rather difficult to model coupling, although it has been used successfully.

Microstrip antennas - Transmission-Line Model

- It was indicated earlier that the transmission-line model is the easiest of all but it yields the *least accurate results* and it lacks the versatility. However, it does shed some physical insight.
- Basically the transmission-line model represents the microstrip antenna by *two slots*, separated by a low-impedance Z_c transmission line of length L .

Fringing Effects:

- Because the dimensions of the patch are finite along the length and width, the fields at the edges of the patch undergo fringing.
- This is illustrated along the length in Figures (a,b) for the two radiating slots of the microstrip antenna.
- The same applies along the width.
- The amount of fringing is a function of the dimensions of the patch and the height of the substrate.

Microstrip antennas - Transmission-Line Model

Fringing Effects:

- For the principal E-plane (xy-plane) fringing is a function of the ratio of the length of the patch L to the height h of the substrate (L/h) and the dielectric constant ϵ_r of the substrate.
- Since for microstrip antennas $L/h \gg 1$, fringing is reduced; however, it must be taken into account because it influences the resonant frequency of the antenna. The same applies for the width.
- For a microstrip line shown in Figure (a), typical electric field lines are shown in Figure (b).
- This is a nonhomogeneous line of two dielectrics; typically the substrate and air.
- As can be seen, most of the electric field lines reside in the substrate and parts of some lines exist in air.

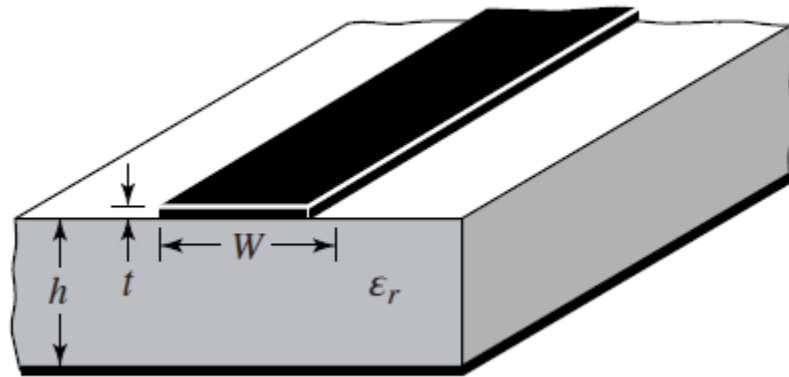
Microstrip antennas - Transmission-Line Model

Fringing Effects:

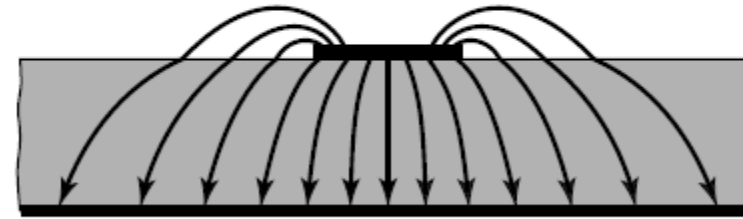
- As $W/h \gg 1$ and $\epsilon_r \gg 1$, the electric field lines concentrate mostly in the substrate.
- Fringing in this case makes the *microstrip line look wider electrically compared to its physical dimensions*.
- Since some of the waves travel in the substrate and some in air, an *effective dielectric constant* ϵ_{reff} is introduced to account for fringing and the wave propagation in the line.
- For a line with air above the substrate, the effective dielectric constant has values in the range of $1 < \epsilon_{\text{reff}} < \epsilon_r$.

$$\epsilon_{\text{reff}} = \frac{\epsilon_r + 1}{2} + \frac{\epsilon_r - 1}{2} \left[1 + \frac{12h}{W} \right]^{-1/2}$$

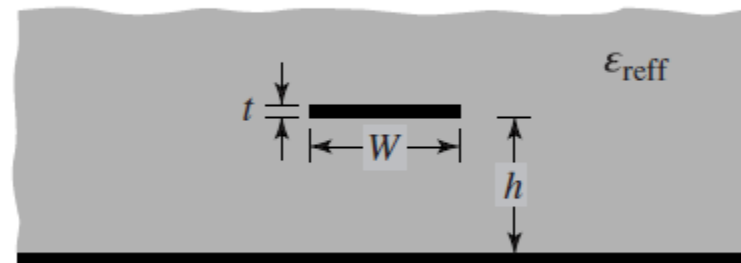
Microstrip antennas - Transmission-Line Model



(a) Microstrip line



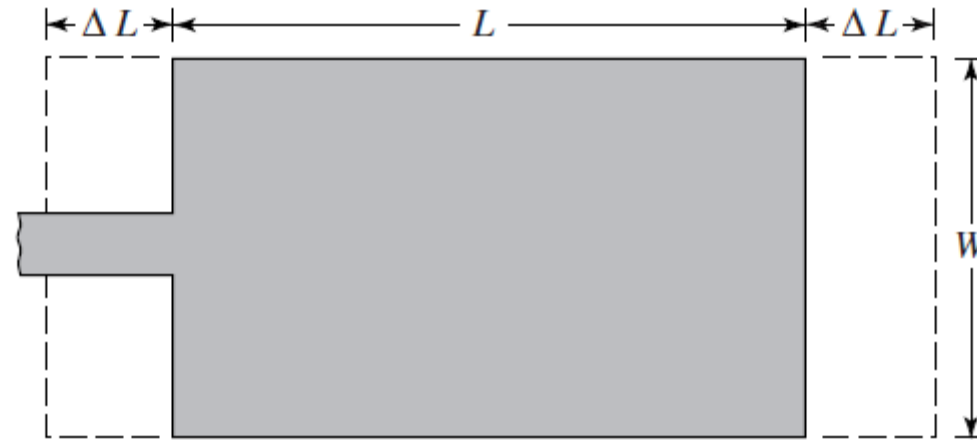
(b) Electric field lines



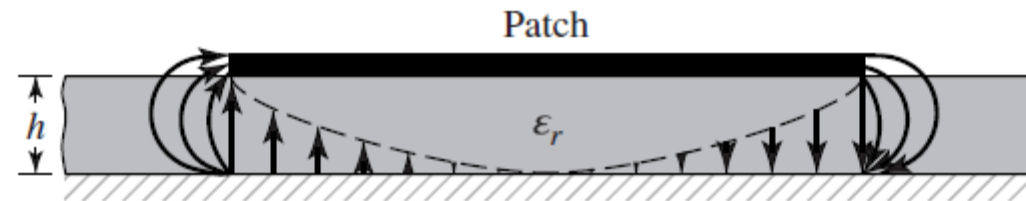
(c) Effective dielectric constant

Microstrip line and its electric field lines, and effective dielectric constant geometry

Microstrip antennas - Transmission-Line Model



(a) Top view



(b) Side view

Physical and effective lengths of rectangular microstrip patch

Microstrip antennas - Transmission-Line Model

Effective Length, Resonant Frequency, and Effective Width:

- Because of the fringing effects, electrically the patch of the microstrip antenna looks greater than its physical dimensions.

$$\frac{\Delta L}{h} = 0.412 \frac{(\epsilon_{reff} + 0.3) \left(\frac{W}{h} + 0.264\right)}{(\epsilon_{reff} - 0.258) \left(\frac{W}{h} + 0.8\right)}$$

$$L_{eff} = L + 2\Delta L$$

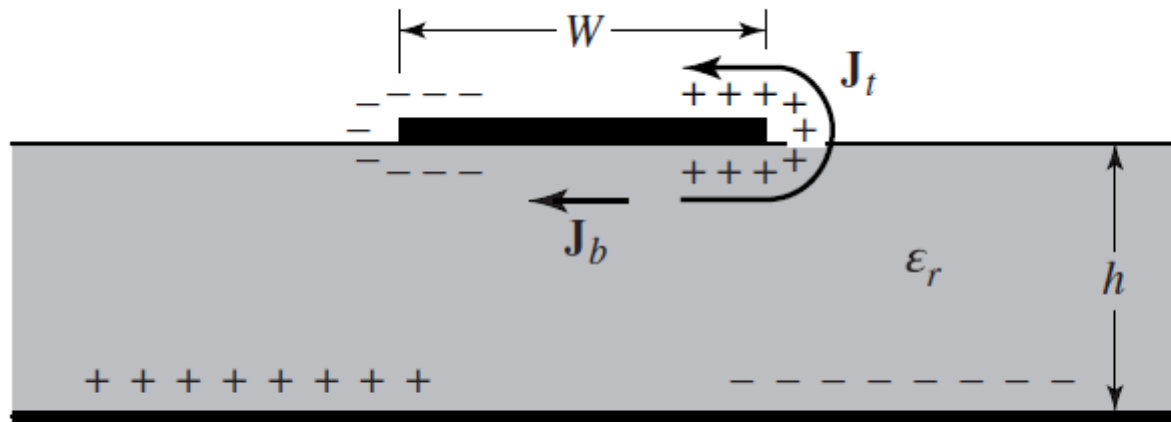
- $L = \lambda/2$ for dominant TM_{010} mode with no fringing.

$$(f_{rc})_{010} = \frac{1}{2L_{eff} \sqrt{\epsilon_{reff}} \sqrt{\mu_0 \epsilon_0}}$$

Microstrip antennas - Cavity Model

- Using the cavity model, a rectangular microstrip antenna can be represented as an array of two radiating narrow apertures (slots), each of width W and height h , separated by a distance L .
- Excitation establishes charge distribution
- This distribution controlled by two mechanisms:
 - An attractive one between the corresponding opposite charges on the bottom side of the patch and the ground plane; and
 - A repulsive one between like charges on the bottom side of the patch
- While the attraction tends to keep charge distribution on the bottom of the patch, the repulsion tends to push some charges from the patch bottom to the top surface, around the edges.
- The consequent movement of charge carriers gives rise to J_b and J_t

Microstrip antennas - Cavity Model



Charge distribution and current density creation on microstrip patch

Microstrip antennas - Cavity Model

- The height-to-width ratio generally being very small, attractive mechanism dominates, thus restricting current flow to bottom surface.
- Thus the tangential magnetic field to the edges is insignificant
- Consequently, *the four side walls can be modeled as PMC* which ideally would not disturb the H-field, and in turn, the E-field beneath the patch.
- This model produces good E and H field distributions beneath the patch.
- With the assumption of *lossless cavity walls* and the filling dielectric, the *cavity would not radiate* and its input impedance would be purely reactive.
- To account for radiation, therefore, a loss mechanism needs to be introduced.
- This is done by introducing an effective loss tangent δ_{eff} .

Microstrip antennas - Cavity Model

- The top and bottom walls are *PEC*
- The four side walls are *PMC*
 - Tangential magnetic fields vanish along these four walls
- In summary, the cavity method is based on the following physical model:
 - The electric field is mainly localized in the cylinder of height h between the patch and the ground plane;
 - The radiation is the result of leakage from the cylindrical cavity via its lateral walls (as the cavity ends are perfectly conducting)

FREQUENCY INDEPENDENT ANTENNAS

CONTENTS

- Principle of frequency independent antennas
 - Spiral antenna
 - Log-Periodic antenna

FREQUENCY INDEPENDENT ANTENNA

- A true frequency independent antenna is physically fixed in size and operates on an instantaneous basis over a wide bandwidth with relatively constant impedance, pattern, polarization and gain.
- Their geometries are specified in angles.
- E.g.: Bi-conical antenna, Spiral antenna and Helical antenna

RUMSEY'S PRINCIPLE

Rumsey's principle is that the impedance and pattern properties of an antenna will be frequency independent if the antenna shape is specified only in terms of angles.

We begin by assuming that an antenna, whose geometry is best described by the spherical coordinates (r, θ, ϕ) , has both terminals infinitely close to the origin and each is symmetrically disposed along the $\theta = 0, \pi$ -axes. It is assumed that the antenna is perfectly conducting, it is surrounded by an infinite homogeneous and isotropic medium, and its surface or an edge on its surface is described by a curve

$$r = F(\theta, \phi) \quad (1)$$

where r represents the distance along the surface or edge. If the antenna is to be scaled to a frequency that is K times lower than the original frequency, the antenna's physical surface must be made K times greater to maintain the same electrical dimensions. Thus, the new surface is described by

$$r' = KF(\theta, \phi) \quad (2)$$

The new and old surfaces are identical; that is, not only are they similar but they are also congruent (if both surfaces are infinite). Congruence can be established only by rotation in ϕ . Translation is not allowed because the terminals of both surfaces are at the origin. Rotation in θ is prohibited because both terminals are symmetrically disposed along the $\theta = 0, \pi$ -axes.

For the second antenna to achieve congruence with the first, it must be rotated by an angle C so that

$$KF(\theta, \phi) = F(\theta, \phi + C) \quad (3)$$

The angle of rotation C depends on K but neither depends on θ or ϕ . Physical congruence implies that the original antenna electrically would behave the same at both frequencies. However, the radiation pattern will be rotated azimuthally through an angle C . For unrestricted values of $K(0 \leq K \leq \infty)$, the pattern will rotate by C in ϕ with frequency, because C depends on K but its shape will be unaltered. Thus, the impedance and pattern will be frequency independent.

To obtain the functional representation of $F(\theta, \phi)$, both sides of (3) are differentiated with respect to C to yield

$$\frac{d}{dC} [KF(\theta, \phi)] = \frac{d}{dC} [F(\theta, \phi + C)] \quad (4)$$

$$F(\theta, \phi) \frac{dK}{dC} = \frac{\partial}{\partial(\phi + C)} [F(\theta, \phi + C)] \quad (5)$$

Differentiate (3) relative to ϕ , we get

$$\frac{\partial}{\partial\phi} [KF(\theta, \phi)] = \frac{\partial}{\partial\phi} [F(\theta, \phi + C)] \quad (6)$$

$$K \frac{\partial}{\partial \phi} [F(\theta, \phi)] = \frac{\partial}{\partial (\phi + C)} [F(\theta, \phi + C)] \quad (7)$$

Comparing (5) and (7), we get

$$F(\theta, \phi) \frac{dK}{dC} = K \frac{\partial F(\theta, \phi)}{\partial \phi} \quad (8)$$

By substituting $r = F(\theta, \phi)$,

$$r \frac{dK}{dC} = K \frac{\partial r}{\partial \phi} \quad (9)$$

$$\frac{1}{K} \frac{dK}{dC} = \frac{1}{r} \frac{\partial r}{\partial \phi} \quad (10)$$

LHS of above equation is independent of θ and ϕ , therefore the general solution is given as

$$r = F(\theta, \phi) = e^{a\phi} f(\theta) \quad (11)$$

where $a = \frac{1}{K} \frac{dK}{dC}$ and $f(\theta)$ is completely arbitrary function

Thus, for any antenna to have frequency independent characteristics, its surface must be described by (11).

FREQUENCY-INDEPENDENT PLANAR LOG SPIRAL ANTENNA

The equation for a logarithmic or log spiral is given by

$$r = a^\theta \quad (12)$$

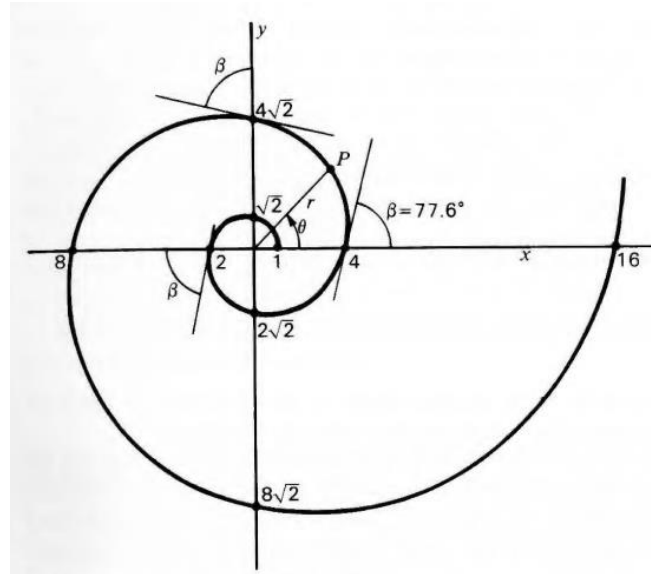
$$\ln r = \ln a^\theta = \theta \ln a \quad (13)$$

where,

r = radial distance to point P on spiral

θ = angle with respect to x-axis

a = constant



Logarithmic spiral or log spiral

From (12), the rate of change of radius with angle is

$$\frac{dr}{d\theta} = a^\theta \ln a = r \ln a \quad (14)$$

The constant a in (14) is related to the angle β between the spiral and a radial line from the origin as given by

$$\ln a = \frac{dr}{rd\theta} = \frac{1}{\tan\beta} \quad (15)$$

Thus, from (14) and (15),

$$\theta = \frac{\ln r}{\ln a} = \tan\beta \ln r \quad (16)$$

The log spiral was constructed so as to make $r = 1$ at $\theta = 0$ and $r = 2$ at $\theta = \pi$. These conditions determine the value of the constants a and β . Thus, from (15) and (16),

$$\tan\beta = \frac{\theta}{\ln r} \Rightarrow \beta = \tan^{-1} \left[\frac{\theta}{\ln r} \right] \quad (17)$$

For $r = 2$ at $\theta = \pi$,

$$\beta = \tan^{-1} \left[\frac{\pi}{\ln 2} \right] = 77.6^\circ$$

$$\ln a = \frac{1}{\tan\beta} \Rightarrow a = \exp \left\{ \frac{1}{\tan 77.6^\circ} \right\} = 1.247$$

Thus, the shape of the spiral is determined by the angle β which is same for all points on the spiral.

Let a second log spiral, identical in form to the one in figure shown below, be generated by an angular rotation of the spiral by a factor δ so that (12) becomes

$$r_2 = a^{\theta-\delta} \quad (18)$$

and a third spiral and fourth spiral given by

$$r_3 = a^{\theta-\pi} \quad (19)$$

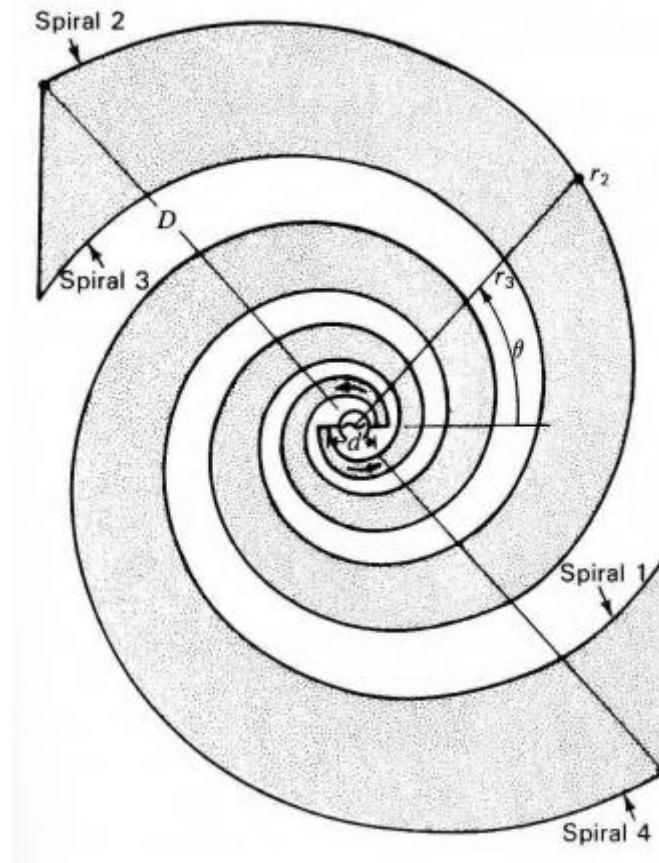
and

$$r_4 = a^{\theta-\pi-\delta} \quad (20)$$

Then, for a rotation $\delta = \frac{\pi}{2}$, we have 4 spirals at 90° angles. Metalizing the areas between the spirals 1 & 4 and 2 & 3, with the other areas open, self-complementary and congruence conditions are satisfied. Connecting a generator or receiver across the inner terminals, we obtain ***Dyson's frequency independent planar spiral antenna.***

- **Polarization:** RHCP radiation from the page and LHCP radiation into the page
- High frequency limit of operation is determined by the spacing d of the input terminal and the low frequency limit by overall diameter D . The ratio D/d for the antenna is about 25 to 1.

- If we take $d = \frac{\lambda}{10}$ at the high frequency limit and $D = \frac{\lambda}{2}$ at the low frequency limit, the antenna bandwidth is 5 to 1.



Frequency Independent Planar Spiral Antenna

- **Spiral slot antenna:** Spiral-shaped slots are cut from a large ground plane and the antenna is fed with a co-axial cable bonded to one of the spiral arms. A dummy cable may be bonded to the other arm of symmetry.
- **Radiation Pattern for spiral antenna:** Bi-directional broadside to the plane of the spiral. The patterns in both directions have a single broad lobe so that the gain is only a few dBi.
- **Input Impedance** depends on δ and a and terminal separation (≈ 50 to 100Ω)
- Ratio K of the radii across any arm, such as between spiral 2 and 3 is given by

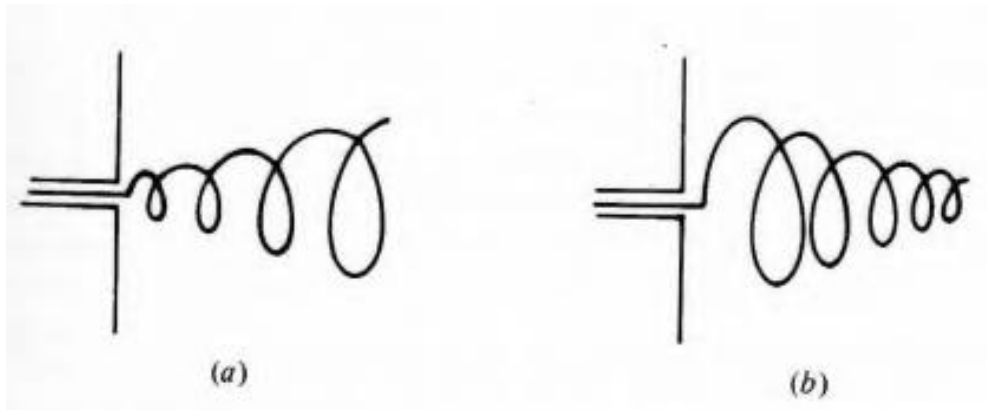
$$K = \frac{r_3}{r_2} = \frac{a^{\theta-\pi}}{a^{\theta-\delta}} = a^{-\pi+\delta}$$

$$\text{For } \delta = \frac{\pi}{2}$$

$$K = \frac{r_3}{r_2} = a^{-\frac{\pi}{2}} = (1.247)^{-\frac{\pi}{2}} = 0.707 = \frac{1}{\sqrt{2}}$$

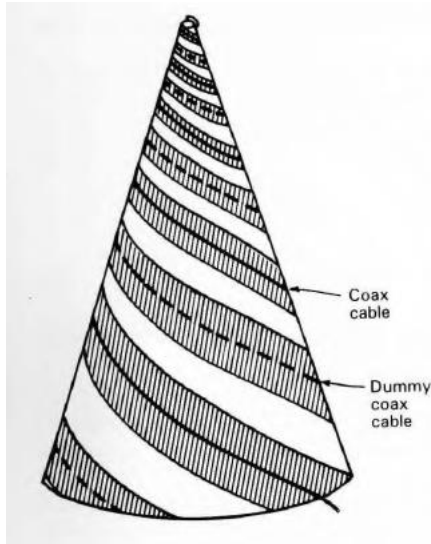
FREQUENCY INDEPENDENT CONICAL SPIRAL ANTENNA

- A tapered helix is a conical spiral antenna in which pitch angle is constant with diameter and turn spacing variable.



Tapered Helix or Conical Spiral (Forward-fire)

- Two arms of the conical spiral are fed at the center point or apex from a co-axial cable bonded to one of the arms, the spiral acting as a balun.
- For symmetry, a dummy cable may be bonded to the other arm.
- According to Dyson, input impedance is between 100 to 150 Ω for pitch angle $\alpha = 17^\circ$ and full cone angles of 20° and 60° . Smaller cone angles (less than 30°) have high F/B ratio.
- Radiation Pattern: Uni-directional with maxima towards the apex.
- $\text{Bandwidth} \propto \frac{\text{Base diameter } (\sim \frac{\lambda}{2} \text{ at low frequency})}{\text{Truncated apex diameter } (\sim \frac{\lambda}{4} \text{ at high frequency})}$



Dyson 2-arm balanced conical spiral (backward-fire antenna). Polarization is RCP. Inner conductor of coax connects to dummy at apex.

LOG-PERIODIC ANTENNA

- Log-Periodic antenna operates over broadband and its size varies with the operating frequency or wavelength. Although, LPDA is not specified in terms of angles yet its geometry is adjusted such that all the electrical properties of the antenna are repeated periodically with the logarithm of the frequency.
- Dwight Isbell demonstrated first LPDA (1960).
- **Basic concept:** A gradually expanding periodic structure array radiates more effectively when the array elements(dipoles) are near resonance so that with change in frequency, the active region moves along the array.
- LPDA– a number of dipole antennas of different lengths are arranged at different spacings, used in array form.
- Dipole lengths increase along the antenna so that the included angle α is constant, and the lengths (l) and spacing(S) of the adjacent elements are scaled so that

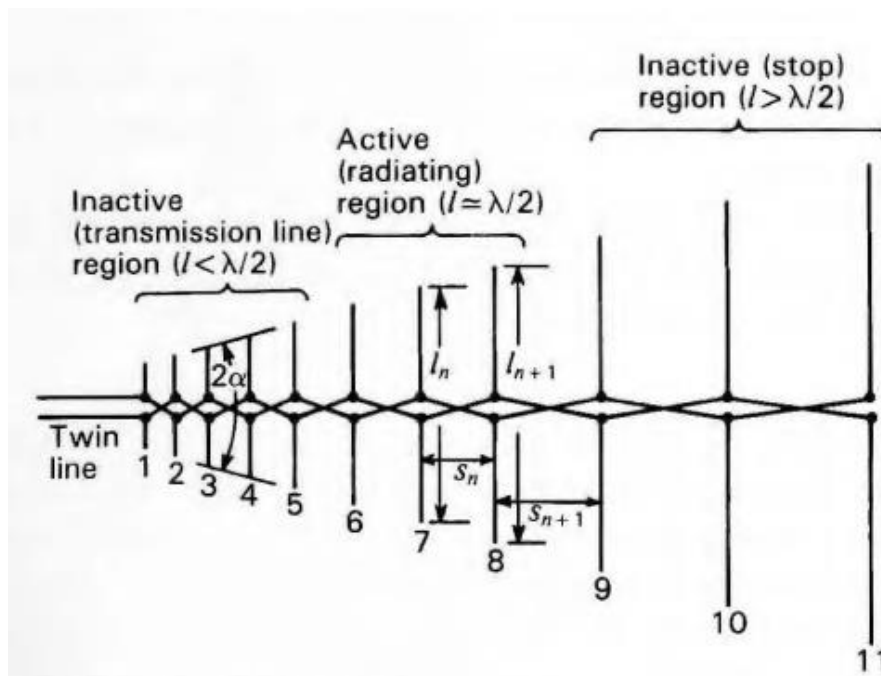
$$\frac{l_{n+1}}{l_n} = \frac{S_{n+1}}{S_n} = k = \frac{1}{\tau} \quad (21)$$

where

$k = \text{constant } (k > 1)$

$\tau = \text{scale factor or design ratio or geometric ratio or periodicity factor } (\tau > 1)$

- Antenna is fed through a balanced twin line in zig-zag form which means the alternate dipole arms are fed through common line.
- The apex angle is formed by the two imaginary straight lines passing through the edges of the dipole arms located on either side.
- Moreover, the spacing between the dipoles near to the apex is smaller as compared to the spacing at the base.
- There are three important regions of LPDA namely
 - i. Inactive region (Transmission line)
 - ii. Active region
 - iii. Inactive region (Stop)



Log-Periodic Dipole Array (LPDA)

- i. Transmission line region ($L \leq \frac{\lambda}{2}$):* At the middle of the operating range, the antenna elements are short with the resonant length, therefore the elements offer large capacitive reactance to the line. Hence, currents in these elements (1, 2, 3, 4, 5) are small and radiation is small.
- ii. Active region ($L = \frac{\lambda}{2}$):* At a wavelength near the middle of the operating range, radiation occurs primarily from the central region of the antenna. This region offers

resistive impedance. Currents in this region have large values and maximum radiation takes place from this region. The current and input RF voltage is in phase. The spacing between the elements are now sufficiently large, causing the phase in the particular element to lead approximately by 90° . For example, by the time the field radiated from the element l_{n+1} reaches l_n , the phase of the element l_n advances by 90° and the field from element l_n add to the field of l_{n+1} element, in phase producing a large resultant field towards left. Hence, there is a strong radiation towards right.

- iii. **Stop region** ($L \leq \frac{\lambda}{2}$): Elements 9,10&11 are almost one wavelength long and carry only small currents (they present a large inductive reactance to the line). Small currents in 9,10 & 11 mean that the antenna is effectively truncated at the right of the active region. Any fields (smaller magnitude) from elements 9,10&11 tend to cancel in both forward and backward directions. However, some radiation may occur broadside since the currents are approximately in phase.

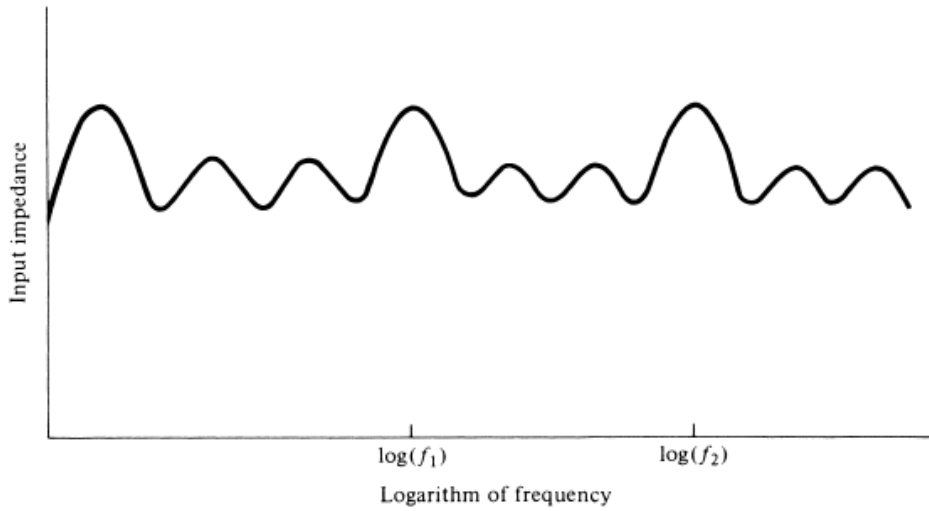
Radiation pattern: Thus, at a wavelength(λ), the radiation occurs from the middle portion where the dipole elements are $\frac{\lambda}{2}$ long. When the wavelength is increased, the radiation zone moves towards the right and when the wavelength is decreased, it moves to the left with maximum radiation toward the apex or feed point of the array.

Log-periodic behaviour:

- If the input impedance of a LPDA is plotted as a function of frequency, it will be repetitive.
- However, if the input impedance is plotted as a function of logarithm of frequency, it will be periodic with each cycle being exactly identical to the preceding one. Hence the name log-periodic, because the variations are periodic with respect to the logarithm of frequency.
- Other parameters that undergo the similar variations are the pattern, directivity, beamwidth and sidelobe level.
- The relationship between two consecutive maxima frequencies and logarithmic frequency period is

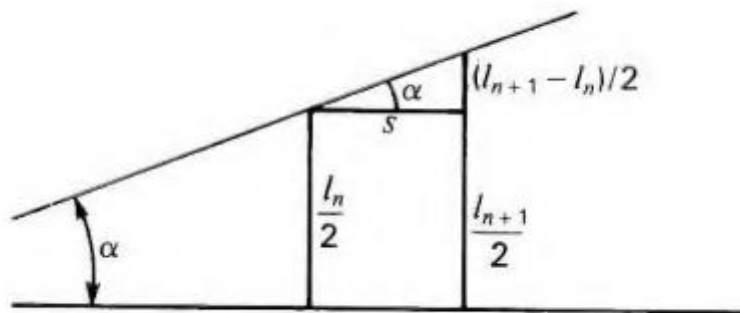
$$\log \frac{f_2}{f_1} = \log \frac{1}{\tau} \Rightarrow \frac{f_2}{f_1} = \frac{1}{\tau} \Rightarrow f_1 = \tau f_2$$

Because $\tau > 1$, $f_1 < f_2$.



Input impedance of LPDA as function of logarithm of frequency

Design equations for LPDA



LPDA geometry or determining the relation of parameters

From the above figure,

$$\tan \alpha = \frac{\left[\frac{l_{n+1} - l_n}{2} \right]}{S} = \frac{\frac{l_{n+1}}{2} \left[1 - \frac{l_n}{l_{n+1}} \right]}{S} \quad (22)$$

$$\tan \alpha = \frac{\frac{l_{n+1}}{2} \left[1 - \frac{1}{k} \right]}{S} \quad (23)$$

where α =apex angle

k = scale-factor

For $l_{n+1} = \frac{\lambda}{2}$ (when active)

$$\tan \alpha = \frac{\left[1 - \frac{1}{k}\right]}{4 \left(\frac{S}{\lambda}\right)} \quad (24)$$

$$\tan \alpha = \frac{\left[1 - \frac{1}{k}\right]}{4S_{\lambda}} \quad (25)$$

From (24),

$$\tan \alpha = \frac{\left[1 - \frac{1}{k}\right]}{4 \left(\frac{S}{\lambda}\right)} = \frac{1 - \tau}{4\sigma} \quad (26)$$

Where

$$\tau = \frac{1}{k}$$

$\sigma = \left(\frac{S}{\lambda}\right)$ = spacing factor

Hence from (26),

$$\alpha = \tan^{-1} \left[\frac{1 - \tau}{4\sigma} \right] \quad (27)$$

- Bandwidth of LPDA (frequency ratio) is $F = \frac{l_{n+1}}{l_1} = k^n = \frac{1}{\tau^n}$. The length l and spacing S for element $n + 1$ is k^n greater than for element 1.

Example:

Design a log periodic dipole array operating from 50 MHz to 200 MHz which has $\tau = 0.822$ and $\sigma = 0.149$. Find out the number of dipoles required to cover this bandwidth.

Solution:

Given: $\tau = 0.822$; $\sigma = 0.149$; $f_{min} = 50$ MHz; $f_{max} = 200$ MHz

The apex angle (α) is defined as

$$\alpha = \tan^{-1} \left[\frac{1 - \tau}{4\sigma} \right] = \tan^{-1} \left[\frac{1 - 0.822}{4 \times 0.149} \right] = 16.62^\circ$$

The lowest and highest wavelength can be obtained as

$$\lambda_L = \frac{c}{f_{max}} = \frac{3 \times 10^8}{200 \times 10^6} = 1.5 \text{ m}$$

$$\lambda_H = \frac{c}{f_{min}} = \frac{3 \times 10^8}{50 \times 10^6} = 6 \text{ m}$$

The maximum and minimum lengths required for dipole antennas are defined as follows:

$$L_{d \text{ max}} = \frac{\lambda_H}{2} = \frac{6}{2} = 3 \text{ m}$$

$$L_{d \text{ min}} = \frac{\lambda_L}{2} = \frac{1.5}{2} = 0.75 \text{ m}$$

The dipole lengths to cover 50 MHz to 200 MHz can be obtained using the following relation:

$$L_n = \tau L_{n+1}$$

$$L_{n+1} = \frac{L_n}{\tau}$$

Now, let us start from the highest frequency of operation, i.e., 200 MHz for which the minimum size of dipole is $L_{d \min} = \frac{\lambda_L}{2} = \frac{1.5}{2} = 0.75$ m.

Therefore, to start with minimum size, we have

$$L_1 = L_{d \min} = 0.75 \text{ m}$$

The other dipole lengths are

$$L_2 = \frac{L_1}{\tau} = \frac{0.75}{0.822} = 0.912 \text{ m}$$

$$L_3 = \frac{L_2}{\tau} = \frac{0.912}{0.822} = 1.109 \text{ m}$$

$$L_4 = \frac{L_3}{\tau} = \frac{1.109}{0.822} = 1.350 \text{ m}$$

$$L_5 = \frac{L_4}{\tau} = \frac{1.350}{0.822} = 1.642 \text{ m}$$

$$L_6 = \frac{L_5}{\tau} = \frac{1.642}{0.822} = 2 \text{ m}$$

$$L_7 = \frac{L_6}{\tau} = \frac{2}{0.822} = 2.433 \text{ m}$$

$$L_8 = \frac{L_7}{\tau} = \frac{2.433}{0.822} = 2.96 \text{ m}$$

$$L_9 = \frac{L_8}{\tau} = \frac{2.96}{0.822} = 3.6 \text{ m}$$

It should be noted that the highest length of the dipole should be greater than or equal to the maximum half wavelength, i.e., $L_{d \max} = \frac{\lambda_H}{2} = \frac{6}{2} = 3$ m

The spacing between n th and $(n + 1)$ th dipole can be derived from the following relationship:

$$S_n = 2\sigma L_n$$

Therefore, the spacing between all sections is given as:

$$S_1 = 2\sigma L_1 = 2 \times 0.149 \times 0.75 = 0.224$$

$$S_2 = 2\sigma L_2 = 2 \times 0.149 \times 0.912 = 0.272$$

$$S_3 = 2\sigma L_3 = 2 \times 0.149 \times 1.109 = 0.33$$

$$S_4 = 2\sigma L_4 = 2 \times 0.149 \times 1.350 = 0.402$$

$$S_5 = 2\sigma L_5 = 2 \times 0.149 \times 1.642 = 0.489$$

$$S_6 = 2\sigma L_6 = 2 \times 0.149 \times 2 = 0.596$$

$$S_7 = 2\sigma L_7 = 2 \times 0.149 \times 2.433 = 0.725$$

$$S_8 = 2\sigma L_8 = 2 \times 0.149 \times 2.96 = 0.882$$

To cover the entire frequency range, nine dipoles are required.

UNIT-III ANTENNA ARRAYS

- Antenna array is system of a similar antennas oriented similarly to get greater directivity in a desired direction.
- Antenna array is a radiating system consisting of several spaced and properly phased (current phase) radiators.

Linear Array:

- An antenna array is said to be linear if the individual antennas of the array are equally spaced along a straight line.
- Individual elements of the array are termed as **Elements**.

Uniform Linear Array:

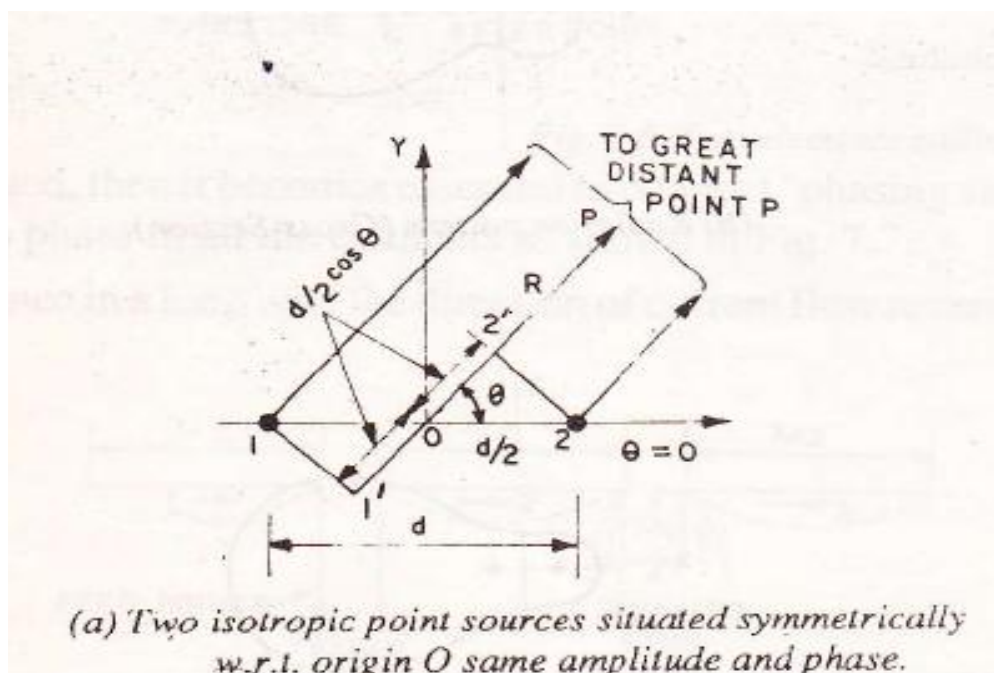
- Uniform linear array is one in which the elements are fed with a current of equal amplitude(magnitude) with uniform progressive phase shift along the line.
- Elements in a multi-element array is generally a $\frac{\lambda}{2}$ dipole antenna.

Factors that shape the radiation pattern of antenna array:

- The geometrical configuration of the array
- The spacing between the elements
- The excitation amplitude of the individual elements
- The excitation phase of the individual elements
- The radiation pattern of the individual elements

Two element array:

- Simplest array is an array of two isotropic point sources separated by a distance d .
- Two isotropic point sources symmetrically situated w.r.t the origin in the Cartesian co-ordinate system is shown in figure



- Consider that the two isotropic point sources are fed with current of equal amplitude and phase
- The fields at a greater distant point at distant R from the origin O can be calculated as follows:

- Origin is taken as reference point for phase calculation. The waves from source1 reaches the point P at a later time than the waves from source 2 because of path difference between the two waves.

Path difference between the two waves is, $d \cos \theta$

$$\Psi = \text{phase angle} = \frac{2\pi}{\lambda} \times \text{path difference} = \frac{2\pi}{\lambda} d \cos \theta$$

$$= \beta d \cos \theta \quad \text{radians}$$

Field component due to source1 (field lags) $= E_1 e^{-j\frac{\Psi}{2}}$

Field component due to source2(field leads) $= E_2 e^{+j\frac{\Psi}{2}}$

Two isotropic point sources are fed with current of equal amplitude and phase.

$$E_1 = E_2 = E_0$$

Total electric field at point P $= E = E_1 + E_2$

$$E = E_1 e^{-j\frac{\Psi}{2}} + E_2 e^{+j\frac{\Psi}{2}} = 2E_0 \cos\left(\frac{\Psi}{2}\right) = 2E_0 \cos\left(\frac{\beta d \cos \theta}{2}\right)$$

$$E_{\max} = 2E_0$$

$$E_{\text{nor}} = \frac{E}{2E_0}$$

$$E_{\text{nor}} = \cos\left(\frac{\beta d \cos \theta}{2}\right)$$

For the case $d = \frac{\lambda}{2}$,

$$E_{\text{nor}} = \cos\left(\frac{\frac{2\pi}{\lambda} \times \frac{\lambda}{2} \cos \theta}{2}\right) = \cos\left(\frac{\pi}{2} \cos \theta\right)$$

Calculation of maximum, minimum and half power direction of the field pattern:

Maxima directions

Normalized total field is maximum when $\cos\left(\frac{\pi}{2} \cos \theta\right) = \pm 1$

$$\left(\frac{\pi}{2} \cos \theta_{\max}\right) = \pm n\pi \quad \text{where } n = 0,1,2,\dots$$

$$\left(\frac{\pi}{2} \cos \theta_{\max}\right) = 0 \quad \text{when } n = 0$$

$$\theta_{\max} = 90^\circ \text{ and } 270^\circ$$

The field is maximum in the directions where $\theta = 90^\circ \text{ and } 270^\circ$

Minima directions

Normalized total field is minimum when $\cos\left(\frac{\pi}{2} \cos \theta\right) = 0$

$$\left(\frac{\pi}{2} \cos \theta_{\min}\right) = \pm(2n+1)\frac{\pi}{2} \quad \text{where } n = 0,1,2,\dots$$

$$\left(\frac{\pi}{2} \cos \theta_{\min}\right) = \pm \frac{\pi}{2} \quad \text{when } n = 0$$

$$(\cos \theta_{\min}) = \pm 1$$

$$\theta_{\min} = 0^\circ \text{ and } 180^\circ$$

The field is minimum in the directions where $\theta = 0^\circ \text{ and } 180^\circ$

Half Power Directions:

At half power points, power is half the maximum and voltage or current is $\frac{1}{\sqrt{2}}$ times the maximum.

$$\text{Normalized total field is } \cos\left(\frac{\pi}{2} \cos \theta\right) = \pm \frac{1}{\sqrt{2}}$$

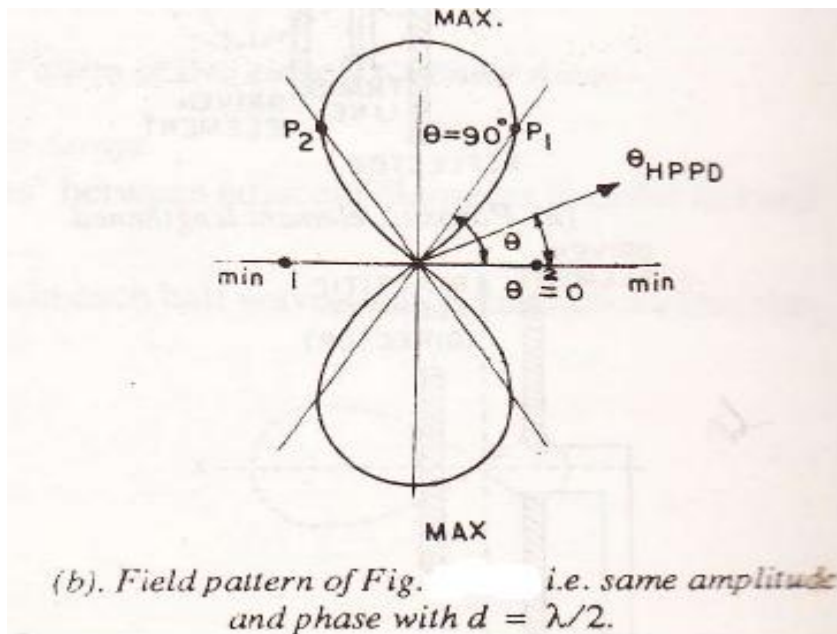
$$\left(\frac{\pi}{2} \cos \theta_{\text{HPPD}}\right) = \pm(2n+1)\frac{\pi}{4} \quad \text{where } n = 0,1,2,\dots$$

$$\left(\frac{\pi}{2} \cos \theta_{\text{HPPD}}\right) = \pm \frac{\pi}{4} \quad \text{when } n = 0$$

$$(\cos \theta_{\text{HPPD}}) = \pm \frac{1}{2}$$

$$\theta_{\text{HPPD}} = 60^\circ \text{ and } 120^\circ$$

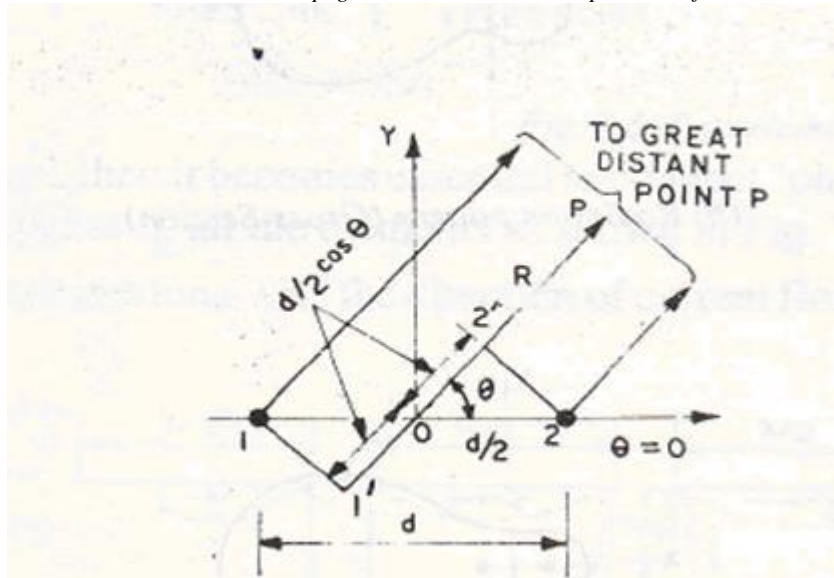
The field is $\frac{1}{\sqrt{2}}$ times the maximum in the directions where $\theta = 60^\circ \text{ and } 120^\circ$



- The radiation pattern is perpendicular to the array axis. This array is referred to as **Broadside array**.

Note: Broadside array is defined as an array in which the principal direction is perpendicular to the array axis

Array of two point source with equal amplitude and opposite phase:



- Consider that the two isotropic point sources are fed with current of equal amplitude and opposite phase
- The fields at a greater distant point at distant R from the origin O can be calculated as follows:
- Origin is taken as reference point for phase calculation. The waves from source 1 reaches the point P at a later time than the waves from source 2 because of path difference between the two waves.

Path difference between the two waves is, $d \cos \theta$

$$\Psi = \text{phase angle} = \frac{2\pi}{\lambda} \times \text{path difference} = \frac{2\pi}{\lambda} d \cos \theta$$

$$= \beta d \cos \theta \quad \text{radians}$$

Field component due to source 1 (field lags) = $-E_1 e^{-j\frac{\Psi}{2}}$

Field component due to source 2 (field leads) = $E_2 e^{+j\frac{\Psi}{2}}$

Two isotropic point sources are fed with current of equal amplitude

$$E_1 = E_2 = E_0$$

Total electric field at point P

$$= E = -E_1 e^{-j\frac{\Psi}{2}} + E_2 e^{+j\frac{\Psi}{2}} = 2 j E_0 \sin\left(\frac{\Psi}{2}\right) = 2 j E_0 \sin\left(\frac{\beta d \cos \theta}{2}\right)$$

$$E_{\max} = 2 j E_0$$

$$E_{\text{nor}} = \frac{E}{2 j E_0}$$

$$E_{\text{nor}} = \sin\left(\frac{\beta d \cos \theta}{2}\right)$$

For the case $d = \frac{\lambda}{2}$,

$$E_{\text{nor}} = \sin\left(\frac{\frac{2\pi}{\lambda} \times \frac{\lambda}{2} \cos \theta}{2}\right) = \sin\left(\frac{\pi}{2} \cos \theta\right)$$

Calculation of maximum, minimum and half power direction of the field pattern:

Maxima directions

Normalized total field is maximum when $\sin\left(\frac{\pi}{2}\cos\theta\right) = \pm 1$

$$\left(\frac{\pi}{2}\cos\theta_{\max}\right) = \pm \frac{(2n+1)\pi}{2} \quad \text{where } n = 0,1,2,\dots$$

$$\left(\frac{\pi}{2}\cos\theta_{\max}\right) = \pm \frac{\pi}{2} \quad \text{when } n = 0$$

$$\theta_{\max} = 0^\circ \quad \text{and } 180^\circ$$

The field is maximum in the directions where $\theta = 0^\circ$ and 180°

Minima directions

Normalized total field is minimum when $\sin\left(\frac{\pi}{2}\cos\theta\right) = 0$

$$\left(\frac{\pi}{2}\cos\theta_{\min}\right) = \pm n\pi \quad \text{where } n = 0,1,2,\dots$$

$$\left(\frac{\pi}{2}\cos\theta_{\min}\right) = 0 \quad \text{when } n = 0$$

$$(\cos\theta_{\min}) = 0$$

$$\theta_{\min} = 90^\circ \quad \text{and } 270^\circ$$

The field is minimum in the directions where $\theta = 90^\circ$ and 270°

Half Power Directions:

At half power points, power is half the maximum and voltage or current is $\frac{1}{\sqrt{2}}$ times the maximum.

Normalized total field is $\sin\left(\frac{\pi}{2}\cos\theta\right) = \pm \frac{1}{\sqrt{2}}$

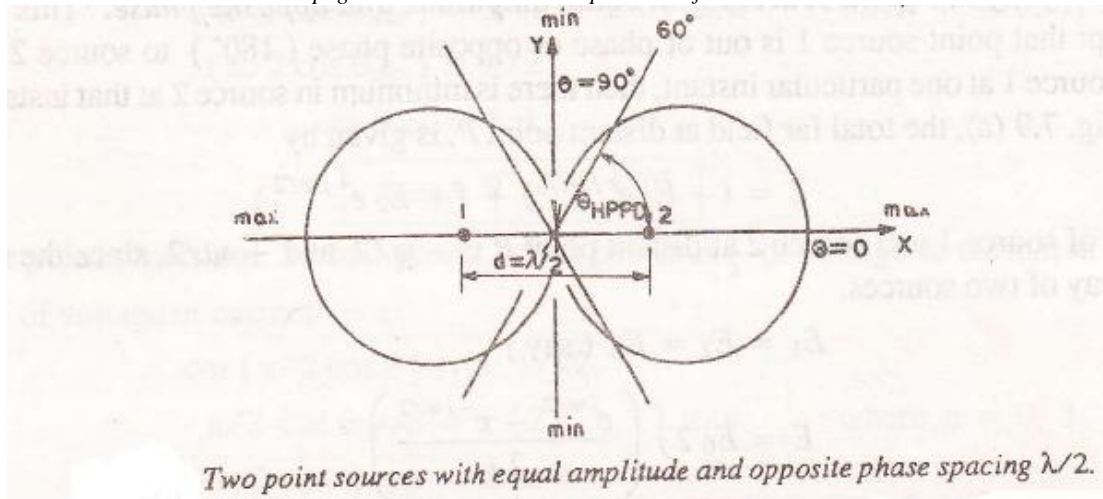
$$\left(\frac{\pi}{2}\cos\theta_{HPPD}\right) = \pm(2n+1)\frac{\pi}{4} \quad \text{where } n = 0,1,2,\dots$$

$$\left(\frac{\pi}{2}\cos\theta_{HPPD}\right) = \pm \frac{\pi}{4} \quad \text{when } n = 0$$

$$(\cos\theta_{HPPD}) = \pm \frac{1}{2}$$

$$\theta_{HPPD} = 60^\circ \quad \text{and } 120^\circ$$

The field is $\frac{1}{\sqrt{2}}$ times the maximum in the directions where $\theta = 60^\circ$ and 120°



- Maximum radiation is along the axis of the array. This array is referred to as **END FIRE** array.

Arrays of point sources with unequal amplitude and any phase

- Let α be the phase difference between the currents.
- The total phase difference between radiations of two sources at a distant point P is given by

$$\Psi = \frac{2\pi}{\lambda} d \cos \theta + \alpha$$

$$\frac{2\pi}{\lambda} d \cos \theta = \text{phase difference due to path difference}$$

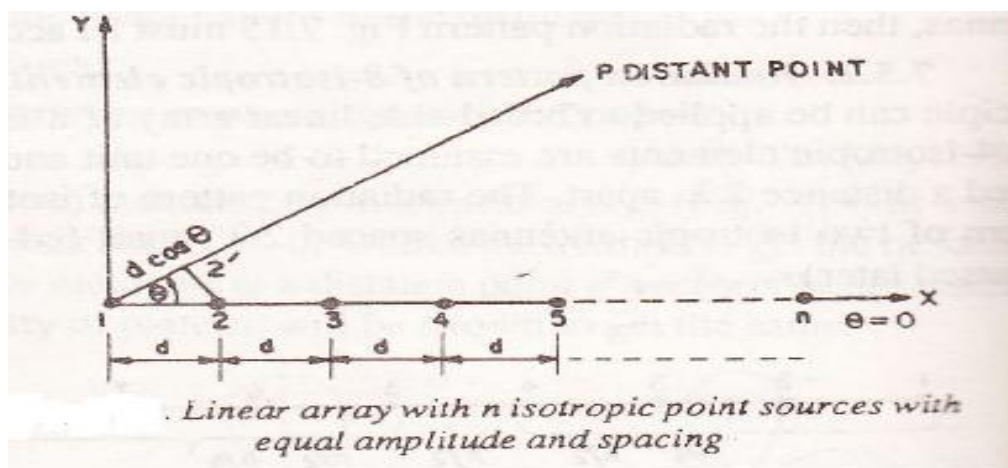
$$E = E_1 e^{j0} + E_2 e^{j\Psi} = E_1 (1 + K e^{j\Psi})$$

$$K = \frac{E_2}{E_1}$$

$$|E| = E_1 \sqrt{(1 + K \cos \Psi)^2 + (K \sin \Psi)^2}$$

$$\text{phase angle} = \tan^{-1} \left(\frac{K \sin \Psi}{(1 + K \cos \Psi)} \right)$$

Linear Array with n isotropic point sources of equal amplitude and spacing



- The point sources are fed with currents of equal amplitude and having an uniform progressive phase shift along the line
- Field at a distant point P is given by,

$$E_t = E_0 e^{j0} + E_0 e^{j\Psi} + E_0 e^{j2\Psi} + E_0 e^{j3\Psi} + \dots + E_0 e^{j(n-1)\Psi}$$

$$E_t = E_0 \{ e^{j0} + e^{j\Psi} + e^{j2\Psi} + e^{j3\Psi} + \dots + e^{j(n-1)\Psi} \} \text{ ----- 1}$$

$$\Psi = \frac{2\pi}{\lambda} d \cos \theta + \alpha$$

$$\frac{2\pi}{\lambda} d \cos \theta = \text{phase difference due to path difference}$$

between two point sources

$\alpha =$ phase shift between two point sources

$$E_t = E_0 e^{j0} + E_0 e^{j\Psi} + E_0 e^{j2\Psi} + E_0 e^{j3\Psi} + \dots + E_0 e^{j(n-1)\Psi}$$

$$E_t = E_0 \{ e^{j0} + e^{j\Psi} + e^{j2\Psi} + e^{j3\Psi} + \dots + e^{j(n-1)\Psi} \} \text{ ----- 1}$$

$$E_t e^{j\Psi} = E_0 \{ e^{j\Psi} + e^{j2\Psi} + e^{j3\Psi} + \dots + e^{jn\Psi} \} \text{ ----- 2}$$

1- 2 gives

$$E_t (1 - e^{j\Psi}) = E_0 \{ 1 - e^{jn\Psi} \}$$

$$E_t = E_0 \frac{(1 - e^{jn\Psi})}{(1 - e^{j\Psi})}$$

$$E_t = E_0 \frac{e^{j\frac{n\Psi}{2}} \left(e^{j\frac{n\Psi}{2}} - e^{-j\frac{n\Psi}{2}} \right)}{e^{j\frac{\Psi}{2}} \left(e^{j\frac{\Psi}{2}} - e^{-j\frac{\Psi}{2}} \right)}$$

$$E_t = E_0 e^{j\frac{(n-1)\Psi}{2}} \frac{\sin\left(\frac{n\Psi}{2}\right)}{\sin\left(\frac{\Psi}{2}\right)}$$

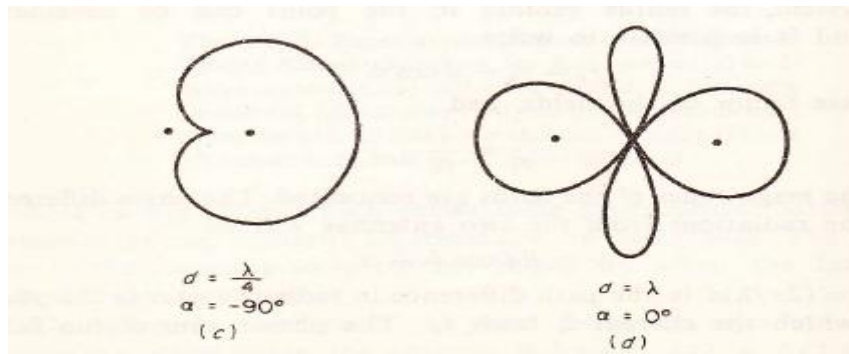
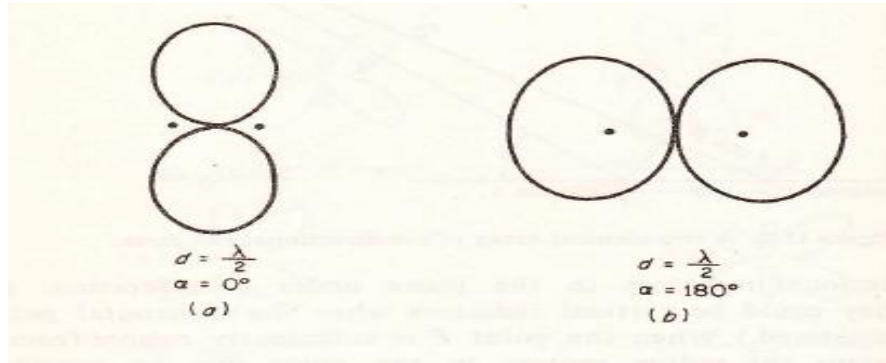
$$E_t = E_0 e^{j\phi} \frac{\sin\left(\frac{n\Psi}{2}\right)}{\sin\left(\frac{\Psi}{2}\right)} \quad \text{where } \phi = \frac{(n-1)\Psi}{2}$$

In this derivation, point source 1 is taken as reference point. In case the reference point is shifted to the center of the array, $\phi = \frac{(n-1)\Psi}{2}$ is automatically eliminated .

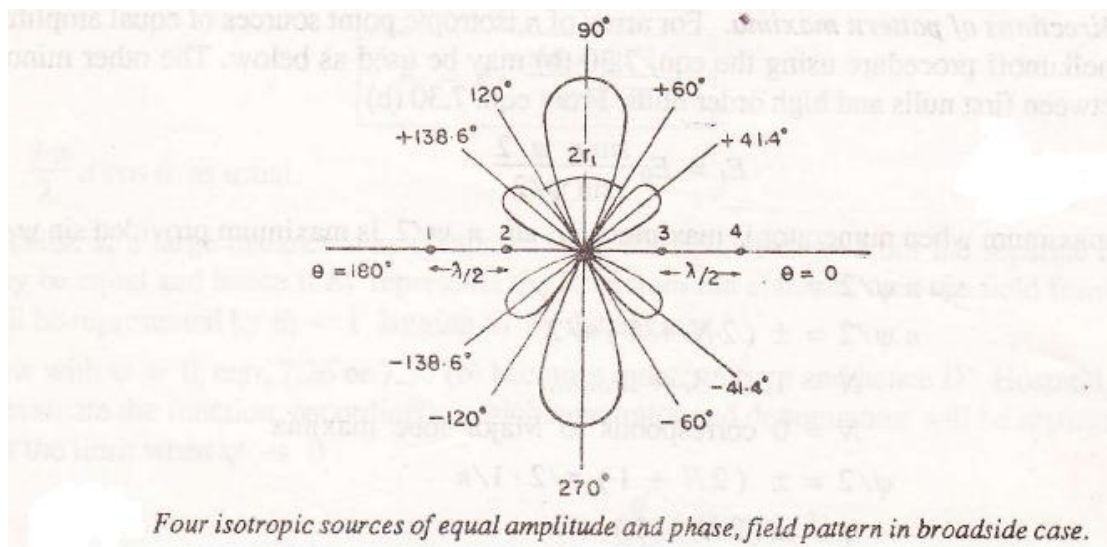
$$E_t = E_0 \frac{\sin\left(\frac{n\Psi}{2}\right)}{\sin\left(\frac{\Psi}{2}\right)}$$

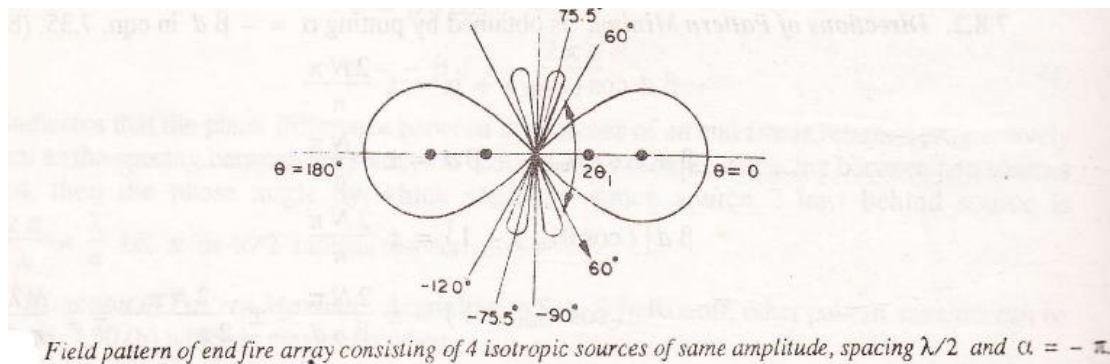
$$E_{t_{\max}} = \lim_{\Psi \rightarrow 0} E_0 \frac{\sin\left(\frac{n\Psi}{2}\right)}{\sin\left(\frac{\Psi}{2}\right)} = \lim_{\Psi \rightarrow 0} E_0 \frac{\frac{n}{2} \cos\left(\frac{n\Psi}{2}\right)}{\frac{1}{2} \cos\left(\frac{\Psi}{2}\right)} = nE_0$$

Radiation pattern of two nondirectional radiators fed with equal currents at the phasings shown.



Radiation pattern of four nondirectional radiators





Principle of pattern multiplication

The field pattern of an array of non-isotropic but similar sources is the product of the pattern of the individual source and the pattern of an array of isotropic point sources having the same locations, relative amplitudes, and phase as the non-isotropic sources.

The total field pattern of an array of non-isotropic but similar sources is the product of individual source pattern and the pattern of an array of isotropic point sources each located at the phase center of the individual source and having the same relative amplitude and phase, while the total phase pattern is the sum of the phase patterns of the individual source and the array of isotropic point sources.

$$E = f(\theta, \phi)F(\theta, \phi) \angle (f_p(\theta, \phi) + F_p(\theta, \phi))$$

$$f(\theta, \phi)F(\theta, \phi) = \text{field pattern of array}$$

$$(f_p(\theta, \phi) + F_p(\theta, \phi)) = \text{phase pattern of array}$$

$$f(\theta, \phi) = \text{field pattern of individual source}$$

$$f_p(\theta, \phi) = \text{phase pattern of individual source}$$

$$F(\theta, \phi) = \text{field pattern of array of isotropic sources}$$

$$F_p(\theta, \phi) = \text{phase pattern of array of isotropic sources}$$

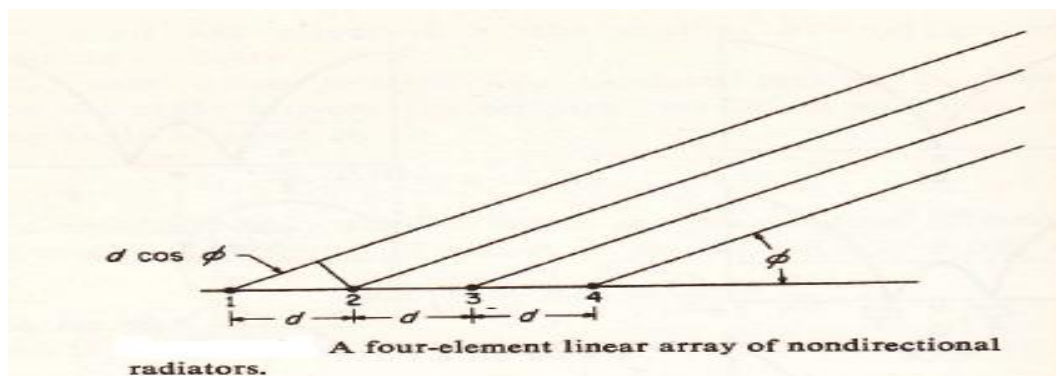
Multiplication of Patterns

- Simple method of obtaining radiation pattern.
- Makes it possible to sketch rapidly, almost by inspection, the patterns of complicated arrays.
- Useful tool in the design of arrays

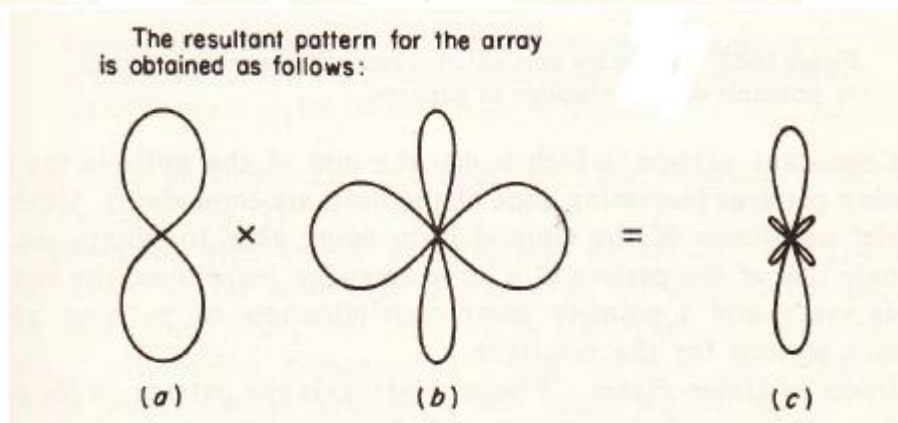
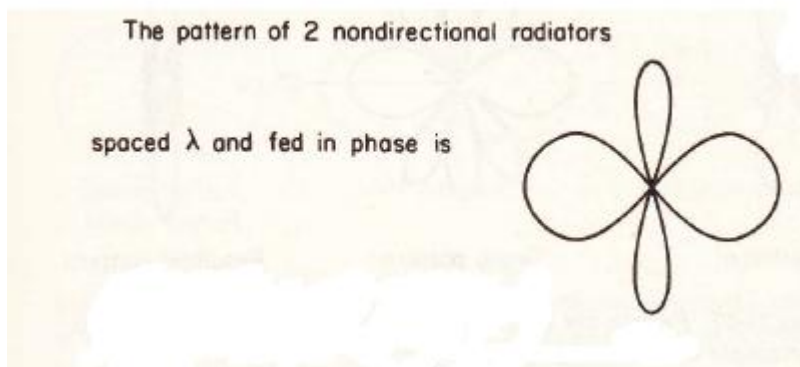
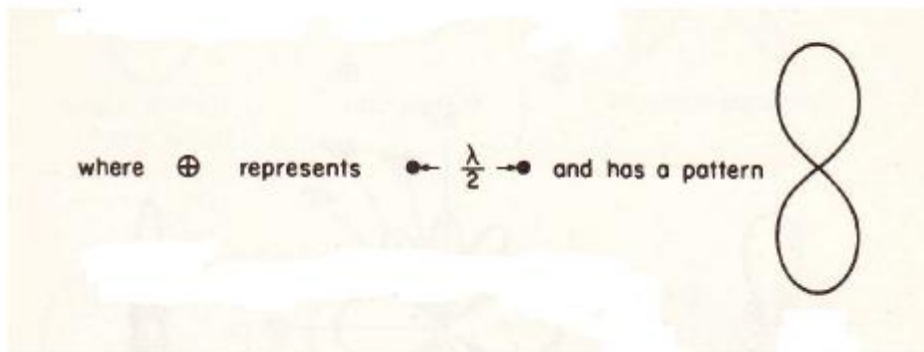
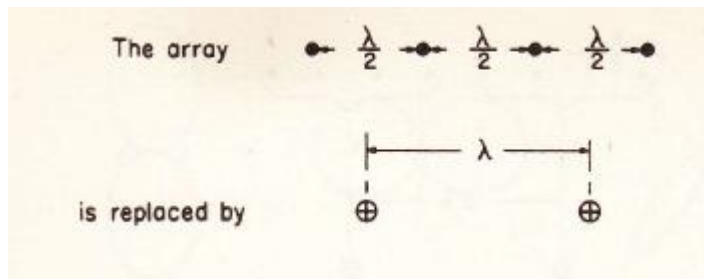
Example:

Determination of radiation pattern of a four element array in which the spacing between units is

$$\frac{\lambda}{2} \text{ and the currents are in phase } (\alpha = 0)$$



Determination of radiation pattern by multiplication of patterns is illustrated below:

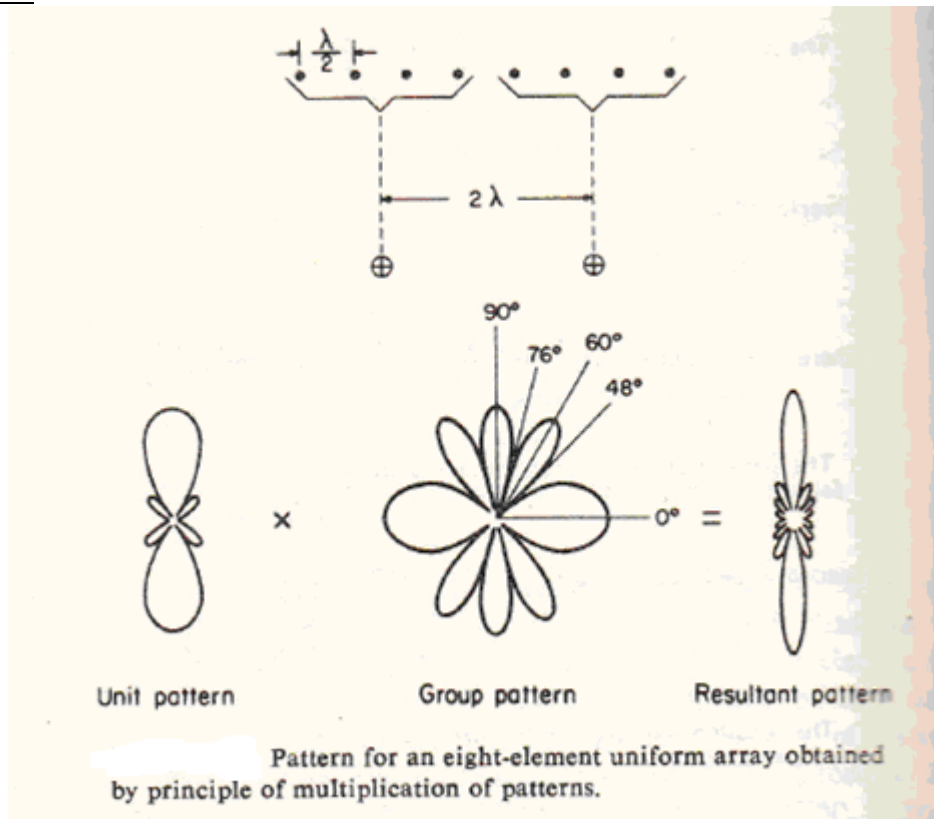


- (a) -- unit pattern
- (b) -- Group pattern
- (c) – resultant pattern

- This procedure provides a means for rapidly determining the radiation pattern of a complicated array without making length calculations.
- The width of the principle lobe (between nulls) is the same as the width of the corresponding lobe of the group pattern.

- The number of secondary lobes can be determined from the number of nulls in the resultant pattern, which is just the sum of the nulls in the unit and group pattern (assuming none of the nulls are coincident).
- Point by point multiplication of patterns yields the exact pattern .

Example 2



Linear Array with uniform spacing, non uniform amplitude:

- (i) Binomial array
- (ii) Dolph-Tschebysheff array (also referred to as Chebyshev array or Tschbyscheff array)

Binomial array:

- Current distribution follows the binomial series.
- The current amplitudes are proportional to the coefficients of the successive terms of the Binomial series.
- Binomial array possess the smallest side lobes. If the spacing is $\frac{\lambda}{2}$ or less than $\frac{\lambda}{2}$, Binomial array has no side lobes.
- Binomial series :

$$(1 + x)^{m-1} = 1 + (m-1)x + \frac{(m-1)(m-2)}{2!}x^2 + \frac{(m-1)(m-2)(m-3)}{3!}x^3 + \dots$$

- For m = 3

$$(1 + x)^{3-1} = 1 + (3-1)x + \frac{(3-1)(3-2)}{2!}x^2 + \frac{(3-1)(3-2)(3-3)}{3!}x^3 + \dots$$

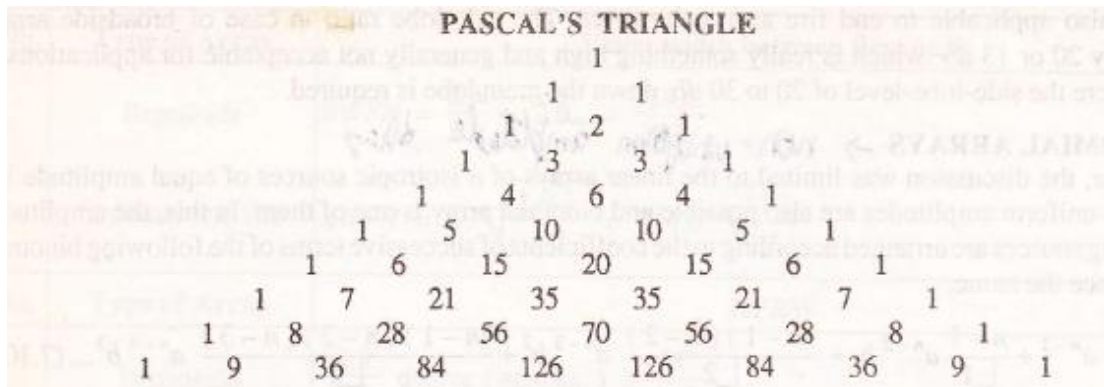
$$= 1 + 2x + x^2$$

Current distribution for 3 element binomial array is 1:2:1

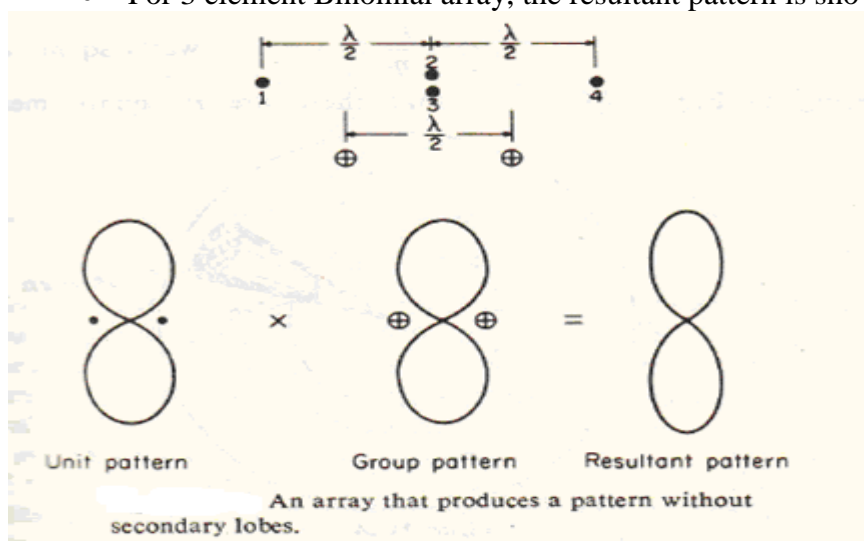
- Current distribution for 4 element binomial array is 1:3:3:1

So on....

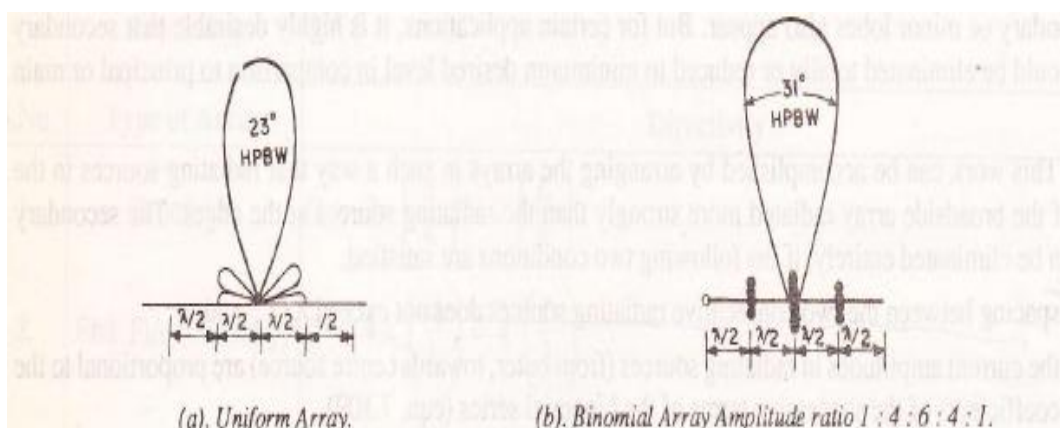
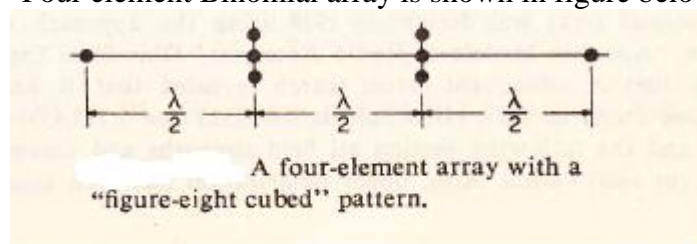
- Current distribution can be determined easily from Pascal' s Triangle



- Principle of pattern multiplication can be used to determine the resultant radiation pattern.
- For 3 element Binomial array, the resultant pattern is shown below:



- Here antenna 2 and 3 coincide, and so they would be replaced by with a single antenna carrying double the current.
- Four element Binomial array is shown in figure below:



Note:

- No side lobes in the radiation pattern of Binomial array
- Half Power Beam width is more

Disadvantages of Binomial array:

- HPBW increases and hence the directivity decreases
- For design of large array, larger amplitude ratio of sources required.

Adaptive array (smart antennas)

- **Smart antennas** (also known as adaptive array antennas, digital antenna arrays) are antenna arrays with smart signal processing algorithms used to identify spatial signal signatures such as the direction of arrival (DOA) of the signal, and use them to calculate beamforming vectors which are used to track and locate the antenna beam on the mobile/target.
- Smart antenna techniques are used notably in acoustic signal processing, track and scan radar, radio astronomy and radio telescopes, and mostly in cellular systems like W-CDMA, UMTS, and LTE.
- Smart antennas have many functions: DOA (Direction of Arrival) estimation, beamforming, interference nulling, and constant modulus preservation.

Beam forming

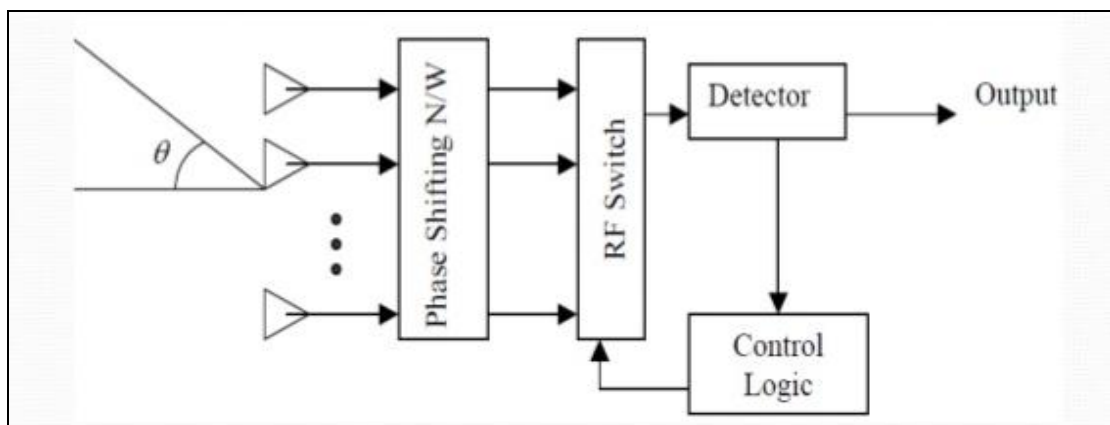
- Beamforming is the method used to create the radiation pattern of the antenna array by adding constructively the phases of the signals in the direction of the targets/mobiles desired, and nulling the pattern of the targets/mobiles that are undesired/interfering targets.
- This can be done with a simple Finite Impulse Response (FIR) tapped delay line filter. The weights of the FIR filter may also be changed adaptively, and used to provide optimal beamforming, in the sense that it reduces the Minimum Mean Square Error between the desired and actual beam pattern formed.
- Typical algorithms are the steepest descent, and Least Mean Squares algorithms. In digital antenna arrays with multi channels use the digital beamforming, usually by DFT or FFT.

Types of Smart antennas

Two main types of smart antennas:

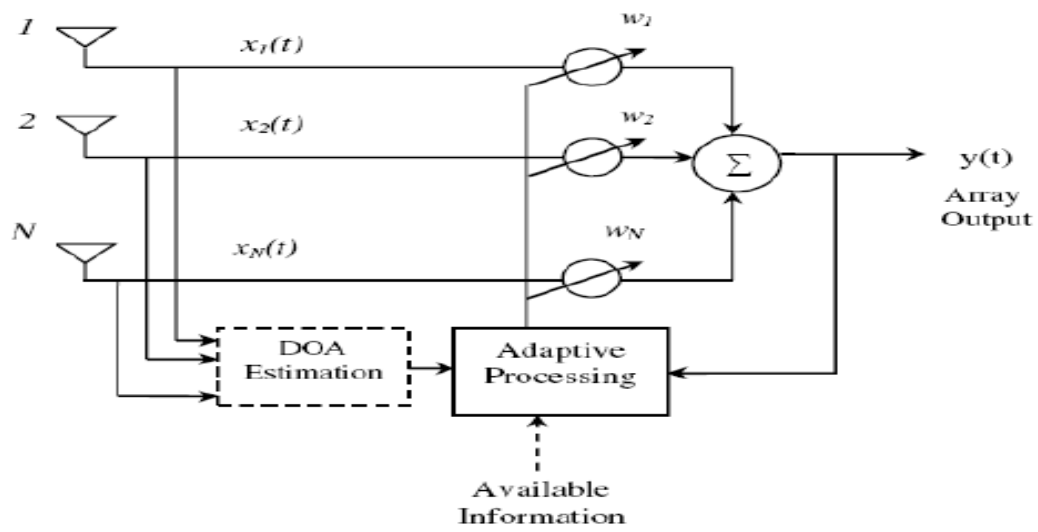
- Switched beam smart antennas
- Adaptive array smart antennas

- Switched beam systems have several available fixed beam patterns. A decision is made as to which beam to access, at any given point in time, based upon the requirements of the system.



Switched beam smart antennas

- Adaptive arrays allow the antenna to steer the beam to any direction of interest while simultaneously nulling interfering signals.



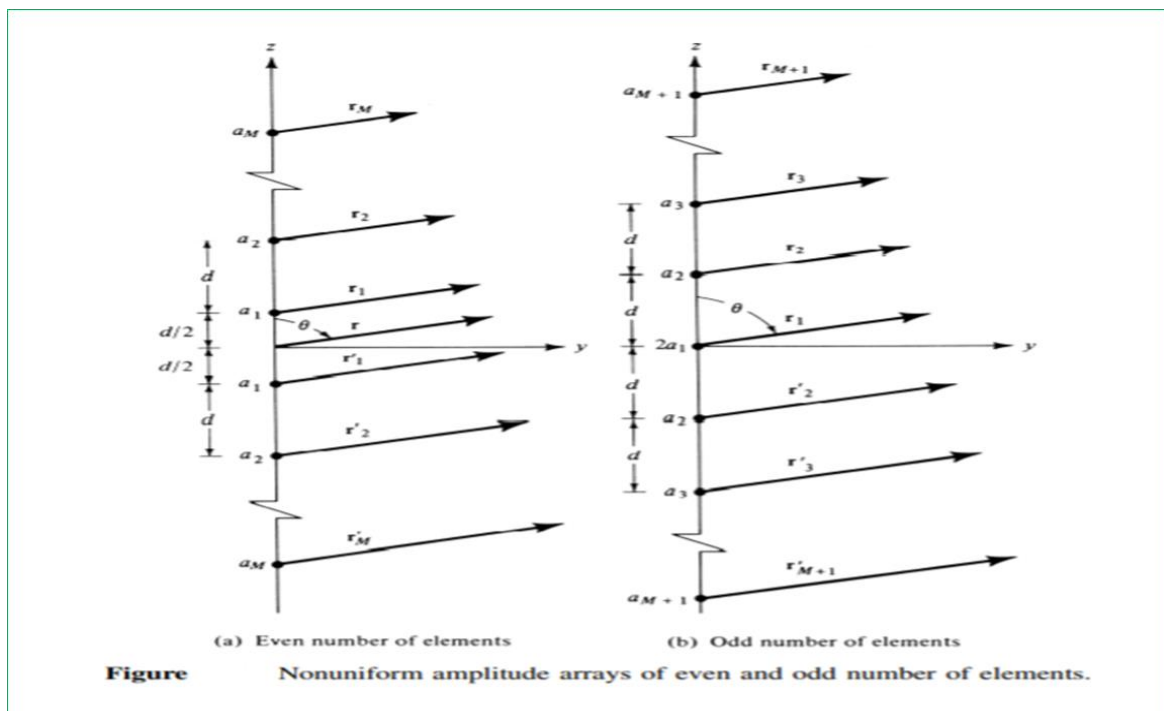
Adaptive array smart antenna

N-ELEMENT LINEAR ARRAY: UNIFORM SPACING, NONUNIFORM AMPLITUDE

- Of the three distributions (uniform, binomial, and Tschebyscheff), a uniform amplitude array yields the smallest half-power beamwidth. It is followed, in order, by the Dolph-Tschebyscheff and binomial arrays.
- In contrast, binomial arrays usually possess the smallest side lobes followed, in order, by the Dolph-Tschebyscheff and uniform arrays.
- Binomial arrays with element spacing equal or less than $\lambda/2$ have no side lobes.
- It is apparent that the designer must compromise between side lobe level and beamwidth.
- A criterion that can be used to judge the relative beamwidth and side lobe level of one design to another is the amplitude distribution (tapering) along the source.
- It has been shown analytically that for a given side lobe level the Dolph-Tschebyscheff array produces the smallest beamwidth between the first nulls. Conversely, for a given beamwidth between the first nulls, the Dolph-Tschebyscheff design leads to the smallest possible side lobe level.

Array Factor of non uniform array

- An array of an even number of isotropic elements $2M$ (where M is an integer) is positioned symmetrically along the z -axis, as shown in Figure (a). The separation between the elements is d , and M elements are placed on each side of the origin.
- If the total number of isotropic elements of the array is odd $2M + 1$ (where M is an integer), the arrangement is as shown in Figure (b)



An array of an even number of isotropic elements $2M$ (where M is an integer) is positioned symmetrically along the z -axis, as shown in Figure (a). The separation between the

elements is d , and M elements are placed on each side of the origin. Assuming that the amplitude excitation is symmetrical about the origin, the array factor for a nonuniform amplitude broadside array can be written as,

$$(AF)_{2M} = a_1 e^{+j\left(\frac{1}{2}\right)kdcos\theta} + a_2 e^{+j\left(\frac{3}{2}\right)kdcos\theta} + \dots + a_M e^{+j\left(\frac{(2M-1)}{2}\right)kdcos\theta} \\ + a_1 e^{-j\left(\frac{1}{2}\right)kdcos\theta} + a_2 e^{-j\left(\frac{3}{2}\right)kdcos\theta} + \dots + a_M e^{-j\left(\frac{(2M-1)}{2}\right)kdcos\theta}$$

$$(AF)_{2M} (even) = \sum_{n=1}^M a_n \cos [(2n - 1)u]$$

Where $u = \frac{\pi d}{\lambda} \cos\theta$

a_n is excitation coefficient

If the total number of isotropic elements of the array is odd $2M + 1$ (where M is an integer), as shown in Figure (b), the array factor can be written as

$$(AF)_{2M+1} = 2a_1 + a_2 e^{+jkdcos\theta} + \dots + a_{M+1} e^{+j(M)kdcos\theta} \\ + a_2 e^{-jkdcos\theta} + \dots + a_{M+1} e^{-j(M)kdcos\theta}$$

$$(AF)_{2M+1} (odd) = \sum_{n=1}^{M+1} a_n \cos [2(n - 1)u]$$

Where $u = \frac{\pi d}{\lambda} \cos\theta$

a_n is excitation coefficient

I. Dolph-Tschebyscheff Array: Broadside

- Array, with many practical applications, is the Dolph-Tschebyscheff array.
- The method was originally introduced by Dolph. It is primarily a compromise between uniform and binomial arrays.
- Its excitation coefficients are related to Tschebyscheff polynomials.
- A Dolph-Tschebyscheff array with no side lobes (or side lobes of $-\infty$ dB) reduces to the binomial design. The excitation coefficients for this case, as obtained by both methods, would be identical.

Array Factor(for odd and even number of elements)

$$(AF)_{2M} (even) = \sum_{n=1}^M a_n \cos [(2n - 1)u]$$

$$(AF)_{2M+1} (odd) = \sum_{n=1}^{M+1} a_n \cos [2(n-1)u]$$

Where $u = \frac{\pi d}{\lambda} \cos\theta$

a_n is excitation coefficient

- The array factor of an array of even or odd number of elements with symmetric amplitude excitation is nothing more than a summation of M or M + 1 cosine terms.
- The largest harmonic of the cosine terms is one less than the total number of elements of the array. Each cosine term, whose argument is an integer times a fundamental frequency, can be rewritten as a series of cosine functions with the fundamental frequency as the argument.

That is,

$$m=0 \quad \cos(mu) = 1$$

$$m=1 \quad \cos(mu) = \cos u$$

$$m=2 \quad \cos(mu) = \cos(2u) = 2 \cos^2 u - 1$$

$$m=3 \quad \cos(mu) = \cos(3u) = 4 \cos^3 u - 3 \cos u$$

$$m=4 \quad \cos(mu) = \cos(4u) = 8 \cos^4 u - 8 \cos^2 u + 1$$

$$m=5 \quad \cos(mu) = \cos(5u) = 16 \cos^5 u - 20 \cos^3 u + 5 \cos u$$

$$m=6 \quad \cos(mu) = \cos(6u) = 32 \cos^6 u - 48 \cos^4 u + 18 \cos^2 u - 1$$

$$m=7 \quad \cos(mu) = \cos(7u) = 64 \cos^7 u - 112 \cos^5 u + 56 \cos^3 u - 7 \cos u$$

$$m=8 \quad \cos(mu) = \cos(8u) = 128 \cos^8 u - 256 \cos^6 u + 160 \cos^4 u - 32 \cos^2 u + 1$$

$$m=9 \quad \cos(mu) = \cos(9u) = 256 \cos^9 u - 576 \cos^7 u + 432 \cos^5 u - 120 \cos^3 u + 9 \cos u$$

The above equations are obtained using Euler's formula:

$$[e^{ju}]^m = (\cos u + j \sin u)^m = e^{jmu} = \cos(mu) + j \sin(mu)$$

$$\text{And } \sin^2 u = 1 - \cos^2 u$$

If we let $z = \cos u$

The above equations can be written as

$$m=0 \quad \cos(mu) = 1 = T_0(z)$$

$$m=1 \quad \cos(mu) = \cos u = z = T_1(z)$$

$$m=2 \quad \cos(mu) = \cos(2u) = 2z^2 - 1 = T_2(z)$$

$$m=3 \quad \cos(mu) = \cos(3u) = 4z^3 - 3z = T_3(z)$$

$$m=4 \quad \cos(mu) = \cos(4u) = 8z^4 - 8z^2 + 1 = T_4(z)$$

$$m=5 \quad \cos(\mu u) = \cos(5u) = 16z^5 - 20z^3 + 5z = T_5(z)$$

$$m=6 \quad \cos(\mu u) = \cos(6u) = 32z^6 - 48z^4 + 18z^2 - 1 = T_6(z)$$

$$m=7 \quad \cos(\mu u) = \cos(7u) = 64z^7 - 112z^5 + 56z^3 - 7z = T_7(z)$$

$$m=8 \quad \cos(\mu u) = \cos(8u) = 128z^8 - 256z^6 + 160z^4 - 32z^2 + 1 = T_8(z)$$

$$m=9 \quad \cos(\mu u) = \cos(9u) = 256z^9 - 576z^7 + 432z^5 - 120z^3 + 9z = T_9(z)$$

and each is related to a Tschebyscheff (Chebyshev) polynomial $T_m(z)$. These relations between the cosine functions and the Tschebyscheff polynomials are valid only in the $-1 \leq z \leq +1$ range.

- Because $|\cos(\mu u)| \leq 1$, each Tschebyscheff polynomial is $|T_m(z)| \leq 1$ for $-1 \leq z \leq +1$. For $|z| > 1$, the Tschebyscheff polynomials are related to the hyperbolic cosine functions.
- The recursion formula for Tschebyscheff polynomials is

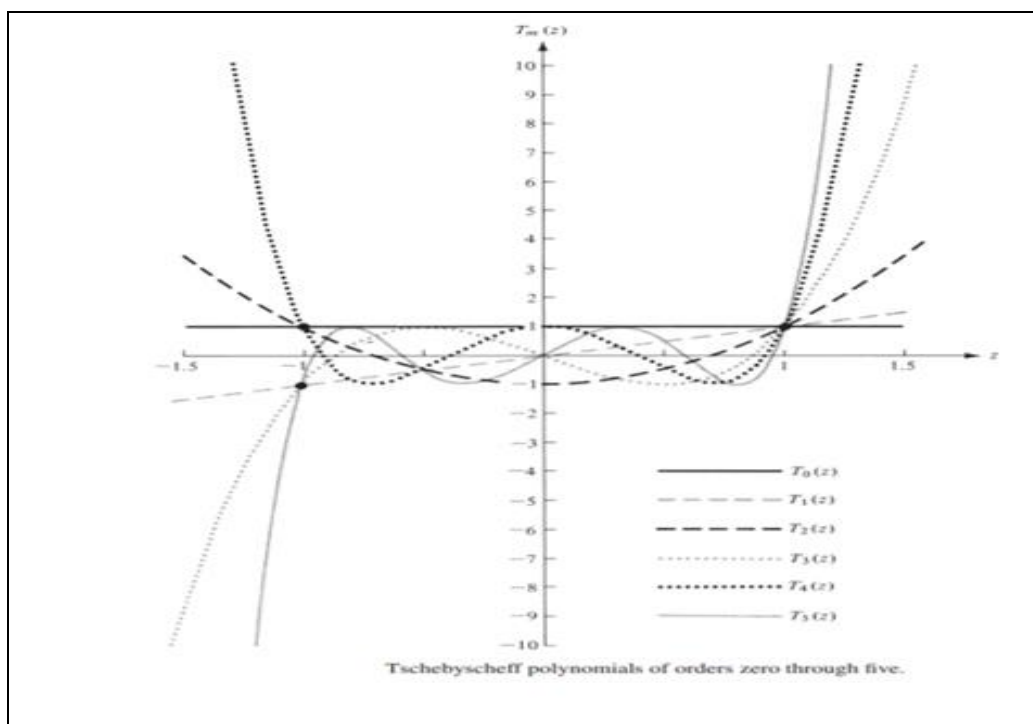
$$T_m(z) = 2zT_{m-1}(z) - T_{m-2}(z)$$

- Each polynomial can also be computed using

$$T_m(z) = \cos[m \cos^{-1}(z)] \quad -1 \leq z \leq +1$$

$$T_m(z) = \cosh[m \cosh^{-1}(z)] \quad z < -1, z > +1$$

In Figure below, the first six Tschebyscheff polynomials have been plotted.



The following properties of the polynomials are of interest:

1. All polynomials, of any order, pass through the point (1, 1).
2. Within the range $-1 \leq z \leq 1$, the polynomials have values within -1 to $+1$.
3. All roots occur within $-1 \leq z \leq 1$, and all maxima and minima have values of $+1$ and -1 , respectively.

Since the array factor of an even or odd number of elements is a summation of cosine terms whose form is the same as the Tschebyscheff polynomials, the unknown coefficients of the array factor can be determined by equating the series representing the cosine terms of the array factor to the appropriate Tschebyscheff polynomial. The order of the polynomial should be one less than the total number of elements of the array.

Outline of the design procedure

Assumptions:

Number of elements, spacing between the elements, and ratio of major-to-minor lobe intensity (R_0) are known. The requirements will be to determine the excitation coefficients and the array factor of a Dolph-Tschebyscheff array.

Problem Statement

Design a broadside Dolph-Tschebyscheff array of $2M$ or $2M + 1$ elements with spacing d between the elements. The side lobes are R_0 dB below the maximum of the major lobe. Find the excitation coefficients and form the array factor.

Procedure

- a. Select the appropriate array factor (for odd or even number of array elements).
- b. Expand the array factor. Replace each $\cos(mu)$ function ($m = 0, 1, 2, 3, \dots$) by its appropriate series expansion.
- c. Determine the point $z = z_0$ such that $T_m(z_0) = R_0$ (voltage ratio). The order m of the Tschebyscheff polynomial is always one less than the total number of elements. The design procedure requires that the Tschebyscheff polynomial in the $-1 \leq z \leq z_1$, where z_1 is the null nearest to $z = +1$, be used to represent the minor lobes of the array. The major lobe of the pattern is formed from the remaining part of the polynomial up to point $z_0(z_1 < z \leq z_0)$.
- d. Substitute $\cos(u) = z/z_0$ in the array factor of step b. The $\cos(u)$ is replaced by z/z_0 , and not by z , so that the equation, $\cos(u) = z/z_0$ would be valid for $|z| \leq |z_0|$. At $|z| = |z_0|$, the equation, $\cos(u) = z/z_0$ attains its maximum value of unity.
- e. Equate the array factor from step b, after substitution of $\cos(u) = z/z_0$, to a $T_m(z)$. The $T_m(z)$ chosen should be of order m where m is an integer equal to one less than the total number of elements of the designed array. This will allow the determination of the excitation coefficients a_n 's.
- f. Write the array factor using the coefficients found in step e.

Example problem

Design a broadside Dolph-Tschebyscheff array of 10 elements with spacing d between the elements and with a major-to-minor lobe ratio of 26 dB. Find the excitation coefficients and form the array factor.

Solution:

1. The array factor is given by

$$(AF)_{2M} (even) = \sum_{n=1}^M a_n \cos [(2n - 1)u]$$

Where, $u = \frac{\pi d}{\lambda} \cos\theta$ and a_n is excitation coefficient

2. When expanded, the array factor can be written as

$$(AF)_{10} = a_1 \cos(u) + a_2 \cos(3u) + a_3 \cos(5u) + a_4 \cos(7u) + a_5 \cos(9u)$$

Replace $\cos(u)$, $\cos(3u)$, $\cos(5u)$, $\cos(7u)$, and $\cos(9u)$ by their series expansions

$$\cos u = z$$

$$\cos(3u) = 4z^3 - 3z$$

$$\cos(5u) = 16z^5 - 20z^3 + 5z$$

$$\cos(7u) = 64z^7 - 112z^5 + 56z^3 - 7z$$

$$\cos(9u) = 256z^9 - 576z^7 + 432z^5 - 120z^3 + 9z$$

$$(AF)_{10} = a_1 [z] + a_2 [4z^3 - 3z] + a_3 [16z^5 - 20z^3 + 5z] + a_4 [64z^7 - 112z^5 + 56z^3 - 7z] + a_5 [256z^9 - 576z^7 + 432z^5 - 120z^3 + 9z]$$

$$(AF)_{10} = z[(a_1 - 3a_2 + 5a_3 - 7a_4 + 9a_5)] + z^3[(4a_2 - 20a_3 + 56a_4 - 120a_5)] + z^5[(16a_3 - 112a_4 + 432a_5)] + z^7[(64a_4 - 576a_5)] + z^9[(256a_5)]$$

3. R_0 (dB) = 26 = $20 \log_{10}(R_0)$ or R_0 (voltage ratio) = 20.

Determine z_0 by equating R_0 to $T_9(z_0)$.

$$\text{Thus } R_0 = 20 = T_9(z_0) = \cosh[9 \cosh^{-1}(z_0)]$$

$$\text{or } z_0 = \cosh[(1/9) \cosh^{-1}(20)] = 1.0851$$

Note: Another equation which can, in general, be used to find z_0 and does not require hyperbolic functions is

$$z_0 = \frac{1}{2} \left[\left(R_0 + \sqrt{R_0^2 - 1} \right)^{\frac{1}{P}} + \left(R_0 - \sqrt{R_0^2 - 1} \right)^{\frac{1}{P}} \right]$$

where P is an integer equal to one less than the number of array elements (in this case P=9)

4. Substitute $\cos(u) = z/z_0 = z/1.0851$ in the array factor found in step 2.

5. Equate the array factor of step 2, after the substitution from step 4, to $T_9(z)$.

$$(AF)_{10} = z[(a_1 - 3a_2 + 5a_3 - 7a_4 + 9a_5)/z_0] + z^3[(4a_2 - 20a_3 + 56a_4 - 120a_5)/z_0^3] + z^5[(16a_3 - 112a_4 + 432a_5)/z_0^5] + z^7[(64a_4 - 576a_5)/z_0^7] + z^9[(256a_5)/z_0^9] = 256z^9 - 576z^7 + 432z^5 - 120z^3 + 9z = T_9(z)$$

Matching similar terms allows the determination of the a_n 's.

That is,

$$(256a_5)/z_0^9 = 256 \rightarrow a_5 = 2.0856$$

$$(64a_4 - 576a_5)/z_0^7 = -576 \rightarrow a_4 = 2.8308$$

$$(16a_3 - 112a_4 + 432a_5)/z_0^5 = 432 \rightarrow a_3 = 4.1184$$

$$(4a_2 - 20a_3 + 56a_4 - 120a_5)/z_0^3 = -120 \rightarrow a_2 = 5.2073$$

$$(a_1 - 3a_2 + 5a_3 - 7a_4 + 9a_5)/z_0 = 9 \rightarrow a_1 = 5.8377$$

In normalized form, the an coefficients can be written as

$a_5 = 1$, $a_4 = 1.357$, $a_3 = 1.974$, $a_2 = 2.496$, $a_1 = 2.798$ normalized with respect to the amplitude of the elements at the edge.

(The values can also be normalized with respect to the amplitude of the center element)

6. Using the set of normalized coefficients, the array factor can be written as

$$(AF)_{10} = 2.798 \cos(u) + 2.496 \cos(3u) + 1.974 \cos(5u) + 1.357 \cos(7u) + \cos(9u)$$

where $u = [(\pi d \lambda) \cos \theta]$.

Binomial array:

- Current distribution follows the binomial series.
- The current amplitudes are proportional to the coefficients of the successive terms of the Binomial series.
- Binomial array possess the smallest side lobes. If the spacing is $\frac{\lambda}{2}$ or less than $\frac{\lambda}{2}$, Binomial array has no side lobes.
- Binomial series :

$$(1+x)^{m-1} = 1 + (m-1)x + \frac{(m-1)(m-2)}{2!}x^2 + \frac{(m-1)(m-2)(m-3)}{3!}x^3 + \dots$$

- For m=3

$$(1+x)^{3-1} = 1 + (3-1)x + \frac{(3-1)(3-2)}{2!}x^2 + \frac{(3-1)(3-2)(3-3)}{3!}x^3 + \dots$$

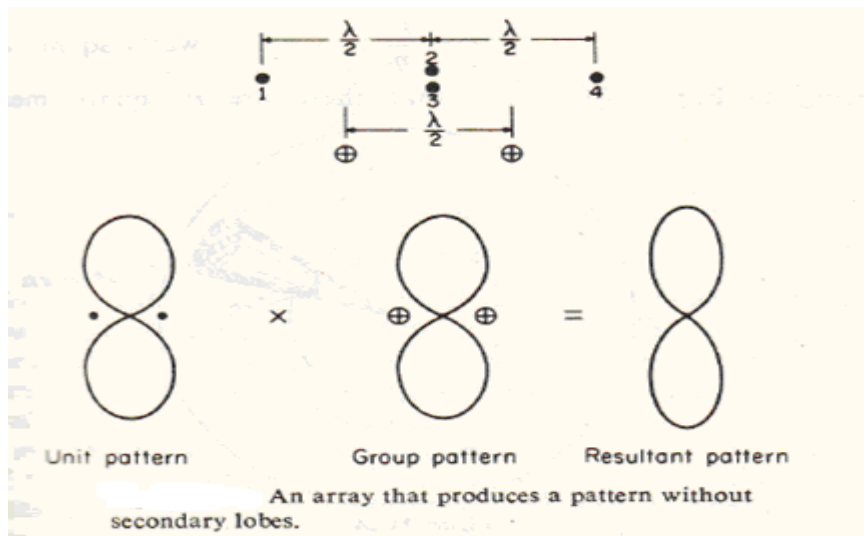
$$= 1 + 2x + x^2$$

Current distribution for 3 element binomial array is 1:2:1

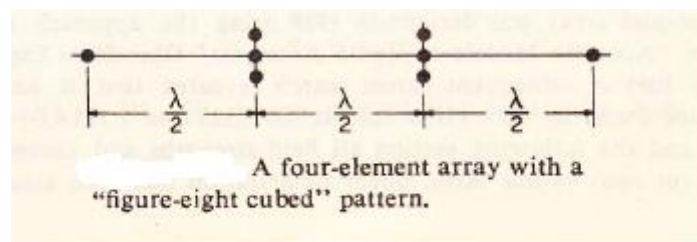
- Current distribution for 4 element binomial array is 1:3:3:1
So on.....
- Current distribution can be determined easily from Pascal's Triangle

PASCAL'S TRIANGLE										
					1					
				1	1					
			1	2	1					
		1	3	3	1					
		1	4	6	4	1				
		1	5	10	10	5	1			
	1	6	15	20	15	6	1			
	1	7	21	35	35	21	7	1		
1	1	8	28	56	70	56	28	8	1	
1	9	36	84	126	126	84	36	9	1	

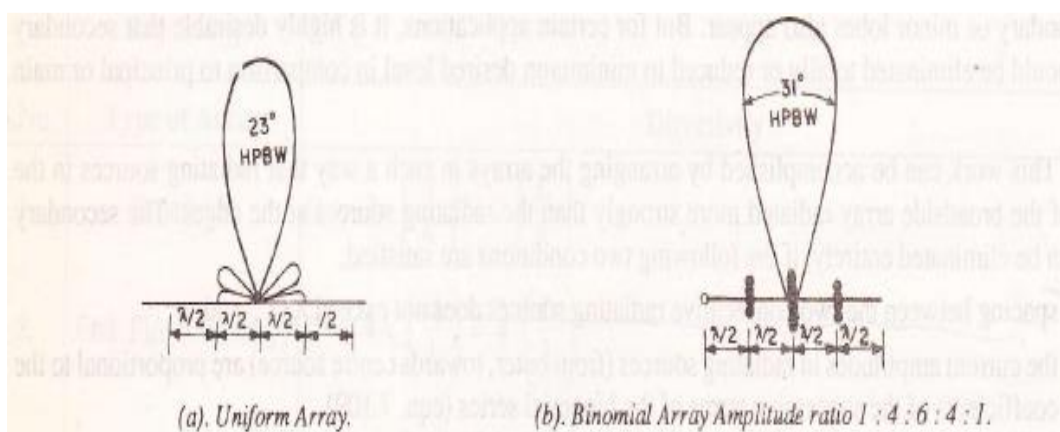
- Principle of pattern multiplication can be used to determine the resultant radiation pattern.
- For 3 element Binomial array, the resultant pattern is shown below:



- Here antenna 2 and 3 coincide, and so they would be replaced by with a single antenna carrying double the current.
 - Four element Binomial array is shown in figure below:



Comparison of Uniform array and Binomial array



Note:

- No side lobes in the radiation pattern of Binomial array
- Half Power Beam width is more

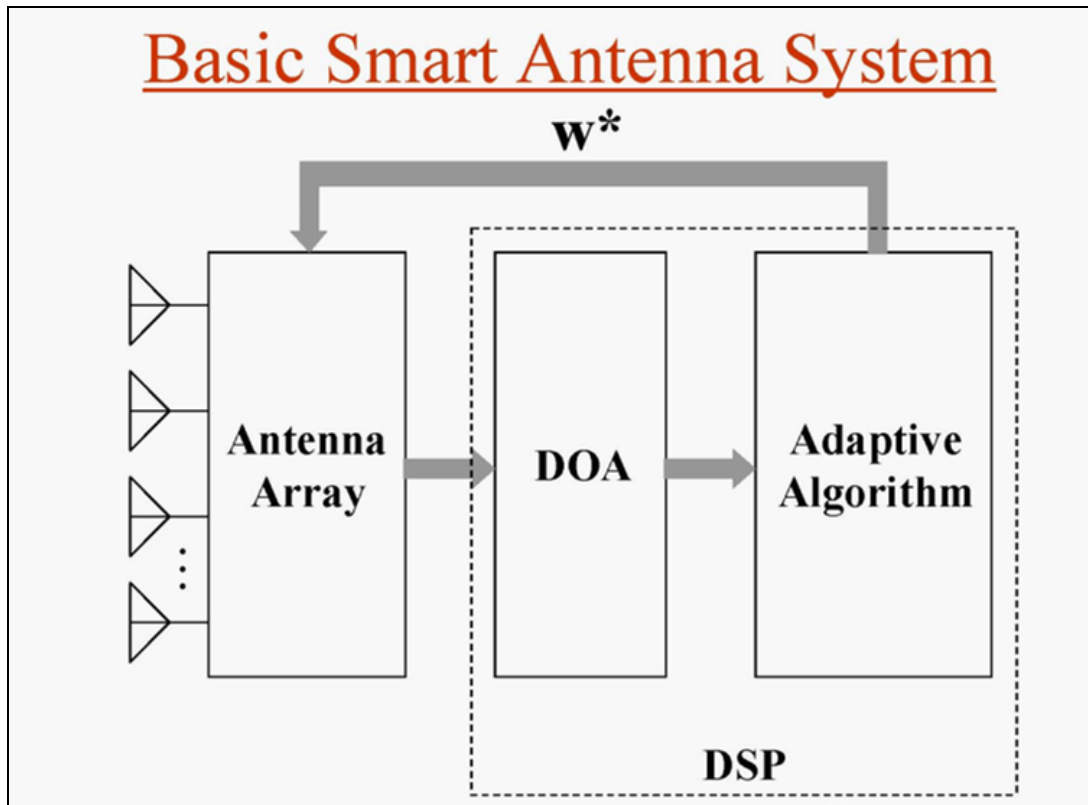
Disadvantages of Binomial array:

- HPBW increases and hence the directivity decreases
- For design of large array, larger amplitude ratio of sources required.

Smart antennas

Smart antenna systems combine:

- (i) Antenna arrays with
- (ii) Digital signal processing algorithms to make the antenna systems smart.



- Smart antennas integrate antenna array technology and Digital Signal Processing(DSP) Techniques to enhance communication system performance including
 - (a) Capacity improvement
 - (b) Range increase
 - (c) Link quality improvement
 - (d) Mitigation of fading
- These are accomplished by
 - (1) Beam steering
Placing Beam maxima toward Signals of Interest (SOI)
 - (2) Null Steering
Placing Beam minima, ideally nulls, toward Interfering signals; Signals Not of Interest (SNOI)
 - (3) Spatially separate signals
Allowing different users to share the same spectral resources(SDMA)

Beam forming in Smart antenna systems

By means of an internal feed back control, smart antenna can generate a **customized radiation pattern** to each remote user. In general, they form a main lobe toward a desired signal and rejects interference outside the main lobe.

There are two types of smart antenna system

- (i) Switched – beam systems
- (ii) Adaptive antenna systems

(a)Switched Beam Systems

- They use a number of fixed beams at the base station. The base station selects one of the pre- determined fixed beam that provides the greatest output power for the desired user.

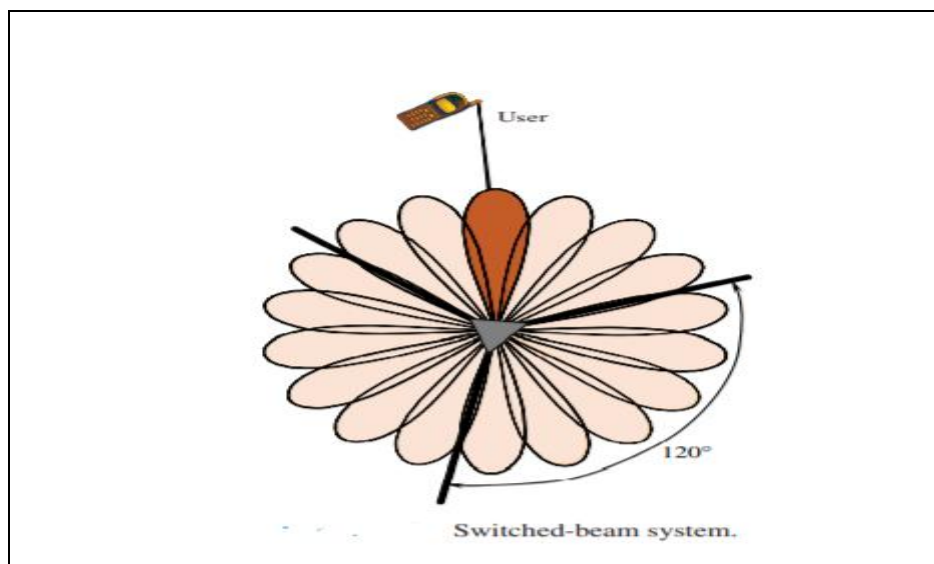


Figure: Concept of Switched Beam systems

- This concept is obviously an extension of cell sectoring as each sector is subdivided into smaller sectors. As the mobile unit moves throughout the cell, the switched-beam system detects the signal strength, chooses the appropriate predefined beam pattern.

Advantages over Adaptive antenna systems:

Low cost

Less complex and easier to retrofit to existing wireless technologies

Dis advantages over Adaptive antenna systems:

Beam Resolution is lower

(b) Adaptive array systems

Adaptive array systems, provide more degrees of freedom since they have the ability to adapt in real time the radiation pattern to the RF signal environment. In other words, they can direct the main beam toward the pilot signal or SOI while suppressing the antenna pattern in

the direction of the interferers or SNOIs. To put it simply, adaptive array systems can customize an appropriate radiation pattern for each individual user.

Adaptive array systems can locate and track signals (users and interferers) and dynamically adjust the antenna pattern to enhance reception while minimizing interference using **signal-processing algorithms**. A functional block diagram of such a system is shown in Figure. This figure shows that after the system down converts the received signals to baseband and digitizes them, it locates the SOI using the direction-of-arrival (DOA) algorithm, and it continuously tracks the SOI and SNOIs by dynamically changing the weights (amplitudes and phases of the signals).

Basically, the DOA computes the direction of arrival of all signals by computing the time delays between the antenna elements, and afterward the adaptive algorithm, using a cost function, computes the appropriate weights that result in an optimum radiation pattern.

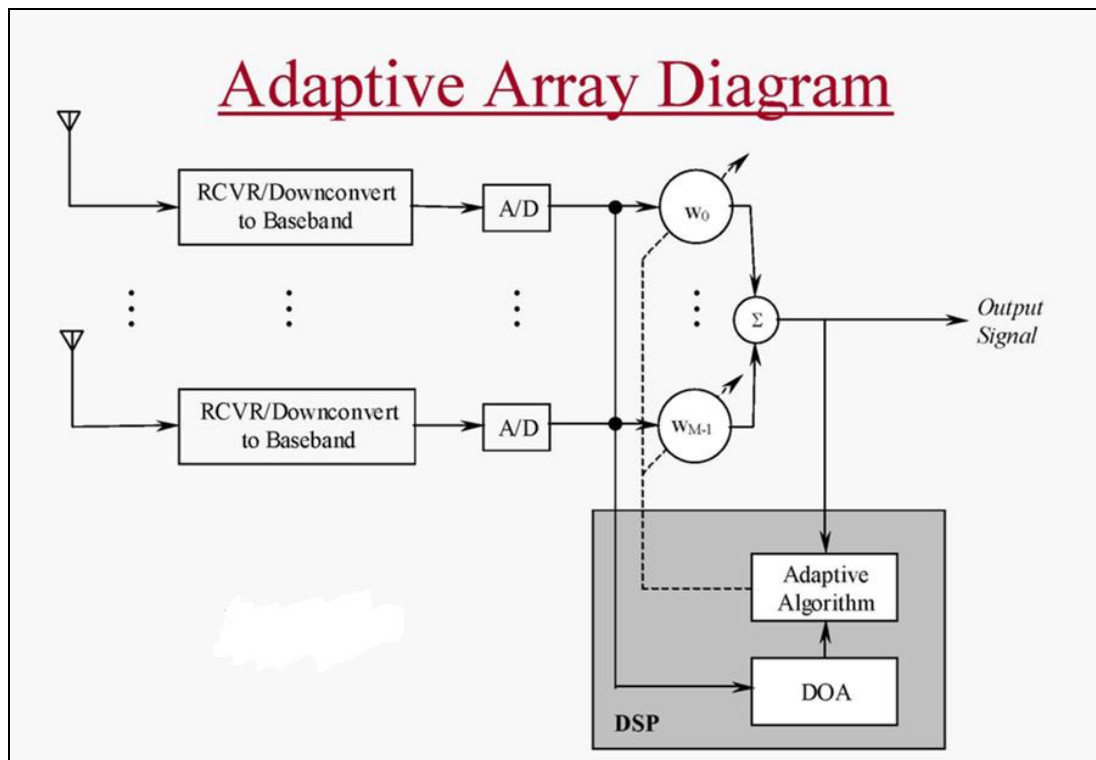


Figure Functional block diagram of an adaptive array system

Optimal Beam forming techniques

In optimal beamforming techniques, a weight vector that minimizes a cost function is determined. Typically, this cost function, related with a performance measure, is inversely associated with the quality of the signal at the array output, so that when the cost function is minimized, the quality of the signal is maximized at the array output.

The most commonly used optimally beamforming techniques or performance measures are the Minimum Mean Square Error (MMSE), Maximum Signal-to-Noise Ratio (MSNR), and Minimum (noise) Variance (MV).

Adaptive algorithm

In practice, the signal environment is dynamic or time varying, and therefore, the weights need to be computed with adaptive methods. One of the simplest algorithms that is commonly used to adapt the weights is the Least Mean Square (LMS) algorithm. The LMS algorithm is a low complexity algorithm that requires no direct matrix inversion and no memory.

Comparison of switched-beam scheme and adaptive array scheme

(i)Minimizing Interference

Adaptive array systems can customize an appropriate radiation pattern for each individual user. This is far superior to the performance of a switched-beam system, as shown in Figure. This figure shows that not only the switched-beam system may not able to place the desired signal at the maximum of the main lobe but also it exhibits the inability to fully reject the interferers.

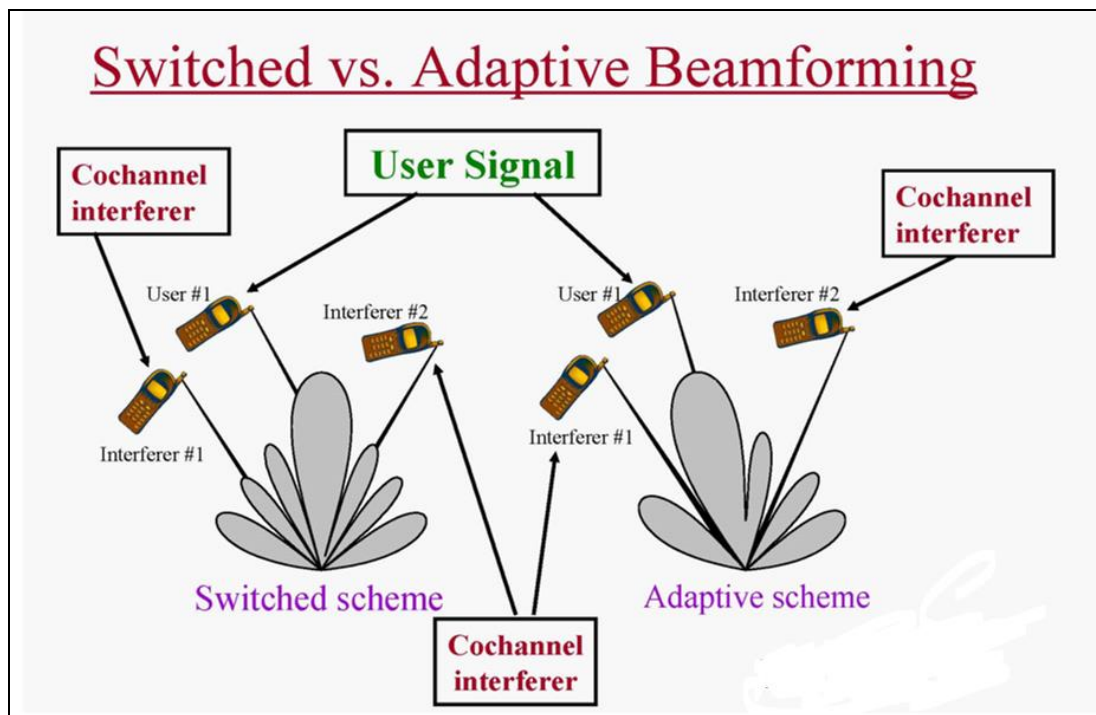


Figure Comparison of (a) switched-beam scheme, and (b) adaptive array scheme.

(ii)Coverage area comparison

Figure shows a comparison, in terms of relative coverage area, of conventional sectorized, switched-beam and adaptive arrays. In the presence of a low-level interference, both types of smart antennas provide significant gains over the conventional sectorized systems. However, when a high-level interference is present, the interference rejection capability of the

adaptive systems provides significantly more coverage than either the conventional or switched-beam system.

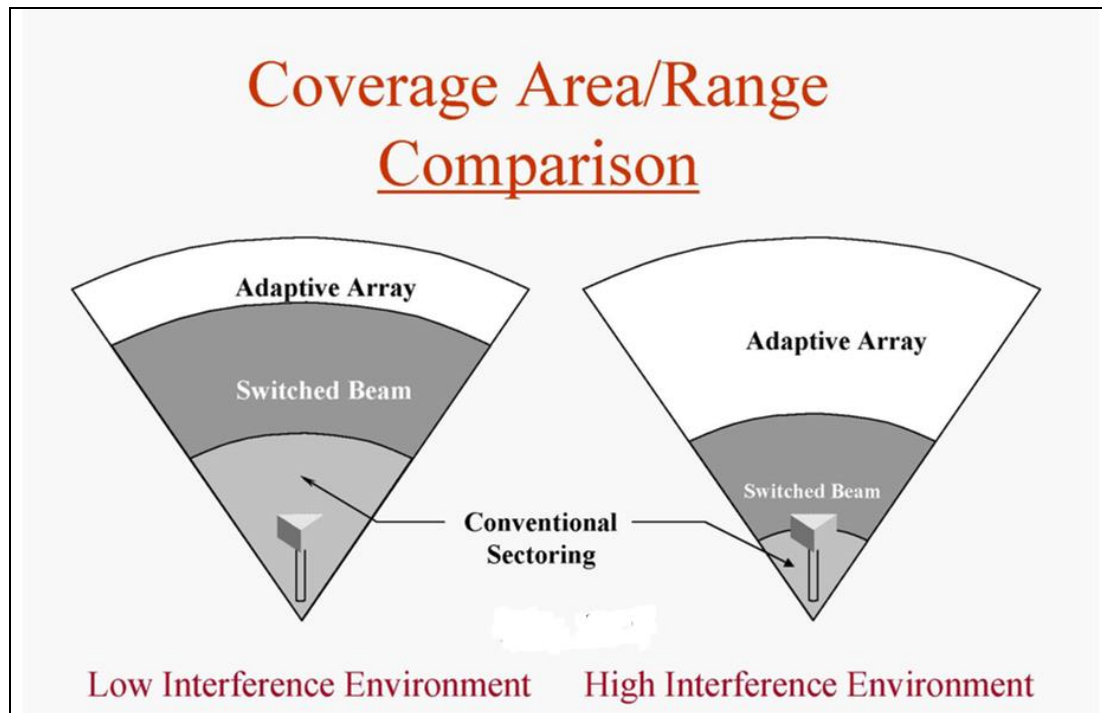


Figure: Relative coverage area comparison among sectorized systems, switched-beam systems, and adaptive array systems in (a) low interference environment, and (b) high interference environment

SMART ANTENNAS' BENEFITS

(i) capacity increase

In densely populated areas, mobile systems are usually interference-limited, meaning that the interference from other users is the main source of noise in the system. This means that the signal-to interference ratio (SIR) is much smaller than the signal-to-noise ratio (SNR).

In general, smart antennas will, by simultaneously increasing the useful received signal level and lowering the interference level, increase the SIR. (there by increase capacity)

(ii) range increase

Another benefit that smart-antenna systems provide is range increase. Because smart antennas are more directional than omnidirectional and sectorized antennas, a range increase potential is available.

In other words, smart antennas are able to focus their energy toward the intended users, instead of directing it in other unnecessary directions (wasting) like omnidirectional antennas do. This means that base stations can be placed further apart, leading to a more cost-efficient development.

Therefore, in rural and sparsely populated areas, where radio coverage rather than capacity is more important, smart-antenna systems are well suited.

(iii) Provide security

Another added advantage of smart-antenna systems is security. Smart antennas make it more difficult to tap a connection, because the intruder must be positioned in the same direction as the user as seen from the base station to successfully tap a connection.

(iv) Helps in location finding

Because of the spatial detection nature of smart-antenna systems, they can be used to locate humans in emergencies or for any other location-specific service.

SMART ANTENNAS' DRAWBACKS

While smart antennas provide many benefits, they do suffer from certain drawbacks.

- Their transceivers are much more complex than traditional base station transceivers. The antenna needs separate transceiver chains for each array antenna element and accurate real-time calibration for each of them.
- The antenna beamforming is computationally intensive, which means that smart-antenna base stations must be equipped with very powerful digital signal processors. This tends to increase the system costs in the short term, but since the benefits outweigh the costs, it will be less expensive in the long run.
- For a smart antenna to have pattern-adaptive capabilities and reasonable gain, an array of antenna elements is necessary.

UNIT IV Microwave Generation

Review of Conventional tubes:

The conventional vacuum tubes are diode, triode, tetrode and Pentode. The vacuum tube diode contains a filament, cathode and anode. The triode contains 3 elements (ie) a cathode, control grid and an anode. The tetrode contains a cathode, two grids and an anode. The Pentode contains 5 elements (ie) a cathode, 3 grids and an anode.

→ The vacuum tube diode is used as a rectifier. However, triode, tetrode and Pentode are used as amplifiers and oscillators.

→ But they are useful at low microwave frequencies as they have some limitations at microwave frequencies.

Limitations of conventional tubes:

→ The conventional tubes - triode, tetrode and Pentode - are useful as signal sources for frequencies less than 1.0 GHz. But, they cannot be used at microwave frequencies as they exhibit a few undesirable effects and limitations.

The main limitations of the conventional tubes are due to the presence of

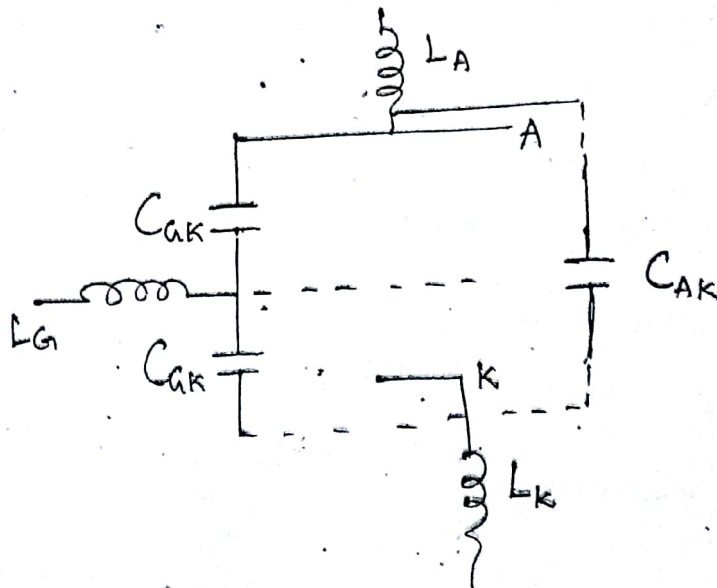
- (1) lead inductance effect
- (2) interelectrode capacitance effect
- (3) transit time effect and
- (4) gain-bandwidth product limitation.

The Lead inductance :

The lead inductance contains 3 components

- Grid lead inductance, L_g
- Cathode lead inductance, L_c
- Plate lead inductance, L_p

These are shown in figure (1) -



Effect of lead inductance : The lead inductance becomes large at microwave frequencies, and it affects the input impedance of the device. This in turn reduces the gain of the tube amplifier.

→ The lead inductance can be minimized by
→ reducing the lead length and
→ the electrode area.

→ The above methods, however, reduce Power handling capacity.

The lead capacitance: The lead capacitance contains 3 components in a triode.

They are

- (1) Plate-grid capacitance
- (2) grid-cathode capacitance and
- (3) Plate-cathode capacitance.

→ Effect of lead capacitance
- At high frequencies, the reactance of each lead capacitance decreases and hence output voltage decreases due to shunt effect.

→ The lead capacitance can be reduced by decreasing the electrode area, increasing the distance between electrodes.

Transit-Time: The transit time is the time taken for an electron to travel from cathode to an anode. It is given by,

$$t_t = \frac{d}{v_b}$$

Here, t_t = transit time

d = distance between anode and cathode

v_e = velocity of the electrons.

→ The transit time effect reduces the efficiency of the tube.

At pulse frequencies, the transit time is large compared to the period of pulse signal. The signal b/w the cathode and grid changes several times during the transit of electron.

→ The positive half-cycle of the grid potential supplies the energy to the electrons. This energy is removed during the negative cycle.

→ This results in the oscillation of electrons in the cathode-grid region.

→ The electrons may also come back to the cathode. This process reduces the efficiency of the tube.

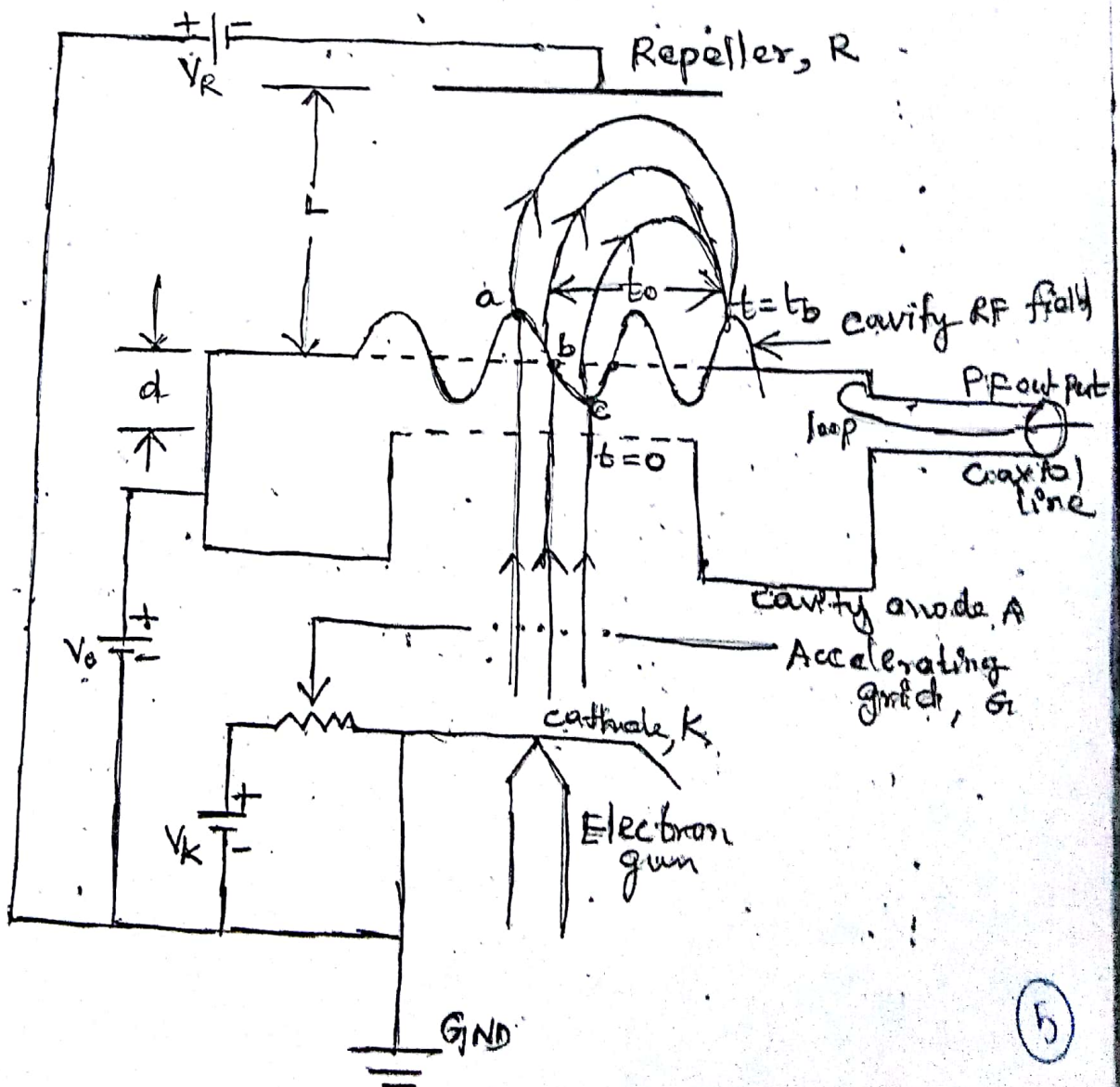
→ The transit time effect can be reduced by reducing the distance b/w the electrodes and increasing the plate-cathode voltage, V_a .

(A)

KLYSTRONS :

→ A klystron is a specialized linear beam vacuum tube, invented by Russel and Sigurd Varian (American) in 1937.

→ There are two basic configurations of klystron tubes, one is called reflex klystron, used as a low-power microwave oscillator and another is called multicavity klystron, used as a low-power microwave amplifier.



(5)

→ The schematic configuration of a reflex klystron tube is shown in the above figure. It uses only a single re-entrant microwave cavity as resonator.

→ It consists of an electron gun producing a collimated beam of electron. The emitted electron is accelerated by the grid, G and passes through the cavity anode AB to the repeller space between the cavity anode and the repeller electrode R .

→ The reflex klystron is an oscillator with a built in feedback mechanism. It uses the same cavity for bunching and for the output cavity.

→ The repeller electrode, R is a negative potential and sends the bunched electron beam back to the resonator cavity. This provides a positive feedback mechanism which supports oscillations.

(a) Mechanism of oscillation:

→ Due to dc voltage in the cavity circuit, RF noise is generated in the cavity.

→ This electromagnetic noise field in the cavity becomes pronounced at cavity resonant frequency. (b)

→ The electrons passing through the cavity gap d experiences this RF field and are velocity modulated in the following manner.

→ The electrons 'a' shown in the figure which encountered the positive half cycle of the RF field in the cavity gap d will be accelerated.

→ Those reference electrons 'b' which encountered zero RF field will pass with unchanged original velocity.

→ And the electrons 'c' which encountered the negative half cycle will be retarded on entering the repeller space.

→ All these velocity modulated electrons will be repelled back to the cavity by the repeller due to its negative potential.

→ The repeller distance L and the voltages can be adjusted to receive all the velocity-modulated electrons at a same time on the positive peak of the cavity RF cycle.

→ Thus, the velocity modulated electrons are bunched together and lose their kinetic energy when they encounter

the positive cycle of the cavity RF field. This loss of energy is thus transferred to the cavity to conserve the total power.

→ If the power delivered by the bunched electrons to the cavity is greater than the power loss in the cavity, the electromagnetic field amplitude at the resonant frequency of the cavity will increase to produce microwave oscillations.

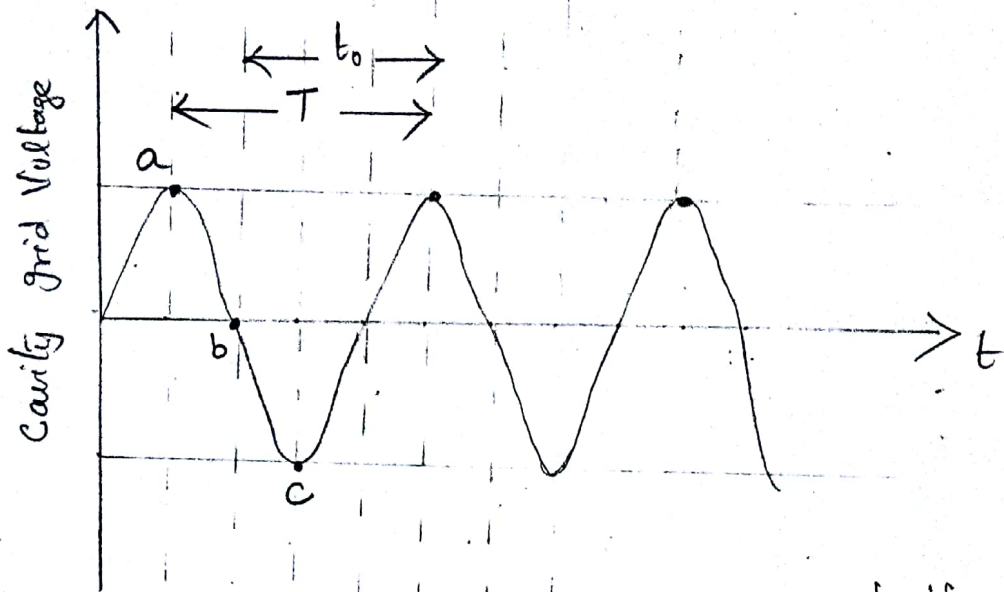
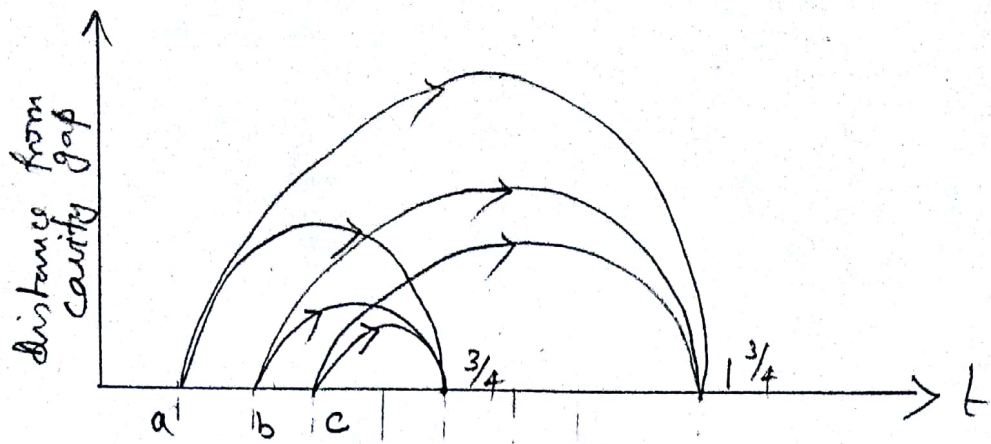
→ The RF power is coupled to the output load by means of a small loop which forms the center conductor of the coaxial line.

→ When the power delivered to the cavity by the electrons becomes equal to the total power loss in the cavity system, a steady microwave oscillation is generated and maintained at resonant frequency of the cavity.

(b). Mode of oscillation:

The bunched electrons in a reflex klystron can deliver maximum power to the cavity at any instant which corresponds to the positive peak of the RF cycle of the cavity oscillation.

(8)



If T is the time period at the resonant frequency, t_0 is the time taken by the reference electron to travel in the repeller space between entering the repeller space at b and the returning to the cavity at positive peak voltage on formation of the bunch, then $t_0 = (n + \frac{3}{4})T$

$$= NT$$

Where, $N = n + \frac{3}{4}$, $n = 0, 1, 2, 3, \dots$ as shown in the above figure.

→ Thus by adjusting repeller voltage for a given dimension of the reflex klystron, the bunching can be made to occur at $N = n + \frac{3}{4}$ Positive half cycle. Accordingly, the mode of oscillation is named as $N = \frac{3}{4}, 1\frac{3}{4}, 2\frac{3}{4}, \dots$ etc, for modes $n = 0, 1, 2, 3, \dots$ respectively.

→ It is obvious that the lowest-order mode $\frac{3}{4}$ occurs for a maximum value of repeller voltage when the transit time of the electrons in the repeller space is minimum.

→ Higher modes occurs at lower repeller voltages.

→ Since at the highest repeller voltage, the acceleration of the bunched electrons on return is maximum, the power output of the lowest mode is maximum.

(c) Power output and efficiency:

The power output of a reflex klystron is maximum if the bunched electrons

on return cross the cavity gap when the gap field is positive maximum. This power can be calculated from various parameters as described below with reference to the above two figures.

Velocity modulation: For ~~the~~ calculation of RF power, we will assume the following:

- (1) Cavity grids and repeller are plane parallel and very large in extent.
- (2) No RF field is excited in the repeller space.
- (3) Electrons are not intercepted by the cavity anode grid.
- (4) No debunching takes place in the repeller space.
- (5) The cavity RF gap voltage amplitude V_1 is small compared to the dc beam voltage V_0

$$V_1 \ll V_0$$

The electron velocity u attained due to the ^{dc} beam voltage V_0 while entering the cavity gap at $t=0$ is uniform, given by,

$$u = u_0$$



$$= \sqrt{\frac{2eV_0}{m}}$$

$$= 5.93 \times 10^5 \sqrt{V_0} \text{ m/sec}$$

where V_0 is in volt and $u = 0$ at cathode.

The instantaneous cavity RF voltage can be written as

$$V(t) = V_1 \sin \omega t, \text{ where } V_1 \ll V_0.$$

The average transit time through the cavity gap d and the transit angle are given by,

$$t_g = \frac{d}{u_0}; \quad \theta_g = \omega t_g$$

$$= \frac{\omega d}{u_0}$$

The average microwave voltage in the cavity gap can be written as,

$$V_{av} = \frac{1}{t_g} \int_0^{t_g} V_1 \sin \omega t \, dt$$

mass of $e^- = 9.1 \times 10^{-31}$

$$= \frac{1}{t_g} V_1 \left[-\frac{\cos \omega t}{\omega} \right]_0^{t_g}$$

$$= \frac{V_1}{\omega t_g} \left[-\cos \omega t_g + 1 \right]$$

$$= \frac{V_1}{\omega t_g} \left[1 - \cos \omega t_g \right]$$

Using the trigonometrical identity

$$\sin^2 \theta = \frac{1 - \cos 2\theta}{2}$$

$$\Rightarrow \text{Var} = \frac{V_1}{\Omega_g} \left[2 \sin^2 \frac{\Omega_g}{2} \right]$$

$$= \frac{V_1 \sin^2 \left(\frac{\Omega_g}{2} \right)}{\left(\frac{\Omega_g}{2} \right)} = \frac{V_1 \sin \left(\frac{\Omega_g}{2} \right) \sin \left(\frac{\Omega_g}{2} \right)}{\left(\frac{\Omega_g}{2} \right)}$$

$$\text{Var} = V_1 \beta_1 \sin \left(\frac{\Omega_g}{2} \right) ; \beta_1 = \frac{\sin \left(\frac{\Omega_g}{2} \right)}{\left(\frac{\Omega_g}{2} \right)}$$

$$\beta_1 = \frac{\sin \left(\frac{\omega d}{2u_0} \right)}{\left(\frac{\omega d}{2u_0} \right)}$$

It is called the beam-coupling Co-efficient of the cavity gap.

Therefore, the coupling between the electron beam and the cavity varies with the cavity gap d is

$\frac{\sin X}{X}$ form. Here, $\beta_1 = 1$ for $d \rightarrow 0$

The exit velocity from the cavity gap after velocity modulation is given by,

$$U(f_g) = \sqrt{\frac{2e(V_0 + \text{Var})}{m}}$$

$$\beta_1 = \frac{0}{0} \text{ form}$$

Applying L'Hospital's rule,

$$\beta_1 = \frac{\cos \frac{\omega d}{2u_0} \times \frac{\omega d}{2u_0}}{\frac{\omega d}{2u_0} \times \frac{\omega d}{2u_0}}$$

$$d \rightarrow 0 ; \beta_1 = 1$$

13

13

$$= \sqrt{\frac{2e (V_0 + V_1 \beta_1 \sin \frac{\phi_g}{2})}{m}}$$

$$= \sqrt{\frac{2e}{m} V_0 \left[1 + \frac{V_1 \beta_1 \sin \frac{\phi_g}{2}}{V_0} \right]} \rightarrow \textcircled{1}$$

where, $\frac{V_1 \beta_1}{V_0}$ is called the depth of modulation.

Equation $\textcircled{1}$ for velocity modulation becomes,

$$u(t_g) = \sqrt{\frac{2eV_0}{m} \left[1 + \frac{V_1 \beta_1 \sin \frac{\phi_g}{2}}{V_0} \right]^{\frac{1}{2}}}$$

$$= u_0 \left[1 + \frac{V_1 \beta_1 \sin \frac{\phi_g}{2}}{2V_0} \right]$$

$$u(t_g) = u_0 \left[1 + \frac{V_1 \beta_1 \sin(\omega t_g - \frac{\phi_g}{2})}{2V_0} \right]$$

Transit Time: The round trip transit time in the repeller space is given by,

$$t_r = \frac{2 \text{ velocity}}{\text{acceleration}}$$

$$= \frac{2u}{a}$$

The factor 2 in the numerator arises because of the to-and-fro journey.

The electron acceleration is given

by, $a = \frac{eE}{m}$

$$= \frac{e}{m} \left[\frac{V_0 + V_R + V_1 \sin \omega t}{L} \right]$$

$$\approx \frac{e}{m} \left(\frac{V_0 + V_R}{L} \right)$$

$$F = ma$$

$$\therefore F_e = qE$$

Here, $q = e$

$$\Rightarrow F_e = eE$$

$$ma = eE$$

$$a = \frac{eE}{m}$$

Therefore, $t_r = \frac{2U_0}{a}$

$$= \frac{2U_0 \left[1 + \frac{\beta_1 V_1}{2V_0} \sin \left(\omega t_0 - \frac{\omega_0}{2} \right) \right]}{e \left(\frac{V_0 + V_R}{mL} \right)}$$

$$\frac{2U_0 (V_0 + V_R)}{mL}$$

$$t_r = \frac{2U_0 mL \left[1 + \frac{\beta_1 V_1}{2V_0} \sin \left(\omega t_0 - \frac{\omega_0}{2} \right) \right]}{e (V_0 + V_R)} \rightarrow (2)$$

Since the reference electron does not undergo any velocity modulation, its transit time in repeller space,

$$t_0 = \frac{2U_0}{a} = \frac{2U_0}{\frac{e(V_0 + V_R)}{mL}}$$

$$= \frac{2U_0 mL}{e(V_0 + V_R)} \rightarrow (3)$$

Since, $t_0 = \left(n + \frac{3}{4} \right) T = NT$;

$$N = \left(n + \frac{3}{4} \right) T$$

$$\textcircled{2} \Rightarrow t_0 = \frac{2U_0 mL}{2(V_0 + V_R)} = NT$$

$$= \frac{2\pi N}{\omega}$$

From equations $\textcircled{2}$ & $\textcircled{3}$,
we get,

$$t_r = t_0 \left[1 + \frac{\beta_1 V_1}{2V_0} \sin \left(\omega t_g - \frac{\phi_g}{2} \right) \right]$$

$$\frac{2\pi N}{2\pi f}$$

$$= NT$$

$$\therefore \frac{1}{f} = T$$

Density modulation and Beam current :

The time of arrival of an electron to the cavity gap can be expressed by,

$$t_b = t_g + t_r$$

$$= t_g + t_0 \left[1 + \frac{\beta_1 V_1}{2V_0} \sin \left(\omega t_g - \frac{\phi_g}{2} \right) \right]$$

$$= t_g + t_0 + t_0 \frac{\beta_1 V_1}{2V_0} \sin \left(\omega t_g - \frac{\phi_g}{2} \right)$$

$$= t_g + \frac{2\pi N}{\omega} + \frac{2\pi N}{\omega} \frac{\beta_1 V_1}{2V_0} \sin \left(\omega t_g - \frac{\phi_g}{2} \right)$$

$$= t_g + \frac{2\pi N}{\omega} + \frac{\pi N}{\omega} \frac{\beta_1 V_1}{V_0} \sin \left(\omega t_g - \frac{\phi_g}{2} \right)$$

$$t_b = t_g + \frac{2\pi N}{\omega} + \frac{X}{\omega} \sin \left(\omega t_g - \frac{\phi_g}{2} \right)$$

(16)

where, $x = \frac{\pi N \beta_1 V_1}{V_0}$ is called the bunching parameter of the reflex klystron. (I)

From eqn (4),

$$\frac{dt_b}{dt_g} = 1 + \left[\frac{x}{\omega} \cos \left(\omega t_g - \frac{\phi_g}{2} \right) \cdot \omega \right]$$

$$= 1 + x \cos \left(\omega t_g - \frac{\phi_g}{2} \right) \rightarrow (5)$$

The bunched electrons on return constitute the bunched beam current $i_b \left(\frac{dq}{dt} \right)$ such that the conservation of charge gives $\frac{dt_g}{|dt_b|} = \frac{i_b}{I_0}$

$$I_0 |dt_g| = i_b |dt_b| \rightarrow (6)$$

where I_0 is the dc beam current.

From eqns, (5) & (6),

$$\therefore \frac{i_b}{I_0} = \left[1 + x \cos \left(\omega t_g - \frac{\phi_g}{2} \right) \right]^{-1}$$

From eqn. (4), since $V_1 \ll V_0$, $x \ll 1$,

$$t_b = t_g + \frac{2\pi N}{\omega}, \text{ so that}$$

$$i_b = I_0 \left[1 + x \cos \left[\omega \left(t_b - \frac{2\pi N}{\omega} \right) - \frac{\phi_g}{2} \right] \right]^{-1}$$

(17)

$$i_b = I_0 \left[1 + x \cos \left(\omega t_b - 2\pi N - \frac{\theta_g}{2} \right) \right] \quad (7)$$

By Fourier expansion, the beam current of a reflex klystron oscillator can be expressed by,

$$\begin{aligned} i_b &= I_0 + 2 I_0 \sum_{n=1}^{\infty} J_n(n x) \cos n \left[\omega t_b - 2\pi N - \frac{\theta_g}{2} \right] \\ &= I_0 + 2 I_0 J_1(x) \cos \left[\omega \left(t_b - \frac{2\pi N}{\omega} - \frac{\theta_g}{2\omega} \right) \right] \\ &\quad + \sum_{n=2}^{\infty} 2 I_0 J_n(n x) \cos n \left[\omega t_b - 2\pi N - \frac{\theta_g}{2} \right] \\ &= I_0 + 2 I_0 J_1(x) \cos \omega \left[t_b - t_0 - \frac{t_g}{2} \right] \\ &\quad + \sum_{n=2}^{\infty} 2 I_0 J_n(n x) \cos n \omega \left[t_b - t_0 - \frac{t_g}{2} \right]; \end{aligned}$$

The fundamental component of the RF induced current in the cavity is, therefore, $t_0 = \frac{2\pi N}{\omega}$

$$i_{RF} = \beta_1 2 I_0 J_1(x) \cos \omega \left(t_b - t_0 - \frac{t_g}{2} \right)$$

or

$$\begin{aligned} i_{RF}(t_b) &= 2 I_0 \beta_1 J_1(x) \cos(\omega t_b - \omega t_0) \\ &= 2 I_0 \beta_1 J_1(x) \cos(\omega t_b - 2\pi N) \end{aligned}$$

where $\frac{t_g}{2} \ll 2\pi N$, is neglected for smallness.

(18) $t_0 = \frac{2\pi N}{\omega}$

Power output: The magnitude of the fundamental RF current in the cavity is given by,

$$|\hat{i}_{RF}| = 2 I_0 \beta_1 J_1(x)$$

The rms RF power delivered to the cavity is

$$P_{RF} = V_1 \frac{\hat{i}_{RF}}{2}$$

$$= V_1 \frac{2 I_0 \beta_1 J_1(x)}{2}$$

$$= V_1 I_0 \beta_1 J_1(x) \rightarrow \textcircled{8}$$

From \textcircled{I} , $\frac{V_1}{V_0} = \frac{x}{\pi N \beta_1} \Rightarrow V_1 = \frac{V_0 x}{\pi N \beta_1}$

\therefore eqn $\textcircled{8}$ reduces to,

$$P_{RF} = \frac{V_0 x}{\pi N \beta_1} I_0 \beta_1 J_1(x)$$

$$= \frac{V_0 I_0 x J_1(x)}{\pi N} \rightarrow \textcircled{9}$$

from \textcircled{II} & \textcircled{III}

$$\frac{2 u_0 m L}{2(V_0 + V_R)} = \frac{2 \pi N}{\omega}$$

$$u_0 = \sqrt{\frac{2eV_0}{m}} ; t_0 = \frac{2 \pi N}{\omega} = \frac{2 u_0 m L}{e(V_0 + V_R)}$$

$\hookrightarrow \textcircled{II}$

$\hookrightarrow \textcircled{III}$

$$\frac{2mL}{e(V_0 + V_R)} \sqrt{\frac{2eV_0}{m}} = \frac{2\pi N}{\omega}$$

$$\pi N = \frac{\omega L}{(V_0 + V_R)}$$

$$\pi N = \frac{\omega L}{(V_0 + V_R)} \sqrt{\frac{2m}{e} V_0} = \sqrt{\frac{2m}{e} V_0}$$

$$\therefore P_{RF} = \frac{V_0 I_0 \times J_1(x)}{\frac{\omega L}{(V_0 + V_R)} \sqrt{\frac{2m}{e} V_0}}$$

$$P_{RF} = \frac{V_0 I_0 \times J_1(x) (V_0 + V_R)}{2\pi f L} \sqrt{\frac{e}{2mV_0}}$$

Efficiency: The dc Power supplied by the beam voltage V_0 is

$$P_{dc} = V_0 I_0$$

Therefore, the electronic efficiency of a reflex klystron oscillator is

$$\eta = \frac{P_{RF}}{P_{dc}} = \frac{V_0 I_0 \times J_1(x)}{\pi N V_0 I_0}$$

$$\eta = \frac{x J_1(x)}{\pi N}$$

where, $x = \frac{\pi N \beta_1 V_1}{V_0}$; $N = n + \frac{3}{4}$,

n , the mode number = 0, 1, 2, 3, ...

$$\beta_1 = \frac{3 \sin \frac{\theta_g}{2}}{\frac{\theta_g}{2}} ; \text{ and } v_0 = \sqrt{\frac{2eV_0}{m}}$$

It can be shown from Bessel's function table, $X J_1(X)$ attains a maximum value of 1.252 at $X = 2.408$. Thus, the optimum values are:

$$P_{RF} = \frac{V_0 I_0 (1.252)}{\pi N}$$

$$= \frac{0.3986 V_0 I_0}{N}$$

$$\therefore \eta = \frac{P_{RF}}{P_{dc}} = \frac{P_{RF}}{V_0 I_0}$$

$$= \frac{0.3986 V_0 I_0}{N \cdot V_0 I_0}$$

$$\eta = \frac{0.3986}{N}$$

It has been observed that it is not possible to get $3/4$ mode in reflex klystron, so that $N = 1 \frac{3}{4}$ mode leads maximum RF Power output and efficiency:

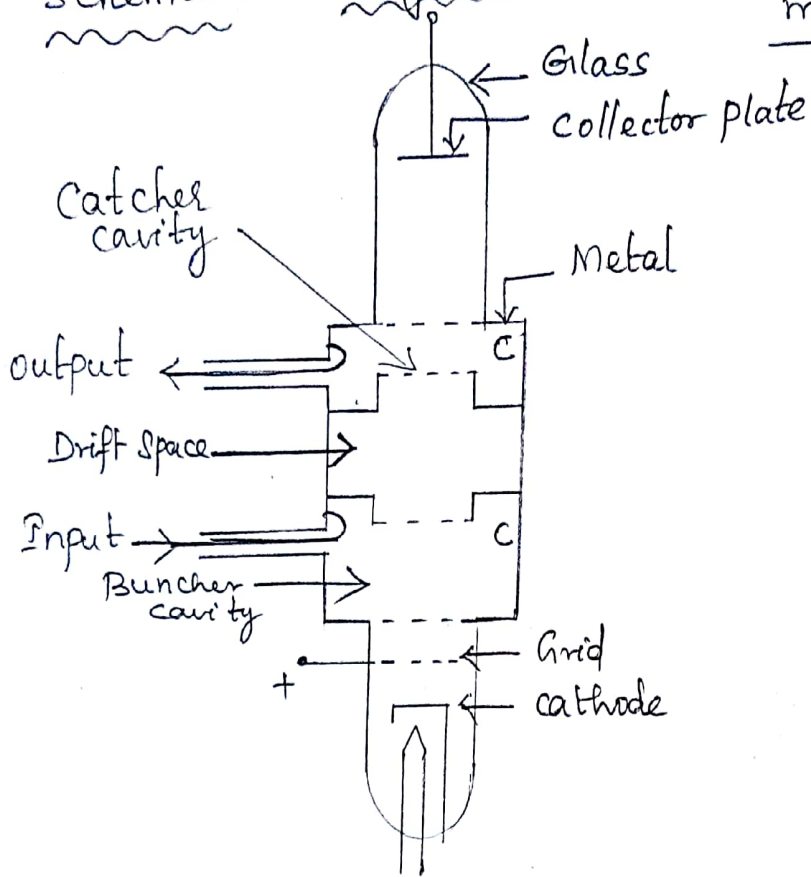
$$P_{RF \max} = 0.227 V_0 I_0$$

$$\eta_{\max} = 29.7\%$$

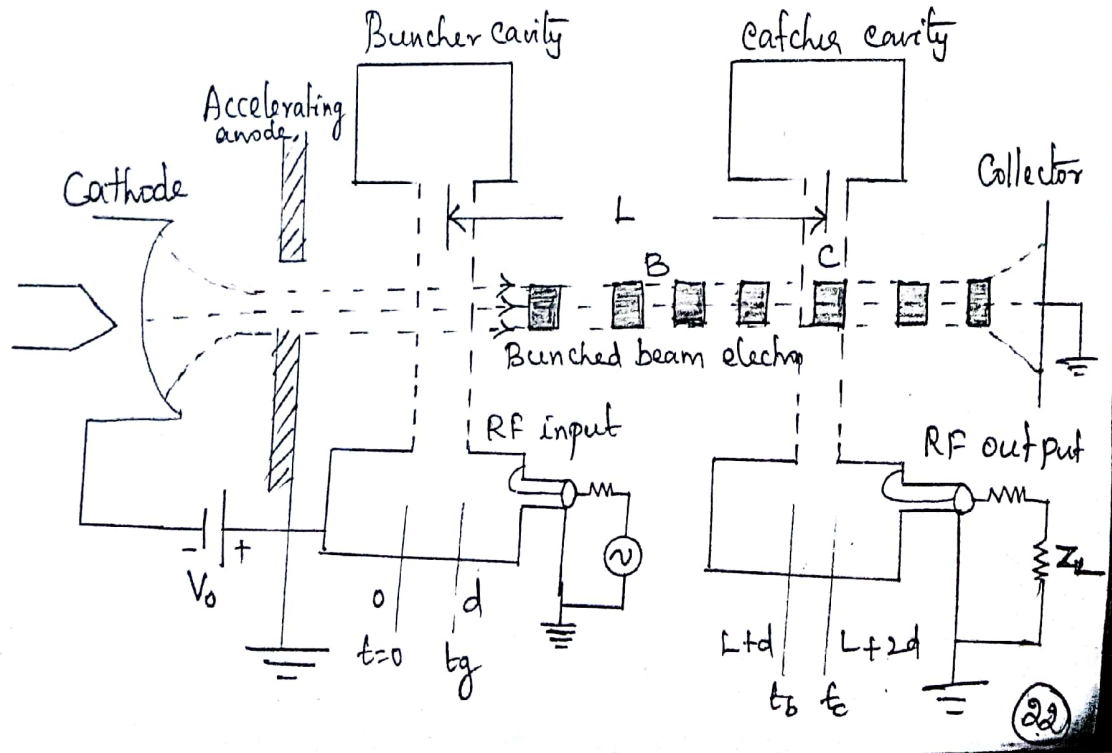
Two-cavity Klystron amplifier:

→ The two-cavity Klystron tubes are widely used for microwave amplification

→ Operated by the Principles of velocity Schematic diagram: and current modulation



Functional Diagram



Mechanism of operation:

A two-cavity klystron amplifier consists of an

(1) input re-entrant cavity resonator (Buncher cavity) [Velocity modulates the e^- beam]

(2) output re-entrant cavity resonator (catcher cavity) [Catches energy from the bunched electron beam]

→ A high velocity electron beam produced by the accelerating anode is successively passed through an input buncher cavity, a field free drift space of length L , an output catcher cavity and finally collected by a collector electrode.

→ The electron beam is focussed to travel axially without spreading during transit by applying an axial magnetic field produced by an external coil current.

→ The input RF signal to be amplified excites the buncher cavity with a coupling loop.

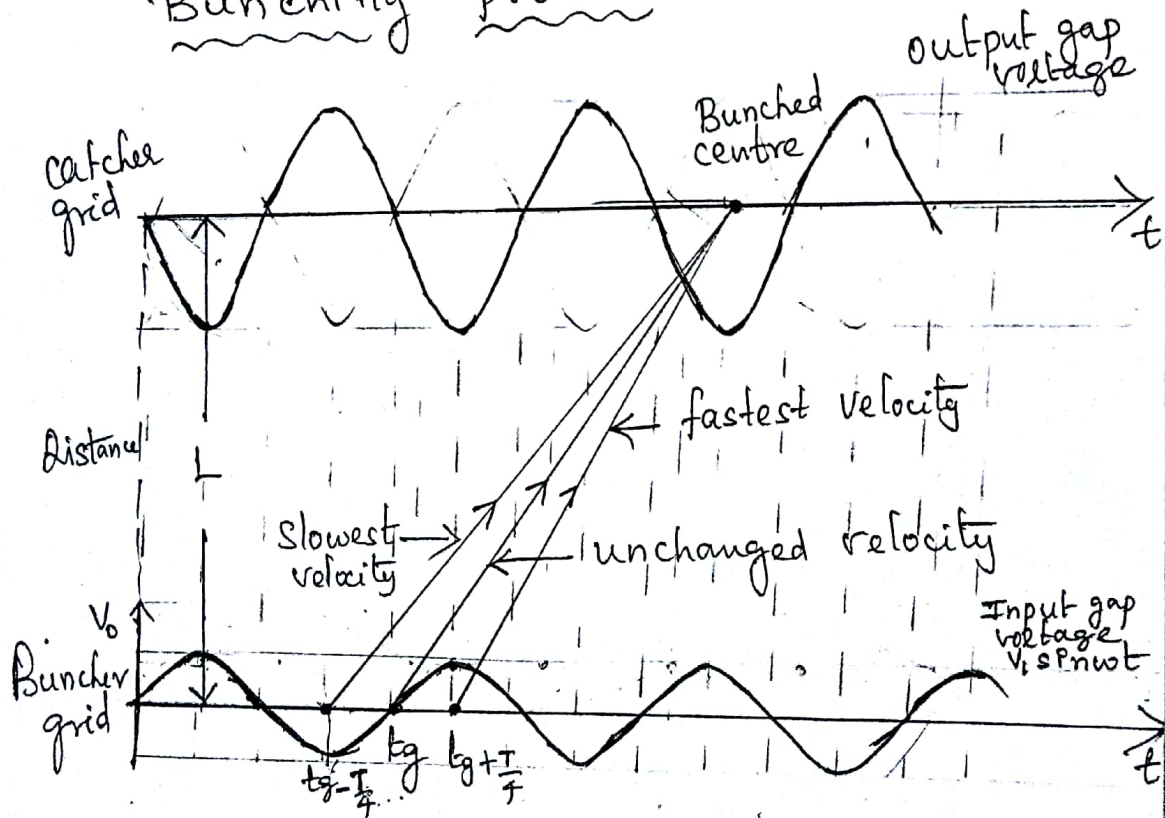
→ The combination of the anode voltage V_0 and the cavity gap width 'd' are such that the transit of electrons through each cavity gap is less than the quarter of the time period T of the input signal cycle.

- All electrons injected from the cathode arrive at the first cavity with uniform velocity.
- Those electrons passing the first cavity gap at zeros of signal voltage pass through with unchanged velocity; those passing through the positive half cycles of the gap voltage or signal voltage undergo an increase in velocity; those passing through the negative half cycles of the gap voltage undergo a decrease in velocity.
- As a result of these actions, the electrons gradually bunch together as they travel down the drift space. The variation in electron velocity in the drift space is known as velocity modulation.
- The density of the electrons in the second cavity gap varies cyclically with time.
- The electron beam contains an ac component and is said to be current-modulated.

→ The bunched electrons undergo density modulation in accordance with the input RF signal cycle.

→ While passing through the catcher cavity grid, this density modulated electron beam induces RF current in the output cavity and thereby excites the RF field in the output cavity at its signal cycle.

Bunching process:



→ The electrons that pass the buncher at $V_s = 0$ travel through with unchanged velocity V_0 and become the bunching center.

→ Those electrons that pass the buncher cavity during the positive half cycles of the microwave input voltage V_s travel faster than the electrons that passed the gap when $V_s = 0$.

→ Those electrons that pass the buncher cavity during the negative half cycles of the voltage V_s travel slower than the electrons that passed the gap when $V_s = 0$.

→ This is shown by the distance-time plot and is called the applegate diagram.

→ From the above diagram, it is shown that the phase of field in the output cavity is opposite to that of the input cavity so that the bunched electrons are retarded by the output gap voltage.

→ The loss of kinetic energy of the electrons on retardation process transfers RF energy to the output cavity continuously at signal cycle.

→ The amplitude of the signal at output cavity attains a steady

large value when the loss of kinetic energy of the bunched electrons compensates the output cavity circuit losses. †

→ The amplified signal is coupled out from the catcher cavity through a current loop to the load.

→ The current induced in the catcher cavity is rich in harmonics up to say, 15. The cavity is tuned to the fundamental or any harmonic as desired.

Analysis : The analysis of two-cavity klystron amplifier is based on the following assumptions:

(1) The transit time in the cavity gap is very small compared to the period of the input RF signal cycle.

(2) The input RF signal amplitude V_1 is very small compared to the dc beam voltage, V_0

(3) The cathode, anode, cavity grids and collector are all parallel and the cavity grids do not intercept any electron while passing.

(4) No space charge or debunching takes place at the bunch point.

(5) The RF fields are totally confined in the cavity gaps so that field is zero in the drift space L .

(6) The electrons leave the cathode with zero initial velocity.

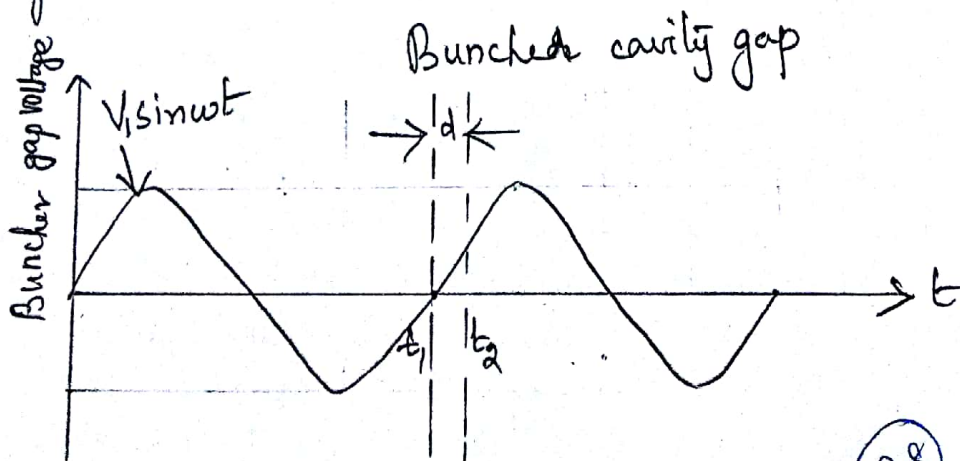
Velocity modulation:

when electrons are first accelerated by the high dc voltage V_0 before entering the buncher grids, their velocity is uniform:

$$v_0 = \sqrt{\frac{2eV_0}{m}} = 5.93 \times 10^5 \sqrt{V_0} \text{ m/sec}$$

when a microwave signal is applied to the input terminal, the gap voltage between the buncher grids appears as, $V_s = V_1 \sin \omega t$

where V_1 is the amplitude of the signal and $V_1 \ll V_0$



(28)

The electron beam enters the buncher cavity at a time $t = t_1$ with velocity u_0 and passes out at $t = t_2$ as shown in the above figure.

The transit time through the buncher gap 'd' is given by

$$t_2 - t_1 = t_g = \frac{d}{u_0}$$

and the transit angle through the buncher gap 'd' is given by

$$\theta_g = \omega t_g = \omega \left(\frac{d}{u_0} \right) \rightarrow (2)$$

Due to input RF signal in the buncher cavity, the average RF voltage in the buncher gap can

be obtained as

$$V_{av} = \frac{1}{t_g} \int_{t_1}^{t_2} V_1 \sin \omega t \, dt$$

$$= \frac{V_1}{t_g} \left[\frac{-\cos \omega t}{\omega} \right]_{t_1}^{t_2}$$

$$= -\frac{V_1}{\omega t_g} \left[\cos \omega t_2 - \cos \omega t_1 \right]$$

$$= \frac{V_1}{\omega t_g} \left[\cos \omega t_1 - \cos \omega t_2 \right]$$

(2)

or

$$V_{av} = V_1 \left(\frac{\sin \frac{\theta_g}{2}}{\frac{\theta_g}{2}} \right) \sin \left(\omega t_1 + \frac{\theta_g}{2} \right)$$

$$= V_1 \beta_1 \sin \left(\omega t_1 + \frac{\theta_g}{2} \right) \rightarrow (3)$$

where, $\beta_1 = \left(\frac{\sin \frac{\theta_g}{2}}{\frac{\theta_g}{2}} \right)$, is called the buncher cavity beam-coupling Co-efficient.

As electrons pass through the buncher gap, their velocities are increased, decreased or unchanged, depending on the positive, negative or zero RF voltage across the grids when they pass through.

At time, $t_{av} = \frac{t_1 + t_2}{2}$, the electron is midway across the gap with velocity say, U_{av} .

$$\text{Then } U_{av} = \sqrt{\frac{2e(V_0 + V_{av})}{m}}$$

$$= \sqrt{\frac{2e}{m} \left(V_0 + V_1 \beta_1 \sin \left(\omega t_1 + \frac{\theta_g}{2} \right) \right)}$$

$$= \sqrt{\frac{2eV_0}{m} \left(1 + \frac{V_1}{V_0} \beta_1 \sin \left(\omega t_1 + \frac{\theta_g}{2} \right) \right)}$$

$$= \sqrt{\frac{2eV_0}{m}} \sqrt{1 + \frac{V_1}{V_0} \beta_1 \sin\left(\omega t_1 + \frac{\phi_g}{2}\right)}$$

$$= U_0 \left(1 + \frac{V_1}{V_0} \beta_1 \sin\left(\omega t_1 + \frac{\phi_g}{2}\right)\right)^{1/2}$$

Using binomial expansion, we get

$$U_{av} = U_0 \left(1 + \frac{V_1}{2V_0} \beta_1 \sin\left(\omega t_1 + \frac{\phi_g}{2}\right)\right)$$

$$U_{av} \approx U_0 \left(1 + \frac{M}{2} \sin\left(\omega t_1 + \frac{\phi_g}{2}\right)\right) \rightarrow \text{⑤}$$

Thus, the electrons in the beam are velocity modulated by the input RF signal with depth of velocity modulation,

$$M = \frac{\beta_1 V_1}{V_0}$$

→ These electrons exit from the buncher gap at time, $t_1 + t_g$ to the field free drift space between the two cavities with velocity,

$$U_e(t_2) \approx U_0 \left[1 + \frac{M}{2} \sin\left(\omega(t_2 - t_g) + \frac{\phi_g}{2}\right)\right]$$

$$\therefore t_1 = t_2 - t_g$$

$$\approx U_0 \left[1 + \frac{M}{2} \sin\left(\omega t_2 - \omega t_g + \frac{\phi_g}{2}\right)\right]$$

$$\approx U_0 \left[1 + \frac{M}{2} \sin\left(\omega t_2 - \phi_g + \frac{\phi_g}{2}\right)\right]$$

$$u_a(t_2) \approx u_0 \left[1 + \frac{M}{2} \sin(\omega t_2 - \frac{\phi_g}{2}) \right] \rightarrow \textcircled{6}$$

Transit Time in the Drift space:

If t_3 is the time when the bunched electrons are at the catcher grid after travelling through the field free drift space L ,

$$t_3 = t_2 + \frac{L}{u(t_2)}$$

$$= t_2 + \frac{L}{u_0 \left[1 + \frac{M}{2} \sin(\omega t_2 - \frac{\phi_g}{2}) \right]}$$

$$\approx t_2 + \left(\frac{L}{u_0} \right) \left(1 + \frac{M}{2} \sin(\omega t_2 - \frac{\phi_g}{2}) \right)^{-1}$$

$$t_3 \approx t_2 + \left(\frac{L}{u_0} \right) \left(1 - \frac{M}{2} \sin(\omega t_2 - \frac{\phi_g}{2}) \right) \rightarrow \textcircled{7}$$

Therefore, the transit time t_d in the drift space is given by,

$$t_d \approx t_3 - t_2$$

$$\approx \left(\frac{L}{u_0} \right) \left[1 - \frac{M}{2} \sin(\omega t_2 - \frac{\phi_g}{2}) \right] \rightarrow \textcircled{8}$$

Density modulation: Because of the difference in velocities of the electrons in the velocity modulated beam, the electrons will form bunches.

(ii) become density modulated, in accordance with the input signal cycle.

From eqn (6), (from the Applegate diagram)

the minimum, maximum and unchanged velocities are given by

$$U_{\min} = U_0 \left[1 - \frac{M}{2} \right]$$

$$U_{\max} = U_0 \left[1 + \frac{M}{2} \right] \text{ and}$$

$$U_{\text{unchange}} = U_0.$$

A maximum degree of bunching takes place when the buncher and catcher cavities are spaced to satisfy the condition,

$$t_d = t_3 - t_2$$

$$= \frac{L}{U(t_2)}$$

$$= \frac{L}{U_0 \left[1 + \frac{M}{2} \sin \left(\omega t_2 - \frac{\alpha_0}{2} \right) \right]}, \text{ from (6)}$$

$$t_d = t_0 \left[1 - \frac{M}{2} \sin \left(\omega t_2 - \frac{\alpha_0}{2} \right) \right] \rightarrow (9)$$

where, $t_0 = \frac{L}{U_0}$, is the dc transit time in drift space.

The corresponding transit angle in the drift space L is given by,

$$Q_d = \omega t_d$$

$$= \omega (t_3 - t_2)$$

$$= \omega \left[t_0 \left(1 - \frac{M}{2} \sin(\omega t_2 - \frac{\phi_g}{2}) \right) \right]$$

$$= \omega t_0 - \frac{\omega M t_0}{2} \sin(\omega t_2 - \frac{\phi_g}{2})$$

$$Q_d = Q_0 - \left(\frac{M}{2} \right) Q_0 \sin(\omega t_2 - \frac{\phi_g}{2}) \rightarrow \textcircled{10}$$

where, $Q_0 = \omega t_0$; the dc transit angle.

→ If N is the number of RF cycles elapsed during the transit time of reference electron with velocity U_0 , at the point of bunching ωt_0 's

$$t_0 = \frac{L}{U_0} = NT = \frac{N}{f} = \frac{2\pi N}{\omega} \rightarrow \textcircled{11}$$

$$\Rightarrow t_0 = \frac{2\pi N}{\omega} \Rightarrow \omega t_0 = 2\pi N$$

Therefore, $Q_0 = \omega t_0 = 2\pi N$.

The catcher cavity is placed so that the electron bunching occurs within the catcher gap.

N = no. of electron transit cycles in the drift space.

Let I_0 = dc beam current in the buncher

i_b = bunched beam current at catcher.

$$i_b = \frac{dQ}{dt_3}$$

Assuming there is no loss of electrons,
total charge

$$I_0 = \frac{dQ}{dt_2}$$

$$Q = \sum dQ$$

$$Q = \sum I_0 dt_2 = \sum i_b dt_3$$

$$\Rightarrow i_b = \frac{I_0}{\sum \left(\frac{dt_3}{dt_2} \right)} \rightarrow \textcircled{11}$$

from eqn $\textcircled{7}$,

$$t_3 = t_2 + \left(\frac{L}{u_0} \right) \left[1 - \frac{M}{2} \sin \left(\omega t_2 - \frac{\phi_0}{2} \right) \right]$$

$$= t_2 + \frac{2\pi N}{\omega} \left[1 - \frac{M}{2} \sin \left(\omega t_2 - \frac{\phi_0}{2} \right) \right]$$

$$= t_2 + \frac{2\pi N}{\omega} - \frac{2\pi N}{\omega} \frac{M}{2} \sin \left(\omega t_2 - \frac{\phi_0}{2} \right)$$

$$= t_2 + \frac{2\pi N}{\omega} - \frac{\cancel{2\pi N}}{\cancel{\omega}} \frac{M}{2} \sin \left(\omega t_2 - \frac{\phi_0}{2} \right)$$

$$t_3 = t_2 + \frac{2\pi N}{\omega} - \frac{M\pi N}{\omega} \sin \left(\omega t_2 - \frac{\phi_0}{2} \right)$$

$$\frac{dt_3}{dt_2} = 1 - \frac{M\pi N}{\omega} \cos \left(\omega t_2 - \frac{\phi_0}{2} \right)$$

$$= 1 - X \cos \left(\omega t_2 - \frac{\phi_0}{2} \right)$$

where, $X = \pi M N = \pi N \frac{\beta_1 V_1}{V_0}$ and is $\textcircled{12}$

defined as the bunching parameter of
a two-cavity klystron.

$\textcircled{35}$

$$(11) \Rightarrow i_b = \frac{I_0}{\sum [1 - x \cos(\omega t_2 - \frac{\phi_0}{a})]} \rightarrow (12)$$

Eqn (12) is the expression of the bunched beam current in the o/p cavity.

$$\therefore t_0 = t_3 - t_2 \Rightarrow t_2 = t_3 - t_0$$

$$i_b = \frac{I_0}{\sum [1 - x \cos(\omega(t_3 - t_0) - \frac{\phi_0}{2})]}$$

$$= \frac{I_0}{\sum [1 - x \cos(\omega t_3 - \underbrace{\omega t_0}_{\omega(\frac{2DN}{\omega})} - \frac{\phi_0}{2})]}$$

$$= \frac{I_0}{\sum [1 - x \cos(\omega t_3 - \frac{\phi_0}{2} - \frac{2DN}{\omega})]}$$

$$= I_0 \sum [1 - x \cos(\omega t_3 - 2DN - \frac{\phi_0}{2})] \rightarrow (13)$$

Expanding as a Fourier series and finding the coefficients.

$$i_b = I_0 + 2I_0 \sum_{n=1}^{\infty} J_n(nx) \cos n(\omega t_3 - \frac{\phi_0}{2})$$

$\rightarrow (14)$

\rightarrow This is the generalized expression of a bunched beam current consisting of a dc component I_0 plus the fundamental ac component of amplitude $2I_0 J_1(x)$ and harmonics of amplitude $2I_0 J_n(nx)$, $n=2,3, \dots$

The klystron is generally tuned to fundamental ac component of current given by,

$$i_f(t_3) = 2 I_0 J_1(x) \cos(\omega t_3 - \alpha_g - \alpha_0) \quad \rightarrow (15)$$

The current can be made maximum when $J_1(x)$ is $\text{max} = 0.582$ at $x = 1.841$ by adjusting the dc beam voltage V_0 .

We know that, $t_0 = \frac{L}{U_0} = \frac{2\pi N}{\omega}$
(ie) (I) \Rightarrow

$$\Rightarrow L = \frac{2\pi N U_0}{\omega} \quad \rightarrow (16)$$

$$(ii) \Rightarrow x = \pi M N = \pi N \left(\frac{\beta_1 V_1}{V_0} \right)$$

$$\Rightarrow N = \frac{x V_0}{\pi \beta_1 V_1}$$

$$\therefore (16) \Rightarrow L = 2\pi \left(\frac{x V_0}{\pi \beta_1 V_1} \right) \frac{U_0}{\omega}$$

It can be shown that the fundamental component of current is maximum for an optimum drift space length.

$$\Rightarrow L_{opt} = \frac{2 (1.841) V_0 U_0}{\omega \beta_1 V_1}$$

$$L_{opt} = \frac{3.682 V_0 U_0}{\omega \beta_1 V_1} \rightarrow (17)$$

Beam Spreading / Debunching: At the point of bunching, the electrostatic force of repulsion between the electrons does not allow electronic collision, but may cause beam spreading or undesirable debunching. This reduces the efficiency of the klystrons.

Power output: The fundamental component of RF beam current passing through the output cavity gap induces a current in the catcher cavity.

$$i_c = i_f \beta_2 \rightarrow (18)$$

where β_2 is the beam-coupling Co-efficient of the catcher cavity gap.

$\beta_2 = \beta_1$ when both buncher and catcher cavities are identical.

From (15) & (18),

$$i_c = 2 I_0 J_1(x) \cos(\omega t_3 - \alpha_g - \alpha_0) \beta_2$$

$$= \frac{I_2}{\beta_2} \cos(\omega t_3 - \alpha_g - \alpha_0); I_2 = 2 I_0 J_1(x) \rightarrow (19)$$

The corresponding RF voltage across the catcher cavity is given

$$\text{by, } V_c = V_2 \cos(\omega t_3 - \omega_0 - \phi) = \beta_2 I_0 R_{ch} \rightarrow (20)$$

where ϕ is the phase angle between i_c and V_c .

The average power delivered to the output cavity is,

$$P_0 = \frac{1}{2\pi} \int_0^{2\pi} i_c V_c d(\omega t_3)$$

$$= \frac{1}{2\pi} \int_0^{2\pi} 2 I_0 \beta_2 J_1(x) V_2 \cos(\omega t_3 - \omega_0 - \phi) d(\omega t_3)$$

$$= \frac{1}{2\pi} 2 I_0 \beta_2 J_1(x) V_2 \int_{\omega t_3=0}^{2\pi} \cos(\omega t_3 - \omega_0 - \phi) d(\omega t_3)$$

$$= \beta_2 I_0 V_2 J_1(x) \cos \phi \rightarrow (21)$$

The coupling between the beam and RF field in the gap is better for higher values of $\beta \leq 1$.

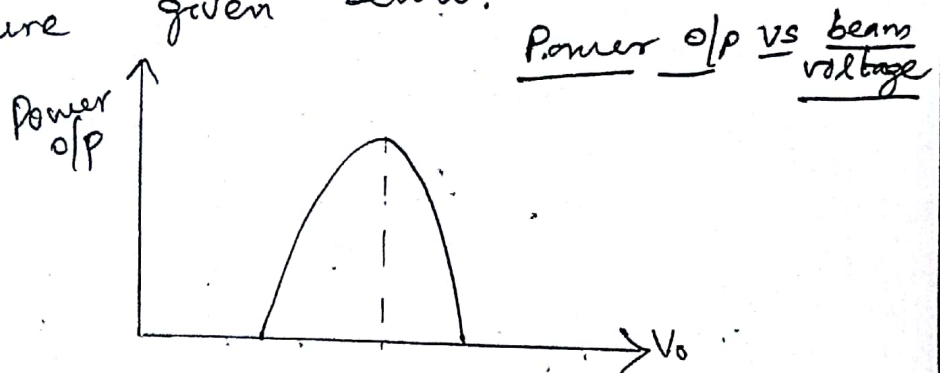
For maximum power output, $\cos \phi = 1$,

$$\phi = 0^\circ; \quad x = 1.841 \text{ and } \beta_2 = 1$$

$$J_1(x) = 0.582$$

Therefore, $P_{\max} = I_0 V_2 J_1(x)$
 $= 0.582 I_0 V_2 \rightarrow (22)$

Since power output is a function of $x = \frac{\pi N \beta_1 V_1}{V_0}$, it can be varied by changing V_0 as shown in figure given below.



from (I), $t_0 = \frac{L}{u_0} = NT = \frac{N}{f}$

$\Rightarrow u_0 N = Lf$
 $u_0 = \frac{Lf}{N}$
 or

$\sqrt{\frac{deV_0}{m}} = \frac{Lf}{N}$

Squaring on both sides, we get

$\frac{deV_0}{m} = \frac{L^2 f^2}{N^2}$

$V_0 = \frac{m}{de} \left(\frac{Lf}{N} \right)^2 \rightarrow (23)$

Efficiency: Efficiency η of a two-cavity klystron can be defined as the ratio of RF o/p power to the dc beam power.

$$\eta = \frac{P_{ac}}{P_{dc}}$$

$$= \frac{\beta_2 I_0 V_2 J_1(x)}{I_0 V_0}$$

$$= \frac{\beta_2 V_2 J_1(x)}{V_0}$$

The efficiency becomes maximum, when $x = 1.841$; $J_1(x) = 0.582$

$$\eta_{max} = \frac{\beta_2 V_2}{V_0} (0.582) \rightarrow (24)$$

If the coupling is perfect, $\beta_2 = 1$

$$I_{cmax} = 2 I_0 J_1(x)$$

$$= 2 I_0 (0.582)$$

and $V_2 = V_0$

$$\Rightarrow \eta_{max} = \frac{(1)(V_0)(0.582)}{V_0}$$

$$= 0.582$$

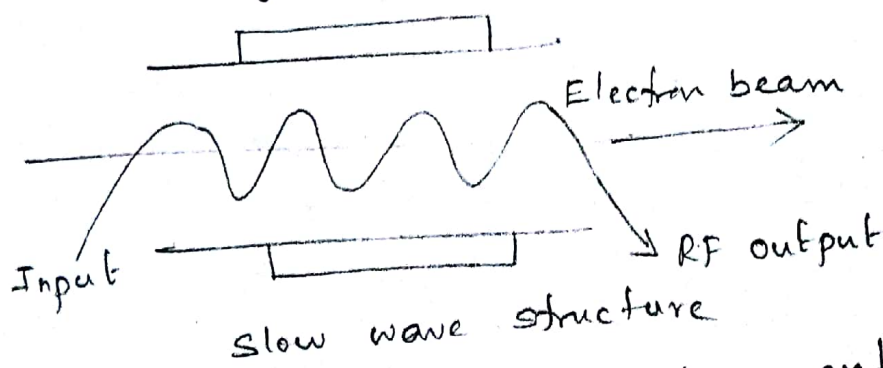
$$\eta_{max} \text{ in } \% = \underline{\underline{58.2 \%}}$$

(41)

Travelling wave Tubes

→ The TWT (Travelling wave Tube) employs a magnetically focussed electron beam and a slow wave structure.

→ Slow wave structures: These are special type of circuits, used in the microwave tubes to reduce the wave velocity in a certain direction so that a prolonged interaction between the electron beam and the signal wave may take place.



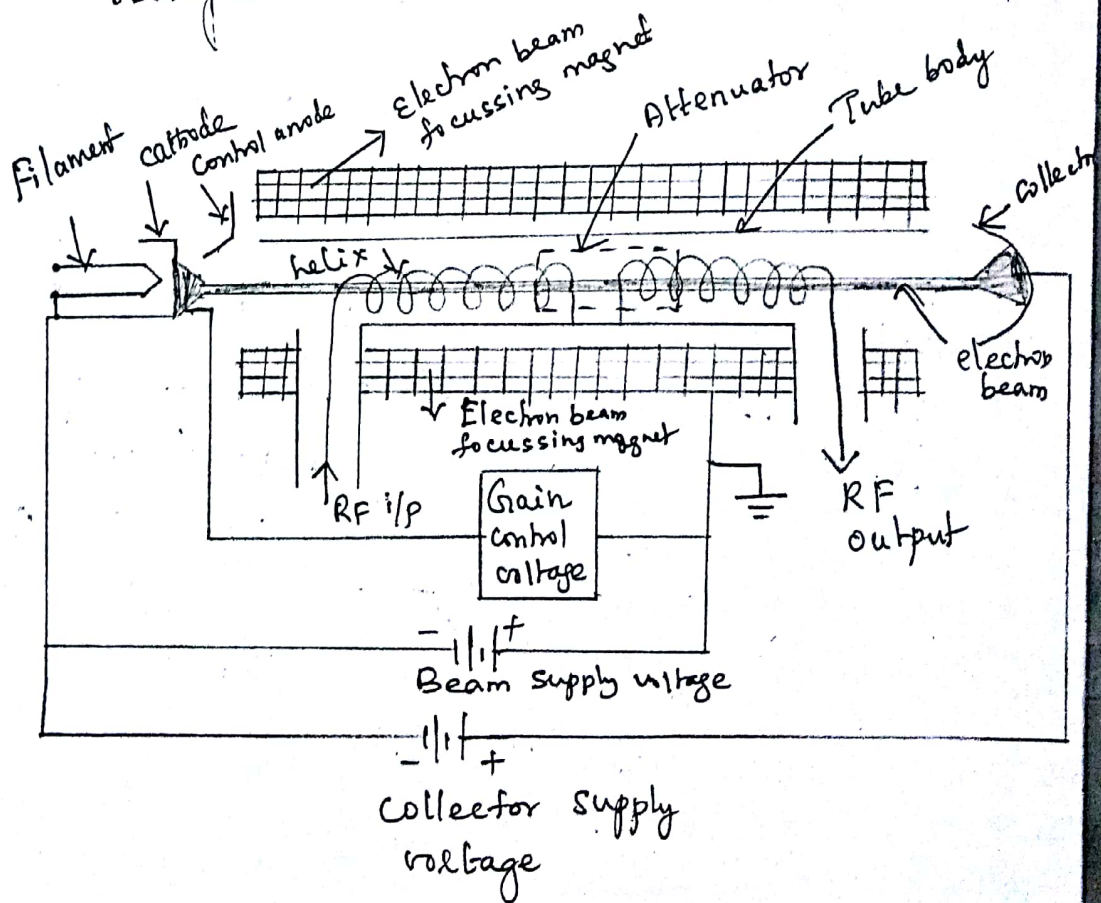
Principle: In order to have extended interaction, the electron beam and the microwave signal must travel together at the same rate of speed. Unfortunately, as we know that the RF signal travels at the speed of light, while the electrons are much slower. With the help of slow wave structure the RF signal is slowed down. The amount of

Slowing down is determined by the pitch of the spiral and thus the phase velocity of the field is given by, $v_p = \frac{p}{t}$, p is the pitch of the helix, t is the transit time to complete one pitch

Travelling wave tube amplifier:

→ TWTA forms a high gain $> 40\text{dB}$, low noise ($\text{NF} < 10\text{dB}$), wide band microwave amplifier.

→ They are designed for the frequency range $0.3 - 50\text{GHz}$.



- The TWT contains an electron gun, which produces and then accelerates an electron beam along the axis of the tube.
- The surrounding static magnet provides a magnetic field along the axis of the tube to focus the electrons into a tight beam.
- A helical slow wave structure is placed at the centre of the tube to slow down the velocity ^{with} which the RF signal travels so that it travels with the velocity of electrons in the beam.
- The RF signal injected at the input end of the helix travels down the helix wire at the speed of light, but the coiled shape causes the wave to travel a much greater total distance than the electron beam.
- Changing the no. of turns or the diameter of the turns in the helix wire, the speed at which the RF signal wave travels in the form of axial \vec{E} field

down the tube, can be varied.

- DC beam voltage is adjusted so that beam velocity is slightly greater than that of the axial field.
- An attenuator is placed over a part of the helix on midway to attenuate any reflected waves generated due to impedance mismatch that could be fed back to the input to cause oscillations.

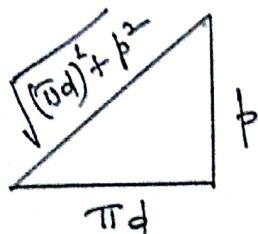
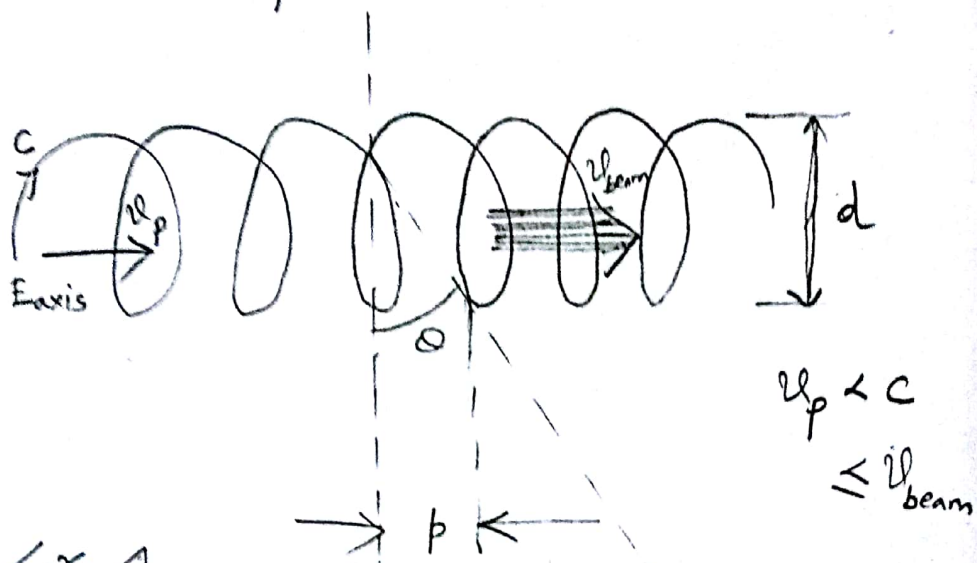
Analysis of TWTA

(1) wave propagation in Helix.

- In TWTA, an axial microwave \vec{E} field component is being created inside a helix slow-wave guide by injecting the input signal into it.
- This field will decelerate the electrons in the axial beam at some sections, and at other sections it will accelerate them resulting in density modulation in the beam in accordance with the input signal.

→ If 'd' is the diameter of the helix and p is the helix pitch, the time taken by the signal along the wire must be equal to the ^{time} taken by the axial wave, so that

$$T = \frac{p \rightarrow \text{height of one complete helix turn measured to the axis}}{v_p} = \frac{\sqrt{(\pi d)^2 + p^2}}{c} \rightarrow \textcircled{1}$$



($2\pi r = \pi d =$ circumference to make a straight line)

$$\textcircled{1} \Rightarrow v_p = \frac{c p}{\sqrt{(\pi d)^2 + p^2}} \approx \frac{c p}{\pi d} \quad ; p \ll \pi d$$

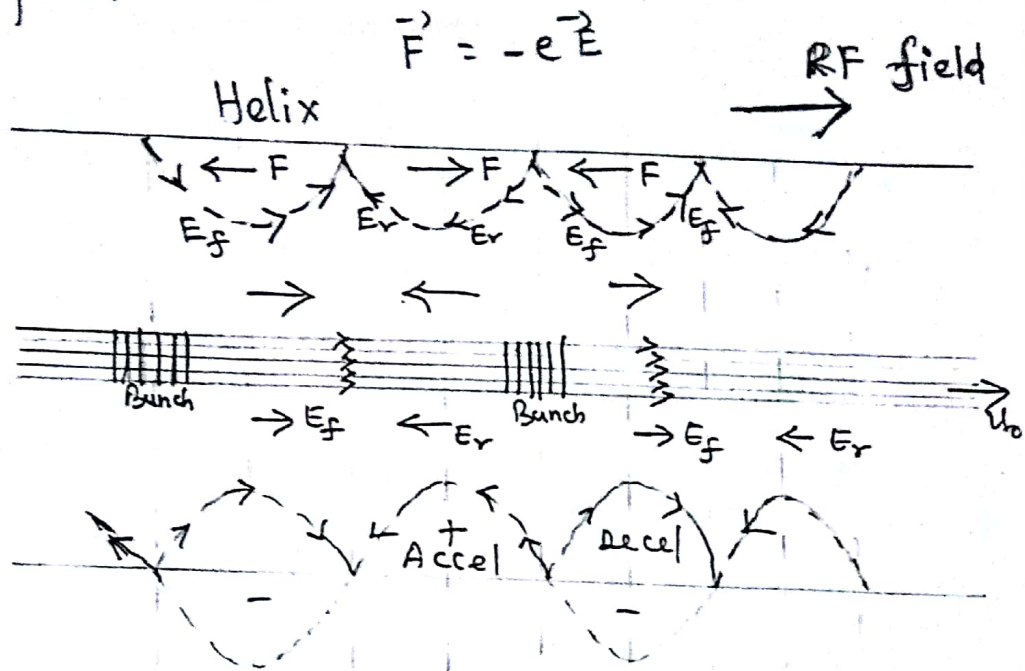
$$= \frac{\omega}{\beta} \rightarrow \textcircled{2} \lambda = \frac{c}{f}$$

$$\frac{2\pi}{\beta} = \frac{c}{f}$$

$$\frac{2\pi f}{\beta} = c$$

Beam and field interaction:

The electromagnetic wave travelling along the helix axis, has a longitudinal component of electric field as shown below.



→ An electric field directed against the electron flow accelerates the electrons and that directed along the electron flow decelerates the electrons.

→ Thus, a density modulation in accordance with the input signal frequency will occur to form bunches.

→ As the dc velocity of the beam is maintained slightly greater than the phase velocity of the travelling wave, more electrons face the retarding field than the accelerating field, and a great amount of kinetic energy is transferred from the beam to the electromagnetic field.

→ Thus the field amplitude increases forming a more compact bunch and a larger amplification of the signal voltage appears at the output of the helix.

(2) Gain characteristics:

The output power gain is

$$\text{defined as, } A_p = 10 \log \left| \frac{\text{output voltage}}{\text{Input voltage}} \right|^2$$

$$= -9.54 + 47.3 \text{ NC dB}$$

↳ (3)

→ The first term -9.54 dB represents a loss due to the fact that the input wave

(48)

divides into three waves of equal magnitude and only one (axial) of these waves is amplified.

→ N is the length of the interaction region in wavelengths,

$$N = \frac{l}{\lambda_e}, \quad l \text{ is the length of the slow-wave structure, } m$$

$$\text{and } \lambda_e = \frac{c}{f}$$

$$= \frac{u_0 \times 2\pi}{f \times 2\pi}$$

$$= \frac{u_0 \cdot 2\pi}{\omega} \quad ; \quad 2\pi f = \omega$$

$$u_0 = \sqrt{\frac{2eV_0}{m}}$$

The factor C is the gain parameter of the circuit defined

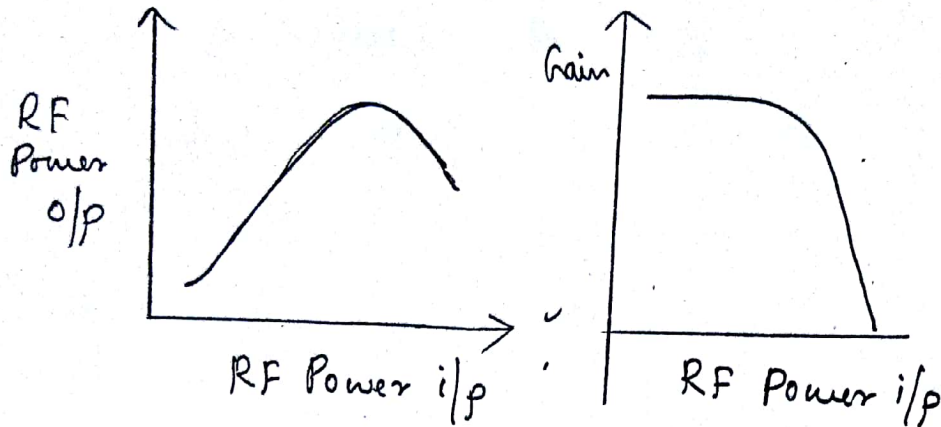
by, $C = \left(\frac{I_0 Z_0}{4V_0} \right)^{1/3}$, where I_0

is the dc beam current, V_0 is the dc beam voltage, and Z_0 is the characteristic impedance of the helix.

The Peak Power o/p of a single helix type tube is limited to about 2kw because of the difficulty

in removing heat due to ohmic loss from the helix conductor.

RF Power o/p and Gain characteristics



→ For low inputs, the small signal gain is almost constant.

→ As the RF Power input is increased, the RF output power does not increase in proportion, but instead attains a maximum and finally starts to decrease.

→ The point at which the o/p Power is maximum is termed as the saturation point, and the gain at this point is the saturation gain.

→ There will be a range of input for which the o/p remains in saturation, and this range defines the overdrive capability of the TWT.

Magnetron oscillator

→ A magnetron oscillator is used to generate high microwave power required in radar and communication systems.

→ Magnetrons are crossed field tubes (M-type) in which the dc magnetic field and the dc electric field are perpendicular to each other.

→ A high power microwave oscillator uses a travelling wave cylindrical magnetron tube as shown in the figs (1) and (2)

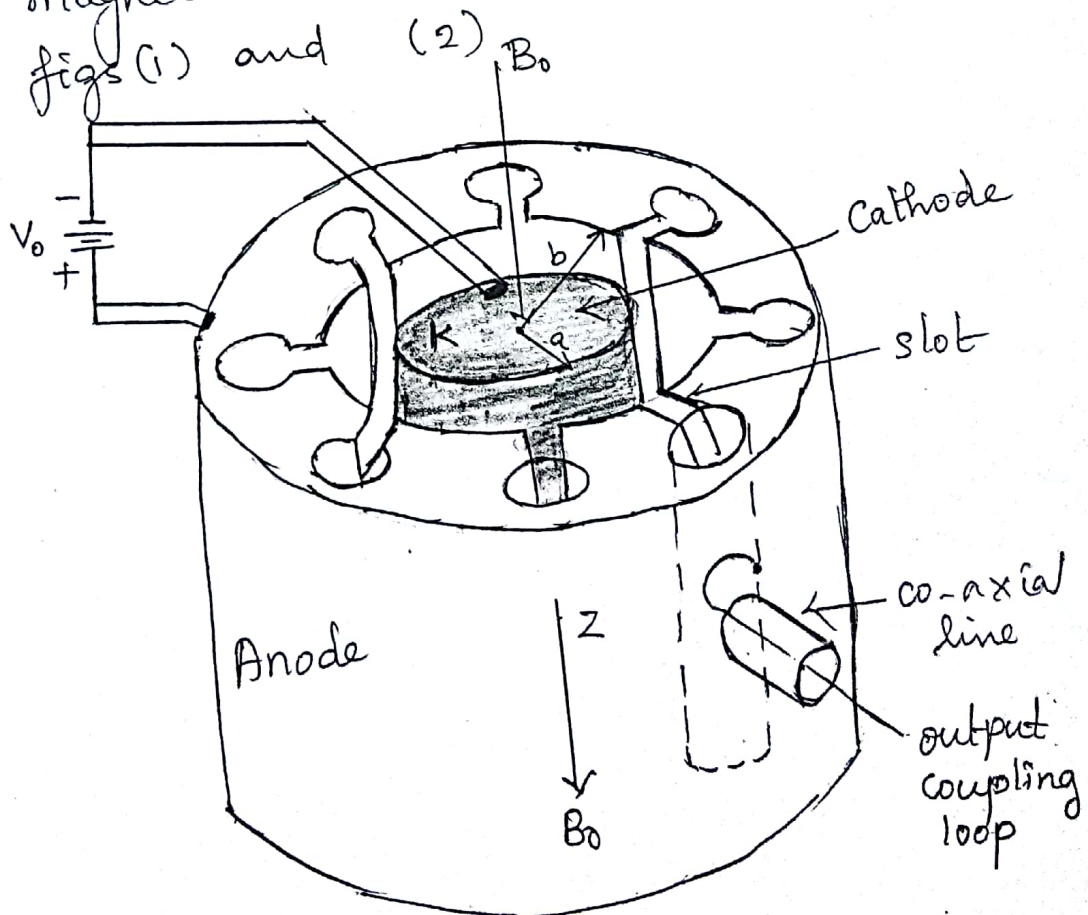


fig (1)

- Magnetron consists of a cylindrical cathode K of finite length, radius 'a' at the centre surrounded by a cylindrical anode A of radius 'b'.
- The anode A is a slow-wave structure consisting of several re-entrant cavities equispaced around the circumference and coupled together through the anode-cathode space by means of slots.
- Radial electric field is established by dc voltage V_0 in between the cathode and the anode.
- An axial dc magnetic flux denoted by B_0 is maintained in the positive z -direction by means of a permanent magnet. or an electromagnet.
- Magnetron theory of operation is based on the motions of electrons under the influence of combined electric and magnetic fields.
- The electrons emitted from the cathode try to travel to the anode. But with the influence of crossed fields \vec{E} and \vec{H} in the space between the anode and the cathode, it

experiences a resultant force,

$$\vec{F} = -e\vec{E} - e(\vec{v} \times \vec{B}),$$

where \vec{v} is the velocity vector of the electron considered and takes a curved trajectory.

→ Due to excitation of the anode cavities by RF noise voltage in the biasing circuit, the RF field lines are fringed out of the cavity slot to the space between the anode and cathode.

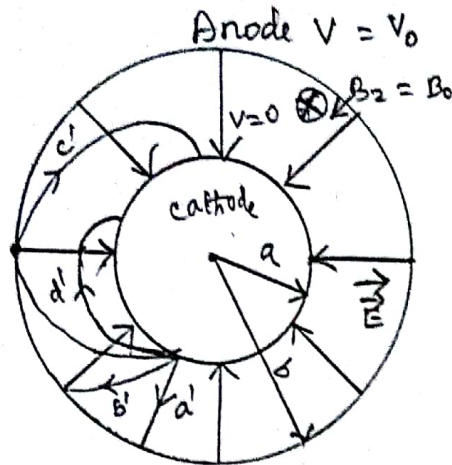
→ The accelerated electrons in the trajectory, when retarded by this RF field, transfer energy from the electron to the cavities to grow RF oscillations.

→ When the system RF losses balance the RF oscillation energy, a stable oscillation is achieved.

→ Output power is extracted through an external line coupled to the cavity.

Equations of Electron trajectory: After emergence from the cathode with zero velocity (say), the electrons

will acquire a velocity \vec{v} having a tangential as well as radial velocity components due to force \vec{F} exerted by the crossed fields \vec{E} and \vec{H} : (ie) $\vec{F} = -e\vec{E} - e(\vec{v} \times \vec{B})$



fig(2) Electron trajectory

Trajectories a' , b' , c' and d' of the electrons are shown for different magnetic field strengths.

- (1) At Zero magnetic field, the e^- s take the straight path a' , by the influence of electric field only and collected by the anode.
- (2) For a given V_0 if the magnetic field is increased, the electrons take the curved path b' due to the above force \vec{F} to reach the anode.

(3) At a critical value of magnetic field B_c , say, the electrons just graze the anode surface at radius 'b' and take the path c' to return to the cathode for a given voltage V_0 . This value B_c is called the cut-off magnetic flux density.

(4) If the magnetic field is greater than B_c , all the electrons return to the cathode as shown by ~~the~~ a typical path d' without reaching the anode.

→ Therefore, when the magnetic field increases from zero to maximum, the anode current decreases from a maximum value to zero.

The average velocity of the electron in the z-direction is constant, and is given by, $v_z = \frac{E_0}{B_0} \rightarrow (1)$

where $E_0 = \frac{V_0}{(b-a)}$ = dc electric field.

B_0 = dc magnetic flux density.

In general, an electron in the trajectory of the radius of curvature 'r' and at velocity v experiences three radial forces:

(1) electric force, $\vec{F}_e = -e\vec{E}$

(2) magnetic force, $\vec{F}_m = -e(\vec{v} \times \vec{B})$

and (3) the centrifugal force, $\frac{mv^2}{r}$, such

that for equilibrium

$$\frac{mv^2}{r} + eE = evB \rightarrow (2)$$

where the electric field is in the radial direction only and is given

by, $E(r) = \frac{-V_0}{r \ln(\frac{b}{a})} \rightarrow (3)$

In the absence of an electric field, the electrons move in a circular path and return to the cathode,

when, $\frac{mv^2}{r} = evB \rightarrow (4)$

$\left[\frac{v}{r} = \frac{e}{m} B = \omega \right]$ called the cyclotron angular frequency.

The equations of motion for electrons in crossed electric and magnetic fields can be given as,

$$\vec{F} = m\vec{a} = m \left(\frac{d\vec{v}}{dt} \right) = -e\vec{E} - e(\vec{v} \times \vec{B}) \rightarrow (5)$$

Since the electron is emitted from the cathode in the direction opposite to \vec{E} , the equation of motion for electrons in cylindrical coordinates, are

$$\frac{dr^2}{dt^2} = r \left(\frac{d\phi}{dt} \right)^2 = \frac{e}{m} E_r - \frac{e}{m} r B_z \frac{d\phi}{dt} \rightarrow \textcircled{7}$$

$$\frac{1}{r} \frac{d}{dt} \left(r^2 \frac{d\phi}{dt} \right) = \frac{e}{m} B_z \frac{dr}{dt} \rightarrow \textcircled{7}$$

where, $\frac{e}{m} = 1.759 \times 10^{11}$ C/kg is the charge-to-mass ratio of the electron and, $B_0 = B_z$ is assumed in the positive z-direction.

$$\textcircled{7} \Rightarrow \frac{d}{dt} \left(r^2 \frac{d\phi}{dt} \right) = \frac{e}{m} B_z r \frac{dr}{dt}$$

$$= \omega_c r \left(\frac{dr}{dt} \right)$$

$$= \frac{1}{2} \omega_c \frac{d}{dt} (r^2) \rightarrow \textcircled{8}$$

where, $\omega_c = \frac{e}{m} B_z =$ cyclotron angular frequency

on integrating eqn $\textcircled{8}$, we get,

$$r^2 \frac{d\phi}{dt} = \frac{1}{2} \omega_c \cdot r^2 + \text{Constant} \rightarrow \textcircled{9}$$

At $r = a$, where a is the radius of the cathode cylinder, and $\frac{d\phi}{dt} = 0$,

$\textcircled{57}$

$$(9) \Rightarrow \frac{1}{2} \omega_c a^2 + \text{Constant} = 0$$

$$\Rightarrow \text{Constant} = -\frac{1}{2} \omega_c a^2$$

$$\therefore (9) \Rightarrow r^2 \frac{d\phi}{dt} = \frac{1}{2} \omega_c r^2 - \frac{1}{2} \omega_c a^2$$

$$\frac{d\phi}{dt} = \frac{1}{2} \omega_c \frac{r^2}{r^2} - \frac{1}{2r^2} \omega_c a^2$$

$$= \frac{1}{2} \omega_c \left(1 - \frac{a^2}{r^2} \right) \rightarrow (10)$$

Since, the magnetic field does no work on the electrons, the kinetic energy of the electron is given by,

$$\frac{1}{2} m v^2 = eV \rightarrow (11)$$

However, the electron velocity has r and ϕ components such as

$$v^2 = \frac{2e}{m} V = v_r^2 + v_\phi^2 = \left(\frac{dr}{dt} \right)^2 + \left(r \frac{d\phi}{dt} \right)^2 \rightarrow (12)$$

At $r = b$, where b is the radius from the center of the cathode to the edge of the anode, $V = V_0$ and

$$\frac{dr}{dt} = 0.$$

when the electrons just graze the anode, eqns (10) & (12) become,

$$\frac{d\phi}{dt} = \frac{1}{2} \omega_c \left(1 - \frac{a^2}{b^2}\right) \rightarrow (13)$$

$$b^2 \left(\frac{d\phi}{dt}\right)^2 = \frac{2e}{m} V_0 \rightarrow (14)$$

Substituting the value of $\frac{d\phi}{dt}$ in (14)

$$(14) \Rightarrow b^2 \left[\frac{1}{2} \omega_c \left(1 - \frac{a^2}{b^2}\right) \right]^2 = \frac{2e}{m} V_0 \rightarrow (15)$$

The electron will acquire a tangential as well as a radial velocity. Whether the electron will just graze the anode and return toward the cathode depends on the relative magnitudes of V_0 and B_0 . The Hull cut-off magnetic equation is obtained from eqn (15) as,

$$b^2 \left[\frac{1}{2} \left(\frac{e}{m} B_{oc}\right)^2 \left(1 - \frac{a^2}{b^2}\right)^2 \right] = \frac{2e}{m} V_0$$

$$b^2 \frac{1}{4} \left(\frac{e}{m}\right)^2 B_{oc}^2 \left(1 - \frac{a^2}{b^2}\right)^2 = \frac{2e}{m} V_0$$

$$B_{oc}^2 = \frac{8 \frac{m}{e} V_0}{b^2 \left(1 - \frac{a^2}{b^2}\right)^2}$$

$$B_{oc} = \frac{\left(8 \frac{m}{e} V_0\right)^{1/2}}{b \left(1 - \frac{a^2}{b^2}\right)} \rightarrow (16)$$

$$\therefore \omega_c = \frac{e B_0}{m}$$

(59)

This means, that if $B_0 > B_{oc}$ for a given V_0 , the electrons will not reach the anode.

Conversely, the cut-off voltage is given by,

$$(16) \Rightarrow V_{oc} = \frac{B_0^2 b^2 \left(1 - \frac{a^2}{b^2}\right)^2}{\left(\frac{8m}{e}\right)}$$

$$V_{oc} = \frac{e}{8m} B_0^2 b^2 \left(1 - \frac{a^2}{b^2}\right)^2 \rightarrow (17)$$

This means, that if $V_0 < V_{oc}$ for a given B_0 , the electrons will not reach the anode.

Equation (17) is often called the Hull cut-off voltage equation.

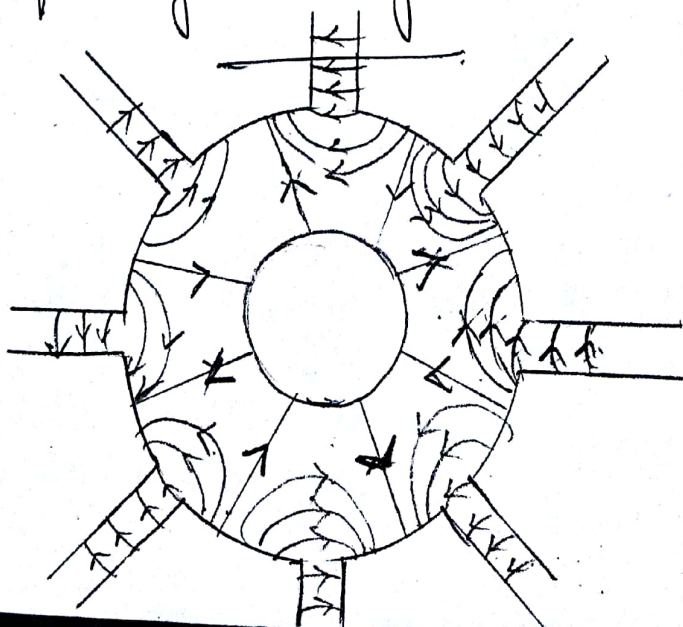
Power o/p and efficiency:

- A magnetron can deliver a peak power output of up to 40 MW with the dc voltage of 50 kV at 10 GHz.
- The average power o/p is of the order of 800 kW.
- The magnetron possesses a very high efficiency ranging from 40 to 70% (6)

→ Magnetrons are commercially available for peak power output from 2 kW and higher.

Applications :

- (1) Magnetrons are widely used on radar transmitters, industrial heating and microwave ovens.
- (2) A microwave oven requires a standard power of 600 W and frequencies of 915 MHz or 2450 MHz.
- (3) For industrial heating, magnetrons generate power of the order of kW ~~at~~ MHz frequencies.
- (4) For radar applications, a magnetron generates peak power of the order of MW in GHz frequency range.

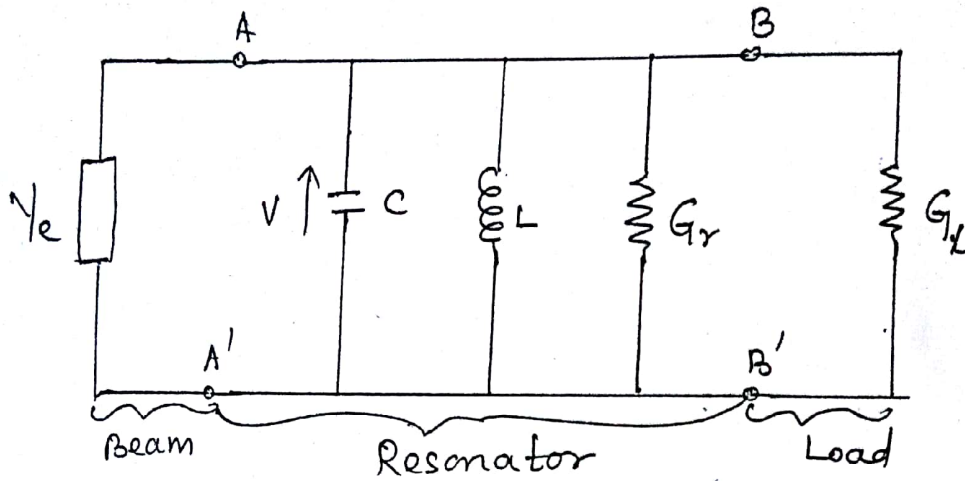


Lines of force
in π mode
of eight-
cavity
magnetron

(61)

Equivalent circuit for one resonator of a Magnetron

Each resonator of the slow-wave structure is taken to comprise a separate resonant circuit as shown in figure below.



where, Y_e = electronic admittance

V = RF voltage across the Vane tips.

C = Capacitance at the vane tips

L = inductance of the resonator

G_r = conductance of the resonator

G_L = load conductance per resonator

unloaded quality factor :

The quality factor Q is a measure of the frequency selectivity of a cavity resonator.

→ when a cavity is assumed to be not connected to any external circuit or load, & accounts for the internal losses and is called the unloaded quality factor Q_{un} .

The unloaded quality factor of the resonator is given by,

$$Q_{un} = \frac{\omega_0 C}{G_r} \rightarrow \textcircled{1} \Rightarrow G_r = \frac{\omega_0 C}{Q_{un}}$$

where, $\omega_0 = 2\pi f_0$ is the angular resonant frequency.

The external quality factor of the load circuit is,

$$Q_{ex} = \frac{\omega_0 C}{G_L} \rightarrow \textcircled{2}$$

Then the loaded Q_L of the resonant circuit is expressed by,

$$Q_L = \frac{\omega_0 C}{G_r + G_L} \rightarrow \textcircled{3}$$

The circuit efficiency is defined as,

$$\eta_c = \frac{G_L}{G_L + G_r} \rightarrow \textcircled{4}$$

Note: In terms of the ckt elements, the Q is given by,

$$Q = \frac{R_L}{\omega_0 L} = R_L \omega_0 C = \frac{\omega_0 C}{G_L}$$

$$= \frac{G_L}{G_{ex}} \rightarrow \textcircled{5}$$

from ②, $G_L = \frac{\omega_0 C}{Q_{ex}}$

from ③, $Q_L = \frac{\omega_0 C}{G_{ex}} ; G_r + G_L = G_{ex}$

$$\Rightarrow G_{ex} = \frac{\omega_0 C}{Q_L}$$

$$\textcircled{4} \Rightarrow \eta_c = \frac{\omega_0 C}{G_{ex}} = \frac{G_L}{G_L \left(1 + \frac{G_r}{G_L}\right)}$$

$$= \frac{1}{1 + \frac{\frac{\omega_0 C}{Q_{un}}}{\frac{\omega_0 C}{Q_{ex}}}}$$

$$\eta_c = \frac{1}{1 + \frac{Q_{ex}}{Q_{un}}} \rightarrow \textcircled{6}$$

The maximum circuit efficiency is obtained when the magnetron is heavily loaded, (ie) for $G_L \gg G_r$.

Heavy loading, however, makes the tube quite sensitive to the load, which is undesirable in some cases.

Therefore, the ratio $\frac{Q_L}{Q_{ex}}$ is often

chosen as a compromise between

the conflicting requirements for high circuit efficiency and frequency stability.

The electronic efficiency is defined as,

$$\eta_e = \frac{P_{gen}}{P_{dc}} = \frac{V_o I_o - P_{lost}}{V_o I_o}$$

where, P_{gen} = RF power induced into the anode circuit

$P_{dc} = V_o I_o$ Power from the dc power supply

V_o = anode voltage

I_o = anode current

P_{lost} = Power lost in the anode circuit.

The RF Power generated by the electrons can be written as,

$$P_{gen} = V_o I_o - P_{lost}$$

EC8701

**ANTENNAS AND MICROWAVE
ENGINEERING**

UNIT IV

PASSIVE AND ACTIVE MICROWAVE DEVICES

1. Microwave Passive components:

- Directional Coupler, Power Divider, Magic Tee, Attenuator, Resonator

2. Principles of Microwave Semiconductor Devices:

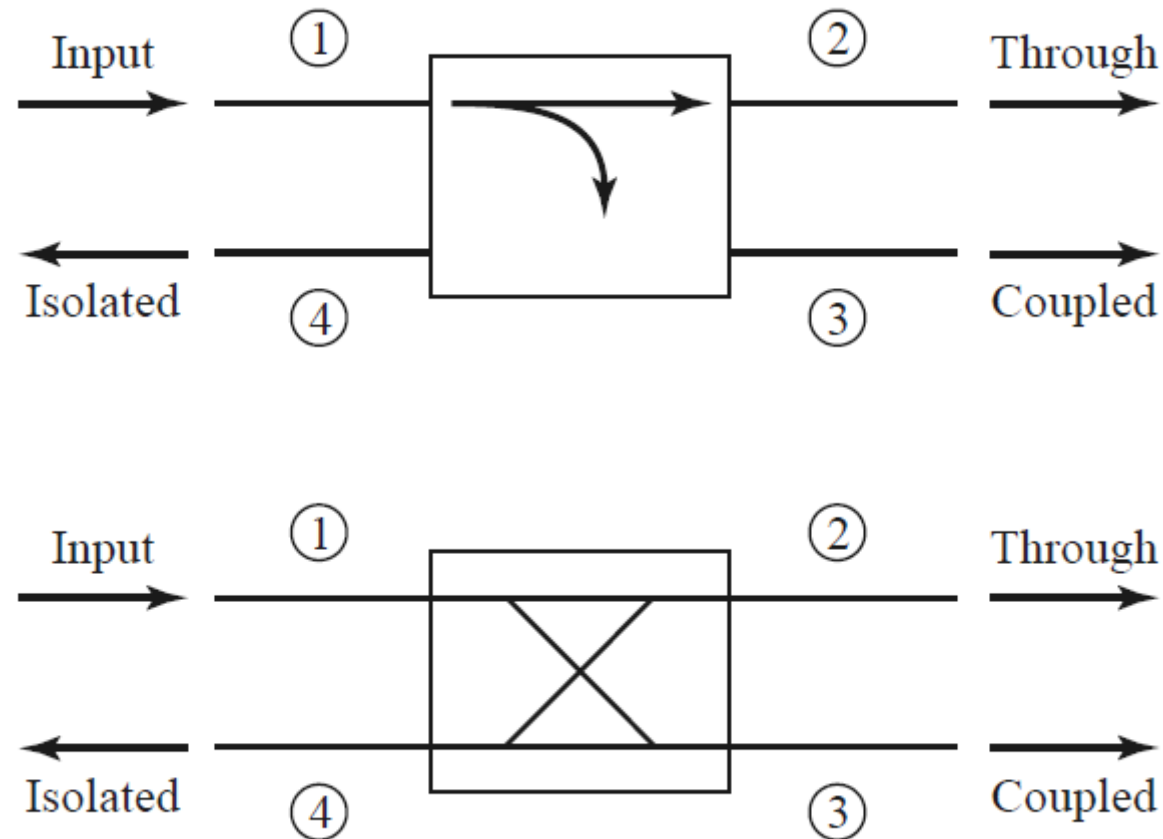
- Gunn Diodes, IMPATT diodes, Schottky Barrier diodes, PIN diodes

3. Microwave tubes:

- Klystron, TWT, Magnetron

Directional Couplers

- A directional coupler is a four-port microwave junction as shown below:



Two commonly used symbols for directional couplers, and power flow conventions

Directional Couplers

- Power supplied to port 1 is coupled to port 3 (the *coupled port*) with the coupling factor $|S_{13}|^2 = \beta^2$, while the remainder of the input power is delivered to port 2 (the *through port*) with the coefficient $|S_{12}|^2 = \alpha^2 = 1 - \beta^2$. In an ideal directional coupler, no power is delivered to port 4 (the *isolated port*).
- The following quantities are commonly used to characterize a directional coupler:

$$\text{Coupling} = C = 10 \log \frac{P_1}{P_3} = -20 \log \beta \text{ dB}$$

$$\text{Directivity} = D = 10 \log \frac{P_3}{P_4} = 20 \log \frac{\beta}{|S_{14}|} \text{ dB}$$

$$\text{Isolation} = I = 10 \log \frac{P_1}{P_4} = -20 \log |S_{14}| \text{ dB}$$

$$\text{Insertion loss} = L = 10 \log \frac{P_1}{P_2} = -20 \log |S_{12}| \text{ dB}$$

Directional Couplers

- The *coupling factor* indicates the fraction of the input power that is coupled to the output port.
- The *directivity* is a measure of the coupler's ability to isolate forward and backward waves (or the coupled and uncoupled ports).
- The *isolation* is a measure of the power delivered to the uncoupled port.
- These quantities are related as

$$I = D + C \text{ dB}$$

- The *insertion loss* accounts for the input power delivered to the through port, diminished by power delivered to the coupled and isolated ports.
- The ideal coupler has infinite directivity and isolation ($S_{14} = 0$). Then both α and β can be determined from the coupling factor, C .

Scattering Matrix of Directional Couplers

- The scattering matrix of a four-port network has the following form:

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{bmatrix}$$

- The scattering matrix of a reciprocal four-port network matched at all ports has the following form:

$$[S] = \begin{bmatrix} 0 & S_{12} & S_{13} & S_{14} \\ S_{12} & 0 & S_{23} & S_{24} \\ S_{13} & S_{23} & 0 & S_{34} \\ S_{14} & S_{24} & S_{34} & 0 \end{bmatrix}$$

where $S_{11} = S_{22} = S_{33} = S_{44} = 0$ and

$$S_{21} = S_{12}; S_{31} = S_{13}; S_{41} = S_{14}; S_{32} = S_{23}; S_{42} = S_{24}; S_{43} = S_{34}$$

Scattering Matrix of Directional Couplers

Unitary Property of S-matrix:

$$\sum_{k=1}^N S_{ki} S_{kj}^* = \delta_{ij} \text{ for all } i, j \text{ where } \begin{cases} \delta_{ij} = 1 \text{ if } i = j \\ \delta_{ij} = 0 \text{ if } i \neq j \end{cases}$$

- The dot product of any column of $[S]$ with the conjugate of that same column gives unity. For $i = j$,

$$\sum_{k=1}^N S_{ki} S_{ki}^* = 1$$

- The dot product of any column of $[S]$ with the conjugate of a different column gives zero (the columns are orthonormal).

$$\sum_{k=1}^N S_{ki} S_{kj}^* = 0 \text{ for } i \neq j$$

Scattering Matrix of Directional Couplers

$$[S] = \begin{bmatrix} 0 & S_{12} & S_{13} & S_{14} \\ S_{12} & 0 & S_{23} & S_{24} \\ S_{13} & S_{23} & 0 & S_{34} \\ S_{14} & S_{24} & S_{34} & 0 \end{bmatrix}$$

- If the network is lossless, 10 equations result from the unitary, or energy conservation, condition. Consider the multiplication of row 1 and row 2, and the multiplication of row 4 and row 3:

$$S_{13}^* S_{23} + S_{14}^* S_{24} = 0 \quad (1)$$

$$S_{14}^* S_{13} + S_{24}^* S_{23} = 0 \quad (2)$$

- Multiply (1) by S_{24}^* , and (2) by S_{13}^* ,

$$S_{13}^* S_{23} S_{24}^* + S_{14}^* S_{24} S_{24}^* = 0 \implies S_{13}^* S_{23} S_{24}^* + S_{14}^* |S_{24}|^2 = 0 \quad (3)$$

$$S_{14}^* S_{13} S_{13}^* + S_{24}^* S_{23} S_{13}^* = 0 \implies S_{14}^* |S_{13}|^2 + S_{24}^* S_{23} S_{13}^* = 0 \quad (4)$$

Scattering Matrix of Directional Couplers

$$S_{13}^* S_{23} S_{24}^* + S_{14}^* S_{24} S_{24}^* = 0 \implies S_{13}^* S_{23} S_{24}^* + S_{14}^* |S_{24}|^2 = 0 \quad (3)$$

$$S_{14}^* S_{13} S_{13}^* + S_{24}^* S_{23} S_{13}^* = 0 \implies S_{14}^* |S_{13}|^2 + S_{24}^* S_{23} S_{13}^* = 0 \quad (4)$$

- Subtract (3) from (4),

$$S_{14}^* (|S_{13}|^2 - |S_{24}|^2) = 0 \quad (5)$$

- Similarly, the multiplication of row 1 and row 3, and the multiplication of row 4 and row 2, gives

$$S_{12}^* S_{23} + S_{14}^* S_{34} = 0 \quad (6)$$

$$S_{14}^* S_{12} + S_{34}^* S_{23} = 0 \quad (7)$$

- Multiply (6) by S_{12} , and (7) by S_{34} ,

$$S_{12} S_{12}^* S_{23} + S_{12} S_{14}^* S_{34} = 0 \implies |S_{12}|^2 S_{23} + S_{12} S_{14}^* S_{34} = 0 \quad (8)$$

$$S_{34} S_{14}^* S_{12} + S_{34} S_{34}^* S_{23} = 0 \implies S_{34} S_{14}^* S_{12} + |S_{34}|^2 S_{23} = 0 \quad (9)$$

Scattering Matrix of Directional Couplers

$$S_{12}S_{12}^*S_{23} + S_{12}S_{14}^*S_{34} = 0 \implies |S_{12}|^2S_{23} + S_{12}S_{14}^*S_{34} = 0 \quad (8)$$

$$S_{34}S_{14}^*S_{12} + S_{34}S_{34}^*S_{23} = 0 \implies S_{34}S_{14}^*S_{12} + |S_{34}|^2S_{23} = 0 \quad (9)$$

- Subtract (9) from (8),

$$S_{23}(|S_{12}|^2 - |S_{34}|^2) = 0 \quad (10)$$

- One way for (5) and (10) to be satisfied is if $S_{14} = S_{23} = 0$, which results in a directional coupler.

$$[S] = \begin{bmatrix} 0 & S_{12} & S_{13} & 0 \\ S_{12} & 0 & 0 & S_{24} \\ S_{13} & 0 & 0 & S_{34} \\ 0 & S_{24} & S_{34} & 0 \end{bmatrix}$$

Scattering Matrix of Directional Couplers

$$[S] = \begin{bmatrix} 0 & S_{12} & S_{13} & 0 \\ S_{12} & 0 & 0 & S_{24} \\ S_{13} & 0 & 0 & S_{34} \\ 0 & S_{24} & S_{34} & 0 \end{bmatrix}$$

- Then the self-products of the rows of the unitary scattering matrix yield the following equations:

$$|S_{12}|^2 + |S_{13}|^2 = 1 \quad (11)$$

$$|S_{12}|^2 + |S_{24}|^2 = 1 \quad (12)$$

$$|S_{13}|^2 + |S_{34}|^2 = 1 \quad (13)$$

$$|S_{24}|^2 + |S_{34}|^2 = 1 \quad (14)$$

$$(11) - (12) \Rightarrow |S_{13}| = |S_{24}|$$

$$(12) - (14) \Rightarrow |S_{12}| = |S_{34}|$$

Scattering Matrix of Directional Couplers

- Further simplification can be made by choosing the phase references on three of the four ports. Thus, we choose $S_{12} = S_{34} = \alpha$, $S_{13} = \beta e^{j\theta}$, and $S_{24} = \beta e^{j\phi}$, where α and β are real, and θ and ϕ are phase constants to be determined (one of which we are still free to choose). The dot product of rows 2 and 3 gives

$$S_{12}^* S_{13} + S_{24}^* S_{34} = 0$$

which yields a relation between the remaining phase constants as

$$\theta + \phi = \pi \pm 2n\pi$$

If we ignore integer multiples of 2π , there are two particular choices that commonly occur in practice:

1. Symmetric Coupler: $\theta = \phi = \pi/2$
2. Antisymmetric Coupler: $\theta = 0, \phi = \pi$

Scattering Matrix of Directional Couplers

1. Symmetric Coupler: $\theta = \phi = \pi/2$,

$$[S] = \begin{bmatrix} 0 & \alpha & j\beta & 0 \\ \alpha & 0 & 0 & j\beta \\ j\beta & 0 & 0 & \alpha \\ 0 & j\beta & \alpha & 0 \end{bmatrix}$$

2. Antisymmetric Coupler: $\theta = 0, \phi = \pi$,

$$[S] = \begin{bmatrix} 0 & \alpha & \beta & 0 \\ \alpha & 0 & 0 & -\beta \\ \beta & 0 & 0 & \alpha \\ 0 & -\beta & \alpha & 0 \end{bmatrix}$$

Note that these two couplers differ only in the choice of reference planes. In addition, the amplitudes α and β are not independent, as (11) requires that

$$\alpha^2 + \beta^2 = 1$$

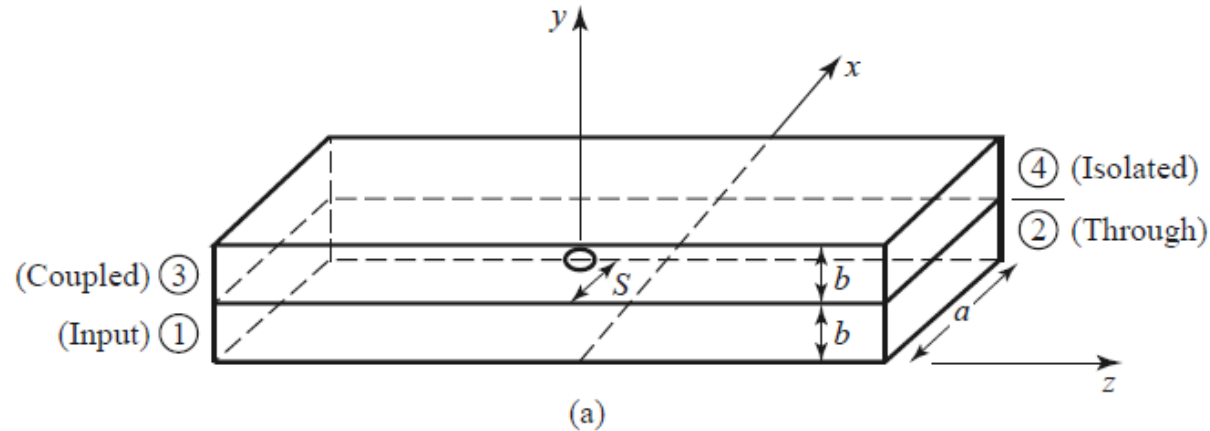
Types of Directional Couplers

1. Bethe hole directional coupler:

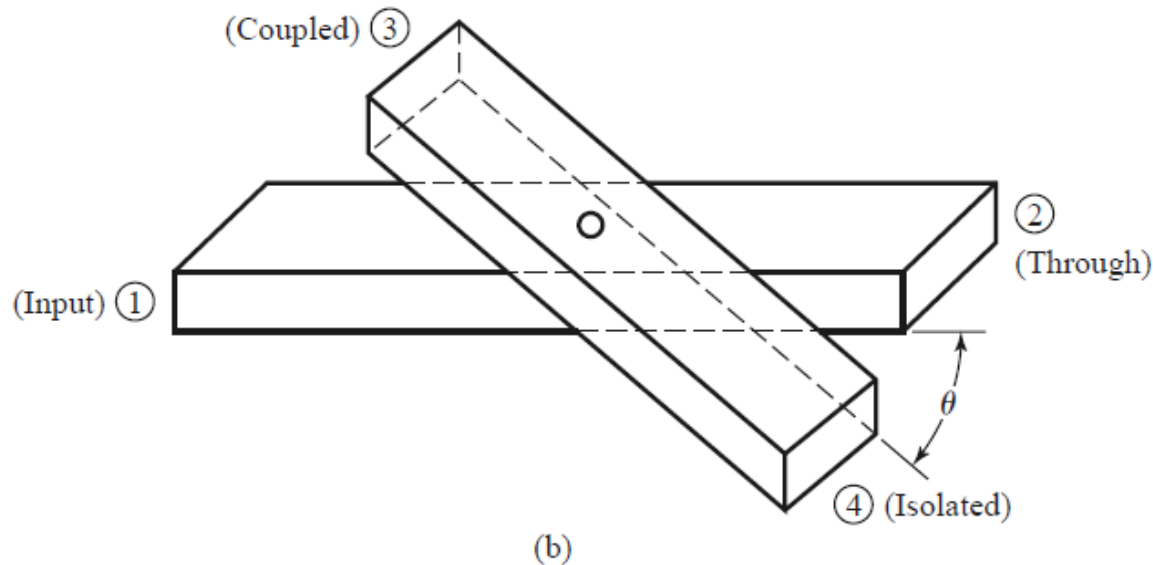
- The directional property of all directional couplers is produced through the use of two separate waves or wave components, which add in phase at the coupled port and are canceled at the isolated port.
- One of the simplest ways of doing this is to couple one waveguide to another through a single small hole in the common broad wall between the two waveguides. Such a coupler is known as a *Bethe hole coupler*, two versions of which are shown in the following figures.
- From the small-aperture coupling theory, an aperture can be replaced with equivalent sources consisting of electric and magnetic dipole moments.
- An incident TE_{10} mode in guide 1, with an amplitude A , produces a normal electric dipole in the aperture plus a tangential magnetic dipole proportional and in the same direction as the magnetic field of the incident wave.

Types of Directional Couplers

Two versions of the Bethe hole directional coupler:



(a) Parallel waveguides



(b) Skewed waveguides

Types of Directional Couplers

1. Bethe hole directional coupler:

- In the upper guide, the normal electric dipole and the axial component of the magnetic dipole radiate symmetrically in both directions.
- The transverse component of the magnetic dipole radiates antisymmetrically.
- The normal electric dipole moment and the axial magnetic dipole moment radiate with even symmetry in the coupled guide, while the transverse magnetic dipole moment radiates with odd symmetry.
- Thus, by adjusting the relative amplitudes of these two equivalent sources, we can cancel the radiation in the direction of the isolated port, while enhancing the radiation in the direction of the coupled port.
- Above figure shows two ways in which these wave amplitudes can be controlled; in the coupler shown in figure(a), the two waveguides are parallel and the coupling is controlled by s , the aperture offset from the sidewall of the waveguide.

Types of Directional Couplers

1. Bethe hole directional coupler:

- For the coupler of figure(b), the wave amplitudes are controlled by the angle, θ , between the two waveguides.
- First consider the configuration of figure (a), with an incident TE_{10} mode into port 1. These fields can be written as

$$E_y = A \sin \frac{\pi x}{a} e^{-j\beta z}$$
$$H_x = \frac{-A}{Z_{10}} \sin \frac{\pi x}{a} e^{-j\beta z}$$
$$H_z = \frac{j\pi A}{\beta a Z_{10}} \cos \frac{\pi x}{a} e^{-j\beta z}$$

where $Z_{10} = k_0 \eta_0 / \beta$ is the wave impedance of the TE_{10} mode .

Types of Directional Couplers

1. Bethe hole directional coupler:

- The amplitudes of the forward and reverse traveling waves in the top guide are

$$A_{10}^+ = \frac{-j\omega A}{P_{10}} \left[\epsilon_0 \alpha_e \sin^2 \frac{\pi S}{a} - \frac{\mu_0 \alpha_m}{Z_{10}^2} \left(\sin^2 \frac{\pi S}{a} + \frac{\pi^2}{\beta^2 a^2} \cos^2 \frac{\pi S}{a} \right) \right]$$
$$A_{10}^- = \frac{-j\omega A}{P_{10}} \left[\epsilon_0 \alpha_e \sin^2 \frac{\pi S}{a} + \frac{\mu_0 \alpha_m}{Z_{10}^2} \left(\sin^2 \frac{\pi S}{a} - \frac{\pi^2}{\beta^2 a^2} \cos^2 \frac{\pi S}{a} \right) \right]$$

where $P_{10} = ab/Z_{10}$ is the power normalization constant.

- Note from above expressions that the amplitude of the wave excited toward port 4 A_{10}^+ is generally different from that excited toward port 3 A_{10}^- , so we can cancel the power delivered to port 4 by setting $A_{10}^+ = 0$.

Types of Directional Couplers

1. Bethe hole directional coupler:

- If we assume that the aperture is round, then the polarizabilities are $\alpha_e = 2r_0^3/3$ and $\alpha_m = 4r_0^3/3$, where r_0 is the radius of the aperture. Then we obtain the following condition $A_{10}^+ = 0$.

$$\sin \frac{\pi s}{a} = \frac{\lambda_0}{\sqrt{2(\lambda_0^2 - a^2)}}$$

- The coupling factor is then given by

$$C = 20 \log \left| \frac{A}{A_{10}^-} \right| \text{ dB}$$

and the directivity by

$$D = 20 \log \left| \frac{A_{10}^-}{A_{10}^+} \right| \text{ dB}$$

Types of Directional Couplers

1. Bethe hole directional coupler:

- For the skewed geometry of figure(b), the aperture may be centered at $s = a/2$, and the skew angle θ adjusted for cancellation at port 4. In this case, the normal electric field does not change with θ , but the transverse magnetic field components are reduced by $\cos \theta$. We can account for the skew angle by replacing α_m in the previous derivation by $\alpha_m \cos \theta$. The wave amplitudes, then become, for $s = a/2$,

$$A_{10}^+ = \frac{-j\omega A}{P_{10}} \left[\epsilon_0 \alpha_e - \frac{\mu_0 \alpha_m}{Z_{10}^2} \cos \theta \right]$$
$$A_{10}^- = \frac{-j\omega A}{P_{10}} \left[\epsilon_0 \alpha_e + \frac{\mu_0 \alpha_m}{Z_{10}^2} \cos \theta \right]$$

Types of Directional Couplers

1. Bethe hole directional coupler:

- Setting $A_{10}^+ = 0$ results in the following condition for the angle θ ,

$$\cos \theta = \frac{k_0^2}{2\beta^2}$$

- The coupling factor then simplifies to

$$C = 20 \log \left| \frac{A}{A_{10}^-} \right| = -20 \log \frac{4k_0^2 r_0^3}{3ab\beta} \text{ dB}$$

- The angular geometry of the skewed Bethe hole coupler is often a disadvantage in terms of fabrication and application. In addition, both coupler designs operate properly only at the design frequency; deviation from this frequency will alter the coupling level and the directivity

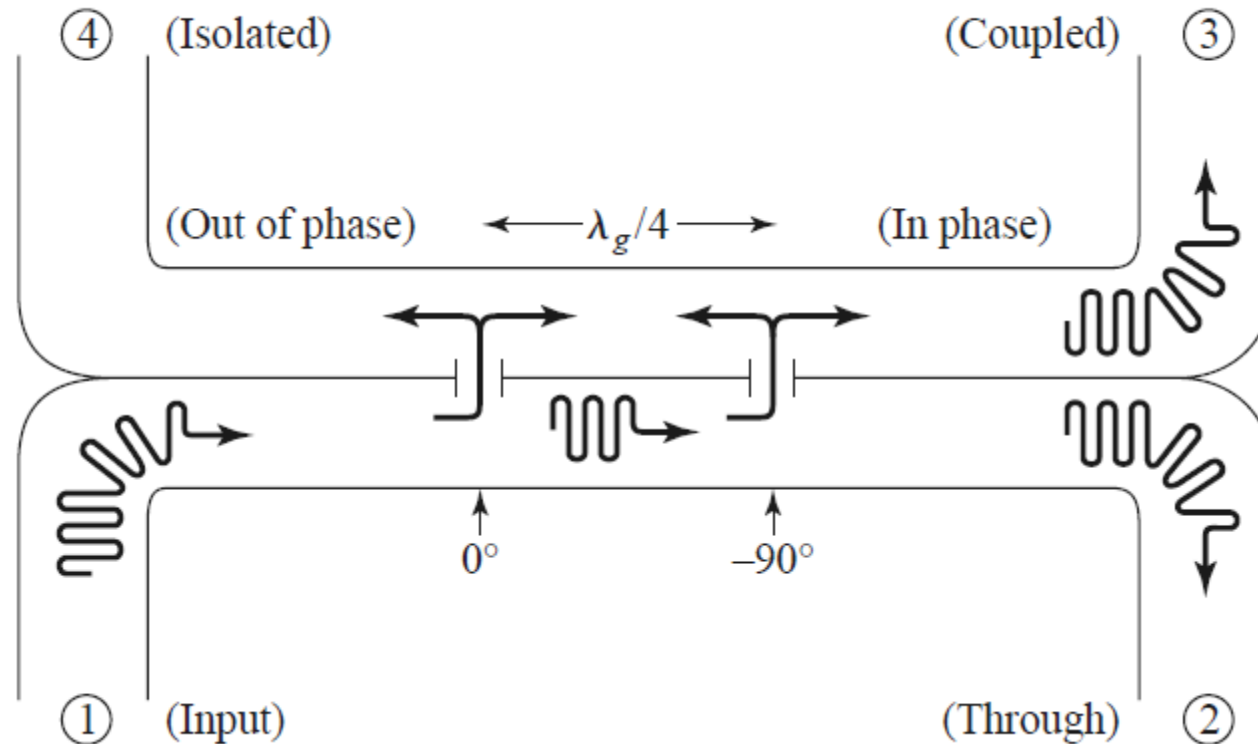
Types of Directional Couplers

2. Multihole Couplers:

- A single-hole coupler has a relatively narrow bandwidth, at least in terms of its directivity. However, if the coupler is designed with a series of coupling holes, the extra degrees of freedom can be used to increase this bandwidth.
- First let us consider the operation of the *two-hole coupler* shown in figure. Two parallel waveguides sharing a common broad wall are shown, although the same type of structure could be made in microstrip line or stripline form.
- Two small apertures are spaced $\lambda_g/4$ apart and couple the two guides. A wave entering at port 1 is mostly transmitted through to port 2, but some power is coupled through the two apertures.
- If a phase reference is taken at the first aperture, then the phase of the wave incident at the second aperture will be -90° .

Types of Directional Couplers

2. Multihole Couplers:



Basic operation of a two-hole directional coupler

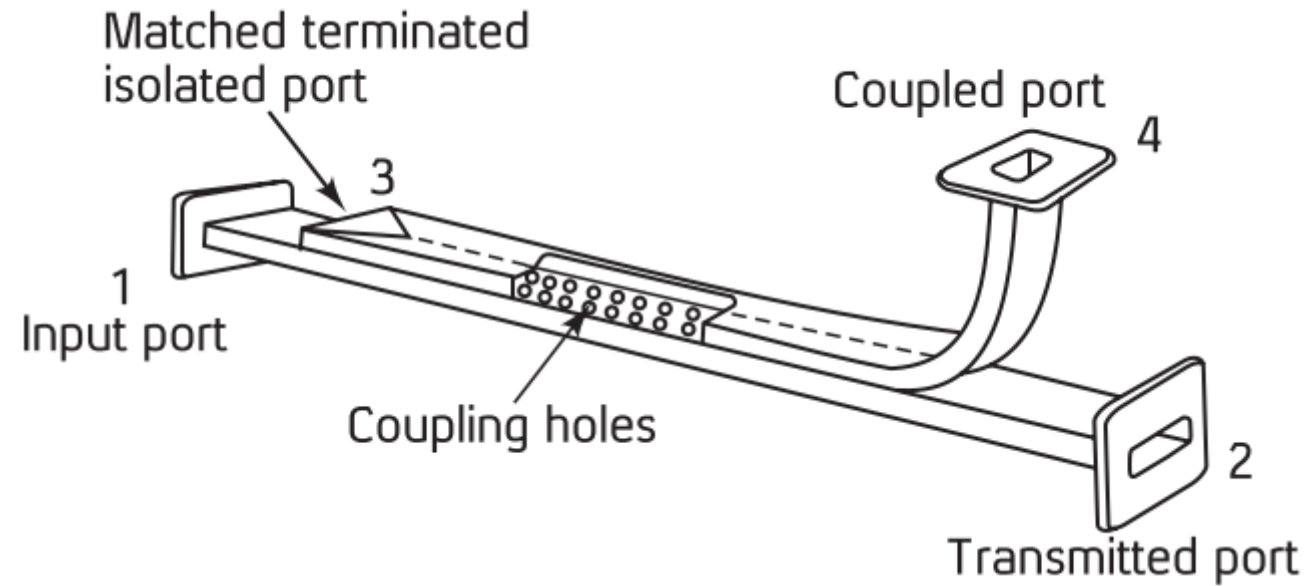
Types of Directional Couplers

2. Multihole Couplers:

- Each aperture will radiate a forward wave component and a backward wave component into the upper guide; in general, the forward and backward amplitudes are different.
- In the direction of port 3, both wave components are in phase because both have traveled $\lambda_g/4$ to the second aperture.
- However, we obtain a cancellation in the direction of port 4 because the wave coming through the second aperture travels $\lambda_g/2$ further than the wave component coming through the first aperture.
- Clearly, this cancellation is frequency sensitive, making the directivity a sensitive function of frequency.
- The coupling is less frequency dependent since the path lengths from port 1 to port 3 are always the same.

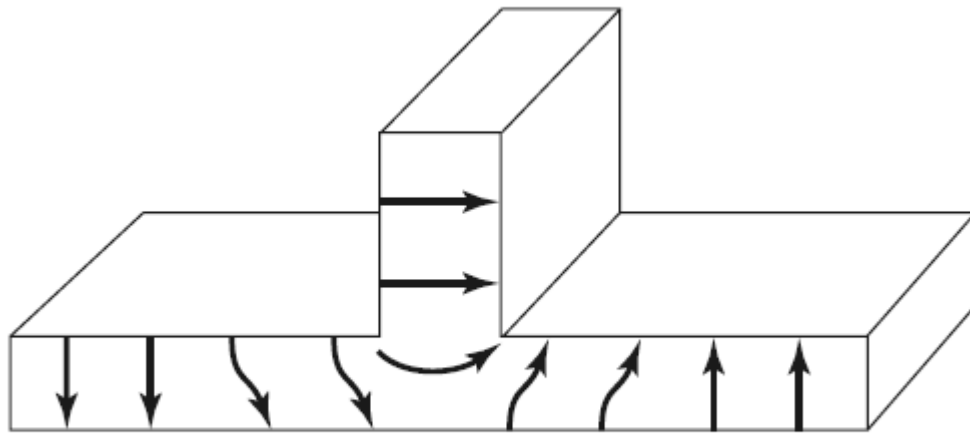
Types of Directional Couplers

2. Multihole Couplers:



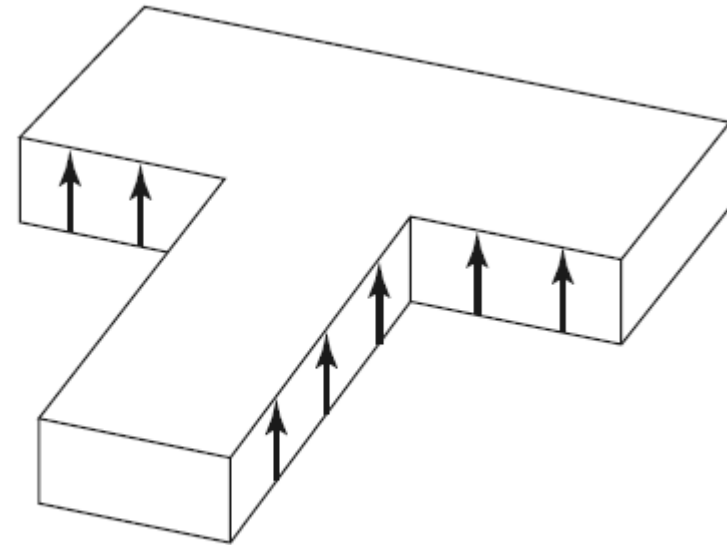
T-Junction Power Divider/ Waveguide Tees

- The T-junction power divider is a simple three-port network that can be used for power division or power combining, and it can be implemented in virtually any type of transmission line medium.



(a)

(a) E-plane waveguide T



(b)

(b) H-plane waveguide T

Scattering Matrix of Three-Port Networks (T-Junctions)

- The simplest type of power divider is a T-junction, which is a three-port network with two inputs and one output. The scattering matrix of an arbitrary three-port network has nine independent elements:

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}$$

- If the device is passive and contains no anisotropic materials, then it must be reciprocal and its scattering matrix will be symmetric ($S_{ij} = S_{ji}$).
- Usually, to avoid power loss, we would like to have a junction that is lossless and matched at all ports. We can easily show, however, that *it is impossible to construct such a three-port lossless reciprocal network that is matched at all ports.*

Scattering Matrix of Three-Port Networks (T-Junctions)

- If all ports are matched, then $S_{ii} = 0$, and if the network is reciprocal, the scattering matrix reduces to

$$[S] = \begin{bmatrix} 0 & S_{12} & S_{13} \\ S_{12} & 0 & S_{23} \\ S_{13} & S_{23} & 0 \end{bmatrix}$$

- If the network is also lossless, then energy conservation requires that the scattering matrix satisfy the unitary properties, which leads to the following conditions:

$$|S_{12}|^2 + |S_{13}|^2 = 1 \quad (1)$$

$$|S_{12}|^2 + |S_{23}|^2 = 1 \quad (2)$$

$$|S_{13}|^2 + |S_{23}|^2 = 1 \quad (3)$$

$$S_{13}^* S_{23} = 0 \quad (4)$$

$$S_{23}^* S_{12} = 0 \quad (5)$$

$$S_{12}^* S_{13} = 0 \quad (6)$$

Scattering Matrix of Three-Port Networks (T-Junctions)

$$|S_{12}|^2 + |S_{13}|^2 = 1 \quad (1)$$

$$|S_{12}|^2 + |S_{23}|^2 = 1 \quad (2)$$

$$|S_{13}|^2 + |S_{23}|^2 = 1 \quad (3)$$

$$S_{13}^* S_{23} = 0 \quad (4)$$

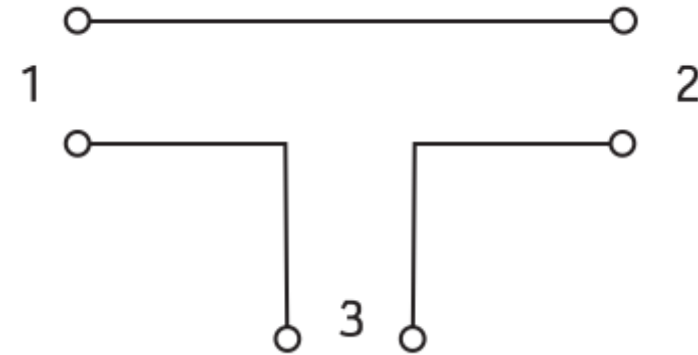
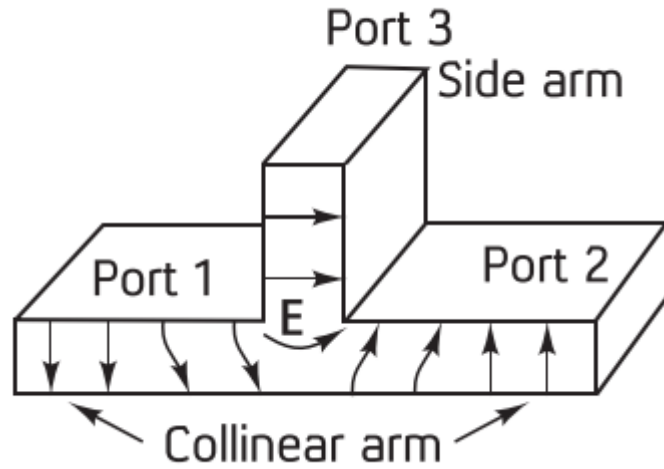
$$S_{23}^* S_{12} = 0 \quad (5)$$

$$S_{12}^* S_{13} = 0 \quad (6)$$

- Equations (4)-(5) show that at least two of the three parameters (S_{12} , S_{13} , S_{23}) must be zero.
- However, this condition will always be inconsistent with one of equations (1)–(3), implying that *a three-port network cannot be simultaneously lossless, reciprocal, and matched at all ports.*
- If any one of these three conditions is relaxed, then a physically realizable device is possible.

E-Plane Tee

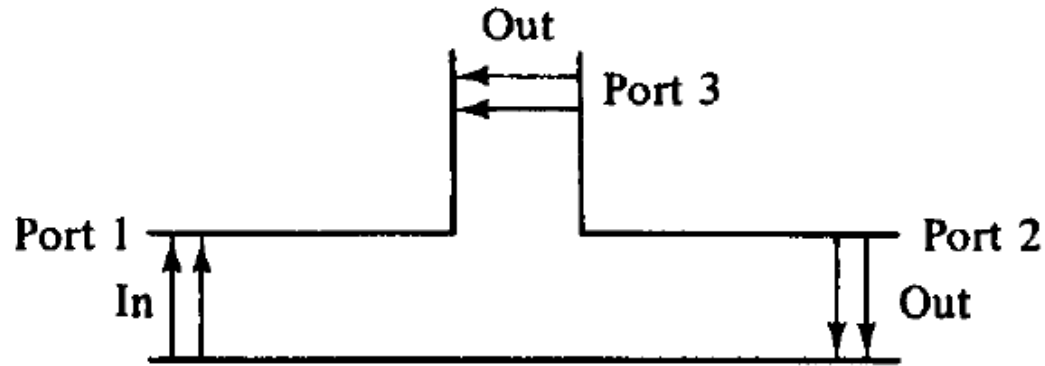
- An E-plane tee is a waveguide tee in which the axis of its side arm is parallel to the E field of the main guide.



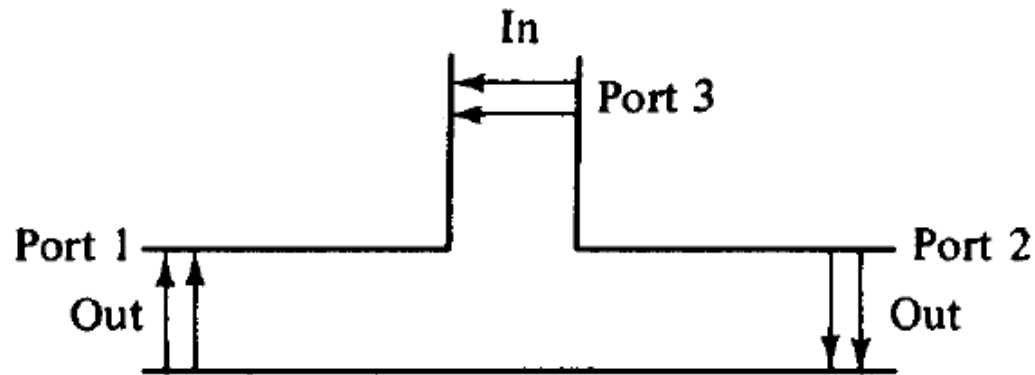
E-plane Tee or Series-Tee

- If the collinear arms are symmetric about the side arm, there are two different transmission characteristics as shown in the following figure.

E-Plane Tee



(a) Input through main arm



(b) Input from side arm

Two way transmission of E-plane Tee

E-Plane Tee

- When waves are fed into the side arm (port 3), the waves appearing at port 1 and port 2 of the collinear arm will be in opposite phase and in the same magnitude. Therefore,

$$S_{13} = -S_{23}$$

- In general, when an E-plane tee is constructed of an empty waveguide, it is poorly matched at the tee junction. Hence $S_{ij} \neq 0$ if $i = j$. However, since the collinear arm is usually symmetric about the side arm, $|S_{13}| = |S_{23}|$ and $S_{22} = S_{11}$. Hence the S-matrix is simplified as

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{11} & -S_{13} \\ S_{13} & -S_{13} & S_{33} \end{bmatrix}$$

E-Plane Tee

- If two in-phase waves are fed into port 1 and 2 of the collinear arm, the output waves at port 3 will be opposite in phase and subtractive. Sometimes, this third port is called the *difference arm*.
- By analogy with the voltage relationship in a series circuit, E-plane tee is also called *Series-T*.
- By suitable matching elements, we can make $S_{33} = 0$ so that $S_{13} = -S_{23}$

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{11} & -S_{13} \\ S_{13} & -S_{13} & 0 \end{bmatrix}$$

E-Plane Tee

- Based on power consideration,

$$S_{13} = \frac{1}{\sqrt{2}} \Rightarrow S_{23} = -\frac{1}{\sqrt{2}}$$

- It can be shown that

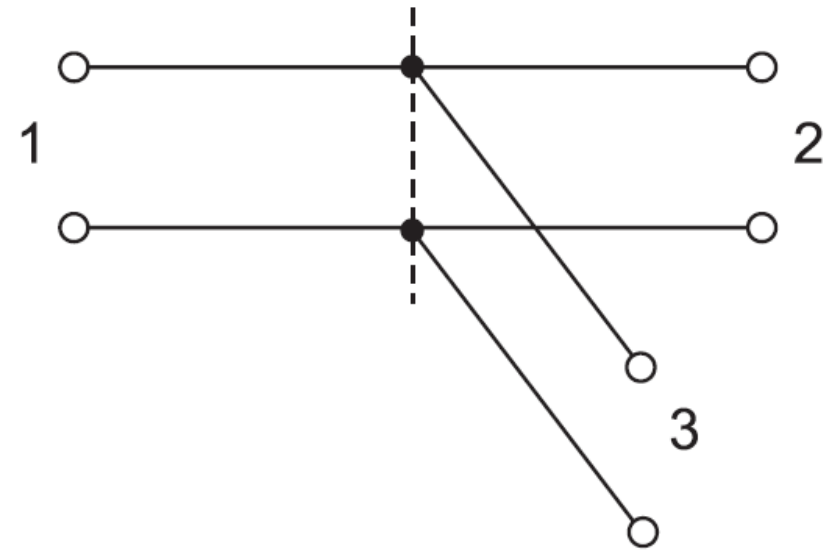
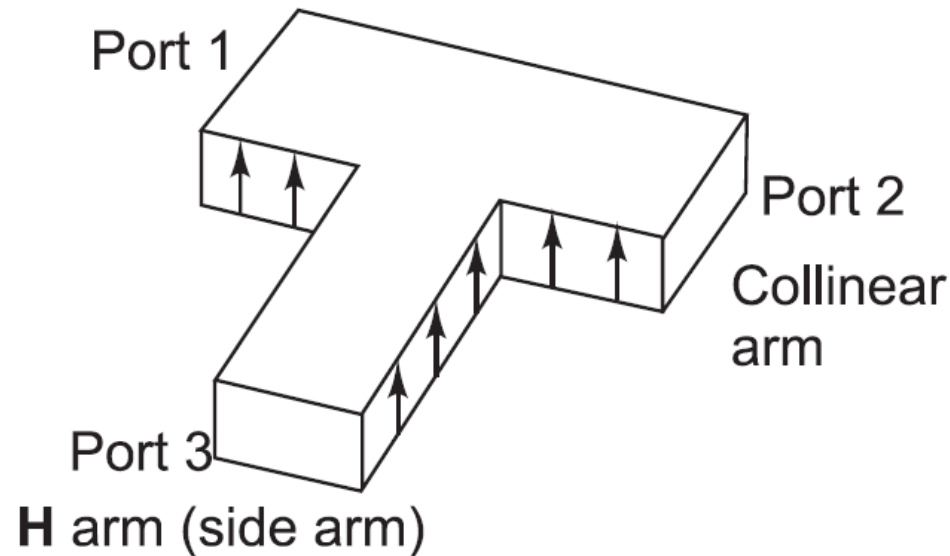
$$S_{11} = S_{22} = \frac{1}{2} \text{ and } S_{12} = S_{21} = \frac{1}{2}$$

- S-matrix of E-plane Tee is

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{11} & -S_{13} \\ S_{13} & -S_{13} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & 1 & \sqrt{2} \\ 1 & 1 & -\sqrt{2} \\ \sqrt{2} & -\sqrt{2} & 0 \end{bmatrix}$$

H-Plane Tee

- In an H-plane tee, if two in-phase input waves are fed into ports 1 and 2 of the collinear arm, the output waves at Port 3 will be in-phase and additive. Because of this, the third port is called the *sum arm*.



H-Tee or Shunt-T

H-Plane Tee

- Conversely, an input wave at Port 3 will be equally divided into ports 1 and 2 in phase. Because the magnetic field loops get divided into two arms 1 and 2 in a manner similar to currents between branches in the parallel circuit, an H-plane junction is also called a *shunt junction*.
- For a symmetrical and lossless junction, in absence of non-linear elements at the H-plane junction, the S-parameters are obtained in a similar manner as in the case of E-plane junction: Since it is a three-port junction, the scattering matrix can be derived as follows:

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix}$$

H-Plane Tee

- Because of plane of symmetry of the junction, the Scattering coefficients are

$$S_{23} = S_{13}$$

- If Port 3 is perfectly matched to the junction, $S_{33} = 0$

- For symmetric property, $S_{ij} = S_{ji}$

$$S_{21} = S_{12}; \quad S_{31} = S_{13}; \quad S_{32} = S_{23} = S_{13}$$

- With the above properties, $[S]$ becomes,

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{13} \\ S_{13} & S_{13} & 0 \end{bmatrix}$$

- For the symmetry and lossless properties,

$$[S][S]^{*t} = [U]$$

H-Plane Tee

$$|S_{11}|^2 + |S_{12}|^2 + |S_{13}|^2 = 1 \quad (1)$$

$$|S_{12}|^2 + |S_{22}|^2 + |S_{13}|^2 = 1 \quad (2)$$

$$|S_{13}|^2 + |S_{13}|^2 = 1 \quad (3)$$

$$S_{13}S_{11}^* + S_{13}S_{12}^* = 0 \quad (4)$$

- From first two equations, we get $S_{22} = S_{11}$
- From third equation, we get $S_{13} = \frac{1}{\sqrt{2}}$
- From last equation, we get

$$S_{13}(S_{11}^* + S_{12}^*) = 0$$

$$\Rightarrow S_{13} = 0 \text{ or } (S_{11}^* + S_{12}^*) = 0$$

$$S_{13} \neq 0 \Rightarrow (S_{11}^* + S_{12}^*) = 0 \Rightarrow S_{12} = -S_{11}$$

- Substituting these values in the first equation, $S_{11} = \frac{1}{2} = S_{22} \Rightarrow S_{12} = -\frac{1}{2}$

H-Plane Tee

$$S_{11} = \frac{1}{2} = S_{22} \Rightarrow S_{12} = -\frac{1}{2}$$

$$S_{13} = \frac{1}{\sqrt{2}}$$

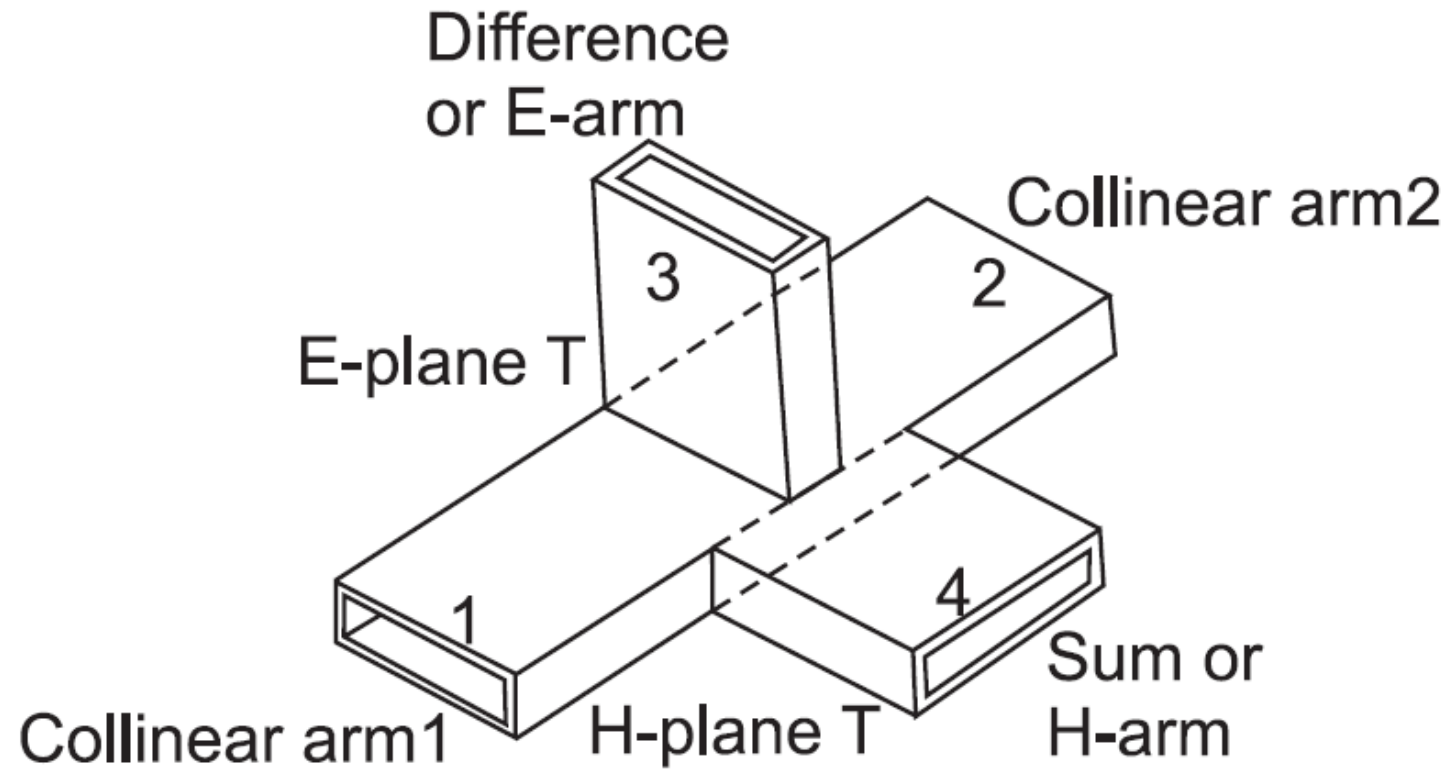
- Substituting these values in S-matrix, we get S-matrix of the H-plane Tee

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{21} & S_{22} & S_{23} \\ S_{31} & S_{32} & S_{33} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{22} & S_{13} \\ S_{13} & S_{13} & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & \frac{1}{\sqrt{2}} \\ -\frac{1}{2} & \frac{1}{2} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}$$

$$[S] = \frac{1}{2} \begin{bmatrix} 1 & -1 & \sqrt{2} \\ -1 & 1 & \sqrt{2} \\ \sqrt{2} & \sqrt{2} & 0 \end{bmatrix}$$

Magic-Tee

- A hybrid tee is formed with the combination of the E-plane and H-plane tees and is called a magic-T. It has four ports as shown in figure:



Magic-Tee

- The magic-T has the following characteristics when all the ports are terminated with a matched loads:
 1. If two waves of equal magnitude and equal phase are fed into ports 1 and 2, the output at Port 3 is subtractive and becomes zero and total output will appear additively at the port 4. Hence, Port 3 is called the difference or E-arm and 4, the sum or H-arm.
 2. A wave incident at Port 3 (E-arm) divides equally between ports 1 and 2 but is opposite in phase with no coupling to Port 4 (H-arm). Thus,

$$S_{13} = -S_{23}, \quad S_{43} = 0$$

3. A wave incident at Port 4 (H-arm) divides equally between ports 1 and 2 in phase with no coupling to port 3 (E-arm). Thus,

$$S_{14} = S_{12} = \frac{1}{\sqrt{2}} = S_{24} = S_{42} \text{ and } S_{34} = 0$$

Magic-Tee

4. A wave fed into one collinear port, 1 or 2, will not appear in the other collinear ports, 2 or 1, respectively. Hence, two collinear ports 1 and 2 are isolated from each other, making

$$S_{12} = S_{21} = 0$$

- A magic-T can be matched by putting tuning screws suitably in the E and H-arms without destroying the symmetry of the junctions. Therefore, for an ideal lossless magic-T matched at ports 3 and 4, $S_{33} = S_{44} = 0$. Therefore, the S-matrix for a magic-T, matched at ports 3 and 4 given by

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{12} & S_{22} & -S_{13} & S_{14} \\ S_{13} & -S_{13} & 0 & 0 \\ S_{14} & S_{14} & 0 & 0 \end{bmatrix}$$

Magic-Tee

$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{21} & S_{22} & S_{23} & S_{24} \\ S_{31} & S_{32} & S_{33} & S_{34} \\ S_{41} & S_{42} & S_{43} & S_{44} \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{12} & S_{22} & -S_{13} & S_{14} \\ S_{13} & -S_{13} & 0 & 0 \\ S_{14} & S_{14} & 0 & 0 \end{bmatrix}$$

- From the unitary property applied to rows 1 and 2, we get

$$|S_{11}|^2 + |S_{12}|^2 + |S_{13}|^2 + |S_{14}|^2 = 1 \quad (1)$$

$$|S_{12}|^2 + |S_{22}|^2 + |S_{13}|^2 + |S_{14}|^2 = 1 \quad (2)$$

- Subtracting these two equations:

$$|S_{11}|^2 + |S_{22}|^2 = 0 \implies S_{11} = S_{22}$$

- Form the unitary property applied to rows 3 and 4, $|S_{13}| = \frac{1}{\sqrt{2}}$ and $|S_{14}| = \frac{1}{\sqrt{2}}$
- Substituting these values in equation(1), $|S_{11}|^2 + |S_{12}|^2 = 0 \implies |S_{11}| = |S_{12}| = 0$

Magic-Tee

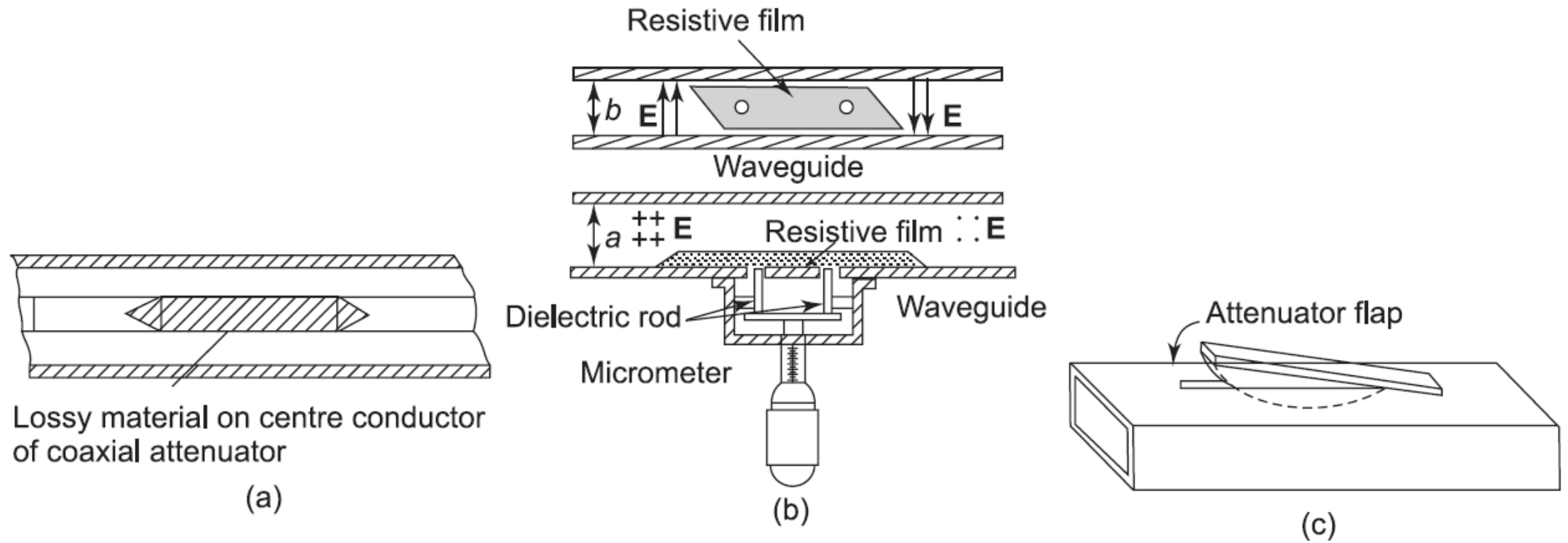
$$[S] = \begin{bmatrix} S_{11} & S_{12} & S_{13} & S_{14} \\ S_{12} & S_{22} & -S_{13} & S_{14} \\ S_{13} & -S_{13} & 0 & 0 \\ S_{14} & S_{14} & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1/\sqrt{2} & 1/\sqrt{2} \\ 0 & 0 & -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \end{bmatrix}$$

$$[S] = \frac{1}{\sqrt{2}} \begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 \\ 1 & -1 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{bmatrix}$$

Attenuator

- Attenuators are passive devices used to control power levels in a microwave system by partially absorbing the transmitted signal wave. Both fixed and variable attenuators are designed using resistive films (aquadag coated dielectric sheet).
- A coaxial fixed attenuator uses a film with losses on the center conductor to absorb some of the power as shown in figure(a).
- The fixed waveguide type [figure (b)] consists of a thin dielectric strip coated with resistive film and placed at the center of the waveguide parallel to the maximum E field.
- Induced current on the resistive film due to the incident wave results in power dissipation, leading to attenuation of microwave energy.
- The dielectric strip is tapered at both ends up to a length of more than half wavelength to reduce reflections. The resistive vane is supported by two dielectric rods separated by an odd multiple of quarter wavelength and perpendicular to the electric field [figure (b)].

Attenuator



Microwave attenuator:

(a) coaxial line fixed attenuator (b) and (c) waveguide attenuators

Attenuator

Variable-type Attenuator:

- A variable-type attenuator can be constructed
 - by moving the resistive vane by means of micrometer screw from one side of the narrow wall to the center where the E-field is maximum [figure(b)] or
 - by changing the depth of insertion of a resistive vane at an E-field maximum through a longitudinal slot at the middle of the broad wall as shown in figure(c).
- A maximum of 90 dB attenuation is possible with VSWR of 1.05.
- The resistance card can be shaped to give a linear variation of attenuation with the depth of insertion.

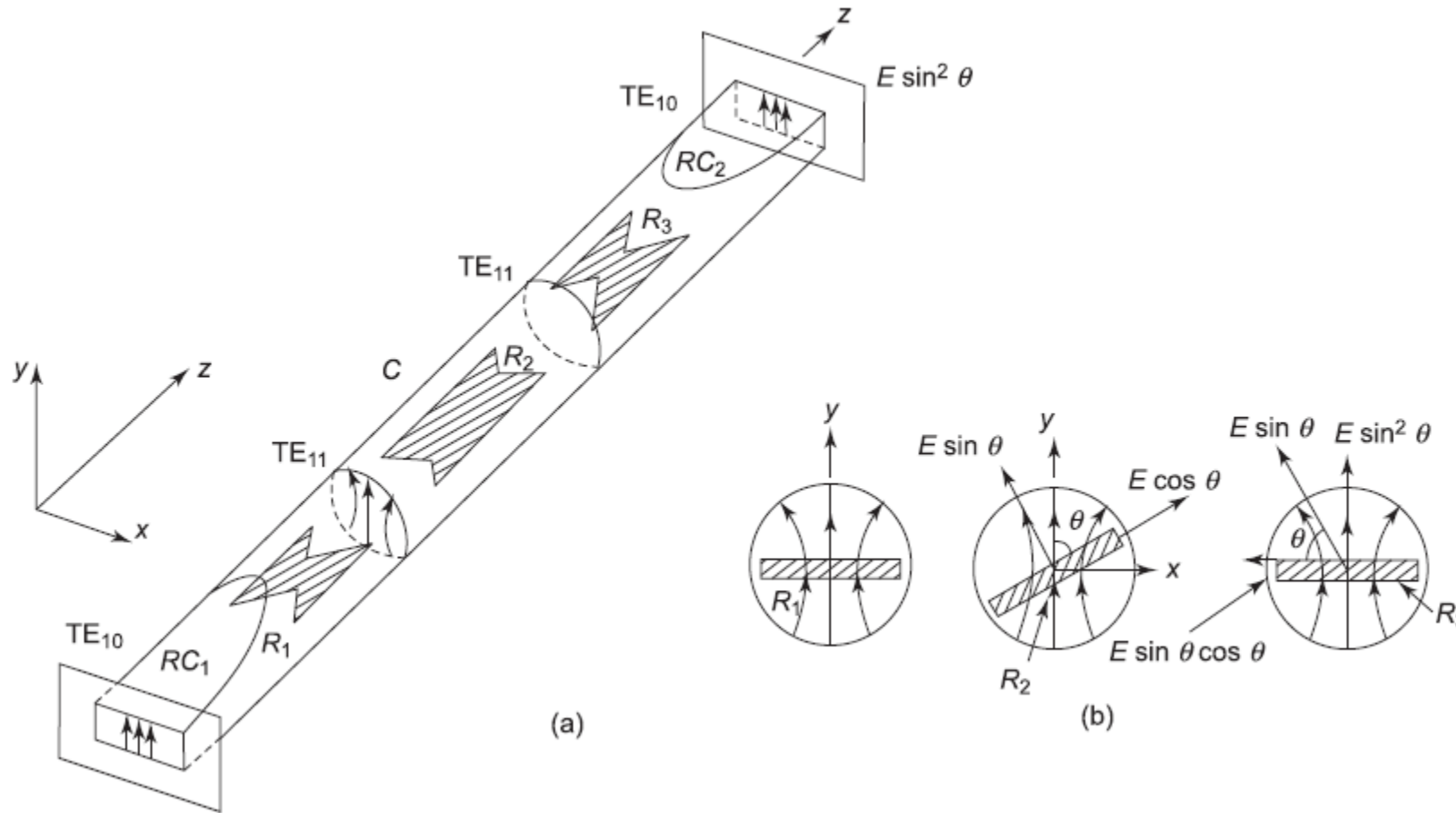
Attenuator

Precision-type Variable Attenuator:

- A precision-type variable attenuator makes use of a circular waveguide section (C) containing a very thin tapered resistive card (R_2), to both sides of which are connected axisymmetric sections of circular-to-rectangular waveguide tapered transitions (RC_1 and RC_2) as shown in figure(a).
- The center circular section with the resistive card can be precisely rotated by 360° with respect to the two fixed sections of circular to rectangular waveguide transitions.
- The induced current on the resistive card R_2 due to the incident signal is dissipated as heat and produces attenuation of the transmitted signal.
- The incident TE_{10} dominant wave in the rectangular waveguide is converted into a dominant TE_{11} mode in the circular waveguide.

Attenuator

Precision-type Variable Attenuator:



R_1, R_2, R_3 : Tapered resistive cards; RC_1 and RC_2 : Rectangular-to-circular waveguide transitions; C : Circular waveguide section

Attenuator

Precision-type Variable Attenuator:

- A very thin tapered resistive card is placed perpendicular to the E-field at the circular end of each transition section so that it has a negligible effect on the field perpendicular to it but absorbs any component parallel to it.
- Therefore, a pure TE_{11} mode is excited in the middle section.
- With reference to figure(b), if the resistive card in the center section is kept at an angle θ relative to the E-field direction of the TE_{11} mode, the component $E \cos \theta$ parallel to the card gets absorbed while the component $E \sin \theta$ is transmitted without attenuation. This later component finally appears as electric field component $E \sin^2 \theta$ in the rectangular output guide.
- Therefore, the attenuation of the incident wave is

$$\alpha = -20 \log |S_{21}| = -20 \log(\sin^2 \theta) = -40 \log(\sin \theta)$$

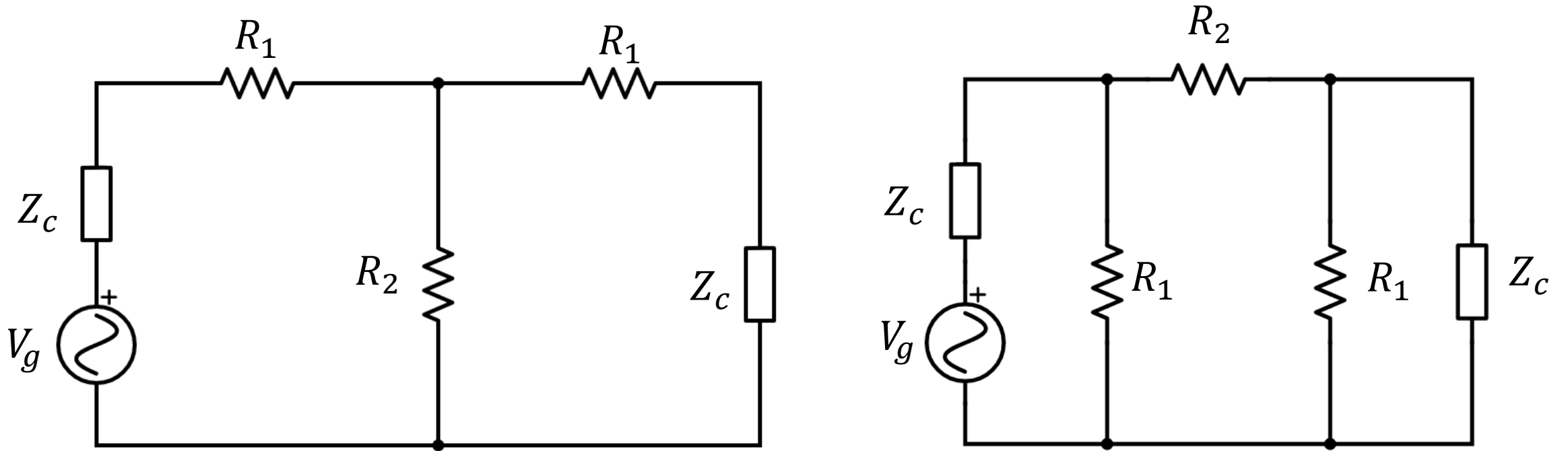
Attenuator

Electronically-Controlled Attenuators:

- For applications in various microwave systems, it is desirable to have an attenuator whose attenuation can be controlled by the application of a suitable signal, such as dc voltage or a bias current.
- Two devices that are suitable for use in an electronically controlled attenuator are the PIN diode and a field-effect transistor.
- These devices can be used as variable resistors whose resistance is controlled by the applied signal.
- The basic attenuator network is a symmetric T or π network as shown in figures(a) and (b).
- Resistor values R_1 and R_2 are chosen so that when the attenuator is terminated in a resistance equal to the transmission-line characteristic impedance Z_c , the input is matched, that is, $Z_{in} = Z_c$, and to provide an output voltage reduction by factor K.

Attenuator

Electronically-Controlled Attenuators:



Two basic attenuator networks

Attenuator

Electronically-Controlled Attenuators:

- For the T-network , we have

$$R_{in} = R_1 + \frac{R_2(R_1 + Z_c)}{R_1 + R_2 + Z_c}$$

- For $R_{in} = Z_c$, We get the equation

$$Z_c = R_1 + \frac{R_2(R_1 + Z_c)}{R_1 + R_2 + Z_c}$$

$$Z_c(R_1 + R_2 + Z_c) = R_1(R_1 + R_2 + Z_c) + R_2(R_1 + Z_c)$$

$$Z_c(R_1 + R_2) + Z_c^2 = Z_c(R_1 + R_2) + R_1^2 + 2R_1R_2$$

$$\Rightarrow Z_c^2 = R_1^2 + 2R_1R_2$$

- For $R_{in} = Z_c$, the Thevenin impedance seen by the load equals Z_c .

Attenuator

Electronically-Controlled Attenuators:

- The Thevenin open-circuit voltage across R_2 is

$$V_{TH} = \frac{R_2}{R_1 + R_2 + Z_c} V_g$$

- The power delivered to the load is

$$P_L = \frac{1}{2} \left| \frac{V_{TH}}{2Z_c} \right|^2 Z_c = \left(\frac{R_2}{R_1 + R_2 + Z_c} \right)^2 \frac{|V_g|^2}{8Z_c}$$

- The available power is $|V_g|^2 / 8Z_c$, so that the power attenuation K^2 is given by

$$K^2 = \left(\frac{R_2}{R_1 + R_2 + Z_c} \right)^2$$

Attenuator

Electronically-Controlled Attenuators:

- When K has been specified, and also requiring that $R_{in} = Z_c$, we have two equations that are readily solved for the required values of R_1 and R_2 . Thus, we find that

$$R_1 = \frac{1 - K}{1 + K} Z_c$$

$$R_2 = \frac{2K}{1 - K^2} Z_c$$

- For 10-dB attenuation in a 50Ω system, $K = \sqrt{0.1}$ which results in

$$R_1 = 25.97 \Omega$$

$$R_2 = 35.14 \Omega$$

Attenuator

Electronically-Controlled Attenuators:

- For 3-dB attenuation in a 50Ω system, $K = \sqrt{0.5}$ which results in

$$R_1 = 8.58 \Omega$$

$$R_2 = 141.4 \Omega$$

- For a π -network, R_1 and R_2 are given by

$$R_1 = \frac{1 + K}{1 - K} Z_c$$

$$R_2 = \frac{1 - K^2}{2K} Z_c$$

Microwave Resonators

Contents

- Series and Parallel Resonant Circuits
- Q-factor (unloaded and loaded)
- Bandwidth
- Transmission Line Resonators
- Waveguide resonators

Introduction

A resonator is a device or circuit that exhibits resonance

In an electrical circuit, resonance condition occurs at a frequency when capacitive and inductive reactances become equal in magnitude and electric energy oscillates between electric field of a capacitor and magnetic field of an inductor.

Microwave resonators are used in a variety of applications:

- Filters
- Oscillators
- Tuned amplifiers
- Frequency meters

At frequencies near resonance, a microwave resonator can be modeled as series or parallel RLC lumped element electric circuit

The basic properties of series and parallel RLC circuits are reviewed first.

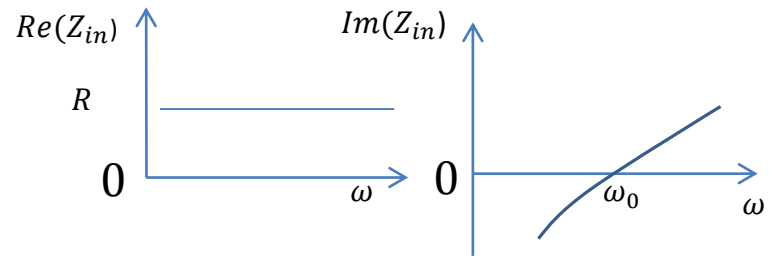
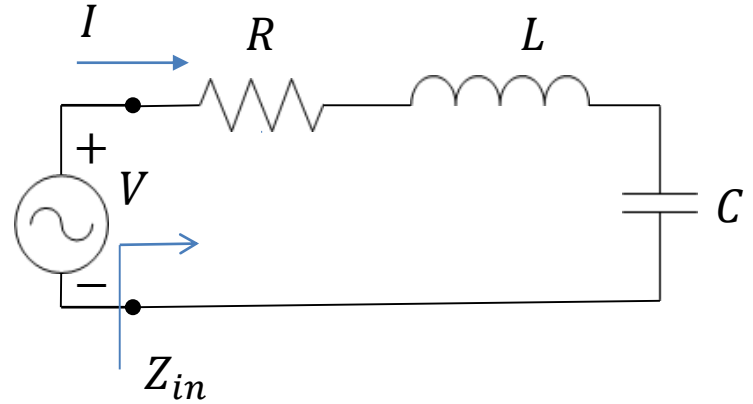
Series RLC Circuit

For the series RLC circuit shown in the figure:

$$\begin{aligned} Z_{in} &= R + j\omega L - j\frac{1}{\omega C} \\ &= R + j\omega L \left(1 - \frac{1}{\omega^2 LC} \right) \end{aligned}$$

$$Z_{in} = R + j\omega L \left(1 - \frac{\omega_0^2}{\omega^2} \right)$$

where $\omega_0 = \frac{1}{\sqrt{LC}}$



Series RLC Circuit

$$|Z_{in}(\omega)| = \sqrt{R^2 + \omega^2 L^2 (1 - \omega_0^2 / \omega^2)^2}$$

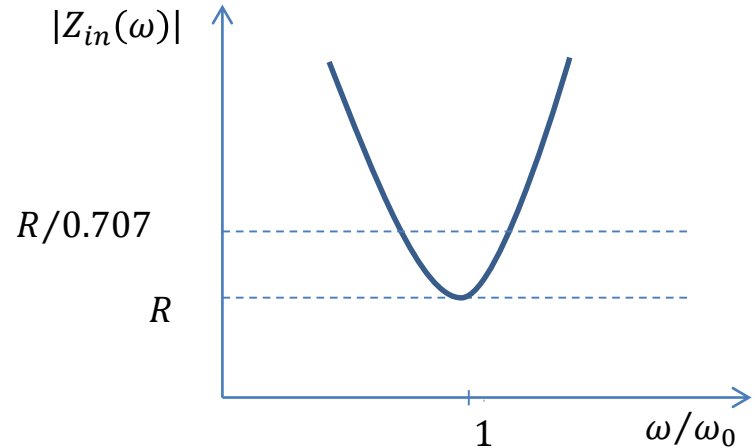
An important parameter of the resonant circuit is its Q which is defined as:

$$Q = \omega \frac{\text{average energy stored}}{\text{average power dissipated}}$$

$$Q = \omega \frac{W_m + W_e}{P_{\text{loss}}}$$

At resonance, $W_m = W_e$

The Q of a resonator itself disregarding the external loading effect is called the unloaded Q and denoted by Q_0



Series RLC Circuit

Therefore, $Q_0 = \omega_0 \frac{2W_m}{P_{\text{loss}}}$

$$Q_0 = \omega_0 \frac{|I|^2 L}{|I|^2 R} = \omega_0 \frac{L}{R}$$

Since $\omega_0^2 = \frac{1}{LC}$

$$Q_0 = \frac{1}{\omega_0 RC}$$

Let us now study the behavior of the input impedance of a series RLC resonator near its resonance

$$Z_{in} = R + j\omega L \left(\frac{\omega^2 - \omega_0^2}{\omega^2} \right)$$

In the vicinity of resonance,

$$\omega^2 - \omega_0^2 =$$

$$(\omega - \omega_0)(\omega + \omega_0) \approx 2\omega\Delta\omega$$

For small $\Delta\omega$,

$$Z_{in} \approx R + j\omega L \frac{2\omega\Delta\omega}{\omega^2} \approx R + j2\Delta\omega L$$

$$Z_{in} \approx R + \frac{j2\Delta\omega R Q_0}{\omega_0}$$

Series RLC Circuit

Let us now consider half power fractional bandwidth of the resonator

$$\text{We have } P_{in} = \frac{1}{2}VI^* = \frac{1}{2}Z_{in} \left| \frac{V}{Z_{in}} \right|^2 \quad \text{Therefore, } Re(P_{in}) = \frac{1}{2}R \left| \frac{V}{Z_{in}} \right|^2$$

$$\text{When } \omega = \omega_0, Z_{in} = R \text{ and } Re(P_{in})|_{\omega=\omega_0} = \frac{|V|^2}{2R}$$

$$\text{When } |Z_{in}|^2 = 2R^2 \text{ that is } |Z_{in}| = \frac{R}{0.707} \quad Re(P_{in}) = \frac{1}{2} Re(P_{in})|_{\omega=\omega_0}$$

$$\text{From } Z_{in} \approx R + \frac{j2\Delta\omega RQ_0}{\omega_0}, \quad |Z_{in}|^2 = R^2 + \frac{4\Delta\omega^2 R^2 Q_0^2}{\omega_0^2} = 2R^2$$

$$\Rightarrow \left(\frac{2\Delta\omega}{\omega_0} \right)^2 = \frac{1}{Q_0^2} \quad \text{Therefore, fractional bandwidth } \frac{2\Delta\omega}{\omega_0} = \frac{1}{Q_0}$$

Parallel RLC Circuit

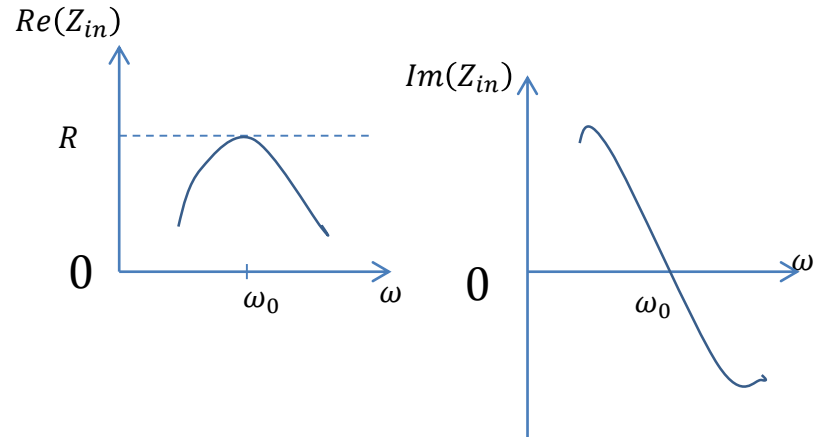
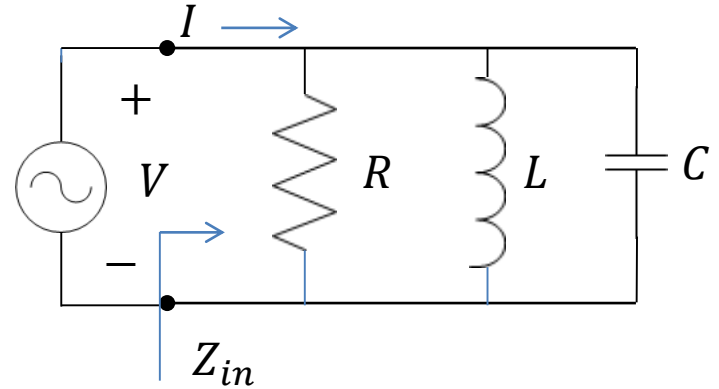
For the parallel RLC circuit shown in the figure:

$$Z_{in} = \frac{1}{\left(\frac{1}{R} + \frac{1}{j\omega L} + j\omega C\right)} = \frac{1}{\left(\frac{1}{R} - \frac{j}{\omega L}(1 - \omega^2 LC)\right)}$$

$$Z_{in} = \frac{\frac{1}{R} + j\left(\frac{1 - \omega^2 LC}{\omega L}\right)}{\left(\frac{1}{R}\right)^2 + \left(\frac{1 - \omega^2 LC}{\omega L}\right)^2}$$

$Re(Z_{in})$ attains its maximum value R at the resonance frequency

$$\omega_0 = \frac{1}{\sqrt{LC}}$$



Parallel RLC Circuit

For such parallel RLC circuit

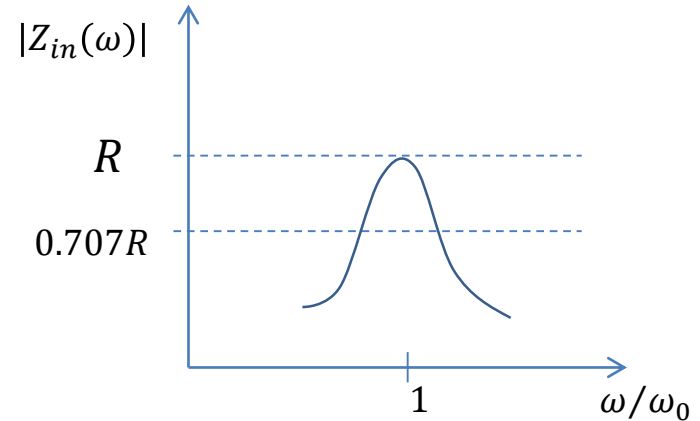
$$Q_0 = \omega_0 RC = \frac{R}{\omega_0 L}$$

Near resonance $\omega = \omega_0 + \Delta\omega$

$$Z_{in} = \left(\frac{1}{R} + j\omega C + \frac{1}{j\omega L} \right)^{-1}$$

$$= \left(\frac{1}{R} + j\omega_0 C + j\Delta\omega C + \frac{1}{j(\omega_0 + \Delta\omega)L} \right)^{-1}$$

$$Z_{in} = \left(\frac{1}{R} + j\omega_0 C + j\Delta\omega C + \frac{1}{j\omega_0 L(1 + \Delta\omega/\omega_0)} \right)^{-1}$$



When $\Delta\omega/\omega_0 \ll 1$ we can use the approximation $1/(1 + \Delta\omega/\omega_0) \approx 1 - \Delta\omega/\omega_0$

$$Z_{in} \approx \left(\frac{1}{R} + j\omega_0 C + j\Delta\omega C + \frac{1}{j\omega_0 L} - \frac{\Delta\omega}{j\omega_0^2 L} \right)^{-1}$$

Parallel RLC Circuit

$$Z_{in} \approx \left(\frac{1}{R} + j\omega_0 C + j\Delta\omega C + \frac{1}{j\omega_0 L} - \frac{\Delta\omega}{j\omega_0^2 L} \right)^{-1}$$

$$Z_{in} = \left(\frac{1}{R} + j\Delta\omega C + \frac{j\Delta\omega}{\omega_0^2 L} \right)^{-1}$$

$$Z_{in} = \left(\frac{1}{R} + j\Delta\omega C + j\Delta\omega C \right)^{-1} = \frac{R}{1 + j2\Delta\omega RC}$$

Since $Q_0 = \omega_0 RC$,

$$Z_{in} = \frac{R}{1 + j2\Delta\omega Q_0/\omega_0}$$

Parallel RLC Circuit

$$Re(P_{in}) = \frac{1}{2} V \left(\frac{V}{R} \right)^* = \frac{1}{2} \frac{|V|^2}{R} = \frac{1}{2} |I|^2 |Z_{in}|^2 \frac{1}{R}$$

At resonance $Re(P_{in})|_{\omega=\omega_0} = \frac{1}{2} |I|^2 R$

Therefore, $\frac{Re(P_{in})}{Re(P_{in})|_{\omega=\omega_0}} = \frac{|Z_{in}|^2}{R^2}$

For $\frac{Re(P_{in})}{Re(P_{in})|_{\omega=\omega_0}}$ to become $\frac{1}{2}$, $\frac{R^2}{2} = |Z_{in}|^2$

From $Z_{in} = \frac{R}{1+j2\Delta\omega Q_0/\omega_0}$, $2\Delta\omega Q_0/\omega_0 = 1$

Therefore, fractional bandwidth $2\Delta\omega/\omega_0 = 1/Q_0$

Loaded Q

The unloaded Q of a circuit Q_0 is the quality factor of the circuit without any external loading

In practice, external circuitry connected to the resonator will produce loading effect.

Let the loading of the external circuit be represented by a load resistance R_L and the Q of the external circuit by Q_e .

Let Q_L be the Q of the loaded circuit.

For series RLC circuit $Q_L = \omega_0 \frac{L}{R+R_L}$ Therefore, $\frac{1}{Q_L} = \frac{R+R_L}{\omega_0 L} = \frac{1}{Q_0} + \frac{1}{Q_e}$

Similarly, for a parallel RLC circuit R and R_L are in parallel and

$Q_L = \frac{RR_L}{\omega_0(R+R_L)L}$ Therefore, $\frac{1}{Q_L} = \frac{\omega_0(R+R_L)L}{RR_L} = \frac{1}{Q_e} + \frac{1}{Q_0}$

Transmission Line Resonators

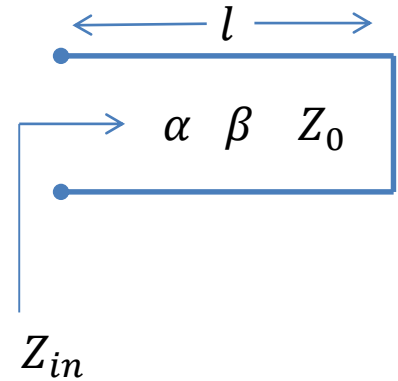
Transmission line sections of various lengths and terminations (open or short) can be used as a resonator.

Let us consider a lossy transmission line of length l terminated to a short circuit at one end. The transmission line is low loss with very small value of attenuation constant α

The transmission line has characteristic impedance Z_0

At $\omega = \omega_0$ $l = \lambda/2$

$$Z_{in} = Z_0 \frac{Z_L + Z_0 \tanh \gamma l}{Z_0 + Z_L \tanh \gamma l}$$



Transmission Line Resonators

For $Z_L = 0$ $Z_{in} = Z_0 \tanh \gamma l = Z_0 \tanh(\alpha + j\beta)l$

Therefore, $Z_{in} = Z_0 \frac{\tanh \alpha l + j \tan \beta l}{1 + j \tan \beta l \tanh \alpha l}$

Since we have considered a low loss line, $\alpha l \ll 1$

$$\tanh \alpha l \approx \alpha l$$

$$\beta l = \frac{\omega l}{v_p} = \frac{\omega_0 l}{v_p} + \frac{\Delta \omega l}{v_p}$$

Since $l = \frac{\lambda}{2}$ at $\omega = \omega_0$, $\frac{\omega_0 l}{v_p} = \frac{2\pi f_0 \lambda}{\lambda f_0 2} = \pi$

Transmission Line Resonators

$$\text{Now, } \beta l = \frac{\omega_0 l}{v_p} + \frac{\Delta\omega l}{v_p} \quad \text{and} \quad \frac{\omega_0 l}{v_p} = \pi$$

$$\text{Therefore, } \beta l = \pi + \frac{\Delta\omega\pi}{\omega_0} \quad \text{and} \quad \tan \beta l = \tan \left(\pi + \frac{\Delta\omega\pi}{\omega_0} \right) \approx \frac{\Delta\omega\pi}{\omega_0}$$

$$\text{Hence, } Z_{in} = Z_0 \frac{\tanh \alpha l + j \tan \beta l}{1 + j \tan \beta l \tanh \alpha l} \approx Z_0 \frac{\alpha l + j \left(\frac{\Delta\omega\pi}{\omega_0} \right)}{1 + j \alpha l \left(\frac{\Delta\omega\pi}{\omega_0} \right)}$$

$$\text{Therefore, } Z_{in} \approx Z_0 \left(\alpha l + j \frac{\Delta\omega\pi}{\omega_0} \right)$$

Comparing with a series resonant circuit for which

$$Z_{in} \approx R + j2\Delta\omega L$$

Transmission Line Resonators

$$R = Z_0 \alpha l \quad \text{and} \quad L = \frac{\pi Z_0}{2\omega_0}$$

Capacitance C can be found from $C = \frac{1}{\omega_0^2 L} = \frac{2}{\pi \omega_0 Z_0}$

Unloaded Q of the resonator $Q_0 = \frac{\omega_0 L}{R} = \frac{\pi}{2\alpha l}$

Transmission Line Resonators

Let us now consider another transmission line resonator which consists of a short-circuited transmission line of length $\lambda/4$.

$$l = \frac{\lambda}{4} \text{ at } \omega = \omega_0$$

We have
$$Z_{in} = Z_0 \frac{\tanh \alpha l + j \tan \beta l}{1 + j \tan \beta l \tanh \alpha l}$$

Multiplying the numerator and denominator by $-j \cot \beta l$

$$Z_{in} = Z_0 \frac{1 - j \tanh \alpha l \cot \beta l}{\tanh \alpha l - j \cot \beta l}$$

Transmission Line Resonators

Let $\omega = \omega_0 + \Delta\omega$

$$\beta l = \frac{\omega_0 l}{v_p} + \frac{\Delta\omega l}{v_p} = \frac{\pi}{2} + \frac{\pi\Delta\omega}{2\omega_0}$$

Therefore, $\cot \beta l = -\tan \frac{\pi\Delta\omega}{2\omega_0} \approx -\frac{\pi\Delta\omega}{2\omega_0}$

We have $Z_{in} = Z_0 \frac{1 - j \tanh \alpha l \cot \beta l}{\tanh \alpha l - j \cot \beta l}$

Therefore, $Z_{in} = Z_0 \frac{1 + j \alpha l \left(\frac{\pi\Delta\omega}{2\omega_0} \right)}{\alpha l + j \frac{\pi\Delta\omega}{2\omega_0}} \approx \frac{Z_0}{\alpha l + j \frac{\pi\Delta\omega}{2\omega_0}}$

Transmission Line Resonators

$$Z_{in} \approx \frac{Z_0}{\alpha l + j \frac{\pi \Delta \omega}{2 \omega_0}} = \frac{1}{\frac{\alpha l}{Z_0} + j \frac{\pi \Delta \omega}{2 \omega_0 Z_0}}$$

For a parallel RLC circuit near resonance,

$$Z_{in} \approx \frac{R}{1 + j 2 \Delta \omega R C} = \frac{1}{\frac{1}{R} + j 2 \Delta \omega C}$$

Therefore, $R = \frac{Z_0}{\alpha l}$ and $C = \frac{\pi}{4 \omega_0 Z_0}$

$$Q_0 = \omega_0 R C = \frac{\pi}{4 \alpha l}$$

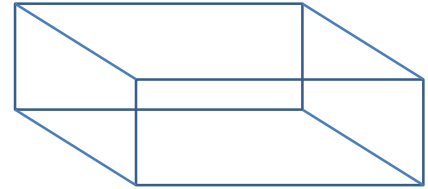
Waveguide Resonators

At higher microwave frequencies transmission line resonators have relatively low value of Q .

Since open ended waveguide can radiate significantly, waveguide resonators are usually short circuited at both ends forming a cavity.

Electric and magnetic energy is stored within the cavity enclosed.

Dissipation of power takes place on the waveguide walls as well as in the dielectric material filling the cavity if the dielectric is lossy.



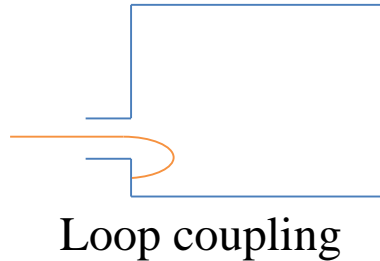
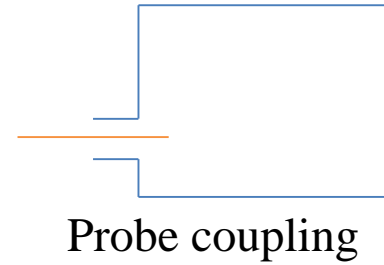
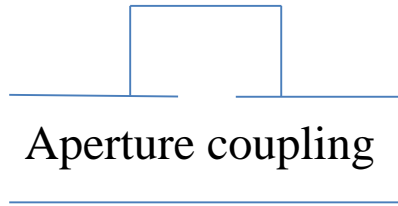
Rectangular cavity



Cylindrical cavity

Waveguide Resonators

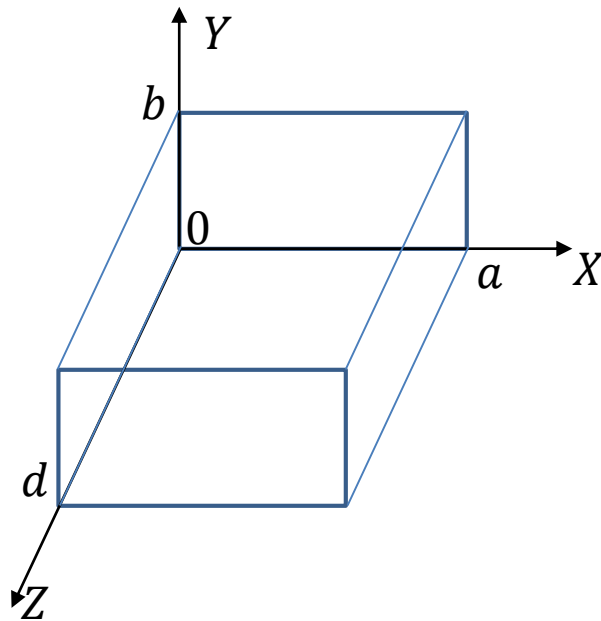
Coupling to cavity resonator may be done using a small aperture or a probe or a loop.



Resonant Frequencies of Rectangular Cavity

We first find the resonant frequencies of the cavity assuming it to be lossless.

The unloaded Q of the cavity is then determined considering small amount of loss on the waveguide walls as well as in the dielectric material.



Resonant Frequencies of Rectangular Cavity

For TE_{mn} or TM_{mn} mode

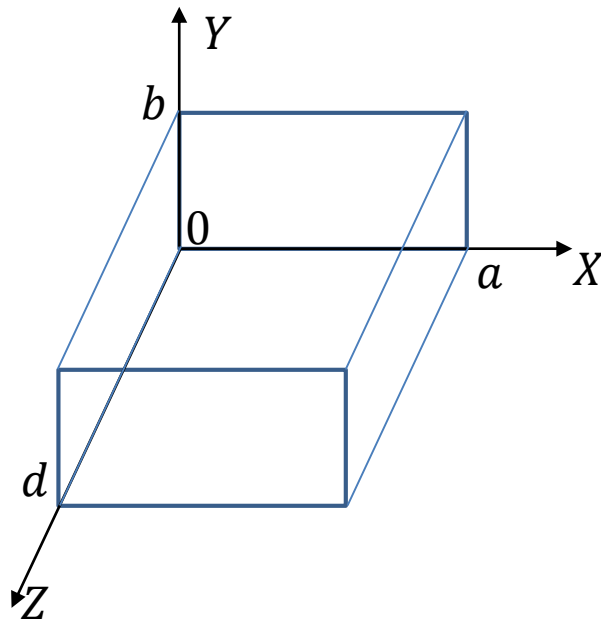
$$\vec{E}_t(x, y, z) = \vec{e}(x, y) (A^+ e^{-j\beta_{mn}z} + A^- e^{j\beta_{mn}z})$$

Transverse variation

Amplitudes of forward
and backward wave

where,

$$\beta_{mn} = \sqrt{k^2 - \left(\frac{m\pi}{a}\right)^2 - \left(\frac{n\pi}{b}\right)^2}$$



Resonant Frequencies of Rectangular Cavity

$$\vec{E}_t = 0 \text{ at } z = 0 \quad \Rightarrow A^+ = -A^-$$

$$\vec{E}_t = 0 \text{ at } z = d$$

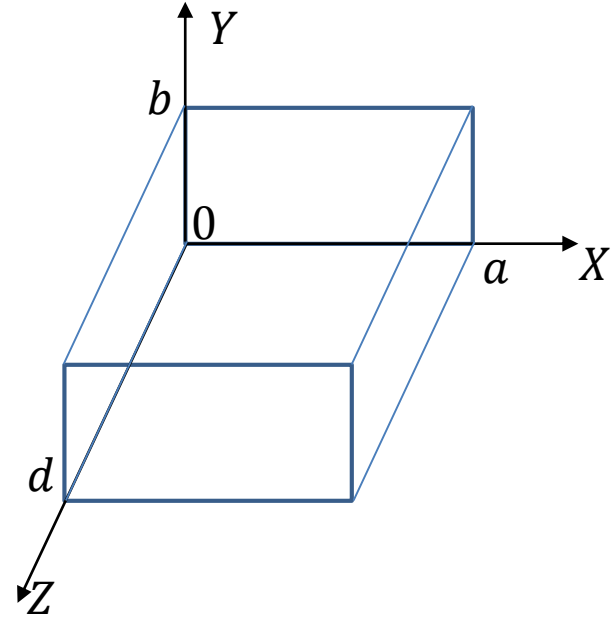
$$\therefore \vec{E}_t(x, y, d) = -\vec{e}(x, y)A^+ 2j \sin \beta_{mn}d = 0$$

For $A^+ \neq 0$

$$\beta_{mn}d = l\pi \quad \text{where } l = 1, 2, 3 \dots$$

\therefore For a rectangular cavity, the wave number

$$k_{mnl} = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 + \left(\frac{l\pi}{d}\right)^2}$$



For $b < a < d$, the dominant resonant mode is TE_{101} and $d = \frac{\lambda_g}{2}$ for TE_{10} mode.

Unloaded Q of TE_{10l} mode

For TE_{10l} mode we can write the field components as follows:

$$E_y = A^+ \sin \frac{\pi x}{a} (e^{-j\beta z} - e^{j\beta z})$$
$$H_x = -\frac{A^+}{Z_{\text{TE}}} \sin \frac{\pi x}{a} (e^{-j\beta z} + e^{j\beta z})$$
$$H_z = \frac{j\pi A^+}{k\eta a} \cos \frac{\pi x}{a} (e^{-j\beta z} - e^{j\beta z})$$

We have seen that for TE₁₀ mode

$$H_z = A_{10} \cos \frac{\pi x}{a} e^{-j\beta z}$$
$$E_y = \frac{-j\omega\mu a}{\pi} A_{10} \sin \frac{\pi x}{a} e^{-j\beta z}$$
$$H_x = \frac{j\beta a}{\pi} A_{10} \sin \frac{\pi x}{a} e^{-j\beta z}$$
$$H_y = E_x = 0$$

$$\therefore \frac{-j\omega\mu a}{\pi} A_{10} = A^+ \qquad \therefore \omega\mu = k\eta$$
$$\Rightarrow A_{10} = jA^+ \frac{\pi}{\omega\mu a} = \frac{jA^+ \pi}{k\eta a}$$

Unloaded Q of TE_{10l} mode

$$E_y = A^+ \sin \frac{\pi x}{a} (e^{-j\beta z} - e^{j\beta z})$$

$$H_x = -\frac{A^+}{Z_{\text{TE}}} \sin \frac{\pi x}{a} (e^{-j\beta z} + e^{j\beta z})$$

$$H_z = \frac{j\pi A^+}{k\eta a} \cos \frac{\pi x}{a} (e^{-j\beta z} - e^{j\beta z})$$

Substituting $E_0 = \frac{2A^+}{j}$, we get

$$E_y = E_0 \sin \frac{\pi x}{a} \sin \frac{l\pi z}{d}$$

$$H_x = -\frac{jE_0}{Z_{\text{TE}}} \sin \frac{\pi x}{a} \cos \frac{l\pi z}{d}$$

$$H_z = \frac{j\pi E_0}{k\eta a} \cos \frac{\pi x}{a} \sin \frac{l\pi z}{d}$$

$$W_e = \frac{\epsilon}{4} \int_V E_y E_y^* dv = \frac{\epsilon abd}{16} |E_0|^2$$

At resonance,

$$W_e = W_m$$

Unloaded Q of TE_{10l} mode

Case I. The dielectric is perfect but cavity walls are slightly lossy

The power loss on the conducting walls can be found as

$$P_c = \frac{R_s}{2} \int_{\text{walls}} |H_t|^2 ds$$
$$R_s = \sqrt{\frac{\omega\mu_0}{2\sigma}}$$

Surface resistivity
of metallic walls



The conductor loss can be found as

$$P_c = \frac{R_s E_0^2 \lambda^2}{8\eta^2} \left(\frac{l^2 ab}{d^2} + \frac{bd}{a^2} + \frac{l^2 a}{2d} + \frac{d}{2a} \right)$$
$$Q_c = \frac{2\omega_0 W_e}{P_c}$$

Unloaded Q of TE_{10l} mode

Case II. The dielectric is lossy but cavity walls are perfectly conducting.

$$\epsilon = \epsilon' - j\epsilon'' = \epsilon_0\epsilon_r(1 - j \tan \delta)$$

Power dissipated within the dielectric volume is

$$\begin{aligned} P_d &= \frac{1}{2} \int_V \vec{J} \cdot \vec{E}^* dv = \frac{\omega\epsilon''}{2} \int_V |E|^2 dv \\ &= \frac{abd\omega\epsilon''|E_0|^2}{8} \end{aligned}$$

Q_d with lossy dielectric but perfectly conducting wall is

$$Q_d = \frac{2\omega \frac{\epsilon'abd}{16} |E_0|^2}{\frac{abd\omega\epsilon''|E_0|^2}{8}} = \frac{\epsilon'}{\epsilon''} = \frac{1}{\tan \delta}$$

Unloaded Q of the cavity is

$$Q_0 = \left(\frac{1}{Q_c} + \frac{1}{Q_d} \right)^{-1}$$

Circular Waveguide Cavity Resonator

Since the dominant mode of circular waveguide is TE_{11} , the dominant mode of the circular waveguide cavity is TE_{111} .

For TM modes, the mode with the lowest cut off frequency is TM_{01} mode.

The resonant frequencies of TE_{nml} and TM_{nml} modes of the circular waveguide cavities can be found as follows:

$$\vec{E}_t(\rho, \phi, z) = \vec{e}(\rho, \phi)(A^+ e^{-j\beta_{nm}z} + A^- e^{j\beta_{nm}z})$$

For TE_{nm} mode

$$\beta_{nm} = \sqrt{k^2 - \left(\frac{p'_{nm}}{a}\right)^2}$$

For TM_{nm} mode

$$\beta_{nm} = \sqrt{k^2 - \left(\frac{p_{nm}}{a}\right)^2}$$

Circular Waveguide Cavity Resonator

$$\vec{E}_t = 0 \text{ at } z = 0$$

We have

$$A^+ = -A^-$$

$$\vec{E}_t = 0 \text{ at } z = d$$

We have

$$\sin \beta_{nm}d = 0$$

$$\beta_{nm}d = l\pi \quad \text{where } l = 1, 2, 3 \dots$$

For the resonant **TE**_{nm l} mode

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_r\epsilon_r}} \sqrt{\left(\frac{p'_{nm}}{a}\right)^2 + \left(\frac{l\pi}{d}\right)^2}$$

For the resonant **TM**_{nm l} mode

$$f_{nml} = \frac{c}{2\pi\sqrt{\mu_r\epsilon_r}} \sqrt{\left(\frac{p_{nm}}{a}\right)^2 + \left(\frac{l\pi}{d}\right)^2}$$

Circular Waveguide Cavity Resonator

Q factor for the cylindrical cavities can be found in the same manner as in rectangular cavities.

Cylindrical cavity operating at TE_{011} mode is often used for frequency meters because of its higher Q

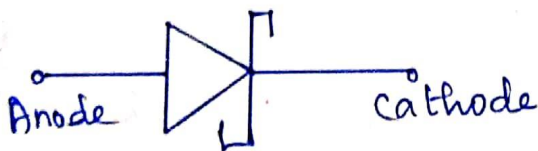


Schottky diode : Schottky diode

is a metal - semiconductor junction diode that has less forward voltage drop than the PN junction diode and can be used in high-speed switching applications.

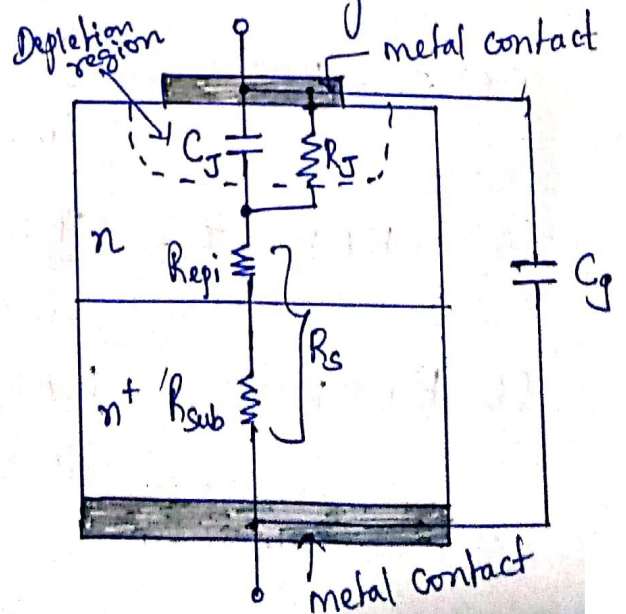
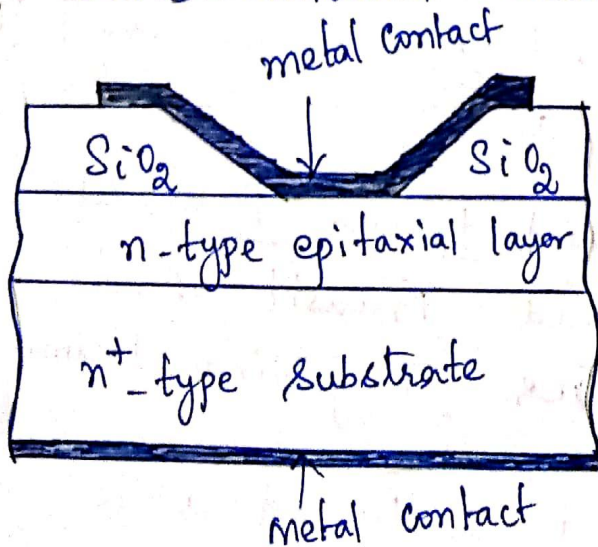
When aluminium or platinum metal is joined with N-type semiconductor, a junction is formed between the metal and N-type semiconductor. This junction is known as a metal - semiconductor junction. This junction creates a barrier or depletion layer known as Schottky barrier.

Symbol



→ In Schottky diode, the metal acts as the anode and n-type semiconductor acts as the cathode.

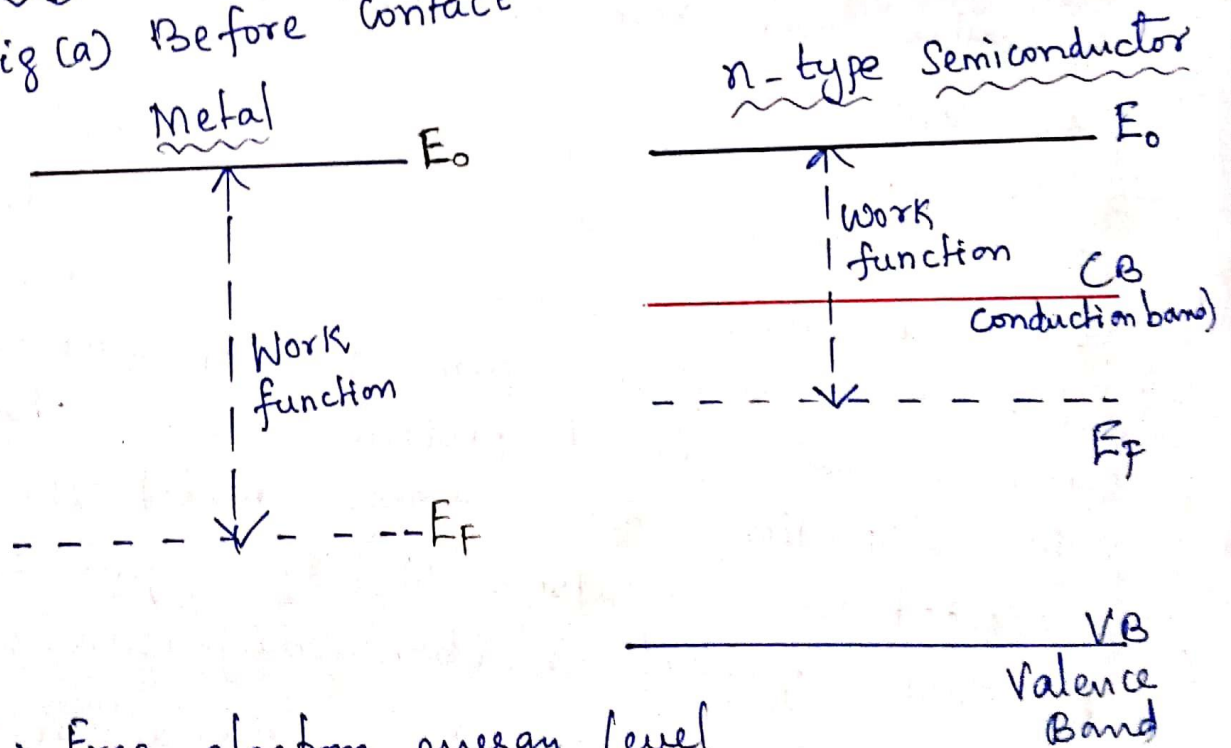
Cross-sectional view of Si Schottky diode



The Schematic diagram of a cross-sectional view of the Schottky diode with the corresponding circuit elements is shown in the above diagram.

→ The metal electrode (tungsten, aluminium, gold, etc.) is in contact with a weakly doped n-semiconductor layer, epitaxially grown on a highly doped n⁺ substrate. The dielectric is assumed to be ideal; that is, the conductance is zero.

Energy band diagram of Schottky diode:
 fig (a) Before contact



E_0 → Free electron energy level

E_F → Fermi level

VB → valence Band

CB → Conduction Band

The work function is defined as the energy required to move an electron from Fermi level E_F to free electron energy level, E_0 .

→ When a metallic electrode is brought in contact with an n-semiconductor, the more familiar behaviour of a PN-junction emerges; a small positive volume charge density is created in the semiconductor due to electron migration from the semiconductor to the metal.

→ This mechanism is due to the fact that the Fermi level is higher in the semiconductor (lower work function) than in the metal (higher work function) when the two materials are apart.

→ However, as both materials are contacted, the Fermi level again has to be the same and band distortions are created.

→ Electrons diffuse from the n-semiconductor and leave behind positive space charge.

→ The depletion zone grows until the electrostatic repulsion of the space charges prevents further electron diffusion.

The current-voltage characteristic is described by the following equation;

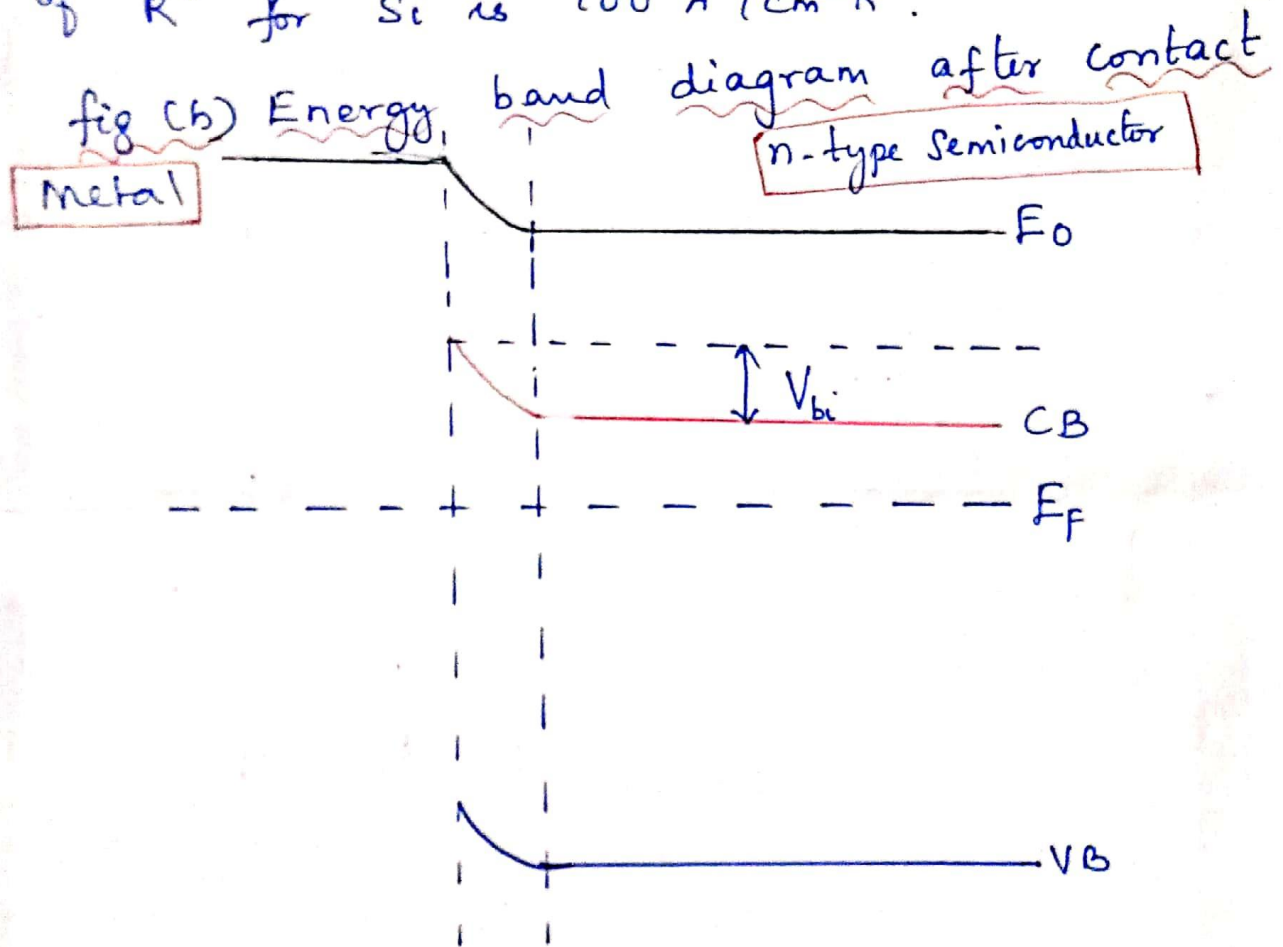
$$I = I_s \left(e^{\frac{(V_A - IR_s)}{V_T}} - 1 \right)$$

where, $V_T = \frac{q}{kT}$ is the thermal voltage.

The reverse-saturation current is given by

$$I_s = A \left(R^* T^2 e^{-\frac{qV_b}{kT}} \right) \text{ and } R^* \text{ is}$$

the so-called Richardson constant for thermionic emission of the majority carrier across the potential barrier. A typical value of R^* for Si is $100 \text{ A/cm}^2\text{K}^2$.

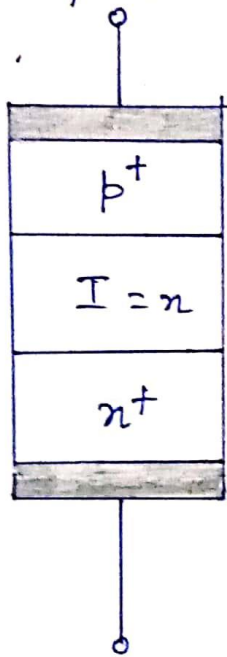


where, V_{bi} is the built-in-voltage for Schottky diode, which is given by the difference between the work functions of a metal and n-type semiconductor. The turn on voltage for Schottky diode is 0.2 to 0.3 volts. So applying a small voltage is enough to produce electric current in the Schottky diode.

PIN Diode:

→ Contains an additional layer of an intrinsic (I-layer) or lightly doped semiconductor sandwiched between highly doped p^+ and n^+ layers.

Simplified Structure of a PIN diode



→ Depending upon application and frequency range, the thickness of the middle layer ranges from 1 to $200\mu\text{m}$, sometimes even more.

→ Under forward bias, the diode behaves as if it possesses a variable resistance controlled by the applied current.

→ Under reverse bias, the lightly doped inner layer creates space charges, whose extent reaches the highly doped outer layers.

→ For instance, a Si-based PIN diode with an internal I-layer of $20\mu\text{m}$ and a surface area of 200 by $200\mu\text{m}$ has a diffusion capacitance on the order of 0.2PF .

→ In the forward direction, and for a weakly doped n-type intrinsic layer, the current through the diode is

$$I = A \left(\frac{q n_i^2 W}{N_D \tau} \right) \left(e^{\frac{V_A}{2V_T}} - 1 \right) \rightarrow \textcircled{1}$$

where, W is the width of the intrinsic layer, τ is the excess minority carrier lifetime, which can be on the order of up to $\tau = 10 \mu s$, N_D is the doping concentration in the middle layer of the lightly doped n-semiconductor.

→ The factor 2 in the exponent takes into account the presence of two junctions.

→ For a pure intrinsic layer, $N_D = n_i$, eqn $\textcircled{1}$ leads to the form,

$$\begin{aligned} \textcircled{1} \Rightarrow I &= A \left(\frac{q n_i^2 W}{n_i \tau} \right) \left(e^{\frac{V_A}{2V_T}} - 1 \right) \\ &= A \left(\frac{q n_i W}{\tau} \right) \left(e^{\frac{V_A}{2V_T}} - 1 \right) \rightarrow \textcircled{2} \end{aligned}$$

→ The total charge can be calculated from the relation, $Q = I \tau$

$$\begin{aligned} \therefore \text{Diffusion capacitance, } C_d &= \frac{dQ}{dV_A} \\ &= \frac{d(I \tau)}{dV_A} = \tau \left(\frac{dI}{dV_A} \right) \end{aligned}$$

Differentiating eqn (2) w.r to V_A , we get

$$\begin{aligned} \frac{dI}{dV_A} &= A \left(\frac{q n_i w}{\gamma} \right) \left(e^{\frac{V_A}{2V_T}} - 1 \right) \\ &= A \left(\frac{q n_i w}{\gamma} \right) \left(e^{\frac{V_A}{2V_T}} - 1 \right) \left(\frac{1}{2V_T} \right) \\ &= I \left(\frac{1}{2V_T} \right) \end{aligned}$$

$$\Rightarrow C_d = \gamma \left(\frac{dI}{dV_A} \right)$$

$$C_d = \gamma \left(\frac{I}{2V_T} \right) \rightarrow (3)$$

→ In the reverse direction, the capacitance is dominated by the parallel plate capacitance of the depleted I-layer; the capacitance, C_J is approximately,

$$C_J = \frac{\epsilon_I A}{w}; \text{ where } \epsilon_I \text{ is the dielectric constant of the intrinsic layer, } w \text{ is the width of the depletion layer.}$$

→ The RF resistance of a PIN diode is found by treating the I-layer as a cylindrical conductor with cross-sectional area A and length w .

$$\begin{aligned} \Rightarrow R_J (I_R) &= \frac{\rho l}{A} & \rho &= \frac{1}{\sigma}; l = w \\ &= \frac{w}{\sigma A} = \frac{w}{qP (n_n + p_p) A} \end{aligned}$$

where, σ , the conductivity can be approximated as $\sigma = qP(\mu_n + \mu_p)$

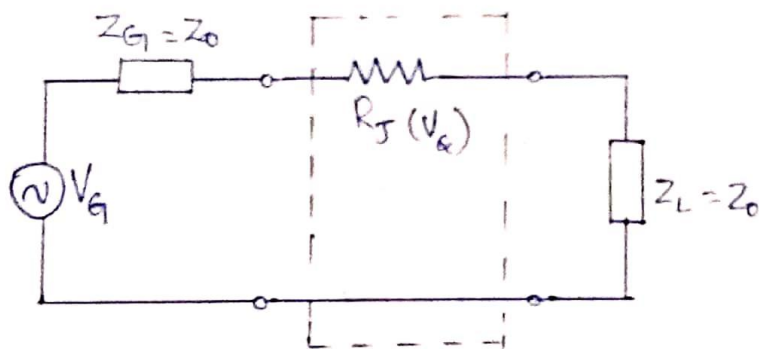
$$R_J = \frac{W^2}{(\mu_n + \mu_p) q I_a}$$

where, I_a is the bias current and $p \approx n$.

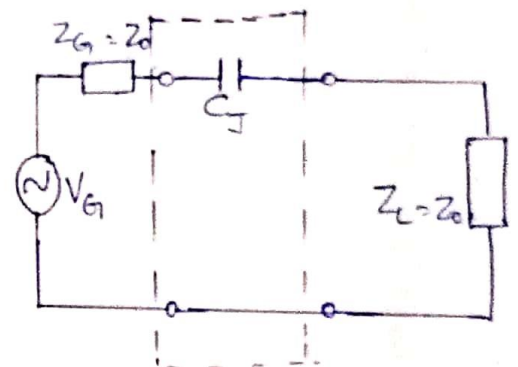
→ Under forward bias ("switch on"), behaviour of the PIN diode is resistive

→ Under reverse bias ("switch off") or isolation, behaviour of the PIN diode is capacitive.

PIN diode in Series connection



(a) Forward bias



(b) Reverse bias (isolation)

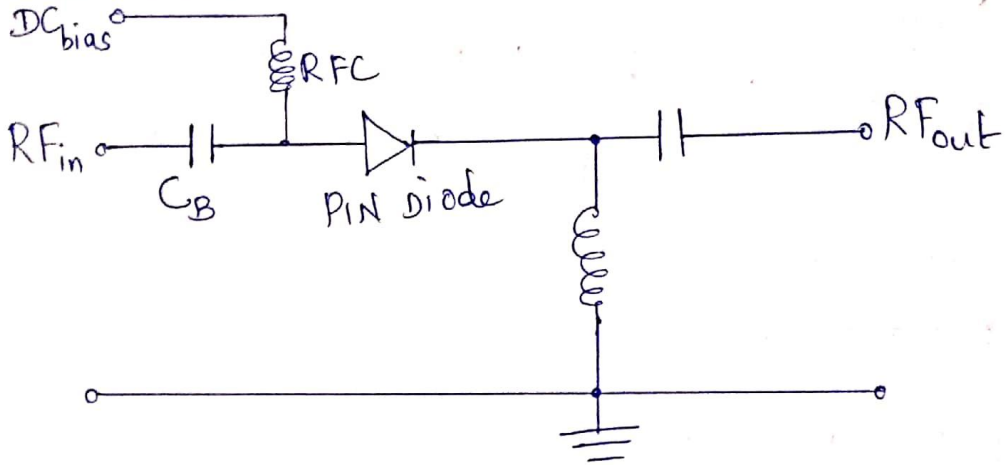
→ In order to operate the PIN diode, the bias point must be set. This has to be provided through a DC circuit that must be separated from the RF signal path.

→ The DC isolation is achieved by a radio frequency choke (RFC), representing a short circuit at DC and an open circuit at high frequency.

→ Blocking capacitors (C_B) represent an open circuit at DC and a short circuit at RF.

Attenuator circuit with biased PIN diode in series and shunt configurations.

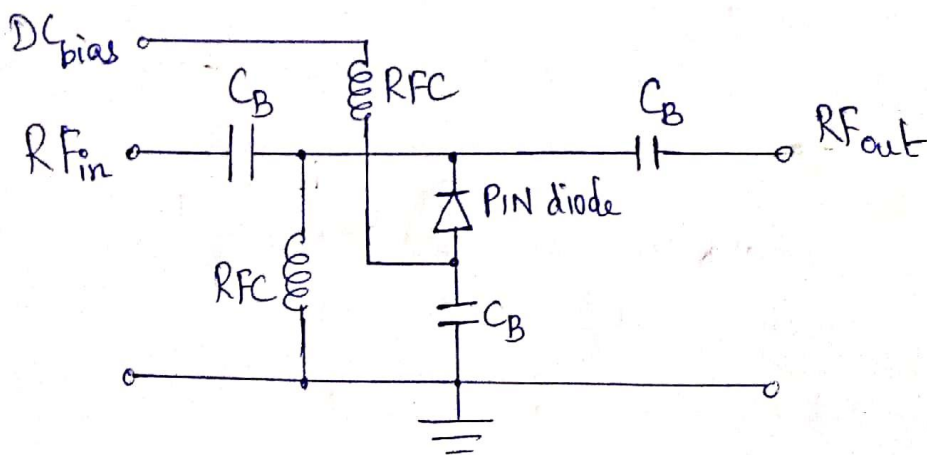
(a) Series Connection of PIN diode



→ For Positive DC bias voltage, the series-connected PIN diode represents a low resistance to the RF signal.

→ For a negative bias condition, where the series-connected PIN diode behaves like a capacitor with high impedance or high insertion loss.

(b) Shunt Connection of PIN diode



→ For positive DC bias voltage, the shunt-connected PIN diode creates a short-circuit condition, permitting only a negligibly small RF signal to appear at the output port. The shunt connection acts like a high attenuation device with high insertion loss.

→ For a negative bias condition, the shunt-connected diode with a high shunt impedance does not affect the RF signal appreciably.

Transducer loss TL in dB is given

$$\text{By, } TL, \text{ dB} = -20 \log |S_{21}|$$

— x —

S_{21} = attenuation of the wave traveling from port 1 to port 2.

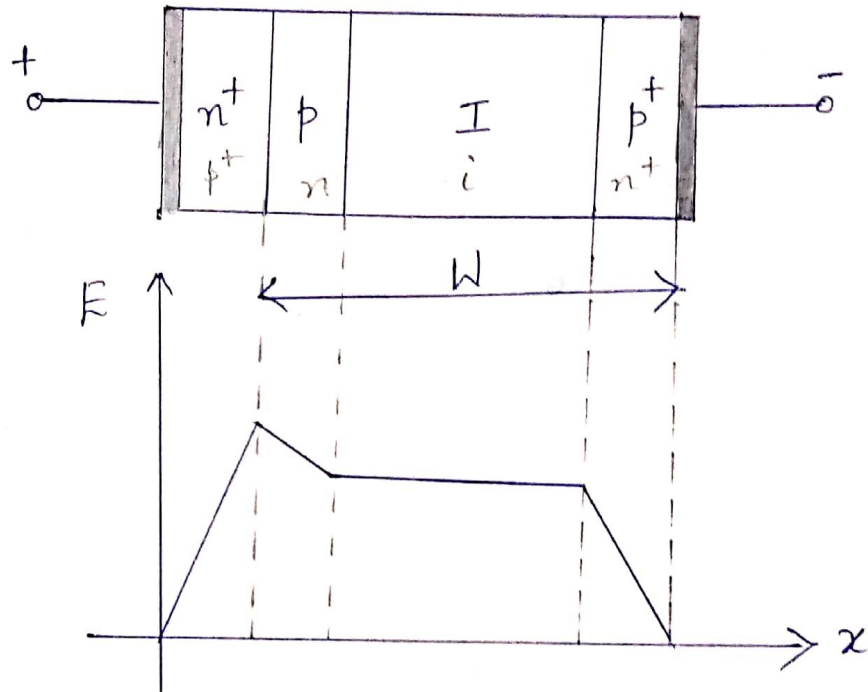
IMPATT Diode :

IMPATT stands for **IMP**act **A**valanche and **ATT** Transit Time diode. It uses the avalanche effect as originally proposed by Read.

→ The key difference between PIN diode and IMPATT diode is the generation of high electric field strength at the interface between the n^+ and p layer,

which results in an avalanche of carriers through impact ionization.

fig(i) Layer structure and electric field profile

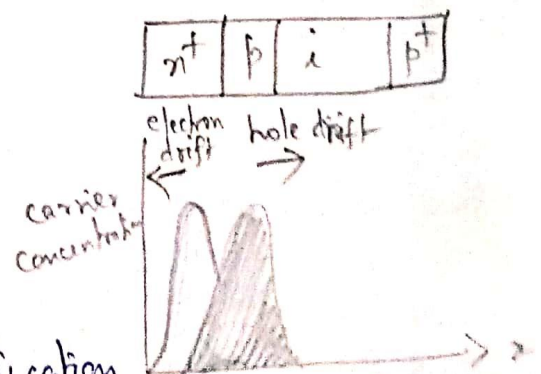


→ For the Purpose of microwave signal generation, there are different types of IMPATT diode structures that can be used.

- (i) $n^+ - p - i - p^+$ and
- (ii) $p^+ - n - i - n^+$

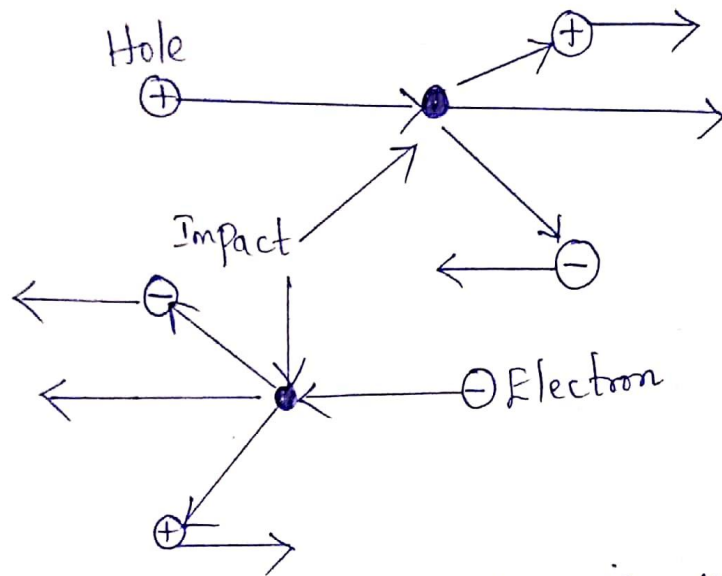
→ There are two distinct spatially separated regions

- (i) the avalanche multiplication region (the forward-biased $n^+ - p$ junction)
- and (ii) the carrier drift region (the i -region which becomes completely depleted at high reverse voltages)



→ The operation of IMPATT diode is explained using the structure n^+p-i-p^+ .

fig(ii) Impact ionization



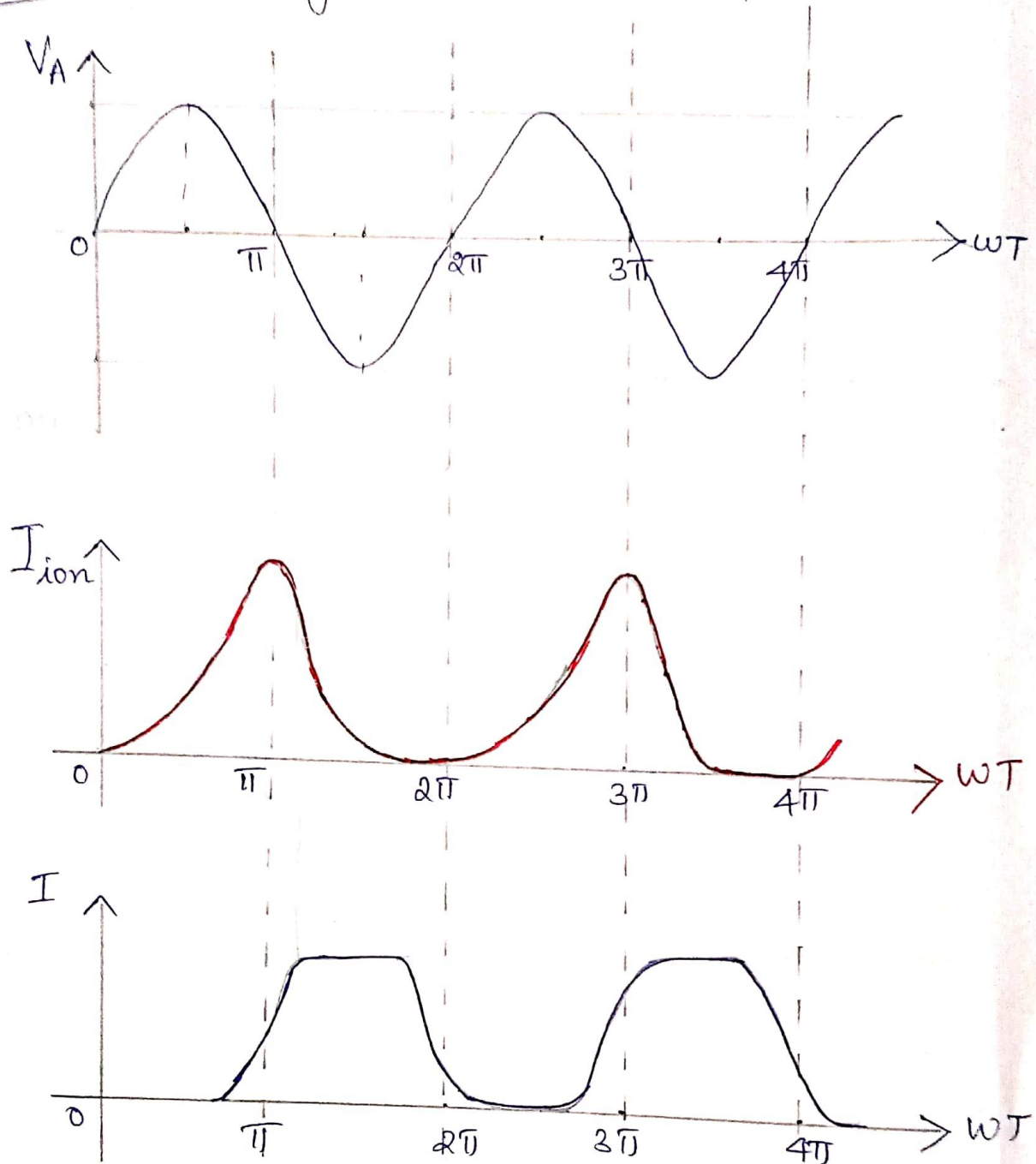
→ The avalanche region is where the injection of carriers (the rapid increase in the number of free charges) occurs.

→ Due to maximally high electric field intensity across the junction, the newly formed electrons are accelerated into the heavily doped n^+ region resulting in impact ionization which, in turn, leads to avalanche carrier multiplication.

→ The holes, on the other hand, are forced through the carrier drift region, where they drift in a constant electric field and then into the p^+ region.

→ Let, V_A be the applied RF voltage. It produces an electric field and when it exceeds the critical threshold level, the additional ionization current I_{ion} is generated.

Fig(m) Applied Voltage, ionization current, and total current



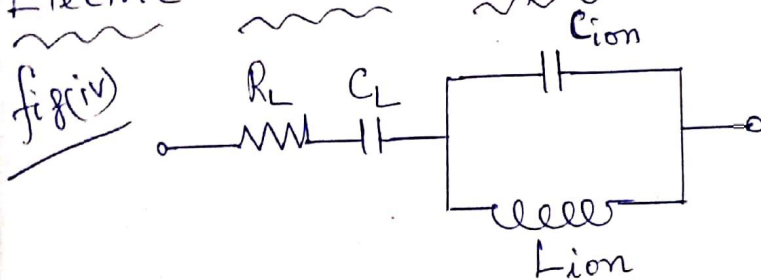
This ionization current I_{ion} slowly decreases during the negative voltage cycle as the excess carriers are removed. The phase shift between this ionization current and the applied voltage can be tailored so as to reach 90° .

→ The total diode current suffers an additional delay since the excess carriers must travel through the intrinsic layer to the p^+ layer.

→ The time constant is dependent on the length and drift velocity.

→ An additional time delay of 90° be created by choosing the intrinsic layer length appropriately, in conjunction with a suitable doping concentration.

Electric circuit representation for the IMPATT diode



→ Reactance is inductive below the diode's resonance frequency f_0 and is capacitive above the resonant frequency.

→ The total resistance is positive for $f < f_0$ and becomes negative for $f > f_0$.

→ The resonance frequency is determined based on the operating current I_0 , dielectric constant, saturation drift velocity v_{max} , and the differential change in the

ionization co-efficient α with respect to the differential change in electric field strength (ie) $\alpha' = \frac{\partial \alpha}{\partial E}$. The resonance frequency is predicted as, $f_0 = \frac{1}{2\pi} \sqrt{2I_q \frac{W d_{max}}{\epsilon} \alpha'}$

The additional circuit parameters are specified as follows:

$$R = R_L + \frac{W d_{max}}{2\pi^2 f_0^2 C_L W \left[1 - \left(\frac{f}{f_0} \right)^2 \right]}$$

$$C_L = \frac{\epsilon A}{W} \quad \text{and} \quad C_{ion} = \frac{\epsilon A}{d}$$

$$\text{and } L_{ion} = \frac{1}{(2\pi f_0)^2 C_{ion}}$$

where R_L is the combined resistance of the semiconductor layers, d is the length of the avalanche region of the p-layer, and W is the total length, as shown in fig(i).

→ The negative resistance of this diode above the resonance frequency can be understood in terms of returning electric energy to the RF or microwave resonance circuit, which means the diode operates as an active device.

→ The efficiency of converting DC to RF power at operating frequencies of 5 to 10 GHz is very low, with typical values in the range of 10 to 15%.

Transferred Electron devices (TEDs)

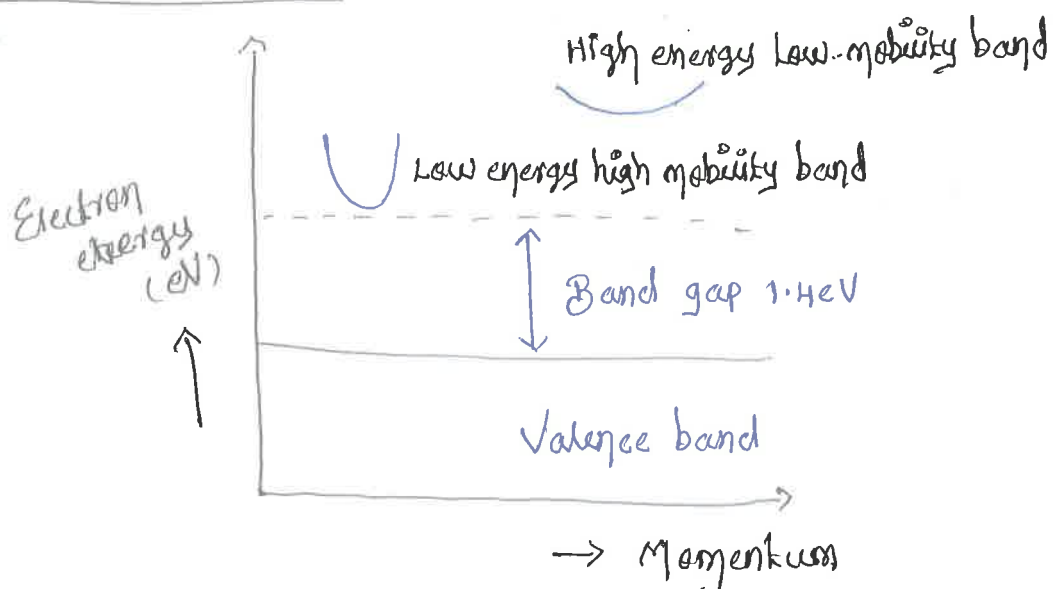
Gunn diode :-

It Make use of 2 terminal devices based on the Phenomenon known as "Transferred electron effect."

Transferred electron effect:-

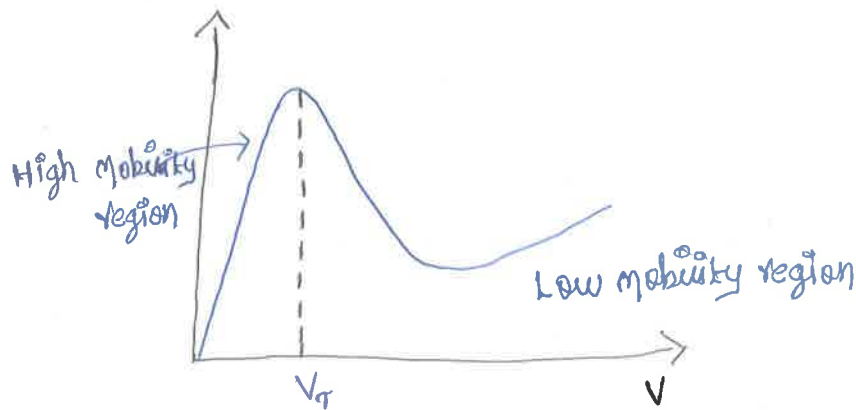
Some materials like GaAs exhibit a Negative differential mobility when biased above a threshold value of the electric field. The electrons in the lower-energy band will be transferred into the higher-energy band. The behaviour is called Transferred electron effect and the device is also called Transferred electron devices (or) Gunn diode.

Energy band diagram:-



↳ In the high energy band the effective electron mass is larger and hence the electron mobility is lower than what it is in the low-energy band.

↳ The conductivity is directly proportional to the mobility, there is a immediate range of electric field strengths for which the fraction of electrons that are transferred into the high energy low mobility conduction band is such that the average mobility, and hence conductivity, decreases with an increase in electric field strength.



↳ Applications

- ① Low power oscillator at microwave frequencies in transmitters
- ② Local oscillator at receiver front ends.

↳ Fabrication Materials

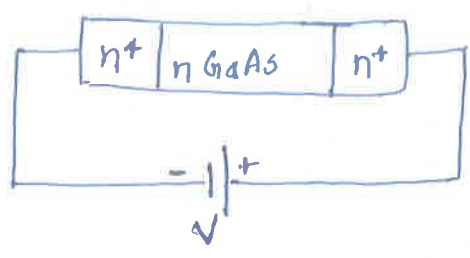
- ① Gallium arsenide (GaAs)
- ② Indium phosphide (InP)
- ③ Cadmium Telluride (CdTe)

↳ Positive resistances absorb power (passive devices), where negative resistances generate power (active devices).

Difference between Microwave Transistors and Transferred Electron devices (TEDs)

- 1). TEDs are bulk devices having no junction (or) gates as compared to microwave transistors which operate with either junction or gates.
- 2). The majority of transistors are fabricated from elemental semiconductors, such as silicon (or) germanium, whereas TEDs are fabricated from compound semiconductors, such as GaAs, InP, CdTe.
- 3). TEDs operate with hot electrons whose energy is very much greater than the thermal energy. Transistors operate with warm electrons whose energy is not much greater than their thermal energy (0.026 eV at room temperature).

Gunn diodes :- (GaAs diode)

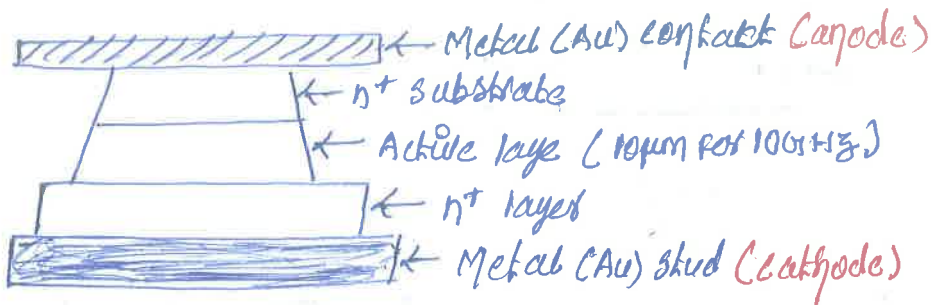


↳ It consists of n-type GaAs semiconductor with regions of high doping (n^+). Even though there is no junction this is called diode with reference to the positive end (anode) and negative end (cathode) of the dc voltage applied across the device.

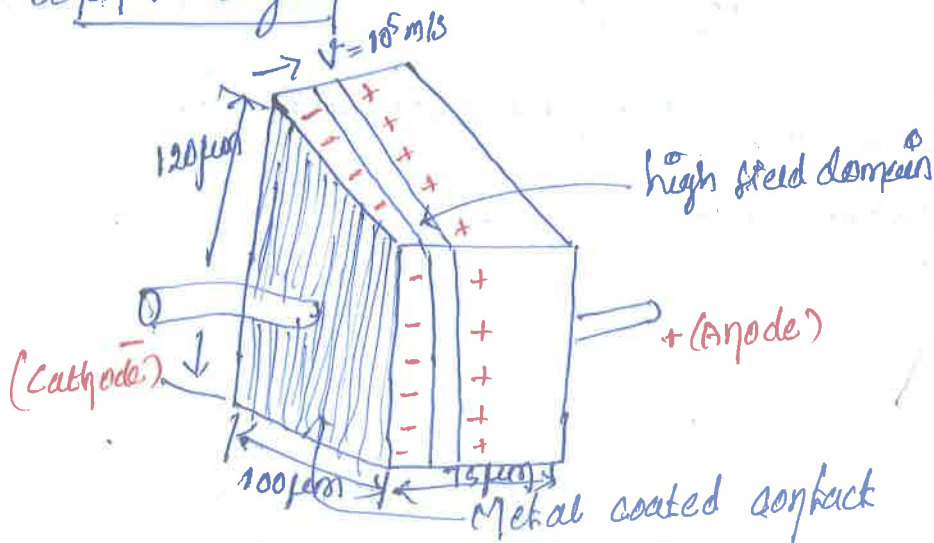
Gunn effect :-

Above some critical voltage, corresponding to an electric field of 2000-4000 Volts/cm, the current in every specimen became a fluctuating function of time. In the GaAs specimens, this fluctuation took the form of a periodic oscillation superimposed upon the pulse current. The frequency of oscillation was determined mainly by the specimen, and not by the external circuit. The period of oscillation was usually inversely proportional to the specimen length and closely equal to the transit time of electrons between the electrodes.

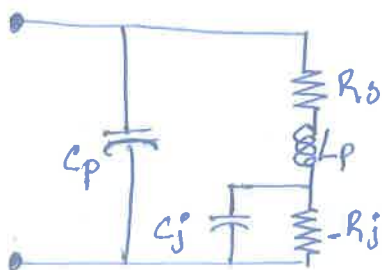
Construction :-



Schematic diagram :-



Equivalent circuit :-



where, C_j - diode capacitance
 R_j - diode resistance
 L_p - package inductance
 C_p - package capacitance
 R_0 - Total resistance of leads, ohmic contact, bulk resistance of diode.

Negative resistance :-

The carrier drift velocity is linearly increased from zero to a maximum when the electric field is varied from zero to a threshold value. When the electric field is beyond the threshold value of 3000V/cm, the drift velocity is decreased and the diode exhibits negative resistance.

Ridley - Watkins - Hixson (RWH) Theory:-

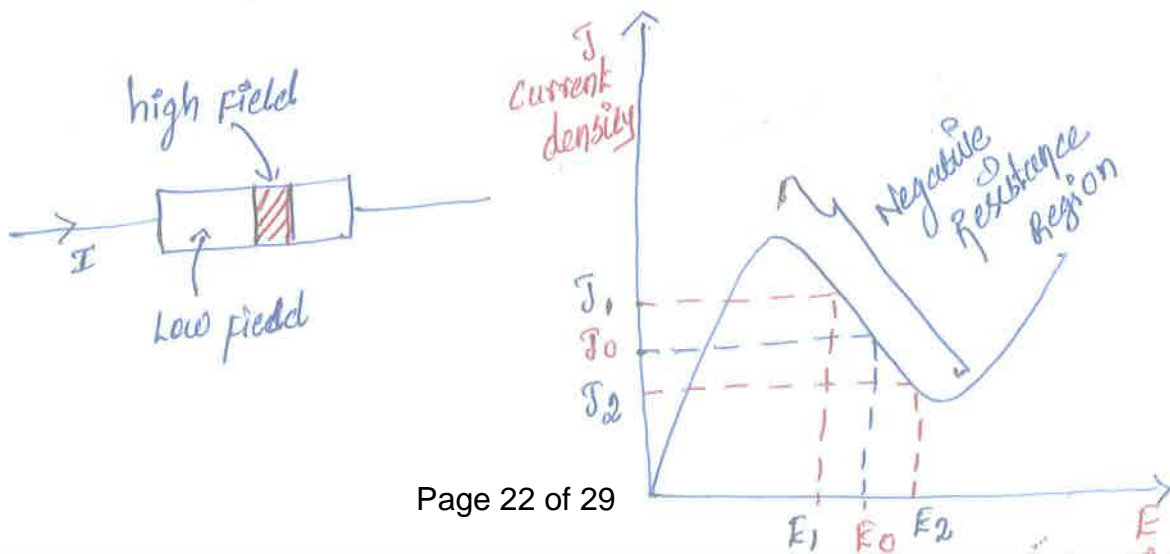
↳ The fundamental concept of RWH theory is the differential negative resistance developed in a bulk [solid state III-V compound] compound semiconductor when electric field is applied to the terminals of the sample.

↳ There are 2 modes of Negative resistance devices

- ① Voltage controlled modes
- ② Current controlled modes.

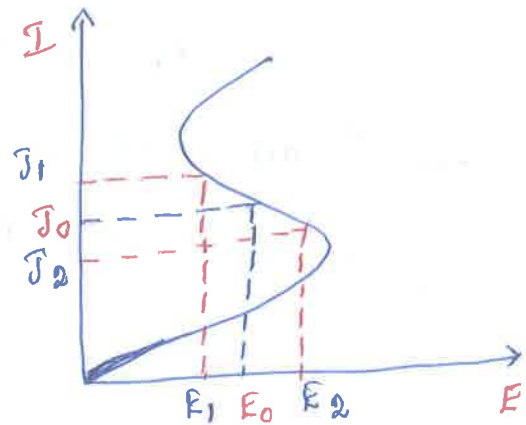
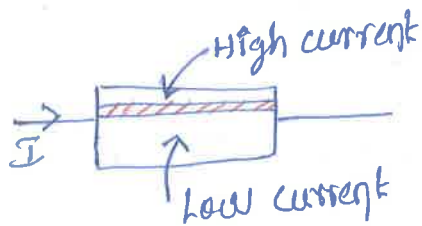
① Voltage controlled mode:-

- ↳ In this mode, the current density can be multivalued.
- ↳ Here, high field domain are formed, separating low-field regions.
- ↳ The interfaces separating low and high field domain lie along equipotentials; thus they are in planes perpendicular to the current direction.



② Current controlled mode:-

- ↳ In this mode, Voltage value can be multivalued.
- ↳ Here, the mode splitting the sample results in high current filaments running along the field direction.



Effect of Negative region (resistance):-

* current density - electric field is to render the sample (specimen) electrically **unstable**.

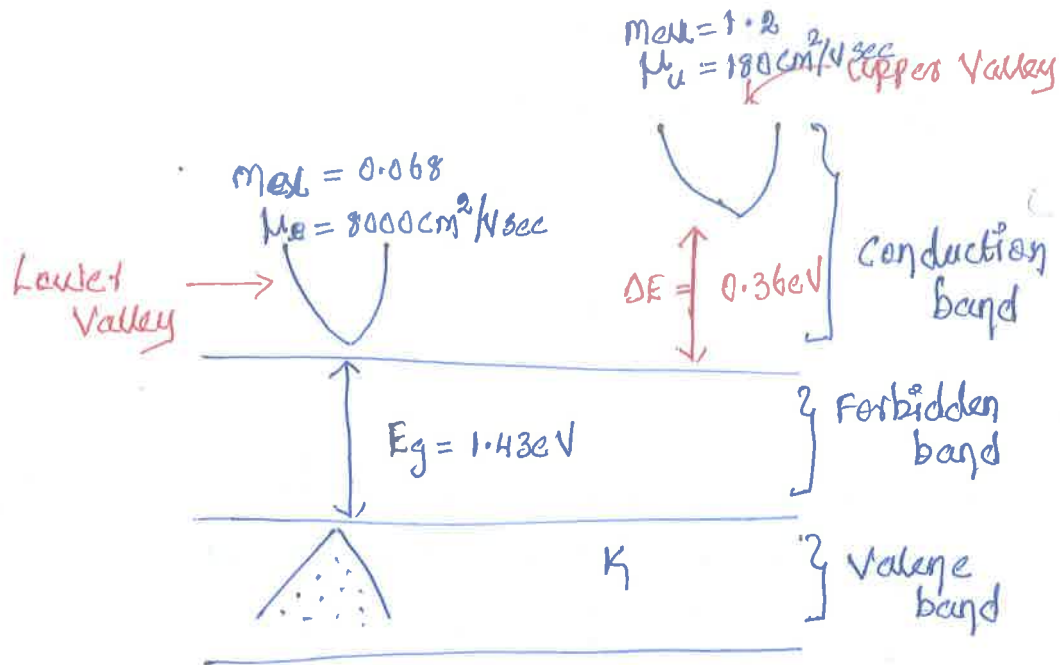
$$\text{Differential Negative resistance} \Rightarrow \frac{dI}{dV} = \frac{dJ}{dE}$$

Note:-

Voltage controlled mode $\Rightarrow E_0 \rightarrow$ sample $\rightarrow J_0$ generated
 \uparrow es $E_2 \rightarrow$ " $\rightarrow J_2$ \downarrow es } inversely proportional $\Rightarrow -ve R$
 \downarrow es $E_1 \rightarrow$ " $\rightarrow J_1$ \uparrow es }

* $-ve R$ generates power
 $+ve R$ dissipates the power

Two-Valley Model Theory :-



→ To design two valley model should consider the following parameters and specification. For eg

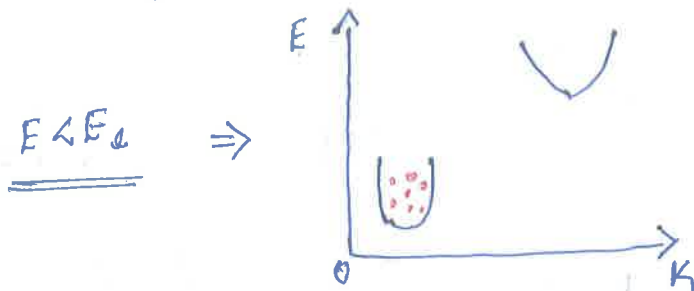
- High in Lower Valley ⇐ (1) Mobility ⇒ separation b/w bottom of Lower and Upper Valley
 Low in upper valley (2) Effective mass
 (3) Density of carrier
- e^- in Lower & upper remain same at equilibrium condition
 Greater than thermal energy (0.025 eV)
 Less in Lower Valley and high in upper valley

eg. Data for 2 Valley in GaAs

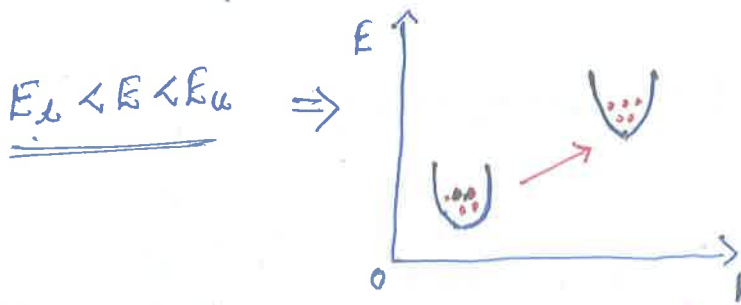
Valley	Effective mass, m_e	Mobility, μ	Separation
Lower	$m_{eL} = 0.068$	$\mu_{eL} = 8000 \text{ cm}^2/\text{V-sec}$	} $\Delta E = 0.36 \text{ eV}$
Upper	$m_{eU} = 1.2$	$\mu_{eU} = 180 \text{ cm}^2/\text{V-sec}$	

↳ Electron densities in the lower and upper valleys remain the same under an equilibrium condition.

↳ When the applied electric field is lower than the electric field of the lower valley ($E < E_L$), no electrons will transfer to the upper valley.



↳ When the applied electric field is higher than that of lower valley and lower than that of the upper valley ($E_L < E < E_U$), electrons will begin to transfer to the upper valley.



↳ When the applied electric field is higher than that of the upper valley ($E_U < E$), all electrons will transfer to the upper valley.



* conductivity of n-type GaAs is,

$$\sigma = e (\mu_L n_L + \mu_U n_U)$$

↳ electron charge = e
electron mobility = μ

where, $n = n_L + n_U$
↳ lower e^- density & higher

On the basis of RWH theory, the band structure of a semiconductor must satisfy 3 criteria in order to exhibit negative resistance.

- ① The separation energy between the bottom of the lower valley and the bottom of the upper valley must be several times larger than the thermal energy at room temperature.

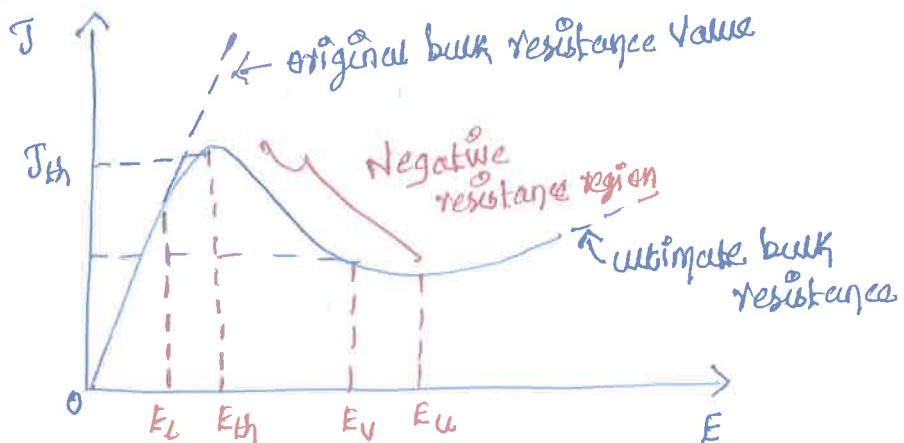
$$\Delta E > kT \quad (\text{or}) \quad \Delta E > 0.026 \text{ eV}$$

- ② The separation energy between the valleys must be smaller than the gap energy between the conduction and valence bands.

$$\Delta E < E_g$$

If $\Delta E > E_g$, semiconductor will break down.

- ③ Electrons in the lower valley must have high mobility, smaller effective mass, and a low density of state, whereas those in the upper valley must have low mobility, large effective mass, and a high density of states.



Classification of Modes:-

↳ The Gunn-effect diodes are basically made from an n-type GaAs, with the concentrations of free electrons ranging from 10^{14} to 10^{17} per cubic cm at room temperature. Its typical dimensions are $150 \times 150 \mu\text{m}$ in cross section and $30 \mu\text{m}$ long.

↳ During the early stages of space-charge accumulation, the time rate of growth of the space-charge layers is given by,

$$Q(x, t) = Q(x - vt, 0) \exp\left(\frac{t}{\tau_d}\right)$$

where, $\tau_d = \frac{\epsilon}{\sigma} \Rightarrow \frac{\epsilon}{en_0|\mu_n|}$

↳ magnitude of -ve dielectric relaxation time

ϵ = semiconductor dielectric permittivity

n_0 = doping concentration

μ_n = negative mobility

e = electron charge

σ = conductivity

↳ It includes 4 basic modes of operation such as,

- ① Gunn oscillation mode
- ② stable amplification mode
- ③ LSA oscillation mode
- ④ Bias-circuit oscillation mode

① Gunn Oscillation mode:-

This mode is defined in the region where the product of frequency multiplied by length is about 10^7 cm/s and the product of doping is multiplied by length is greater than $10^{12}/\text{cm}^2$. In this region the device is unstable because of the cyclic formation of either the accumulation layer (or) high field domain.

② Stable Amplification mode:-

This mode is defined in the region where the product of frequency times length is about 10^7 cm/s and the product of doping times length is between 10^{11} and $10^{12}/\text{cm}^2$.

③ LSA Oscillation mode:- (Limited space-charge accumulation)

This mode is defined in the region where the product of frequency times length is above 10^7 cm/s and the quotient of doping divided by frequency is between 2×10^4 and 2×10^5 .

④ Bias-circuit Oscillation mode:-

This mode occurs only when there is either Gunn (or) LSA mode, and it is usually at the region

where the product of frequency times length is too small. When a bulk diode is biased to threshold, the average current drops as Gunn oscillation begins, typically 1kHz to 100MHz (oscillation).

EC8701

**ANTENNAS AND MICROWAVE
ENGINEERING**

UNIT V

MICROWAVE DESIGN PRINCIPLES

- Impedance transformation
- Impedance Matching
- Microwave Filter Design
- RF and Microwave Amplifier Design
- Microwave Power amplifier Design
- Low Noise Amplifier Design
- Microwave Mixer Design
- Microwave Oscillator Design

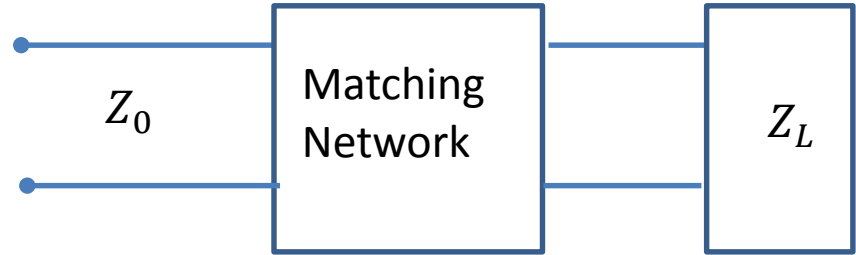
- L-section impedance matching
- Single and double stub matching
- Quarter wave transformer

Matching Network

An impedance matching network is placed between a load impedance and a transmission line.

The matching network is ideally lossless, to avoid any loss of power

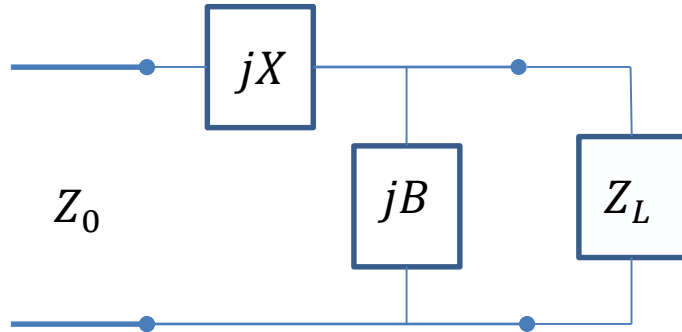
It is designed in such a way that the impedance seen looking into the matching network is Z_0 .



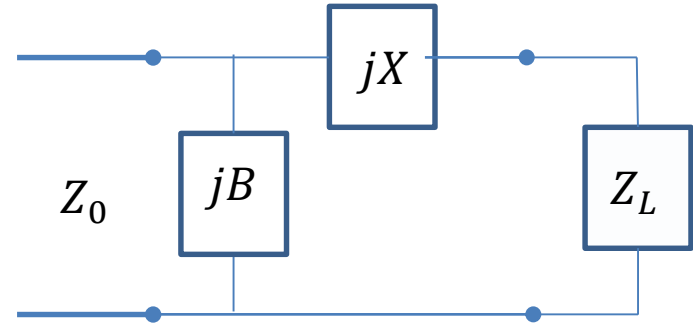
A matching network can be found as long as Z_L has positive real part
Factors that are considered while selecting a particular matching network are:
Complexity, bandwidth, Implementation and adjustability

L-section impedance matching network

Uses two reactive elements to match an arbitrary load to a transmission line



Used when $z_L = Z_L/Z_0$ is inside the $1 + jx$ circle in the smith chart



Used when $z_L = Z_L/Z_0$ is outside the $1 + jx$ circle in the smith chart

L-section impedance matching network

For this case $R_L > Z_0$

For impedance matching, we must have

$$Z_0 = jX + \frac{1}{jB + 1/(R_L + jX_L)}$$

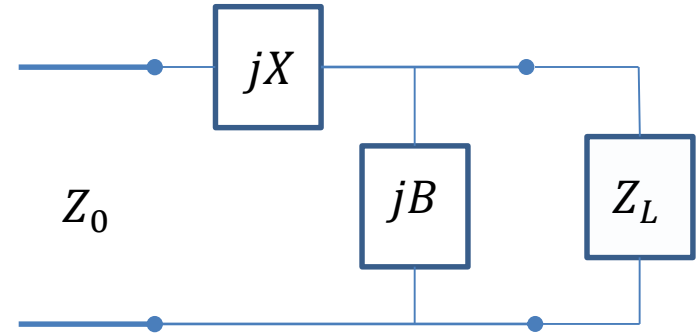
$$Z_0 = jX + \frac{R_L + jX_L}{jBR_L - BX_L + 1}$$

$$\begin{aligned} Z_0(jBR_L - BX_L + 1) \\ = jX(jBR_L - BX_L + 1) + R_L + jX_L \end{aligned}$$

Equating the real part from both sides

$$-Z_0BX_L + Z_0 = -XBR_L + R_L$$

$$B(XR_L - Z_0X_L) = R_L - Z_0$$



Used when $z_L = Z_L/Z_0$ is inside the $1 + jx$ circle in the smith chart

L-section impedance matching network

$$Z_0(jBR_L - BX_L + 1) = jX(jBR_L - BX_L + 1) + R_L + jX_L$$

Equating the imaginary parts we can write

$$\begin{aligned} Z_0BR_L &= -XBX_L + X + X_L \\ \Rightarrow X(1 - BX_L) &= Z_0BR_L - X_L \end{aligned}$$

We therefore have a set of two equations

$$B(XR_L - Z_0X_L) = R_L - Z_0$$

and

$$X(1 - BX_L) = Z_0BR_L - X_L$$

from which B and X are to be determined

L-section impedance matching network

From $X(1 - BX_L) = Z_0BR_L - X_L$,

$$X = \frac{Z_0BR_L - X_L}{(1 - BX_L)}$$

Substituting X in $B(XR_L - Z_0X_L) = R_L - Z_0$

$$Z_0B^2R_L^2 - BR_LX_L - BZ_0X_L + B^2Z_0X_L^2 = (R_L - Z_0) - (R_L - Z_0)BX_L$$

$$B^2Z_0(R_L^2 + X_L^2) - 2BZ_0X_L - (R_L - Z_0) = 0$$

L-section impedance matching network

$$B^2 Z_0 (R_L^2 + X_L^2) - 2B Z_0 X_L - (R_L - Z_0) = 0$$

$$B = \frac{2Z_0 X_L \pm \sqrt{4Z_0^2 X_L^2 + 4Z_0 (R_L^2 + X_L^2)(R_L - Z_0)}}{2Z_0 (R_L^2 + X_L^2)}$$

$$B = \frac{X_L \pm \sqrt{X_L^2 + (R_L^2 + X_L^2)(R_L/Z_0 - 1)}}{(R_L^2 + X_L^2)}$$

$$B = \frac{X_L \pm \sqrt{R_L/Z_0} \sqrt{R_L^2 + X_L^2 - R_L Z_0}}{(R_L^2 + X_L^2)}$$

L-section impedance matching network

Since we use the matching network for $R_L > Z_0$, the term $R_L^2 + X_L^2 - R_L Z_0$ is always positive and therefore there exist a real valued solution for B .

Once B is calculated, X can be calculated from

$$X = \frac{Z_0 B R_L - X_L}{(1 - B X_L)}$$

Example: L-section matching

Let an impedance of $Z_L = (100 - j50)\Omega$ is to be matched to a 50Ω line using a L-section matching network at an operating frequency of 500 MHz. Let us design the matching network.

We have

$$B = \frac{X_L \pm \sqrt{R_L/Z_0} \sqrt{R_L^2 + X_L^2 - R_L Z_0}}{(R_L^2 + X_L^2)}$$
$$B = \frac{-50 \pm \sqrt{100/50} \sqrt{100^2 + 50^2 - 100 \times 50}}{(100^2 + 50^2)} = \begin{cases} 0.0058 \Omega^{-1} \\ -0.0138 \Omega^{-1} \end{cases}$$

$$X = \frac{Z_0 B R_L - X_L}{(1 - B X_L)} = \begin{cases} 61.2372 \Omega \\ -61.2372 \Omega \end{cases}$$

Example: L-section matching

We have two solutions which are as follows:

Solution 1

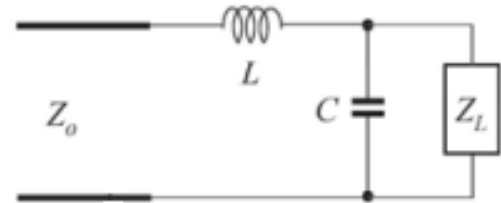
$$C = \frac{B}{2\pi f} = 1.85 \text{ pF}$$

$$L = \frac{X}{2\pi f} = 19.49 \text{ nH}$$

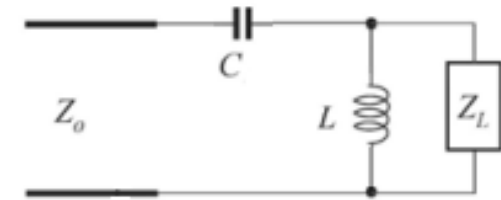
Solution 2

$$C = -\frac{1}{2\pi f X} = 5.2 \text{ pF}$$

$$L = -\frac{1}{2\pi f B} = 23.1 \text{ nH}$$



Network for solution 1



Network for solution 2

L-section matching

For the matching network as shown, we have

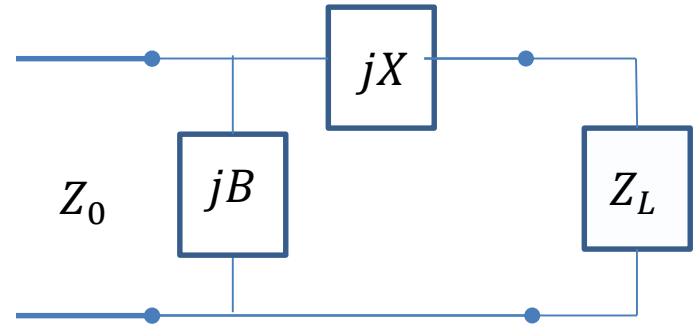
$$R_L < Z_0$$

$$\frac{1}{Z_0} = jB + \frac{1}{jX + R_L + jX_L}$$
$$\frac{1}{Z_0} = jB + \frac{1}{R_L + j(X + X_L)}$$

X and B can be found as

$$X = \pm \sqrt{R_L(Z_0 - R_L)} - X_L$$

$$B = \pm \frac{\sqrt{(Z_0 - R_L)/R_L}}{Z_0}$$



Used when $z_L = Z_L/Z_0$ is outside the $1 + jx$ circle in the smith chart

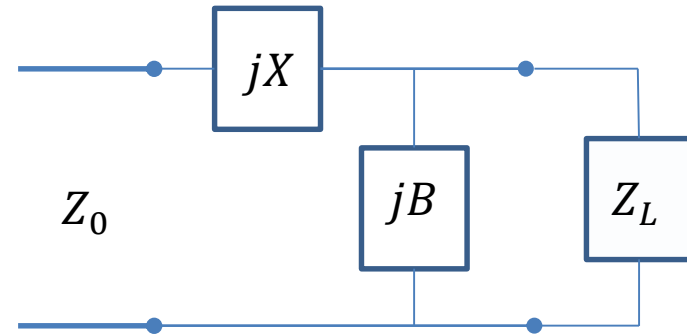
Example: L-section matching with Smith chart

Let us now consider an example how L-impedance can be done using a Smith chart.

Let us discuss the matching of a $100\ \Omega$ load with a transmission line of characteristic impedance $50\ \Omega$ at $100\ \text{MHz}$. We use Smith chart to do this matching.

When we consider the matching in Smith chart our starting point is normalized $z_L = 2 + j0$ and after matching we reach $z = 1 + j0$

Since $R_L > Z_0$, we use the following circuit



We mark $z_L = 2 + j0$ on the smith chart.

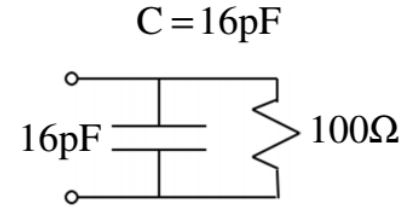
Since we need to add an admittance first, we first find $y = 0.5 + j0$ and add a mirror of $r = 1$ circle.

We add $jb = j0.5$ to reach rotated $r = 1$ circle

Therefore,

$$\frac{0.5}{50} = 2\pi \times 10^8 \times C$$

$$C = 16 \text{ pF}$$



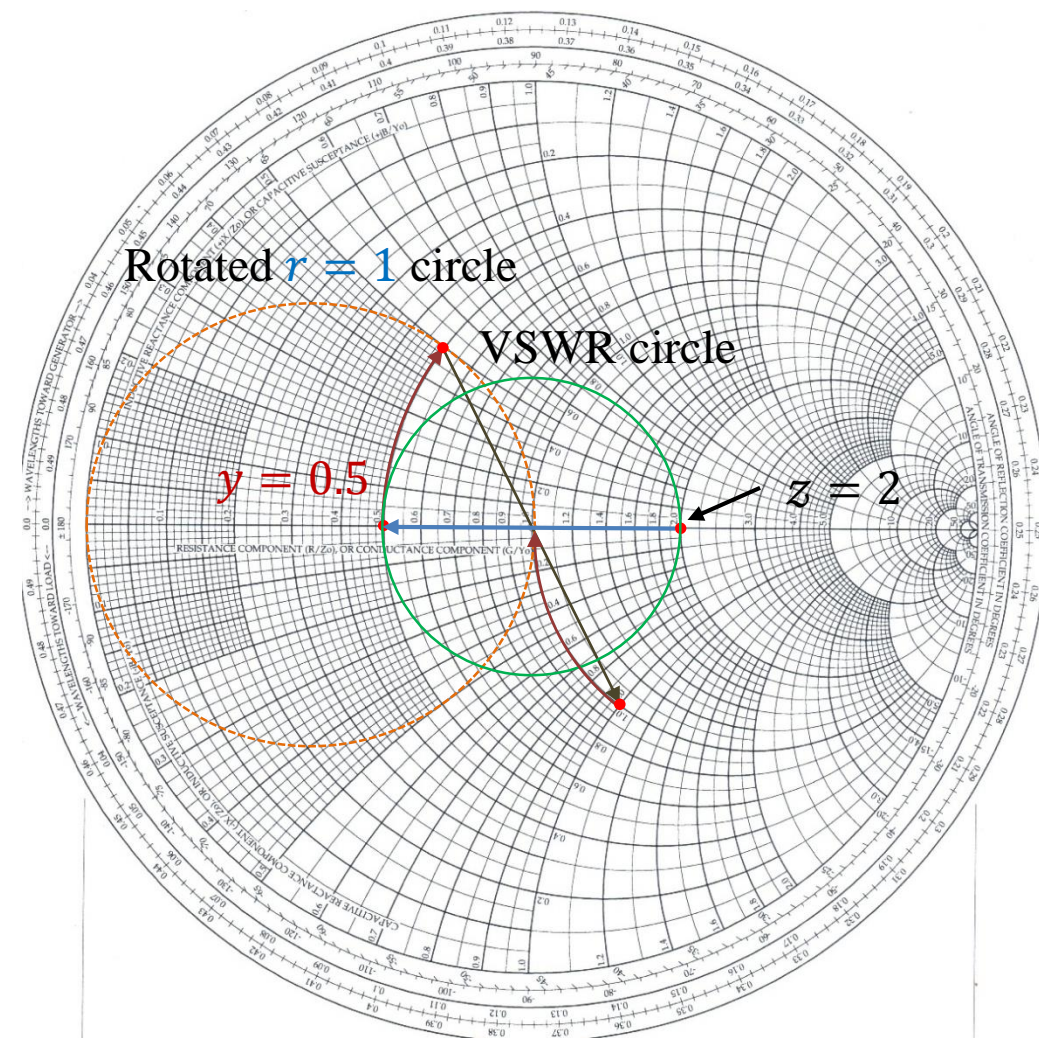
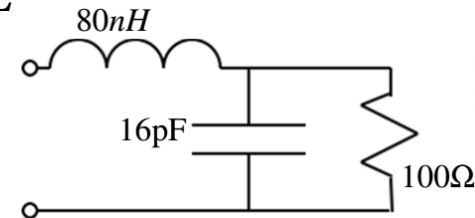
Having determined we come back to $r = 1$ circle and we land at the point $(1 - j1)$.

We add a reactance of $jx = j1$ to move to the centre of the Smith chart.

Therefore,

$$50 = 2\pi \times 10^8 \times L$$

$$L = 80 \text{ nH}$$

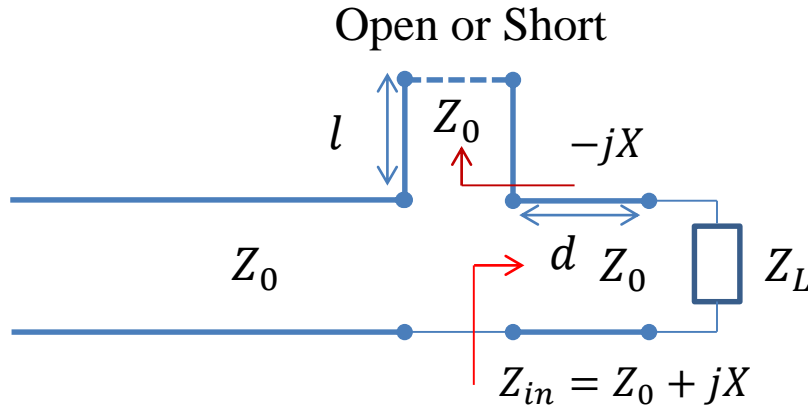


Stub Matching

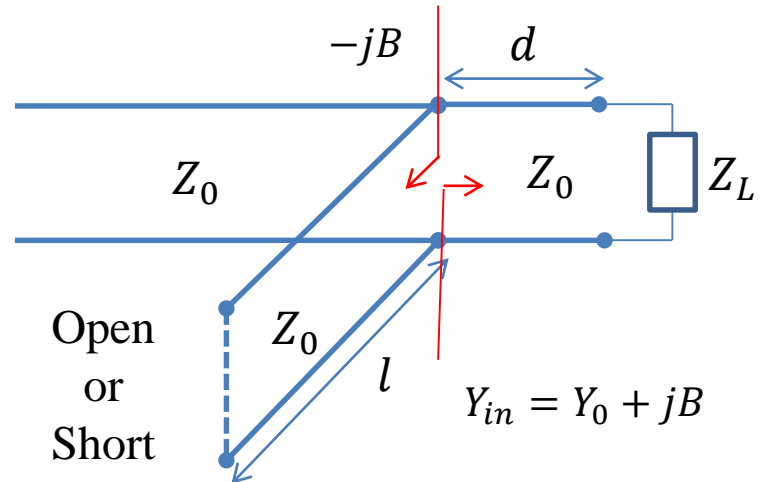
A stub is a short section of transmission line which is either short circuited or open circuited at one end.

A single stub matching circuit consists of a series or shunt stub as shown in the figures below:

The design parameters are the distance of the stub d from the load and length of the stub l



Series Stub Matching



Shunt Stub Matching

Series Stub Matching

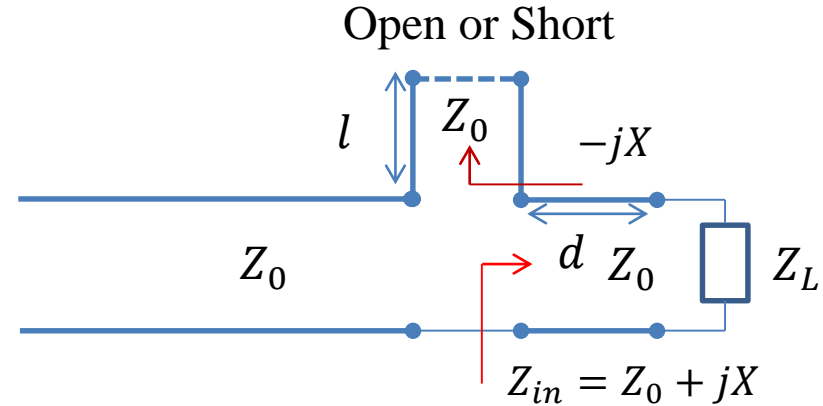
Analytical solution

The distance of the stub location d is so chosen that $Z_{in} = Z_0 + jX$

The stub length l is then so chosen for a short or open stub that input impedance of the stub is $-jX$. This results in matching.

$$Z_{in} = Z_0 \frac{Z_L + jZ_0 \tan \beta d}{Z_0 + jZ_L \tan \beta d}$$

We equate $Re(Z_{in})$ to Z_0 and find solution for d .



Series Stub Matching

Analytical solution

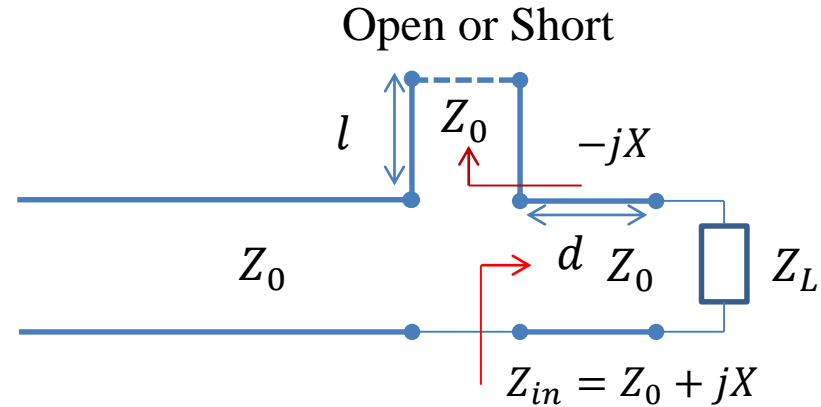
For the computed value of d we calculate X .

The stub length l is then found out for a short or open stub to provide $-jX$.

Let us now derive the closed form expressions

Let $Z_L = R_L + jX_L$

$$Y_L = \frac{1}{Z_L} = G_L + jB_L$$



Series Stub Matching

Let $t = \tan \beta d$

$$Y_{in} = \frac{1}{Z_{in}} = Y_0 \frac{Y_L + jY_0 \tan \beta d}{Y_0 + jY_L \tan \beta d}$$

$$= Y_0 \frac{(G_L + jB_L) + jY_0 t}{Y_0 + j(G_L + jB_L)t}$$

$$Y_{in} = Y_0 \frac{G_L + j(B_L + Y_0 t)}{(Y_0 - B_L t) + jG_L t}$$

$$Z_{in} = R + jX = \frac{1}{Y_{in}}$$

$$R = \frac{G_L(1 + t^2)}{G_L^2 + (B_L + Y_0 t)^2}$$

$$X = \frac{G_L^2 t - (Y_0 - tB_L)(B_L + tY_0)}{Y_0(G_L^2 + (B_L + Y_0 t)^2)}$$

Series Stub Matching

$$\text{From } R = \frac{G_L(1+t^2)}{G_L^2 + (B_L + Y_0 t)^2}$$

$$Y_0(G_L - Y_0)t^2 - 2B_L Y_0 t + (G_L Y_0 - G_L^2 - B_L^2) = 0$$

$$\text{If } G_L = Y_0, \quad t = -B_L / (2Y_0)$$

else

$$t = \frac{B_L \pm \sqrt{G_L[(Y_0 - G_L)^2 + B_L^2]}/Y_0}{(G_L - Y_0)}$$

Series Stub Matching

We get two solutions for d which are given by

$$\frac{d}{\lambda} = \begin{cases} \frac{1}{2\pi} \tan^{-1} t & t \geq 0 \\ \frac{1}{2\pi} (\pi + \tan^{-1} t) & t < 0 \end{cases}$$

With the values of t calculated, we calculate the values of X . Necessary stub reactance $X_S = -X$.

If l_o and l_s respectively denote the lengths for the open and short circuited stubs, then

$$\frac{l_s}{\lambda} = \frac{1}{2\pi} \tan^{-1} \frac{X_S}{Z_0} = -\frac{1}{2\pi} \tan^{-1} \frac{X}{Z_0} \quad \text{and} \quad \frac{l_o}{\lambda} = -\frac{1}{2\pi} \tan^{-1} \frac{Z_0}{X_S} = \frac{1}{2\pi} \tan^{-1} \frac{Z_0}{X}$$

If any of the lengths comes out to be negative, $\lambda/2$ is added.

Example: Impedance Matching Series Stub

Let us consider an example where $Z_L = 100 + j50 \Omega$ is to be matched to a 50Ω line. By applying the analytical solutions we get:

$t = -0.333 = t_1$ and $t = 1.0 = t_2$. We get two solutions for d

$$\frac{d_1}{\lambda} = \frac{1}{2\pi} (\pi + \tan^{-1} t_1) = 0.45$$

$$\frac{d_2}{\lambda} = \frac{1}{2\pi} \tan^{-1} t_2 = 0.125$$

We get two solutions for X as $X_1 = 50$ and $X_2 = -50$

Let us now find the lengths of the open circuited stubs to complete the solution

$$\frac{l_{o1}}{\lambda} = \frac{1}{2\pi} \tan^{-1} \frac{Z_0}{X_1} = 0.125 \quad \text{and} \quad \frac{l_{o2}}{\lambda} = 0.5 + \frac{1}{2\pi} \tan^{-1} \frac{Z_0}{X_2} = 0.375$$

Solution: 1

$$\frac{d}{\lambda} = 0.338 - 0.213 = 0.125$$

$$x = -1$$

$$x_S = 1$$

$$\frac{l_0}{\lambda} = 0.375$$

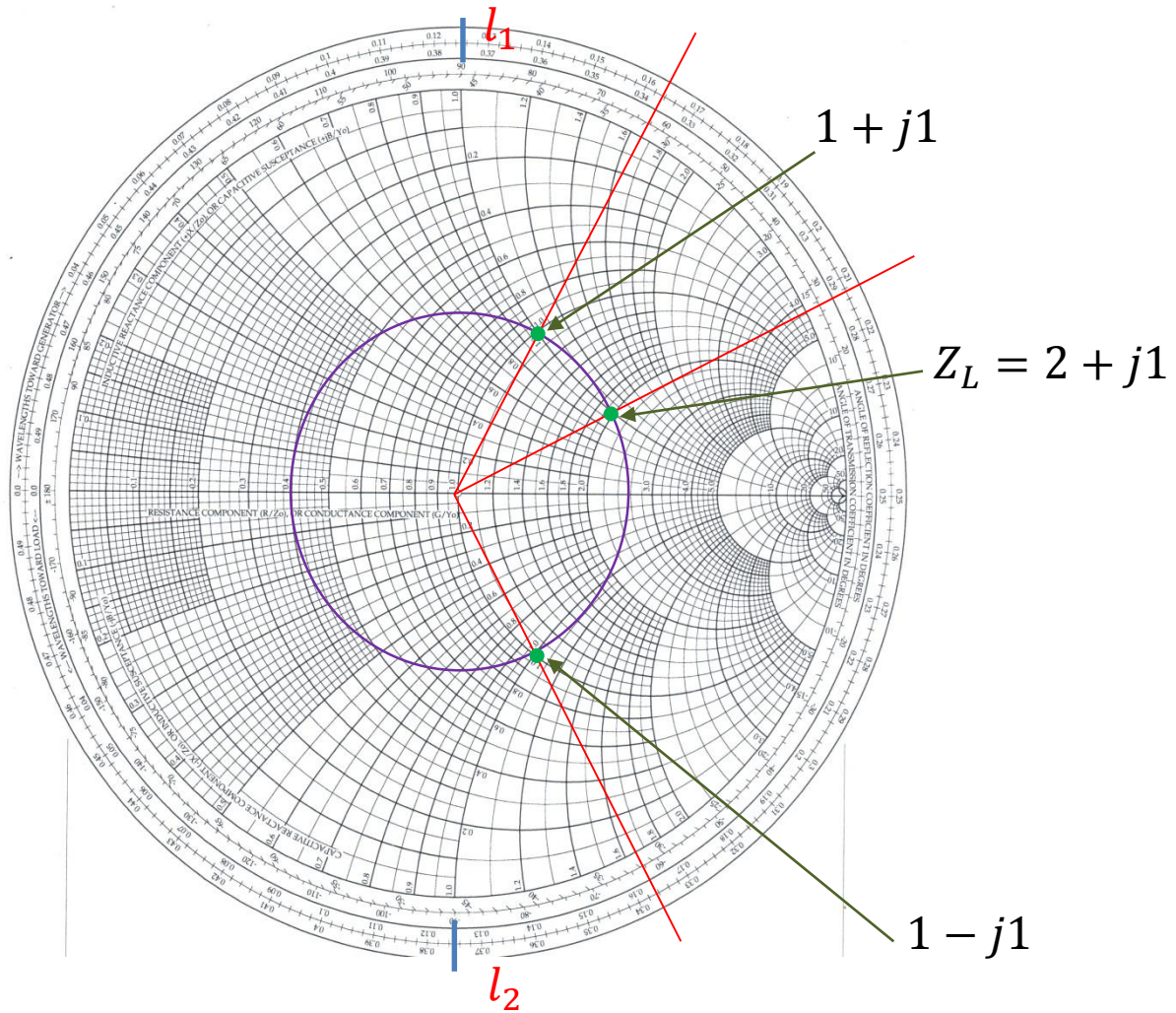
Solution: 2

$$\begin{aligned} \frac{d}{\lambda} &= 0.5 - (0.213 - 0.164) \\ &= 0.451 \end{aligned}$$

$$x = 1$$

$$x_S = -1$$

$$\frac{l_0}{\lambda} = 0.125$$



Shunt Stub Matching

Analytical solution

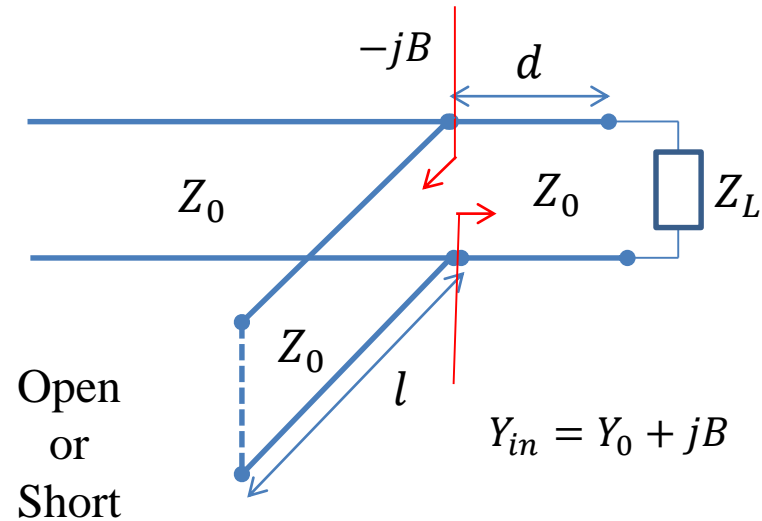
The distance of the stub location d is so chosen that $Y_{in} = Y_0 + jB$

The stub length l is then so chosen for a short or open stub that input susceptance of the stub is $-jB$. This results in matching.

$$Z_{in} = Z_0 \frac{Z_L + jZ_0 \tan \beta d}{Z_0 + jZ_L \tan \beta d}$$

$$Y_{in} = G + jB$$

We equate $Re(Y_{in}) = G$ to Y_0 and find solution for d .



Shunt Stub Matching

Similar to series stub matching, let
 $t = \tan \beta d$

$$Z_{in} = Z_0 \frac{Z_L + jZ_0 \tan \beta d}{Z_0 + jZ_L \tan \beta d}$$

$$= Z_0 \frac{(R_L + jX_L) + jZ_0 t}{Z_0 + j(R_L + jX_L)t}$$

$$Z_{in} = Z_0 \frac{R_L + j(X_L + Z_0 t)}{(Z_0 - X_L t) + jR_L t}$$

$$Y_{in} = G + jB = \frac{1}{Z_{in}}$$

$$G = \frac{R_L(1 + t^2)}{R_L^2 + (X_L + Z_0 t)^2}$$

$$B = \frac{R_L^2 t - (Z_0 - tX_L)(X_L + tZ_0)}{Z_0(R_L^2 + (X_L + Z_0 t)^2)}$$

Shunt Stub Matching

From $G = \frac{R_L(1+t^2)}{R_L^2 + (X_L + Z_0 t)^2}$

$$Z_0(R_L - Z_0)t^2 - 2X_L Z_0 t + (R_L Z_0 - R_L^2 - X_L^2) = 0$$

If $R_L = Z_0$, $t = -X_L / (2Z_0)$

else

$$t = \frac{X_L \pm \sqrt{R_L [(Z_0 - R_L)^2 + X_L^2]} / Z_0}{(R_L - Z_0)}$$

Shunt Stub Matching

We get two solutions for d which are given by

$$\frac{d}{\lambda} = \begin{cases} \frac{1}{2\pi} \tan^{-1} t & t \geq 0 \\ \frac{1}{2\pi} (\pi + \tan^{-1} t) & t < 0 \end{cases}$$

With the values of t calculated, we calculate the values of B . Necessary stub reactance $B_S = -B$.

If l_o and l_s respectively denote the lengths for the open and short circuited stubs, then

$$\frac{l_o}{\lambda} = \frac{1}{2\pi} \tan^{-1} \frac{B_S}{Y_0} = -\frac{1}{2\pi} \tan^{-1} \frac{B}{Y_0} \quad \text{and} \quad \frac{l_s}{\lambda} = -\frac{1}{2\pi} \tan^{-1} \frac{Y_0}{B_S} = \frac{1}{2\pi} \tan^{-1} \frac{Y_0}{B}$$

If any of the lengths comes out to be negative, $\lambda/2$ is added.

Example: Impedance Matching- Shunt Stub

Let us consider an example where $Z_L = 100 + j60 \Omega$ is to be matched to a 50Ω line. By applying the analytical solutions we get:

$t = 3.4091 = t_1$ and $t = -1.0091 = t_2$. We get two solutions for d

$$\frac{d_1}{\lambda} = \frac{1}{2\pi} \tan^{-1} t_1 = 0.2046$$

$$\frac{d_2}{\lambda} = \frac{1}{2\pi} (\pi + \tan^{-1} t_2) = 0.3743$$

We get two solutions for B as $B_1 = 0.0221$ and $B_2 = -0.0221$

Let us now find the lengths of the open circuited stubs to complete the solution

$$\frac{l_{o1}}{\lambda} = \frac{1}{2\pi} \tan^{-1} \frac{B_1}{Y_0} = 0.3671 \quad \text{and} \quad \frac{l_{o2}}{\lambda} = 0.5 + \frac{1}{2\pi} \tan^{-1} \frac{B_2}{Y_0} = 0.1329$$

Shunt Stub Matching using Open Stub

From Smith Chart

$$\frac{d_1}{\lambda} = (0.5 - 0.46) + 0.164 = 0.204$$

$$\frac{d_2}{\lambda} = (0.5 - 0.46) + 0.336 = 0.376$$

$$l_{o1}/\lambda = 0.132$$

$$l_{o2}/\lambda = 0.368$$

$$Z_L = 100 + j60$$

Analytical

$$t1 = 3.4091$$

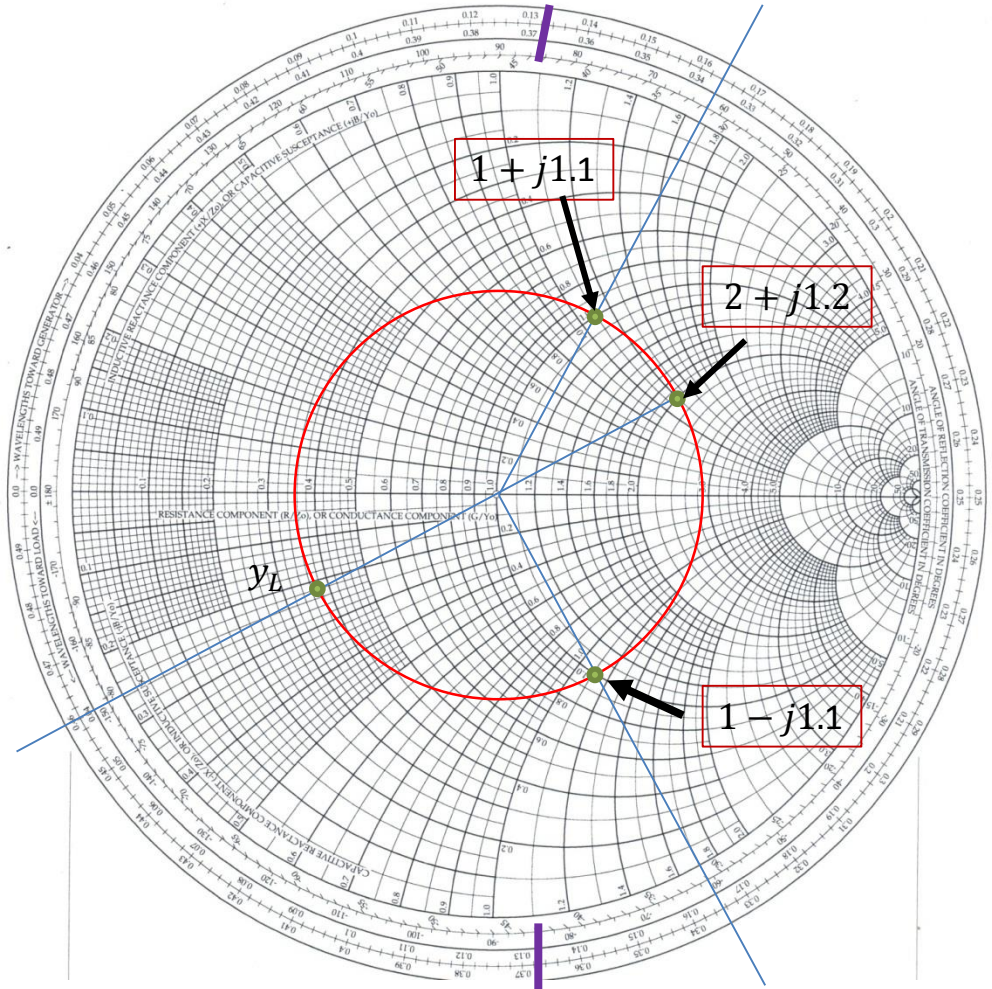
$$t2 = -1.0091$$

$$d_1/\lambda = 0.20459$$

$$d_2/\lambda = 0.37428$$

$$l_{o1}/\lambda = 0.36710$$

$$l_{o2}/\lambda = 0.13290$$



Double Stub Matching

As shown in Fig.1, the load is at an arbitrary distance d from the first stub.

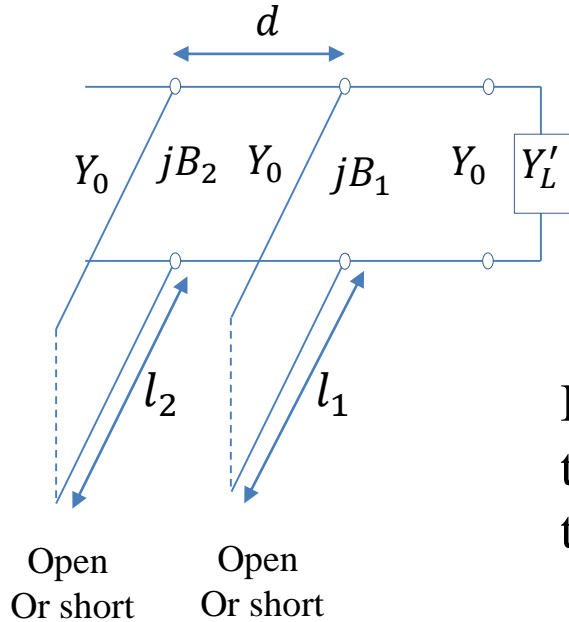


Fig.1

In Fig.2, the load Y'_L is transformed to the position of the first stub as Y_L

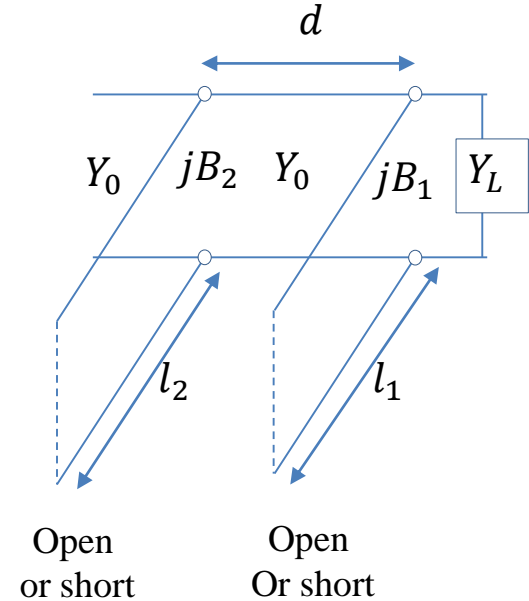


Fig.2

Double Stub Matching

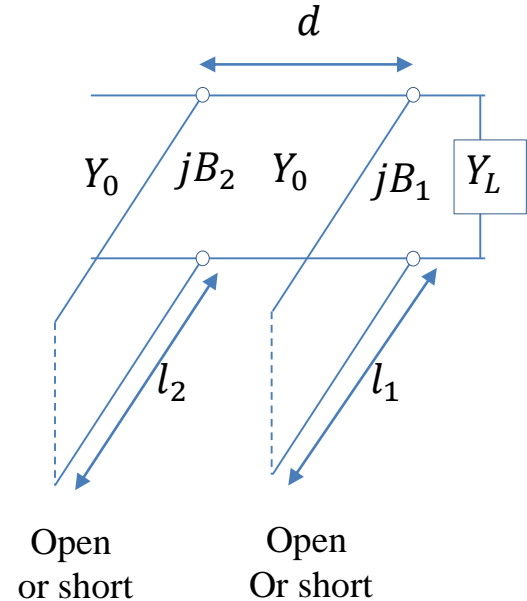
Analytical solution

From the figure, we have

$$\begin{aligned} Y_1 &= Y_L + jB_1 = G_L + jB_L + jB_1 \\ &= G_L + j(B_L + B_1) \end{aligned}$$

$$Y_2 = Y_0 \frac{Y_1 + jY_0 \tan \beta d}{Y_0 + jY_1 \tan \beta d}$$

We equate $Re(Y_2)$ to Y_0 and find solution for d .



Double Stub Matching

Similar to earlier assumptions, let $t = \tan \beta d$

$$Y_2 = Y_0 \frac{Y_1 + jY_0 \tan \beta d}{Y_0 + jY_1 \tan \beta d}$$

$$= Y_0 \frac{[G_L + j(B_L + B_1)] + jY_0 t}{Y_0 + j[G_L + j(B_L + B_1)]t}$$

$$Y_2 = Y_0 \frac{G_L + j(B_L + B_1 + Y_0 t)}{(Y_0 - B_L t - B_1 t) + jG_L t}$$

Double Stub Matching

On equating $Re(Y_2)$ to Y_0

$$G_L^2 + G_L Y_0 \frac{1+t^2}{t^2} + \frac{(Y_0 - B_L t - B_1 t)^2}{t^2} = 0$$
$$G_L = Y_0 \frac{1+t^2}{t^2} \left[1 \pm \sqrt{1 - \frac{4t^2(Y_0 - B_L t - B_1 t)^2}{Y^2(1+t^2)^2}} \right]$$

$\therefore G_L$ is real,

$$0 \leq \frac{4t^2(Y_0 - B_L t - B_1 t)^2}{Y^2(1+t^2)^2} \leq 1$$

$$0 \leq G_L \leq Y_0 \frac{1+t^2}{t^2} = Y_0 \frac{1 + \tan^2 \beta d}{\tan^2 \beta d} = \frac{Y_0}{\sin^2 \beta d}$$

Double Stub Matching

$$B_1 = -B_L + \frac{Y_0 \pm \sqrt{(1 + t^2)G_L Y_0 - G_L^2 t^2}}{t}$$
$$B_2 = \pm \frac{Y_0 \sqrt{(1 + t^2)G_L Y_0 - G_L^2 t^2} + G_L Y_0}{G_L t}$$

If l_o and l_s respectively denote the lengths for the open and short circuited stubs

$$\frac{l_o}{\lambda} = -\frac{1}{2\pi} \tan^{-1} \frac{B}{Y_0}$$

and

$$\frac{l_s}{\lambda} = \frac{1}{2\pi} \tan^{-1} \frac{Y_0}{B}$$

$$B = B_1 \text{ or } B_2$$

Double Stub Matching Using Smith Chart

$$Y_i = Y_B + Y_{SB} = Y_0$$

In normalized form, $1 = y_B + y_{SB}$

Since y_{SB} is purely imaginary we must have,

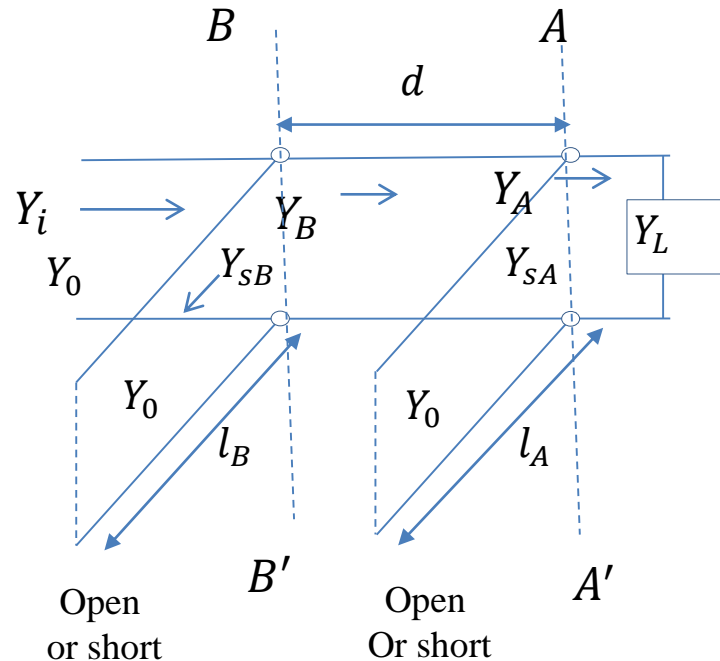
$$y_B = 1 + jb_B \text{ and } y_{SB} = -jb_B$$

Therefore, in the Smith chart y_B must lie in the $g = 1$ circle.

To meet this requirement y_A at AA' must lie on the $g = 1$ circle rotated by $\frac{4\pi d}{\lambda}$ counter clockwise direction.

Since y_{SA} is purely imaginary, the real part of y_A must be contributed solely by real part of y_L i.e. g_L .

The solution of double stub matching is then determined by the intersection of g_L circle with rotated $g = 1$ circle.



Procedure

Plot $g = 1$ circle. y_B should be located on this circle

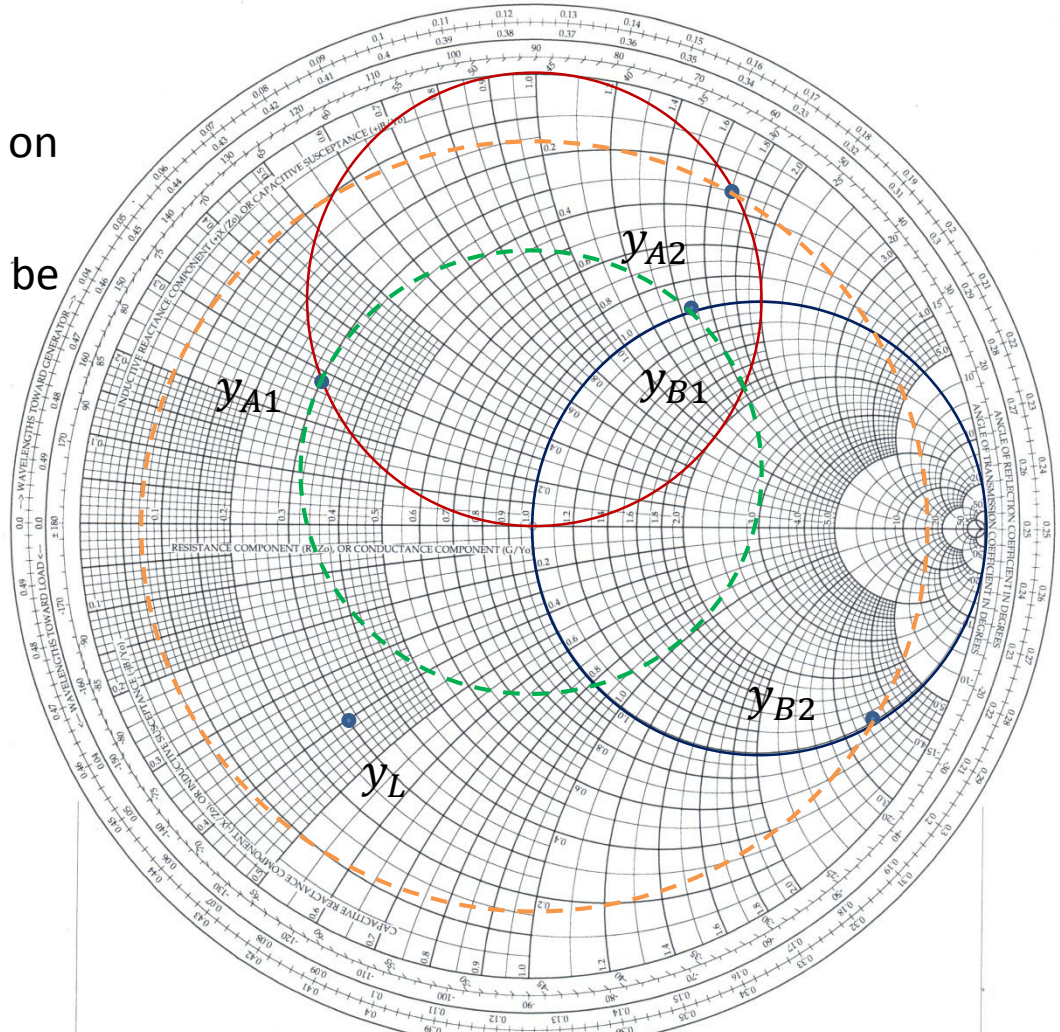
Plot the rotated circle where y_A should be located

Plot y_L

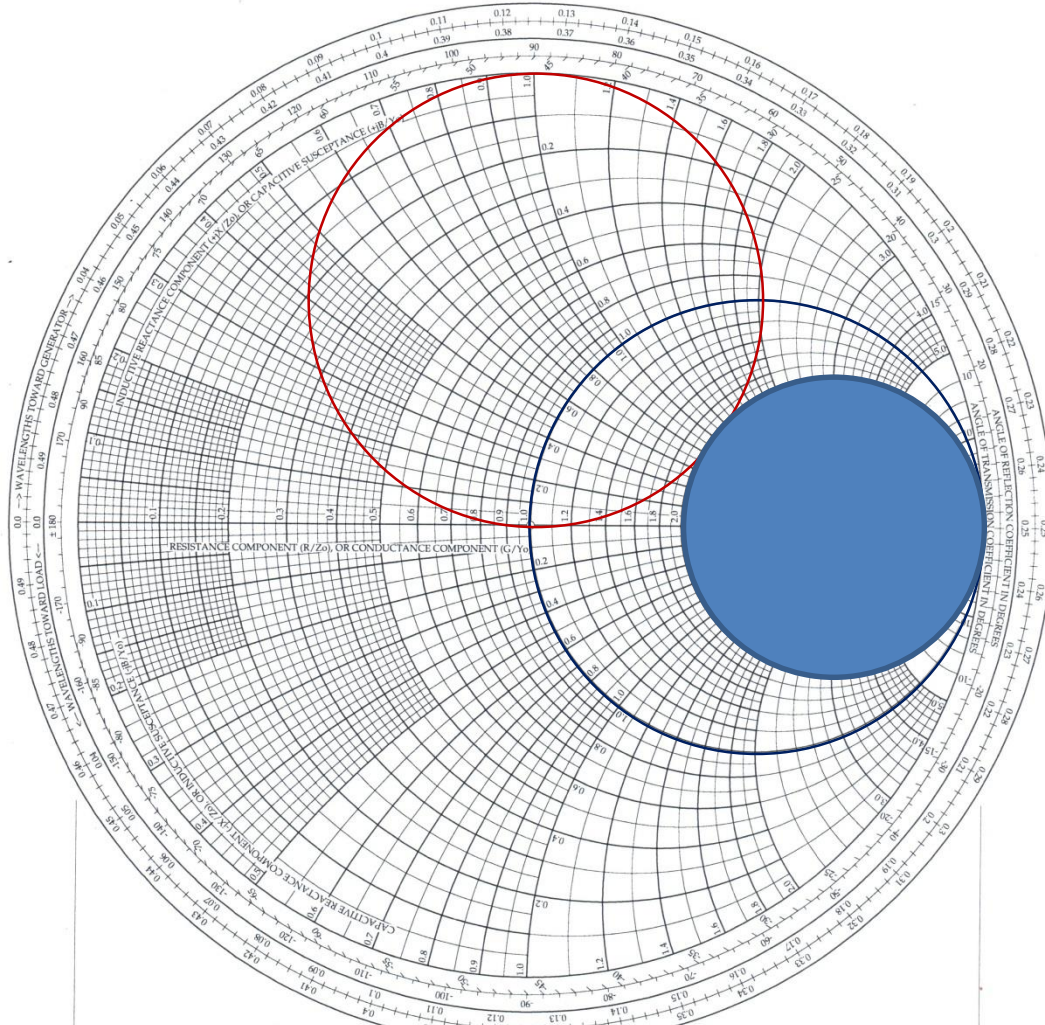
Find intersection of $g = g_L$ circle with rotated $g = 1$ circle at y_{A1} & y_{A2}

Find y_B points on $g = 1$ circle:
 y_{B1} and y_{B2}

Determine the stub lengths l_A and l_B



The shaded region is the forbidden range of load admittances that can not be matched with the given double stub tuner



Quarter-wave Transformer

A quarter-wave transformer is transmission line section of length $\frac{\lambda}{4}$ having characteristic impedance Z_1 and used to match a real load R_L to a transmission line of characteristic impedance Z_0 , as shown in the figure.

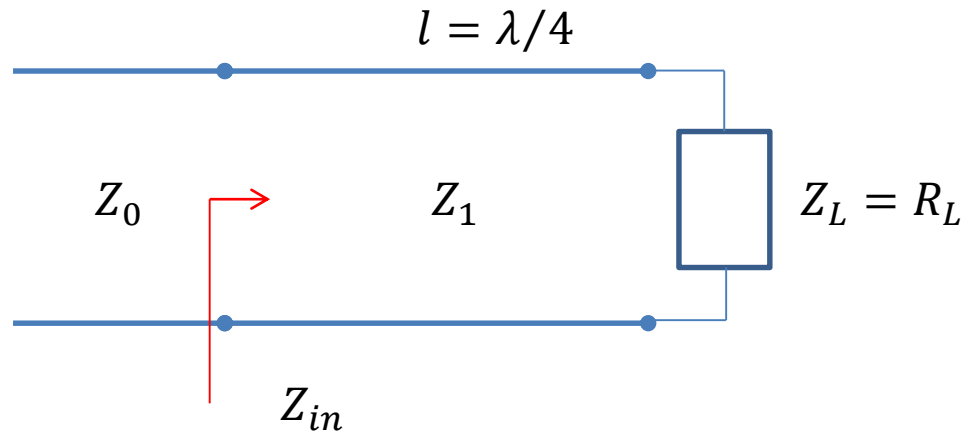
We know that

$$Z_{in} = Z_1 \frac{R_L + jZ_1 \tan \beta l}{Z_1 + jR_L \tan \beta l}$$

Dividing the numerator and denominator by $\tan \beta l$ and take the limit as $\beta l \rightarrow \pi/2$, we can

write $Z_{in} = \frac{Z_1^2}{R_L}$. Equating Z_{in} to

Z_0 we get $Z_1 = \sqrt{R_L Z_0}$



Quarter-wave Transformer

We note that matching is obtained at the frequency at which the transformer is quarter wavelength long and at all odd harmonics where the length corresponds to $(2n + 1) \lambda/4$.

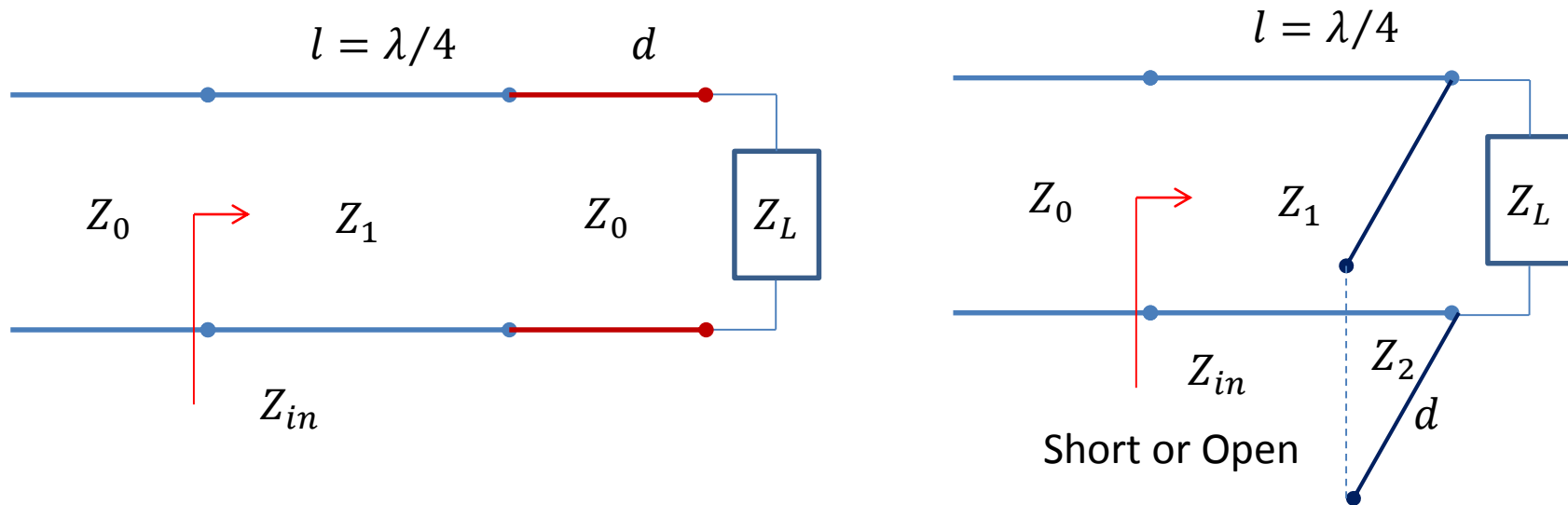
The fractional bandwidth of such quarter-wave transformer can be found as:

$$\frac{\Delta f}{f_0} = 2 - \left(\frac{4}{\pi}\right) \cos^{-1} \left[\frac{\Gamma_m}{\sqrt{1 - \Gamma_m^2}} \frac{2\sqrt{Z_1 R_L}}{|R_L - Z_0|} \right]$$

Γ_m is the magnitude of the acceptable value of reflection coefficient

Use of Quarter-wave Transformer

Quarter-wave transformers can also be used in design of matching network for matching complex load impedance to a transmission line. The examples of such networks are shown:



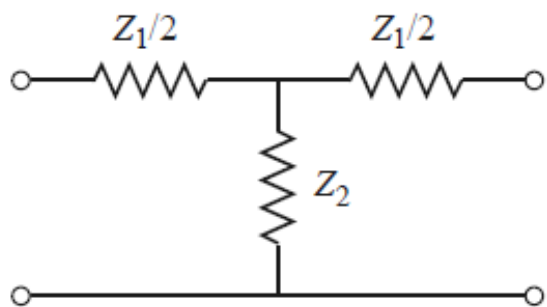
Microwave Filter Design

- A filter is a two-port network used to control the frequency response at a certain point in an RF or microwave system by providing transmission at frequencies within the passband of the filter and attenuation in the stopband of the filter.
- Typical frequency responses include *low-pass*, *high-pass*, *bandpass*, and *band-reject characteristics*.
- Applications can be found in virtually any type of RF or microwave communication, radar, or test and measurement system.
- Filters designed using the *image parameter method* consist of a cascade of simpler two-port filter sections to provide the desired cutoff frequencies and attenuation characteristics but do not allow the specification of a particular frequency response over the complete operating range. Thus, although the procedure is relatively simple, the design of filters by the image parameter method often must be iterated many times to achieve the desired results.

Microwave Filter Design

- A more modern procedure, called the *insertion loss method*, uses network synthesis techniques to design filters with a completely specified frequency response. The design is simplified by beginning with low-pass filter prototypes that are normalized in terms of impedance and frequency. Transformations are then applied to convert the prototype designs to the desired frequency range and impedance level.
- Both the image parameter and insertion loss methods of filter design lead to circuits using lumped elements (capacitors and inductors).
- For microwave applications such designs usually must be modified to employ distributed elements consisting of transmission line sections. The *Richards transformation* and the *Kuroda identities* provide this step.

Image Parameters for T- and π -Networks



T-Network

ABCD parameters:

$$A = 1 + Z_1/2Z_2$$

$$B = Z_1 + Z_1^2/4Z_2$$

$$C = 1/Z_2$$

$$D = 1 + Z_1/2Z_2$$

Z parameters:

$$Z_{11} = Z_{22} = Z_2 + Z_1/2$$

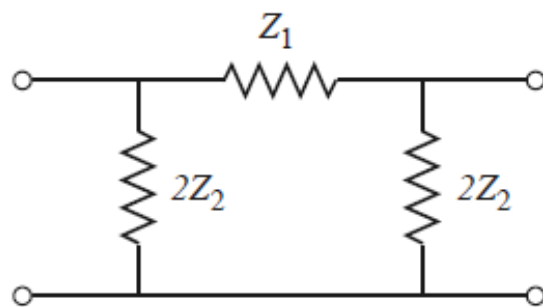
$$Z_{12} = Z_{21} = Z_2$$

Image impedance:

$$Z_{iT} = \sqrt{Z_1 Z_2} \sqrt{1 + Z_1/4Z_2}$$

Propagation constant:

$$e^\gamma = 1 + Z_1/2Z_2 + \sqrt{Z_1/Z_2 + Z_1^2/4Z_2^2}$$



π -Network

ABCD parameters:

$$A = 1 + Z_1/2Z_2$$

$$B = Z_1$$

$$C = 1/Z_2 + Z_1/4Z_2^2$$

$$D = 1 + Z_1/2Z_2$$

Y parameters:

$$Y_{11} = Y_{22} = 1/Z_1 + 1/2Z_2$$

$$Y_{12} = Y_{21} = 1/Z_1$$

Image impedance:

$$Z_{i\pi} = \sqrt{Z_1 Z_2} / \sqrt{1 + Z_1/4Z_2} = Z_1 Z_2 / Z_{iT}$$

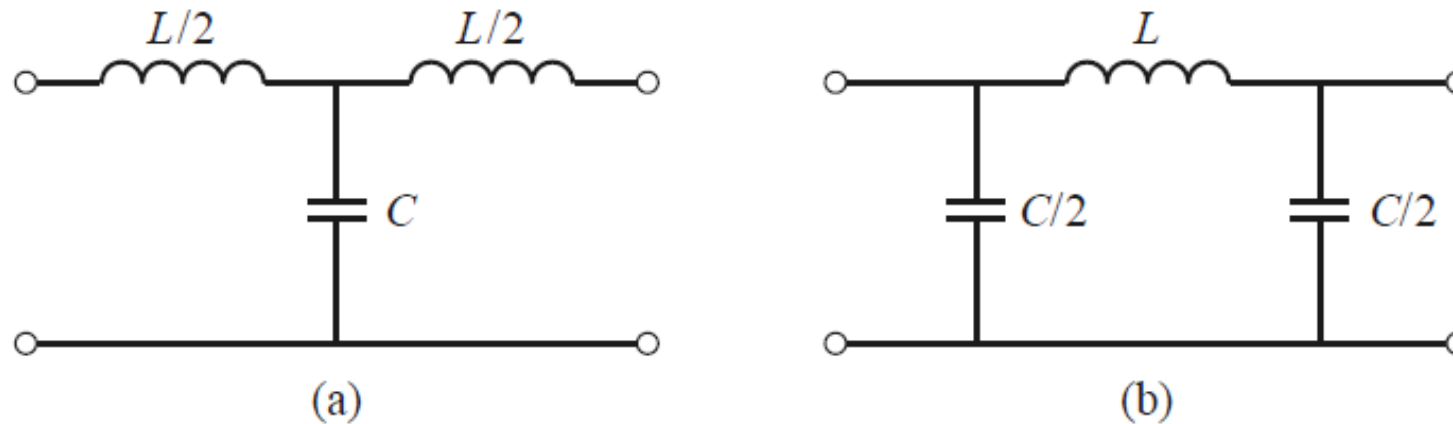
Propagation constant:

$$e^\gamma = 1 + Z_1/2Z_2 + \sqrt{Z_1/Z_2 + Z_1^2/4Z_2^2}$$

Filter Design by the Image Parameter Method

Constant-k Filter Sections:

- First consider the T-network shown in figure. Intuitively, we can see that this is a low-pass filter network because the series inductors and shunt capacitor tend to block high-frequency signals while passing low-frequency signals.



Low-pass constant-k filter sections in T and π forms. (a) T-section. (b) π -section.

Constant-k Filter Sections

- We have $Z_1 = j\omega L$ and $Z_2 = 1/j\omega C$, so the image impedance is

$$Z_{iT} = \sqrt{Z_1 Z_2} \sqrt{1 + \frac{Z_1}{4Z_2}} = \sqrt{\frac{j\omega L}{j\omega C}} \sqrt{1 + \frac{j\omega L}{4\left(\frac{1}{j\omega C}\right)}} = \sqrt{\frac{L}{C}} \sqrt{1 - \frac{\omega^2 LC}{4}}$$

- If we define a cutoff frequency ω_c as $\omega_c = 2/\sqrt{LC}$ and a nominal characteristic impedance, R_0 as $R_0 = \sqrt{L/C} = k$, where k is a constant, then we can rewrite Z_{iT} as

$$Z_{iT} = R_0 \sqrt{1 - \frac{\omega^2}{\omega_c^2}}$$

- Then $Z_{iT} = R_0$ for $\omega = 0$.

Constant-k Filter Sections

- The propagation factor is

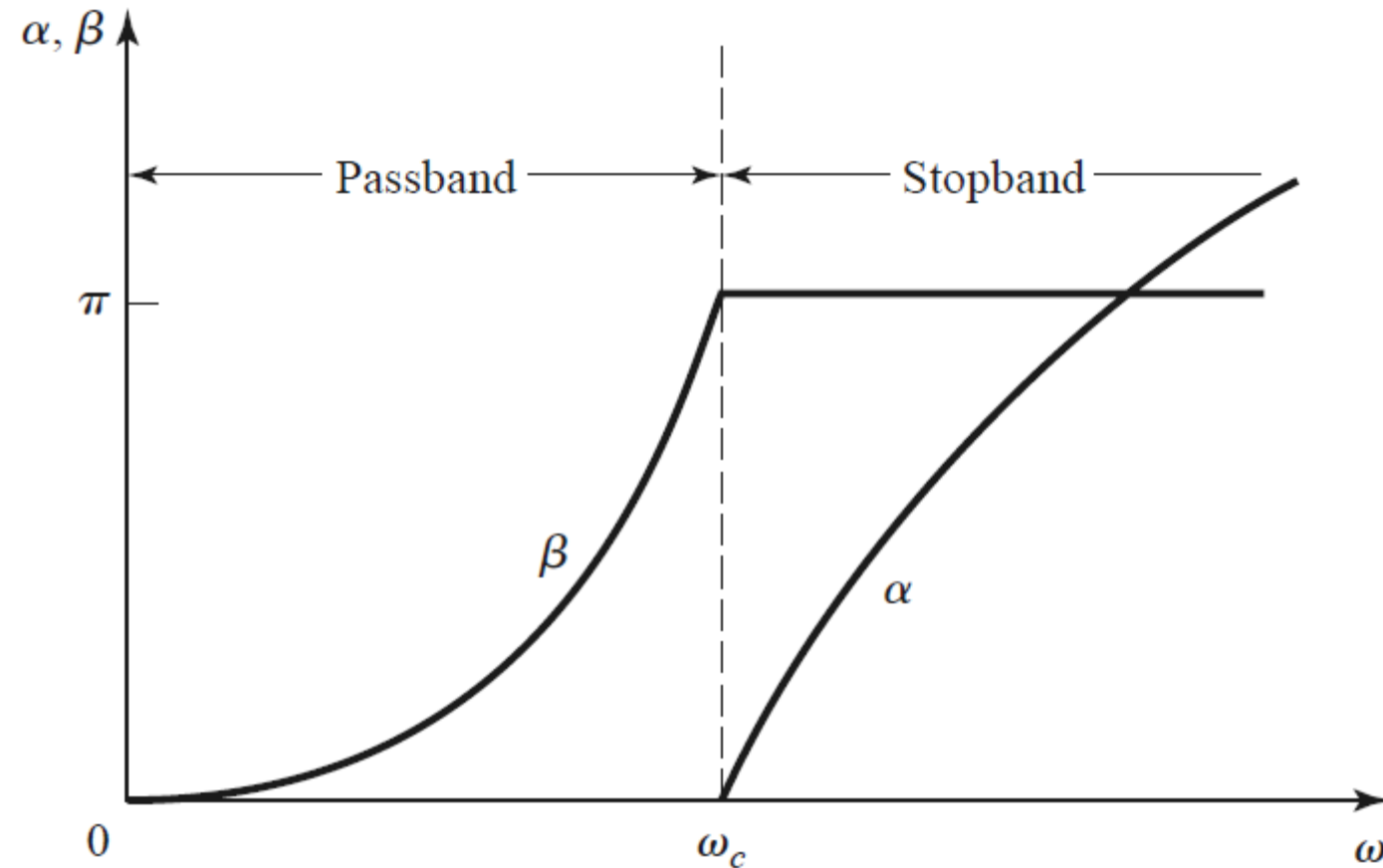
$$e^{\gamma} = 1 + \frac{Z_1}{2Z_2} + \sqrt{\frac{Z_1}{Z_2} + \frac{Z_1^2}{4Z_2^2}} = 1 + \frac{j\omega L}{2\left(\frac{1}{j\omega C}\right)} + \sqrt{\frac{j\omega L}{\left(\frac{1}{j\omega C}\right)} + \frac{(j\omega L)^2}{4\left(\frac{1}{j\omega C}\right)^2}}$$

$$e^{\gamma} = 1 - \frac{\omega^2 LC}{2} + \sqrt{-\omega^2 LC + \frac{(\omega^2 LC)^2}{4}} = 1 - \frac{\omega^2 LC}{2} + \sqrt{\omega^2 LC \left(\frac{\omega^2 LC}{4} - 1\right)}$$

- Since $\omega_c = 2/\sqrt{LC}$,

$$e^{\gamma} = 1 - \frac{2\omega^2}{\omega_c^2} + \frac{2\omega}{\omega_c} \sqrt{\left(\frac{\omega^2}{\omega_c^2} - 1\right)}$$

Constant-k Filter Sections



Typical passband and stopband characteristics of the low-pass constant-k sections

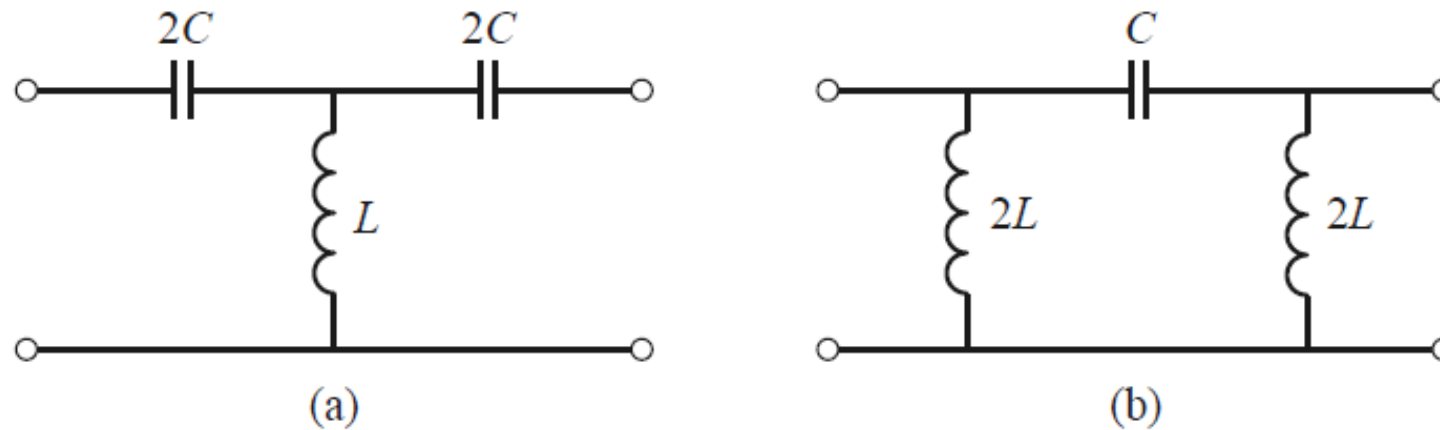
Constant-k Filter Sections

Now consider two frequency regions:

- For $\omega < \omega_c$: This is the *passband of the filter* section. Z_{iT} is real, γ is imaginary, since $\frac{\omega^2}{\omega_c^2} - 1$ is negative and $|e^\gamma| = 1$
- For $\omega > \omega_c$: This is the *stopband of the filter* section. Z_{iT} is imaginary, e^γ is real and $-1 < e^\gamma < 0$ (as seen from the limits as $\omega \rightarrow \omega_c$ and $\omega \rightarrow \infty$). The attenuation rate for $\omega \gg \omega_c$ is 40 dB/decade.
- Typical phase and attenuation constants are sketched in previous figure. Observe that the attenuation, α , is zero or relatively small near the cutoff frequency, although $\alpha \rightarrow \infty$ as $\omega \rightarrow \infty$. This type of filter is known as a *constant-k low-pass prototype*. There are only two parameters to choose (L and C), which are determined by ω_c , the cutoff frequency, and R_0 , the image impedance at zero frequency.

Constant-k Filter Sections

- The above results are valid only when the filter section is terminated in its image impedance at both ports. This is a *major weakness* of the design because *the image impedance is a function of frequency*, and is not likely to match a given source or load impedance. This disadvantage, as well as the fact that *the attenuation is rather low near cutoff*, can be remedied with the modified m-derived sections.

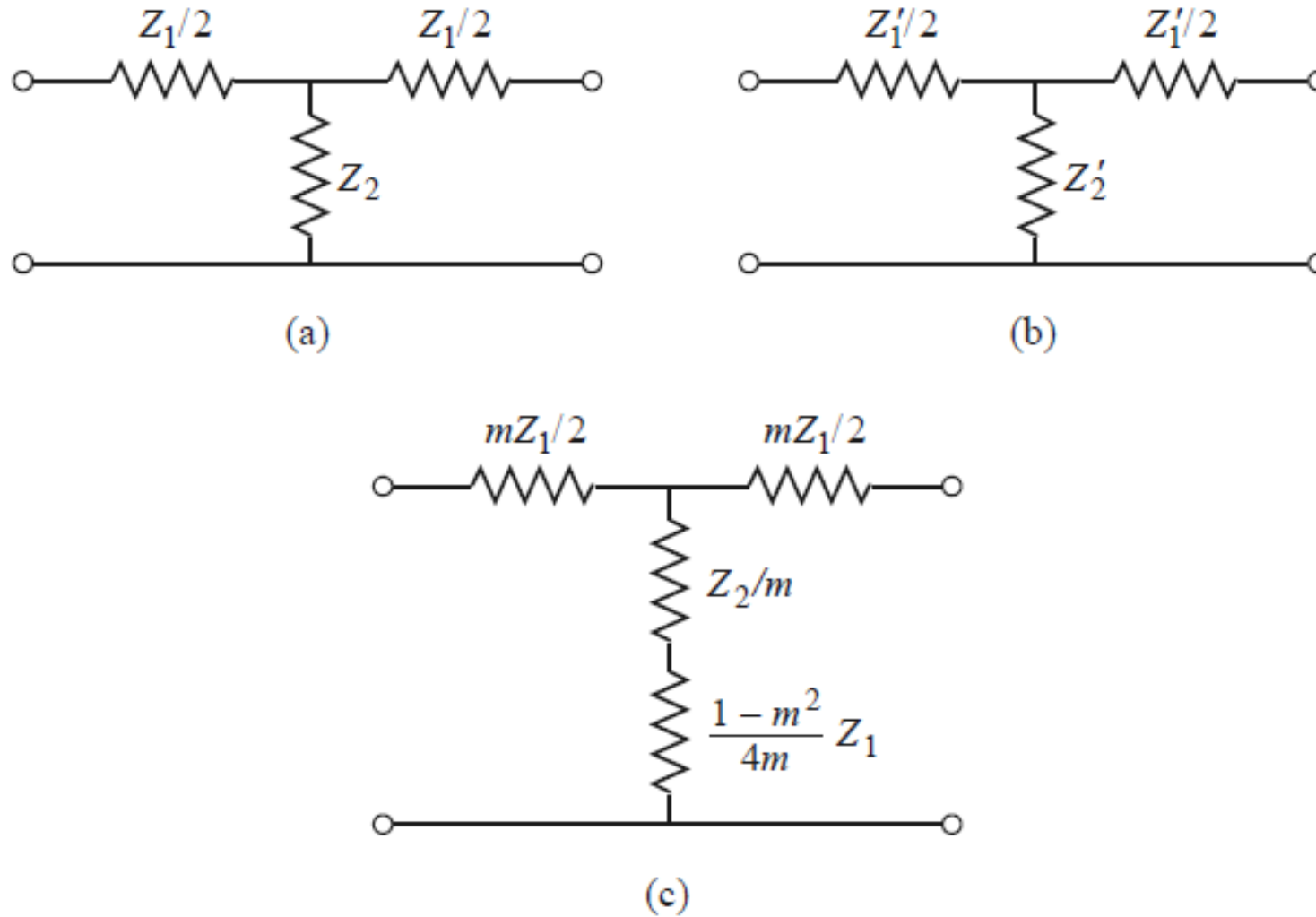


High-pass constant-k filter sections in T and π forms. (a) T-section. (b) π -section.

m-Derived Filter Sections

- We have seen that the constant-k filter section suffers from the disadvantages of a relatively slow attenuation rate past cutoff, and a nonconstant image impedance.
- The m-derived filter section is a modification of the constant-k section designed to overcome these problems.

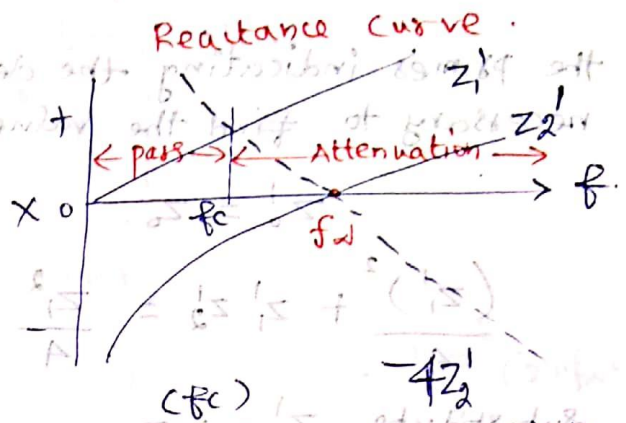
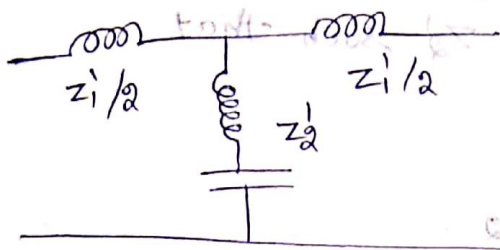
Development of an m-derived filter section from a constant-k section



(a) Constant-k section. (b) General m-derived section. (c) Final m-derived section

m-derived T-section :-

Derivation of a low-pass section having a sharp cut-off action :-



- * $Z_1 = -4Z_2'$ is the cut-off frequency of the filter.
- * At the resonant frequency f_0 which is beyond f_c , the shunt arm appears as a 'short' circuit on the network. Hence, the attenuation becomes infinite. This frequency where attenuation is infinite is called f_d instead of f_0 .
- * In order to pass the signal below f_c without attenuation, the shunt arm (Z_2') should appear as a capacitance, thus presenting a high reactance so that the signals below f_c will pass without much attenuation.
- * For this to happen, the resonance frequency f_d will always be higher than f_c so that the series resonant shunt arm presents a capacitive reactance below f_c . If f_d is chosen to be close to f_c then the attenuation near the cut-off may be made high.
- * At frequencies beyond f_d in the stop band the shunt arm will behave more like an inductance with low value inductive reactance and hence the attenuation will fall to low values.

The network may be derived by assuming that

$$z_1' = m z_1$$

the primes indicating the derived section. It is then necessary to find the value of z_2' such that

$$z_0' = z_0$$

$$\frac{(z_1')^2}{4} + z_1' z_2' = \frac{z_1^2}{4} + z_1 z_2$$

Substitute $z_1' = m z_1$,

$$\frac{(m z_1)^2}{4} + m z_1 z_2' = \frac{z_1^2}{4} + z_1 z_2$$

$$m z_1 z_2' = \frac{z_1^2}{4} (1 - m^2) + z_1 z_2$$

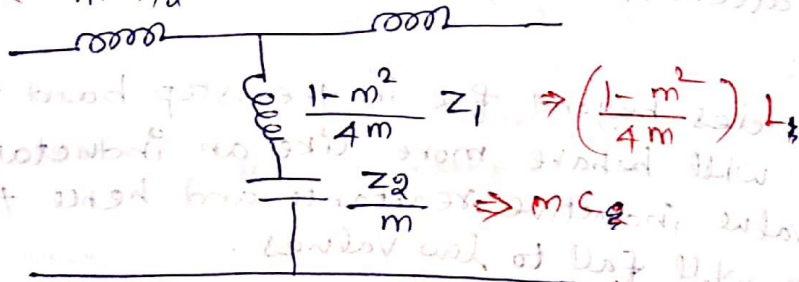
$$z_2' = \frac{(1 - m^2)}{4m} z_1 + \frac{z_2}{m}$$

Now the shunt arm z_2' consists of two impedances in series. As required, the characteristic impedance and f_c remain equal to those of the T-section prototype containing z_1 and z_2 values.

m-derived low pass filter

$$\frac{mL}{2} \leftarrow m z_1/2$$

$$m z_1/2 \Rightarrow mL/2$$



* $\frac{(1 - m^2)}{4m}$ must be positive.

\Rightarrow '1 - m²' & 'm' must be positive.

Thus m must always be chosen so that

$$0 < m < 1.$$

Filter sections obtained in this manner are called *m-derived sections*.

The shunt arm is to be chosen so that it is resonant at some frequency f_d above f_c . This means that at the resonant frequency

$$\left| \frac{Z_2}{m} \right| = \left| \frac{1-m^2}{4m} Z_1 \right|$$

and for the low-pass filter,

$$Z_2 = \frac{1}{j2\pi f_d C} \quad \Rightarrow \quad \frac{1}{(2\pi f_d C)m} = \frac{1-m^2}{4m} (2\pi f_d L)$$

$$Z_1 = j2\pi f_d L$$

$$\Rightarrow \frac{1}{4\pi^2 f_d^2 LC} = \frac{1-m^2}{4m}$$

$$\Rightarrow f_d^2 = \frac{1}{4\pi^2 LC(1-m^2)}$$

$$f_d = \frac{1}{\pi \sqrt{(1-m^2)LC}}$$

since the cut-off frequency for the low-pass filter is

$$f_c = \frac{1}{\pi \sqrt{LC}}$$

\therefore The frequency of infinite attenuation will be

$$f_d = \frac{f_c}{\sqrt{1-m^2}}$$

$$\Rightarrow f_d^2 = \frac{f_c^2}{1-m^2} \Rightarrow f_d^2 - m^2 f_d^2 = f_c^2$$

$$\Rightarrow \frac{f_d^2 - f_c^2}{f_d^2} = m^2$$

$$\Rightarrow m = \sqrt{1 - \left(\frac{f_c}{f_d} \right)^2}$$

This equation determines the m to be used for a particular f_d .

Design a composite Low pass filter to meet the following specifications:

The Composite Low pass - Filter is to be terminated in 500 ohms resistance. It must have a cut-off frequency of 1000 cycles, with very high attenuation at 1065, 1250 and ∞ cycles.

Sol:-

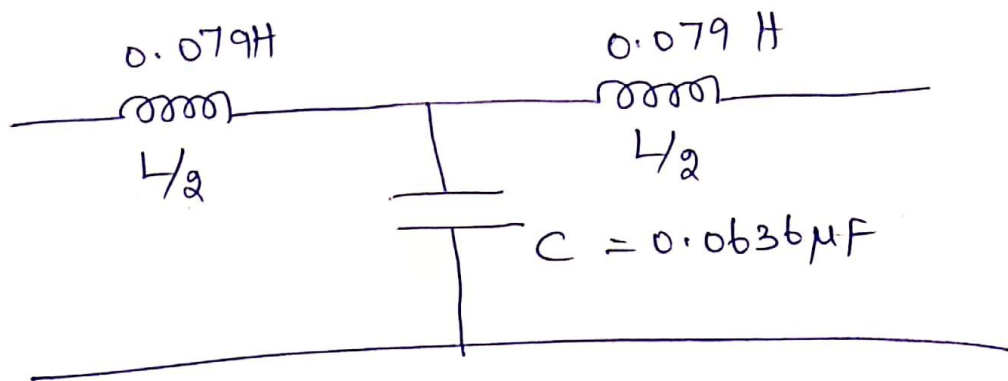
(i) Design of proto-type (constant-K filter) :-

The proto-type is designed first as

$$L = \frac{R}{\pi f_c} = \frac{500}{\pi \times 1000} = 0.159 \text{ H}$$

$$C = \frac{1}{\pi f_c R} = \frac{1}{\pi \times 1000 \times 500} = 0.636 \mu\text{F}$$

This prototype section meets the specification for high attenuation at infinity.



(ii) m-derived section to provide high attenuation at $f_d = 1065$:-

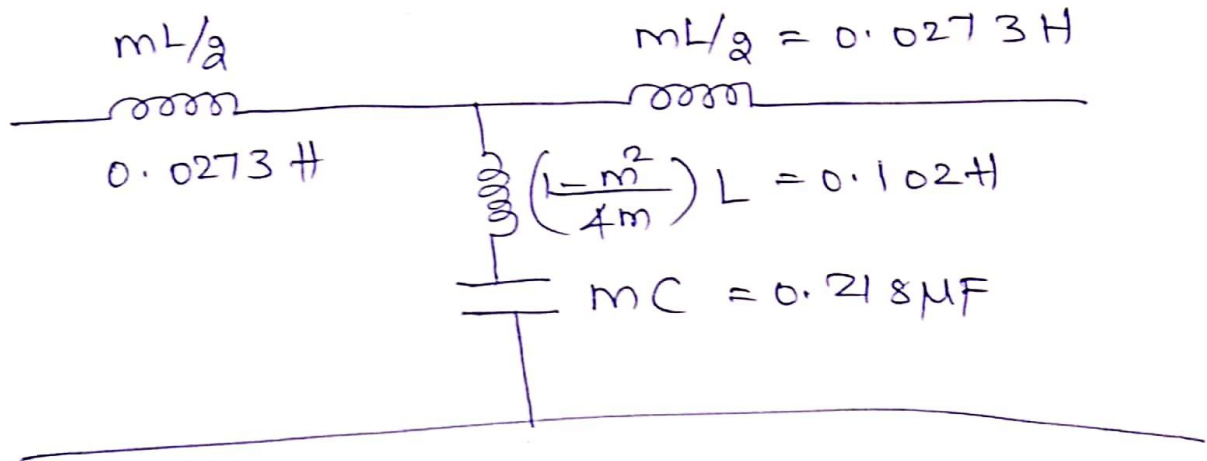
$$m' = \sqrt{1 - \left(\frac{f_c}{f_d}\right)^2} = \sqrt{1 - \left(\frac{1000}{1065}\right)^2}$$

$$= \sqrt{1 - 0.882} = 0.343$$

$$\frac{mL}{2} = \frac{0.343 \times 0.159}{2} = 0.0273 \text{ H}$$

$$\left(\frac{1-m^2}{4m}\right)L = \frac{(1-0.118)}{1.372} \times 0.159 = 0.102 \text{ H}$$

$$mC = 0.343 \times 0.636 = 0.218 \mu\text{F}$$



(iii) m-derived section to provide high attenuation at $f_d = 1250$:-

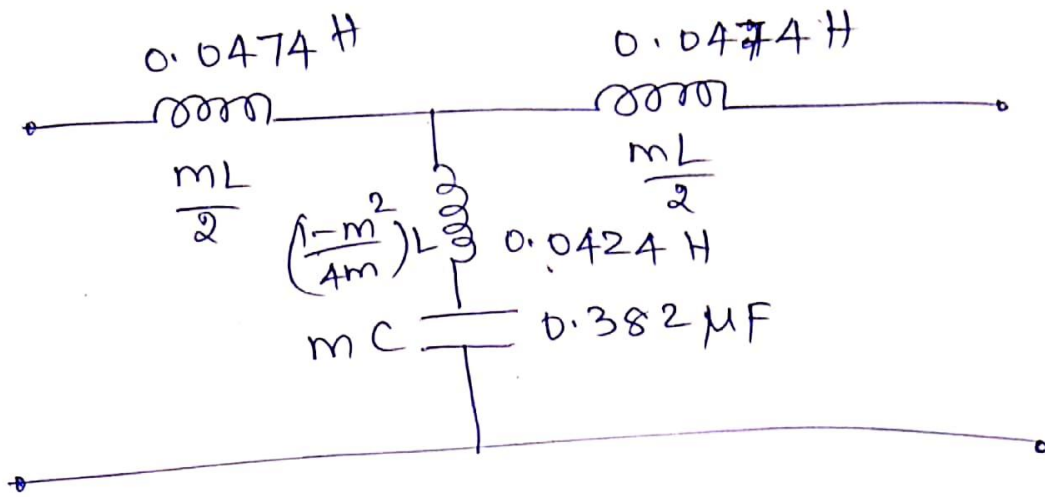
$$m = \sqrt{1 - \left(\frac{1000}{1250}\right)^2} = \sqrt{1 - 0.640} = 0.6$$

$m = 0.6$ is appropriate for use in the terminal half sections.

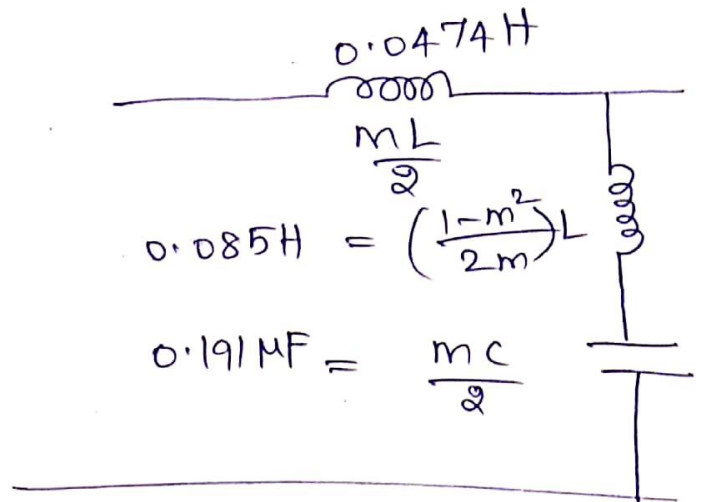
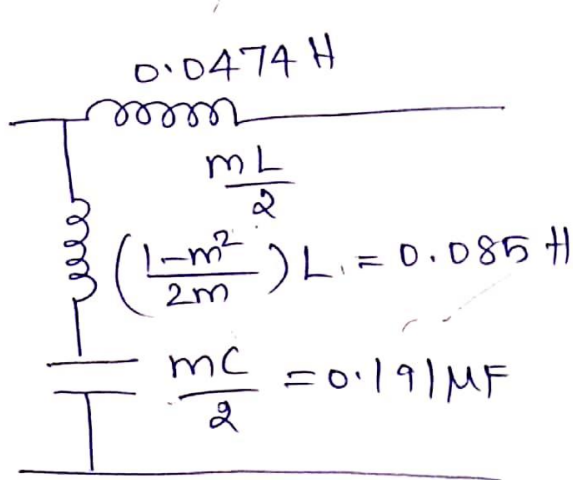
$$\frac{mL}{2} = \frac{0.6 \times 0.159}{2} = 0.0474 \text{ H}$$

$$\left(\frac{1-m^2}{4m}\right)L = \left(\frac{1-0.36}{4 \times 0.6}\right) 0.159 = 0.0424 \text{ H}$$

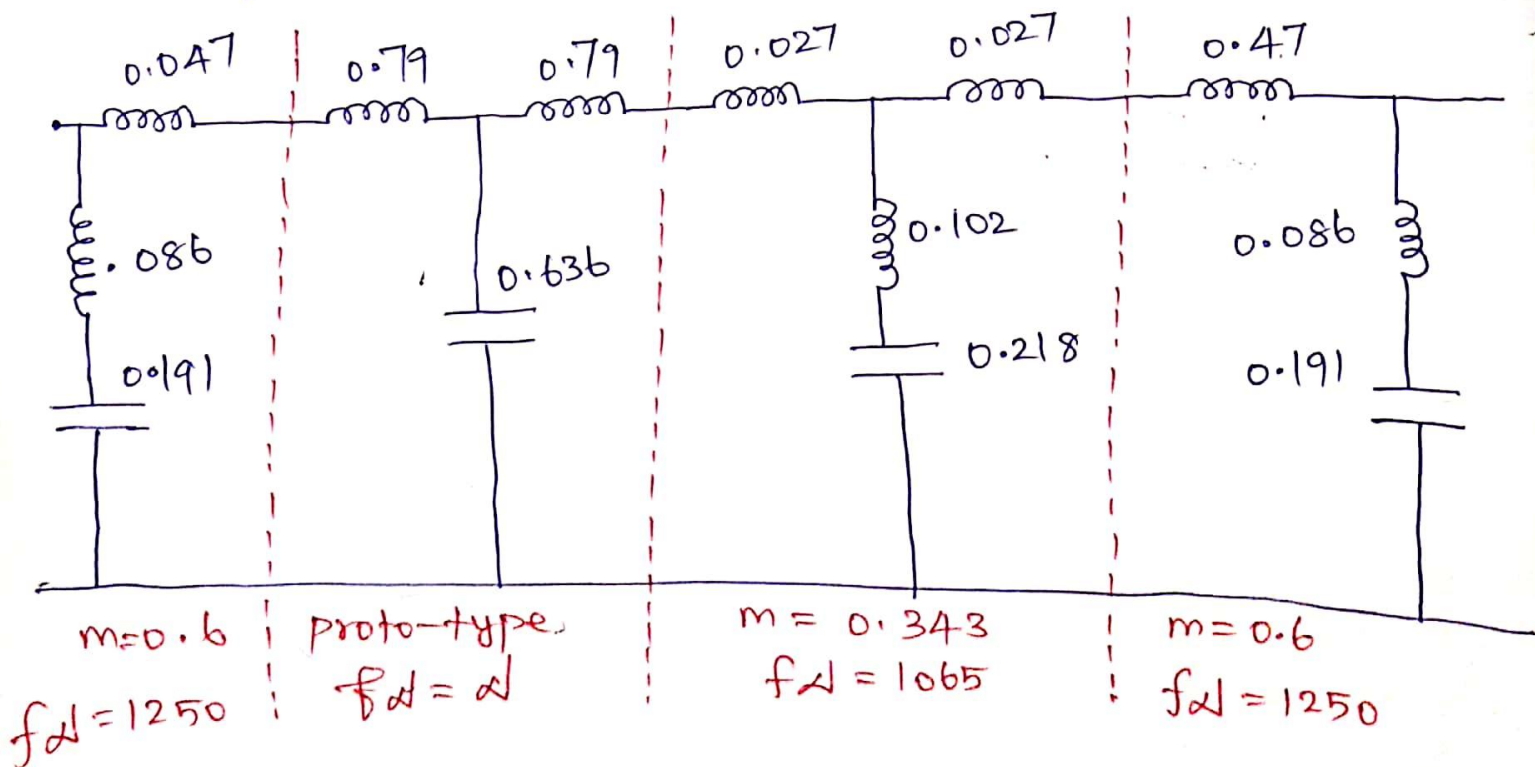
$$mC = 0.6 \times 0.636 \mu = 0.382 \mu\text{F}$$



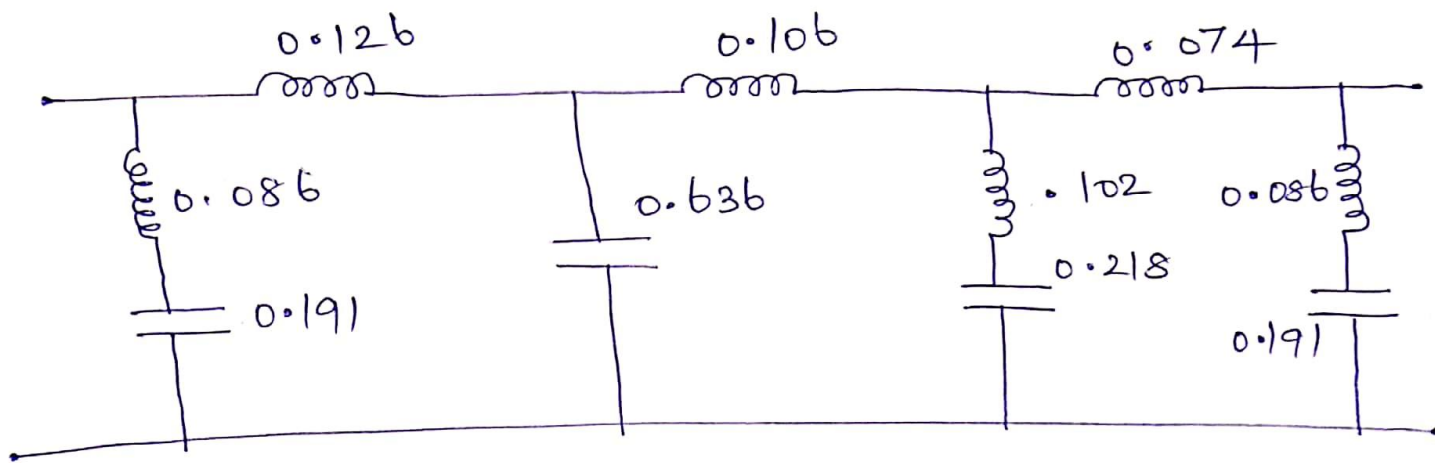
Above T-section is split into two half-sections



combining all the T-sections,



For economy, it is customary to combine elements whenever possible. The series inductors may then be added, resulting in the final design.



Design a composite high pass filter to operate into a load of $600\ \Omega$ and have a cut-off frequency of $1.2\ \text{kHz}$. The filter is to have one constant- k section, one m -derived section with $f_d = 1.1\ \text{kHz}$ and suitable termination half section.

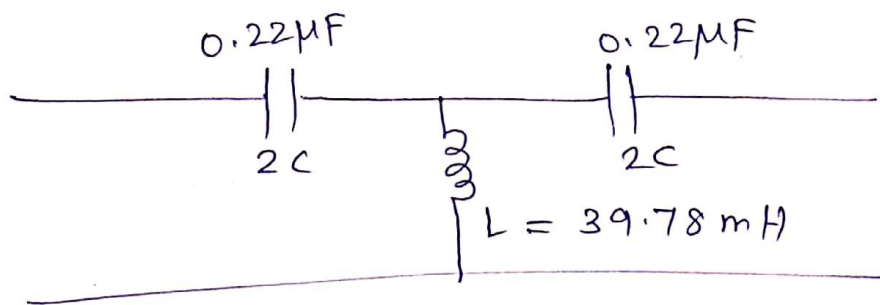
Sol:-

$$R_0 = 600\ \Omega, \quad f_c = 1.2\ \text{kHz}, \quad f_d = 1.1\ \text{kHz}.$$

(i) Constant- k T-section:-

$$L = \frac{R_0}{4\pi f_c} = \frac{600}{4\pi \times 1200} = 39.78\ \text{mH}.$$

$$C = \frac{1}{(4\pi f_c)R_0} = \frac{1}{(4\pi \times 1200) \times 600} = 0.11\ \mu\text{F}$$



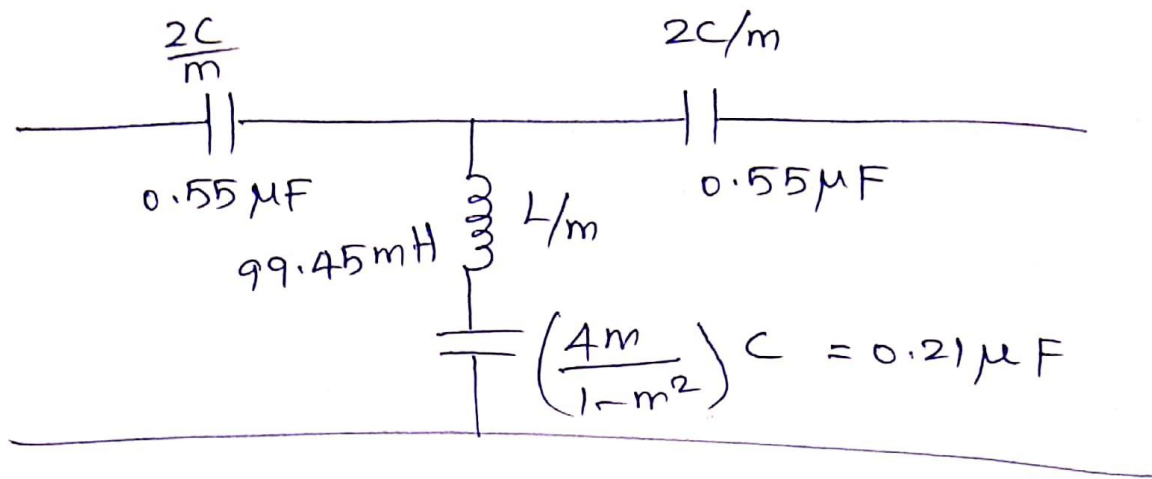
(ii) m -derived T-section:-

$$m = \sqrt{1 - \left(\frac{f_d}{f_c}\right)^2} = \sqrt{1 - \left(\frac{1100}{1200}\right)^2} \approx 0.4$$

$$\frac{2C}{m} = \frac{0.22\ \mu\text{F}}{0.4} = 0.55\ \mu\text{F}$$

$$\frac{L}{m} = \frac{39.78\ \text{mH}}{0.4} = 99.45\ \text{mH}$$

$$\left(\frac{4m}{1-m^2}\right) C = \left[\frac{0.4 \times 4}{1 - (0.4)^2}\right] \times 0.11\ \mu\text{F} = 0.21\ \mu\text{F}$$

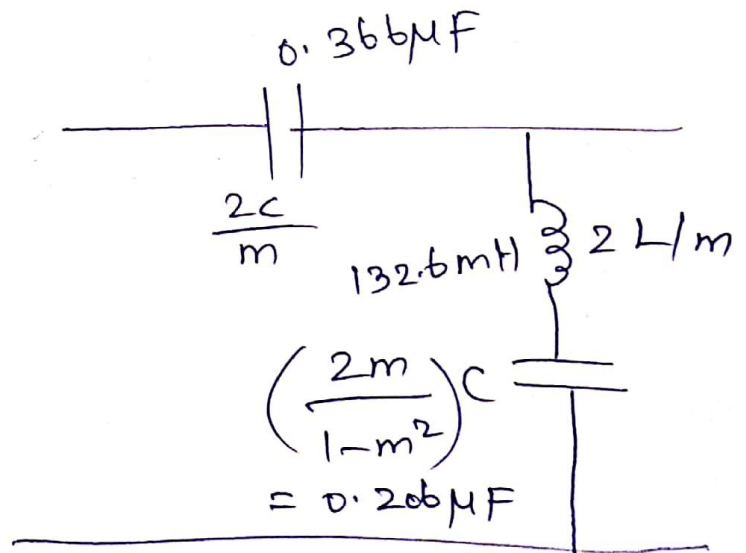
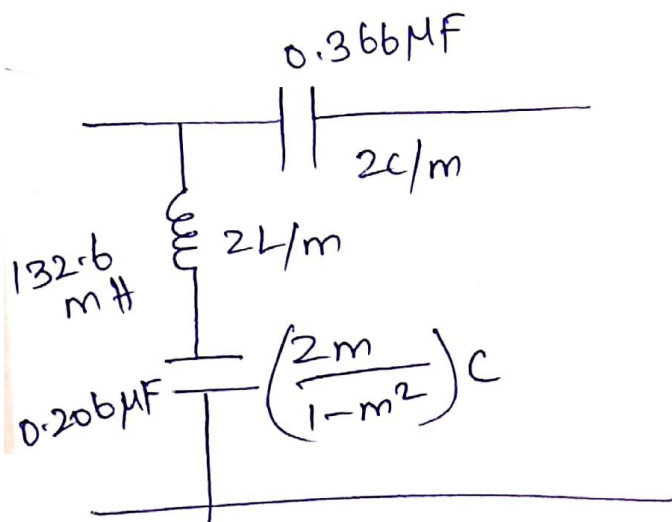


(iii) Half sections :- $m = 0.6$

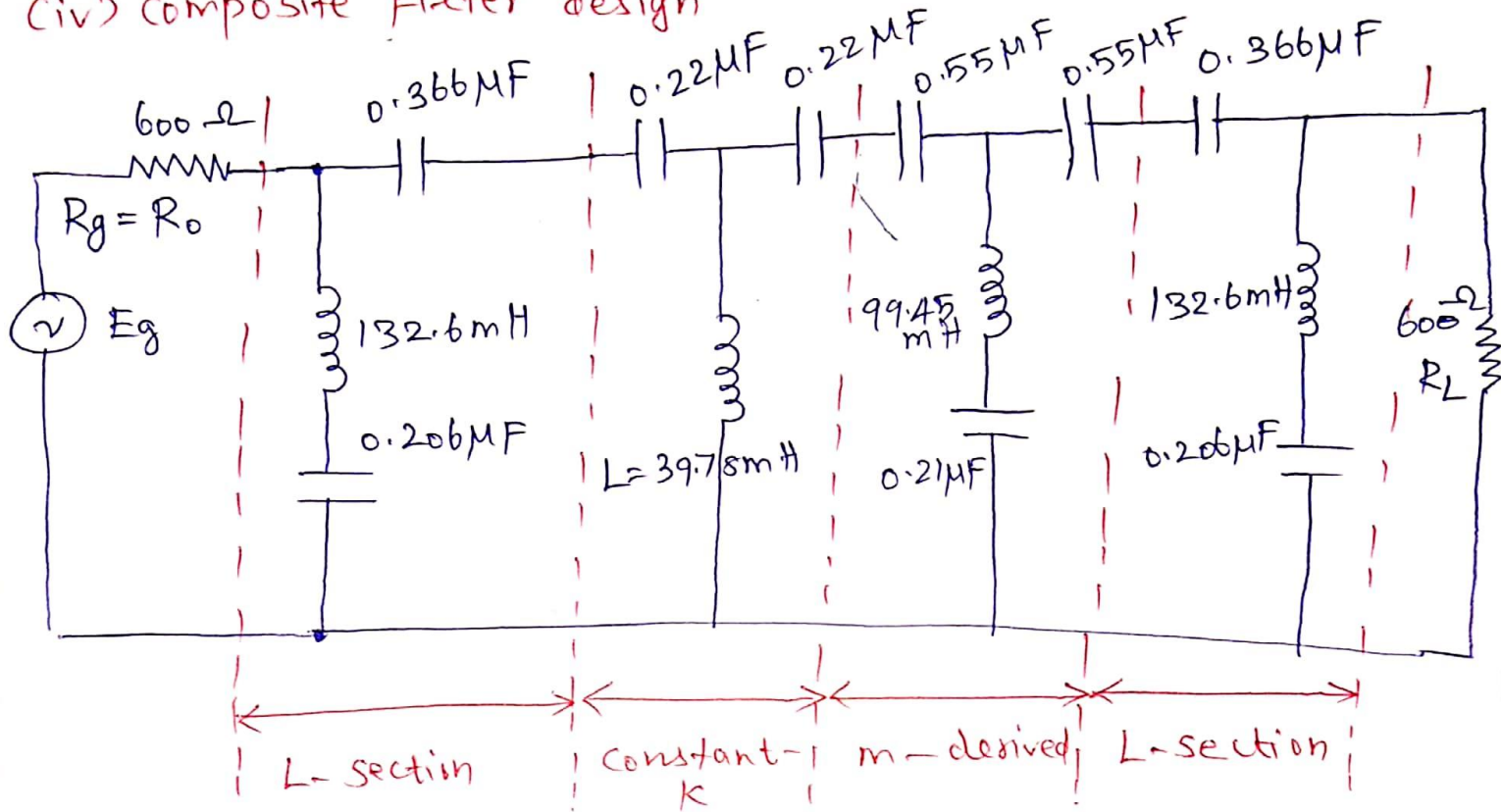
$$\frac{2C}{m} = \frac{2 \times 0.11 \times 10^{-6}}{0.6} = 0.366 \mu F$$

$$\frac{2L}{m} = \frac{2 \times 39.78 \times 10^{-3}}{0.6} = 132.6 \text{ mH}$$

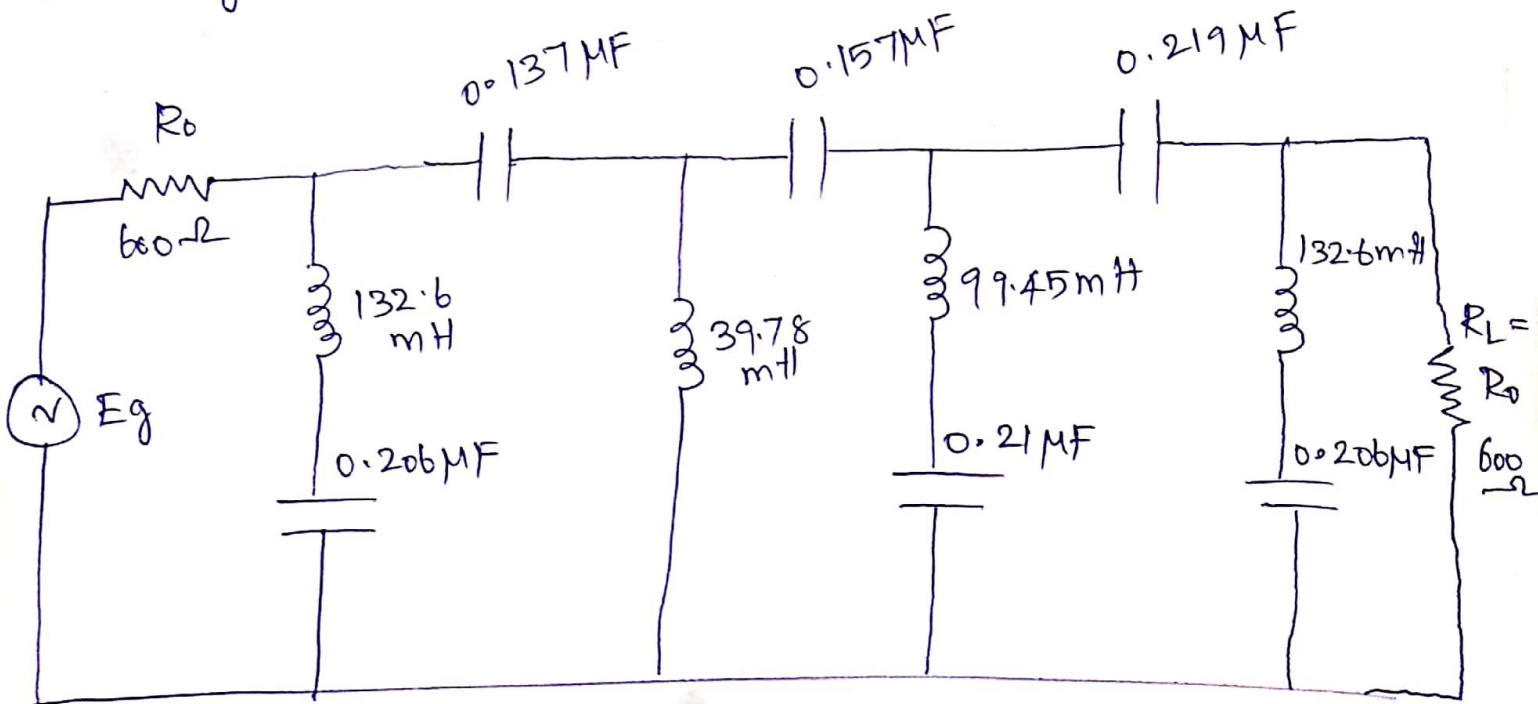
$$\left(\frac{2m}{1-m^2} \right) C = 0.206 \mu F$$



(iv) Composite Filter design



Combining the capacitances,



Microwave Filters

An ideal filter provides:

- Perfect transmission for all frequencies in certain passband region.
- Infinite attenuation in stopband region.

Typical filter responses are:

- Low pass** : Transmits all signals between zero frequency and some upper limit and attenuates all frequencies above the cut off value ω_c .
- High pass** : Transmits all frequencies above some lower cut off frequency and attenuates all frequencies below the cut off value ω_c .
- Band pass** : Transmits all frequencies in the range ω_1 and ω_2 and attenuates all frequencies outside the range.
- Band reject** : Attenuates signals over a band of frequencies

Microwave Filters

Filter design problems at microwave frequencies where distributed parameters must be used is quite complicated.

Two commonly used low frequency filter synthesis techniques are:

- a. **The image parameter method:** Filter with the required passband and stopband characteristics can be synthesized, but without exact frequency characteristics over each region.
- b. **The insertion loss method:** A systematic way to synthesize the desired response with a higher degree of control over the passband and stopband amplitude and phase characteristics.

Design of Microwave Filters by Insertion Loss

Some of the design trade-offs for microwave filter synthesis using the insertion loss method are:

1. A binomial response is used when obtaining a minimum insertion loss is the priority.
2. A Chebyshev response satisfies the requirement for the sharp cutoff.
3. A linear phase filter design is used in cases where the attenuation rate can be sacrificed for a better phase response.

Characterization by Power Loss Ratio

The insertion loss or the power loss ratio in a filter network can be defined as:

$$P_{LR} = \frac{\text{Power available from the source}}{\text{Power delivered to the load}} \\ = \frac{P_{in}}{P_{Load}} = \frac{1}{1 - |\Gamma(\omega)|^2}$$

The insertion loss in dB is given by

$$IL = 10 \log P_{LR}$$

Since $|\Gamma(\omega)|^2$ is an even function of ω . We can express $|\Gamma(\omega)|^2$ as a polynomial in ω^2 .

$$|\Gamma(\omega)|^2 = \frac{M(\omega^2)}{M(\omega^2) + N(\omega^2)}$$

where, M and N are real polynomials in ω^2 .

$$\therefore P_{LR} = \frac{1}{1 - \frac{M(\omega^2)}{M(\omega^2) + N(\omega^2)}} \\ = \frac{M(\omega^2) + N(\omega^2)}{N(\omega^2)} \\ \therefore P_{LR} = 1 + \frac{M(\omega^2)}{N(\omega^2)}$$

Some Practical Filter Responses

Maximally flat: Such filters are also known as binomial or Butterworth filters. For a low pass filter, power loss ratio is specified as

$$P_{LR} = 1 + k^2 \left(\frac{\omega}{\omega_c} \right)^{2N}$$

where, N is the order of the filter and ω_c is the cut off frequency.

The power loss at the band edge is $1 + k^2$.

For $\omega > \omega_c$, the attenuation increases monotonically with frequency.

For $\omega \gg \omega_c$,

$$P_{LR} \simeq k^2 \left(\frac{\omega}{\omega_c} \right)^{2N}$$

The insertion loss increases at the rate $20N$ dB/decade.

Some Practical Filter Responses

Equal ripple: Such filter response is also known as Chebyshev response. For a low pass filter power loss ratio is given by

$$P_{LR} = 1 + k^2 T_N^2 \left(\frac{\omega}{\omega_c} \right)$$

For $\omega < \omega_c$, $T_N^2 \left(\frac{\omega}{\omega_c} \right)$ will oscillate between ± 1 .

The passband response has ripples of amplitude $1 + k^2$.

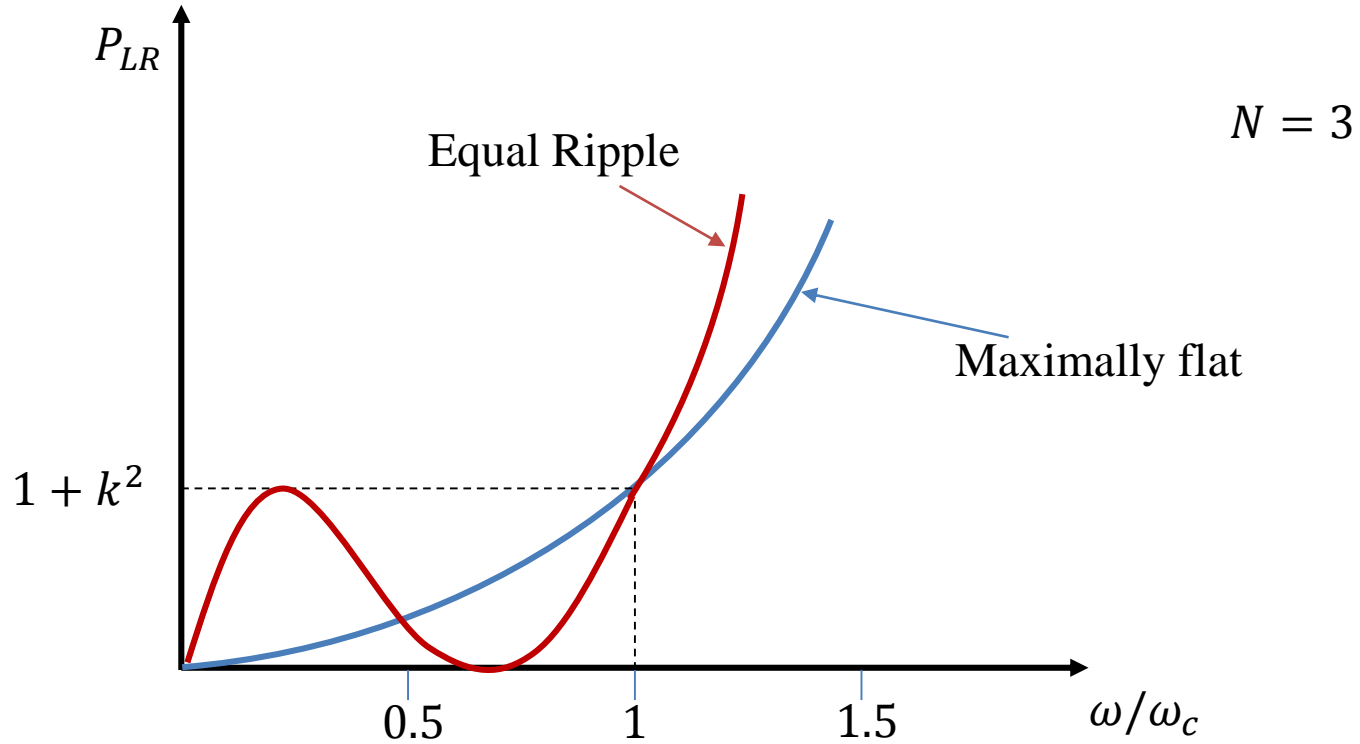
For $\omega \gg \omega_c$,

$$T_N^2 \left(\frac{\omega}{\omega_c} \right) \approx \frac{1}{2} \left(\frac{2\omega}{\omega_c} \right)^{2N}$$

$$\therefore P_{LR} \approx \frac{k^2}{4} \left(\frac{2\omega}{\omega_c} \right)^{2N}$$

The insertion loss increases at the rate $20N$ dB/decade similar to binomial. However, the insertion loss in Chebyshev response is $\frac{(2^{2N})}{4}$ greater than binomial response at $\omega \gg \omega_c$.

Some Practical Filter Responses



Some Practical Filter Responses

Linear phase: In some applications, a linear phase response is desirable in the passband.

A linear phase response can be achieved with the following phase response

$$\phi(\omega) = A\omega \left[1 + p \left(\frac{\omega}{\omega_c} \right)^{2N} \right]$$

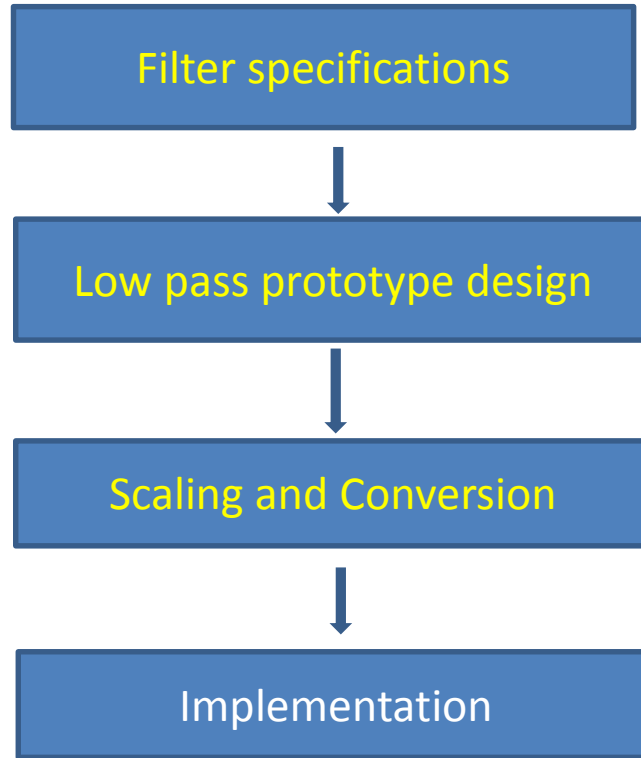
where, $\phi(\omega)$ is the phase of the voltage transfer function of the filter and p is a constant.

The group delay is defined as

$$\begin{aligned} \tau_d &= \frac{d\phi}{d\omega} \\ &= A \left[1 + p(2N + 1) \left(\frac{\omega}{\omega_c} \right)^{2N} \right] \end{aligned}$$

Group delay is a maximally flat response

Steps involved in filter design by insertion loss method



Maximally Flat Low-Pass Filter Prototype

Let the source impedance be 1Ω and
 $\omega_c = 1$ rad/sec.

For $N = 2$,

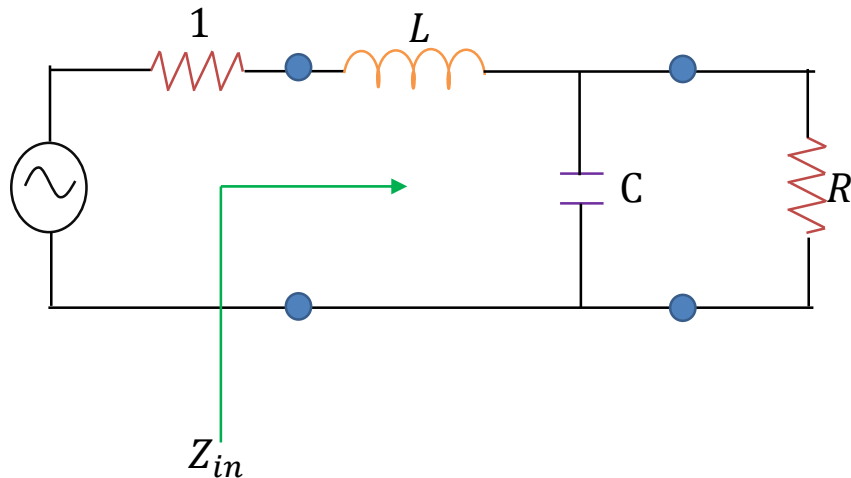
$$P_{LR} = 1 + \omega^4$$

and

$$Z_{in} = j\omega L + \frac{R(1 - j\omega RC)}{1 + \omega^2 R^2 C^2}$$

$$\Gamma = \frac{Z_{in} - 1}{Z_{in} + 1}$$

$$\therefore P_{LR} = \frac{1}{1 - |\Gamma|^2} = \frac{1}{1 - \left[\frac{(Z_{in} - 1)(Z_{in}^* - 1)}{(Z_{in} + 1)(Z_{in}^* + 1)} \right]} = \frac{|Z_{in} + 1|^2}{2(Z_{in} + Z_{in}^*)}$$



Maximally Flat Low-Pass Filter Prototype

Now

$$Z_{in} + Z_{in}^* = j\omega L + \frac{R(1 - j\omega RC)}{1 + \omega^2 R^2 C^2} - j\omega L + \frac{R(1 + j\omega RC)}{1 + \omega^2 R^2 C^2} = \frac{2R}{1 + \omega^2 R^2 C^2}$$

and

$$\begin{aligned} |Z_{in} + 1|^2 &= \left| j\omega L + \frac{R(1 - j\omega RC)}{1 + \omega^2 R^2 C^2} + 1 \right|^2 \\ &= \left| \left(\frac{R}{1 + \omega^2 R^2 C^2} + 1 \right) + j \left(\omega L - \frac{\omega R^2 C}{1 + \omega^2 R^2 C^2} \right) \right|^2 \\ &= \left(\frac{R}{1 + \omega^2 R^2 C^2} + 1 \right)^2 + \left(\omega L - \frac{\omega R^2 C}{1 + \omega^2 R^2 C^2} \right)^2 \end{aligned}$$

Maximally Flat Low-Pass Filter Prototype

$$\begin{aligned}\therefore P_{LR} &= \frac{|Z_{in} + 1|^2}{2(Z_{in} + Z_{in}^*)} = \frac{1 + \omega^2 R^2 C^2}{4R} \left[\left(\frac{R}{1 + \omega^2 R^2 C^2} + 1 \right)^2 + \left(\omega L - \frac{\omega R^2 C}{1 + \omega^2 R^2 C^2} \right)^2 \right] \\ &= \frac{1}{4R} [R^2 + 2R + 1 + \omega^2 R^2 C^2 + \omega^2 L^2 + \omega^4 L^2 R^2 C^2 - 2\omega^2 LCR^2] \\ &= 1 + \frac{1}{4R} [(1 - R)^2 + \omega^2 (R^2 C^2 + L^2 - 2LCR^2) + \omega^4 L^2 R^2 C^2]\end{aligned}$$

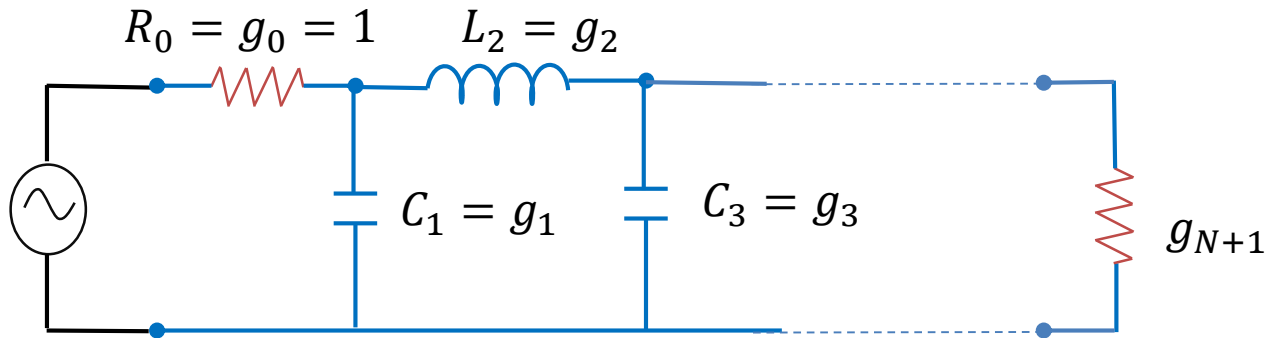
On comparing with $P_{LR} = 1 + (0)\omega^2 + (1)\omega^4$, we get

$$\begin{aligned}1 - R &= 0 \Rightarrow R = 1, \\ C^2 + L^2 - 2LC &= 0 \Rightarrow (L - C)^2 = 0 \Rightarrow L = C\end{aligned}$$

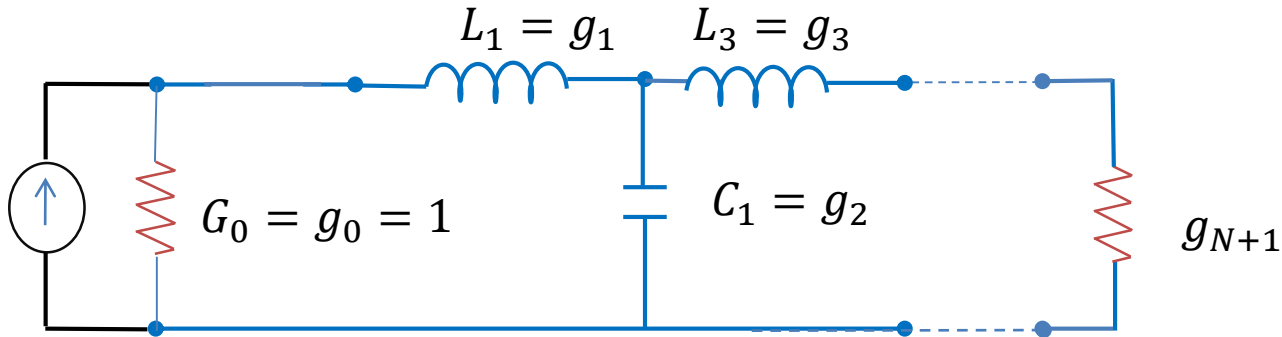
and

$$\frac{L^2 R^2 C^2}{4R} = 1 \Rightarrow \frac{L^2 C^2}{4} = 1 \Rightarrow \frac{C^2 C^2}{4} = 1 \Rightarrow L = C = \sqrt{2}$$

Ladder circuit for a lowpass prototype



Prototype beginning
with a shunt element



Prototype beginning
with a series element

Element Values for Maximally Flat Low-Pass Filter Prototypes

$$g_0 = 1, \omega_c = 1$$

N	g_1	g_2	g_3	g_4	g_5	g_6
1	2.0000	1.0000				
2	1.4142	1.4142	1.0000			
3	1.0000	2.0000	1.0000	1.0000		
4	0.7654	1.8478	1.8478	0.7654	1.0000	
5	0.6180	1.6180	2.0000	1.6180	0.6180	1.0000

For practical filter, it will be necessary to determine the order of the filter. This is usually dictated by a specification on the insertion loss at some frequency in the stop-band of the filter.

Impedance and frequency scaling

The prototype filter has $R_s = 1$ and $\omega_c = 1$. Also for a maximally flat response, the prototype has $R_L = 1$

A source resistance of R_0 can be obtained by multiplying all the impedances of the prototype design by R_0

The change of cutoff frequency from unity to ω_c requires scaling of frequency dependence of filter which is accomplished by replacing ω by ω/ω_c

Therefore, when both impedance and frequency scaling is applied,

$$L'_k = \frac{R_0 L_k}{\omega_c} \qquad C'_k = \frac{C_k}{R_0 \omega_c}$$

Note that with impedance scaling, the scaled values of source and load resistances become R_0 and $R_0 R_L$

Example: Design of a Low Pass Butterworth Filter

Let us consider a maximally flat filter that has cutoff frequency of 2 GHz and the filter provides at least 15 dB attenuation at 4 GHz. The source and load impedances are 50Ω

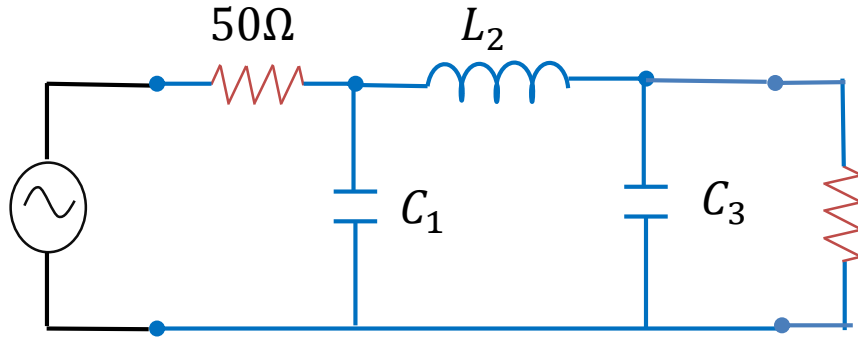
First we need to determine the order of the filter. From the expression of P_{LR} , we find that at angular frequency ω , the attenuation of the filter in dB is

$$10 \log_{10} \left(1 + \left(\frac{\omega}{\omega_c} \right)^{2N} \right)$$

Therefore, $1.5 = \log_{10}(1 + 2^{2N})$. So we get $N = 2.47$ i.e. we use $N = 3$

From the table, $g_1 = 1$ $g_2 = 2$ $g_3 = 1$

Example: Continued



After impedance and frequency scaling

$$C_1 = 1.5915 \text{ pF}$$

$$L_2 = 7.9577 \text{ nH}$$

$$C_3 = 1.5915 \text{ pF}$$

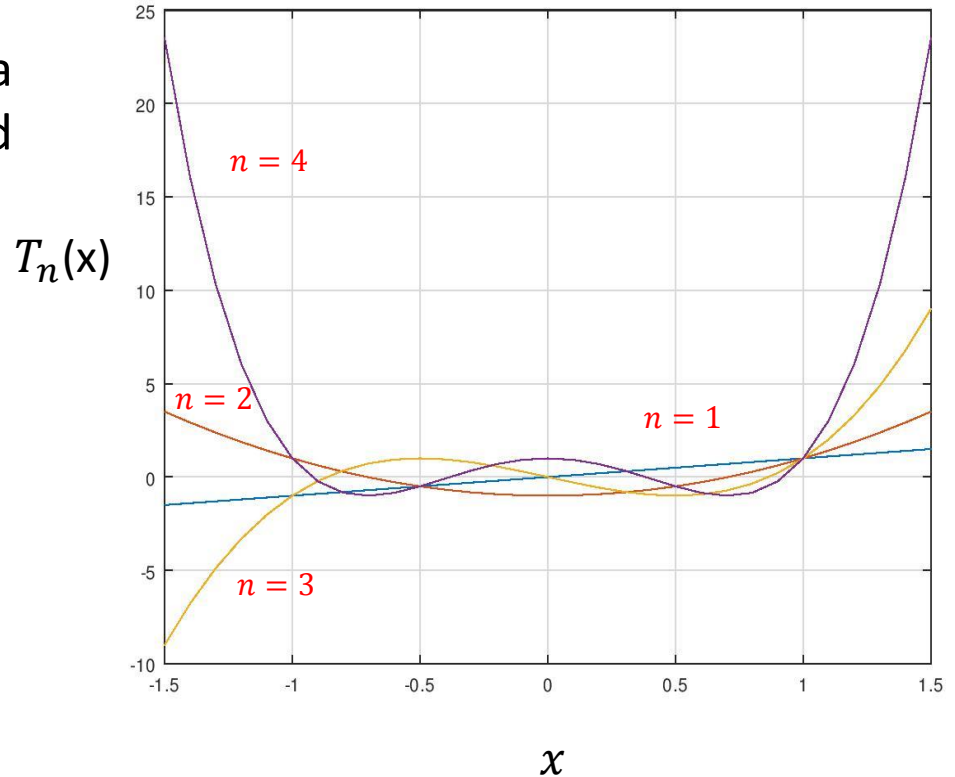
Chebyshev Polynomials

n th order Chebyshev polynomial is a polynomial of degree n and denoted by $T_n(x)$

$$T_1(x) = x$$

$$T_2(x) = 2x^2 - 1$$

$$T_n(x) = 2xT_{n-1}(x) - T_{n-2}(x)$$



Equal Ripple Low-Pass Filter Prototype

Let the cutoff frequency be $\omega_c = 1$ rad/sec.

$$P_{LR} = 1 + k^2 T_N^2(\omega)$$

Since,

$$T_N(0) = \begin{cases} 0 & \text{for } N \text{ odd} \\ \pm 1 & \text{for } N \text{ even} \end{cases}$$

Therefore, at $\omega = 0$,

$$P_{LR} = \begin{cases} 1 & \text{for } N \text{ odd} \\ 1 + k^2 & \text{for } N \text{ even} \end{cases}$$

As

$$T_2(x) = 2x^2 - 1$$

For $N = 2$,

$$P_{LR} = 1 + k^2 T_2^2(\omega)$$

$$= 1 + k^2 (2\omega^2 - 1)^2$$

$$= 1 + k^2 (4\omega^4 - 4\omega^2 + 1)$$

Equal Ripple Low-Pass Filter Prototype

For the two element network,

we have seen

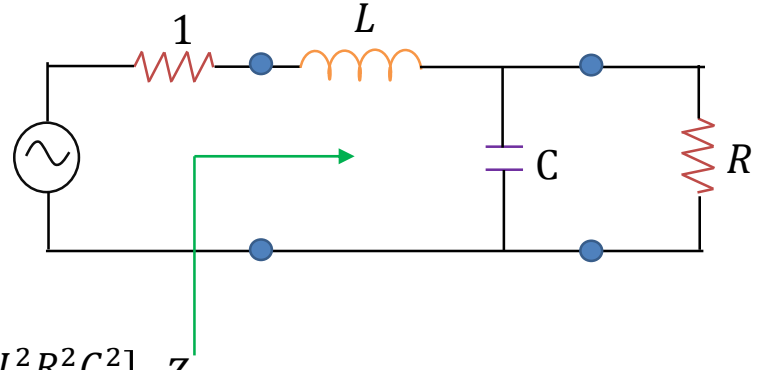
$$P_{LR} = 1 + \frac{1}{4R} [(1 - R)^2 + \omega^2(R^2C^2 + L^2 - 2LCR^2) + \omega^4L^2R^2C^2] \quad Z_{in}$$

$$\text{Also } P_{LR} = 1 + k^2T_2^2(\omega) = 1 + k^2(4\omega^4 - 4\omega^2 + 1)$$

For $\omega = 0$,

$$P_{LR} = 1 + k^2 = 1 + \frac{1}{4R} [(1 - R)^2]$$

$$4k^2 = \frac{(1-R)^2}{4R} \Rightarrow R = 1 + 2k^2 \pm 2k\sqrt{1 + k^2} \text{ (for } N \text{ even)}$$



Equal Ripple Low-Pass Filter Prototype

Equating the two expressions for P_{LR}

$$1 + k^2(4\omega^4 - 4\omega^2 + 1) = 1 + \frac{1}{4R} [(1 - R)^2 + \omega^2(R^2C^2 + L^2 - 2LCR^2) + \omega^4L^2R^2C^2]$$

We get

$$-4k^2 = \frac{R^2C^2 + L^2 - 2LCR^2}{4R}$$

and

$$4k^2 = \frac{L^2R^2C^2}{4R}$$

The values of L and C can be obtained by solving these equations.

Note that value for R (for N even) is not unity, so there will be an impedance mismatch if the load has a unity (normalized) impedance; this can be corrected with a quarter-wave transformer, or by using an additional filter element to make N odd. For odd N , it can be shown that $R = 1$.

Element Values for Equi-ripple Low-Pass Filter Prototypes

$$g_0 = 1, \omega_c = 1, 0.5 \text{ dB ripple}$$

N	g_1	g_2	g_3	g_4	g_5	g_6
1	0.6986	1.0000				
2	1.4029	0.7071	1.9841			
3	1.5963	1.0967	1.5963	1.0000		
4	1.6703	1.1926	2.3661	0.8491	1.9841	
5	1.7058	1.2296	2.5408	1.2296	1.7058	1.0000

For practical filter, it will be necessary to determine the order of the filter. This is usually dictated by a specification on the insertion loss at some frequency in the stop-band of the filter.

Low Pass to High Pass Transformation

The low pass prototype filter designs can be transformed to high pass, band pass or band reject response.

Loss pass to high pass transformation is achieved by the frequency substitution:

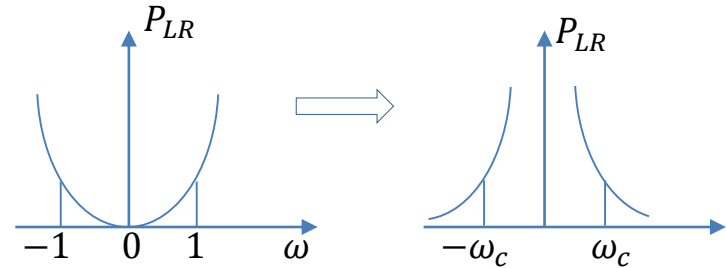
$$\frac{-\omega_c}{\omega} \text{ for } \omega$$

When this transformation is applied

$$j\omega L_k \text{ becomes } -j \frac{\omega_c}{\omega} L_k = \frac{1}{j\omega C'_k}, \text{ where } C'_k = \frac{1}{\omega_c L_k}$$

and

$$j\omega C_k \text{ becomes } -j \frac{\omega_c}{\omega} C_k = \frac{1}{j\omega L'_k}, \text{ where } L'_k = \frac{1}{\omega_c C_k}$$



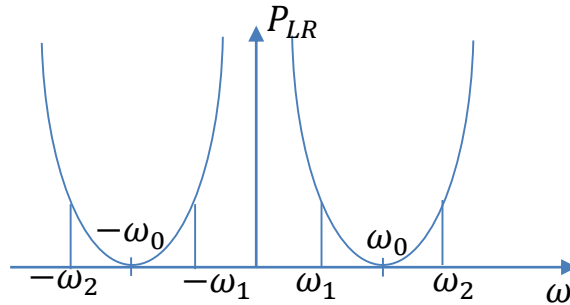
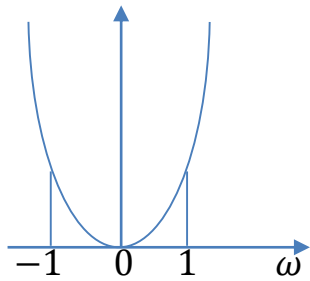
After performing impedance scaling

$$C'_k = \frac{1}{R_0 \omega_c L_k}$$

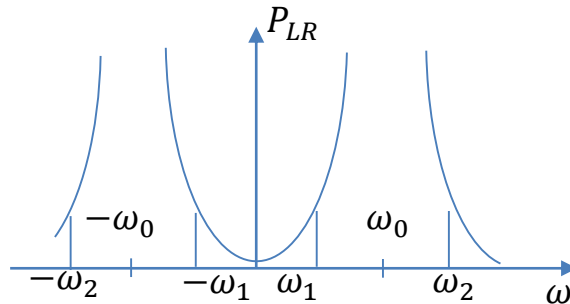
and

$$L'_k = \frac{R_0}{\omega_c C_k}$$

Low Pass to Band Pass and Band Stop Transformation



Band Pass



Band Stop

For LPF to BPF

$$\omega \leftarrow \frac{\omega_0}{\omega_2 - \omega_1} \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right)$$

$$= \frac{1}{\Delta} \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right)$$

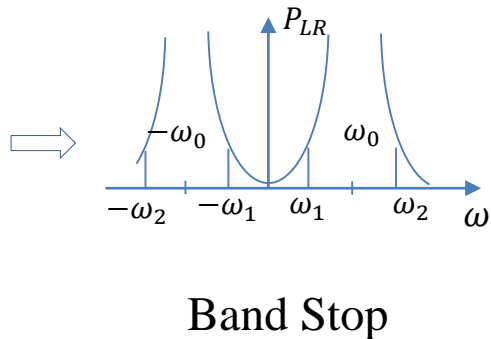
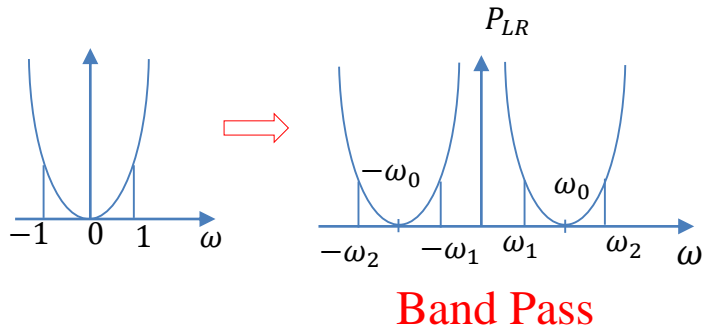
where,

$$\Delta = \frac{\omega_2 - \omega_1}{\omega_0}$$

Simpler equations are obtained when

$$\omega_0 = \sqrt{\omega_1 \omega_2}$$

Low Pass to Band Pass and Band Stop Transformation



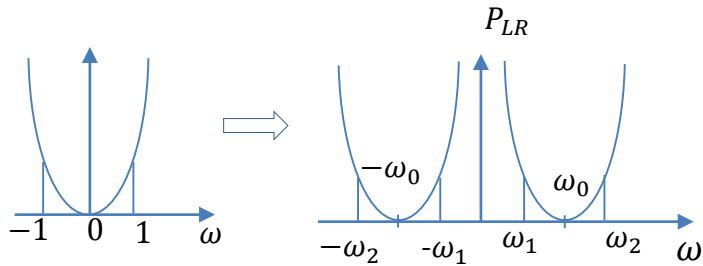
When this transformation is applied,
A series inductor L_k is transformed to a **series LC** circuit with

$$L'_k = \frac{L_k}{\Delta\omega_0} \quad C'_k = \frac{\Delta}{\omega_0 L_k}$$

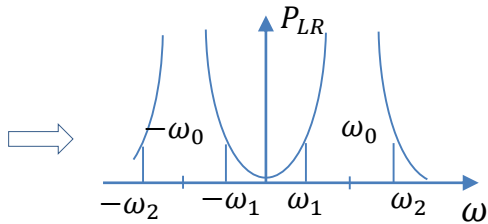
And the shunt capacitor C_k is transformed to a **shunt LC** circuit with elements

$$L'_k = \frac{\Delta}{\omega_0 C_k} \quad C'_k = \frac{C_k}{\Delta\omega_0}$$

Low Pass to Band Pass and Band Stop Transformation



Band Pass



Band Stop

For LPF to Band Stop

$$\omega \leftarrow \Delta \left(\frac{\omega}{\omega_0} - \frac{\omega_0}{\omega} \right)^{-1}$$

A series inductors are transformed to **parallel LC circuit** with

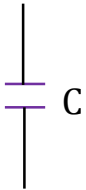
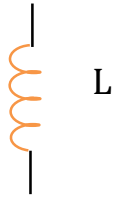
$$L'_k = \frac{\Delta L_k}{\omega_0} \quad C'_k = \frac{1}{\omega_0 \Delta L_k}$$

And the shunt capacitors are transformed to a **series LC circuit** with elements

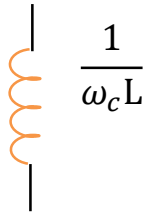
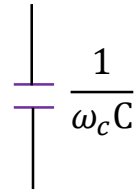
$$L'_k = \frac{1}{\omega_0 \Delta C_k} \quad C'_k = \frac{\Delta C_k}{\omega_0}$$

Summary of prototype filter transformations

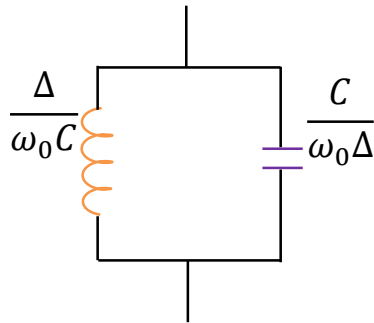
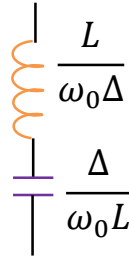
Low Pass



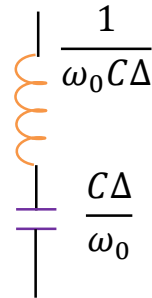
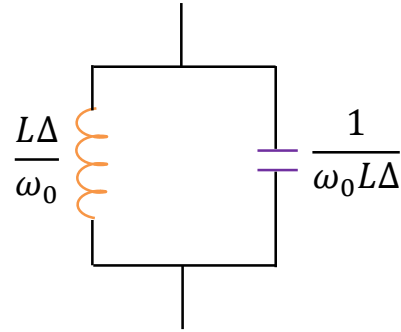
High Pass



Band Pass



Band Stop



Filter Implementation

- ❑ Lumped element filter design discussed so far works well at lower frequencies.
- ❑ At higher RF frequencies lumped element inductors and capacitors are generally available for limited range of values.
- ❑ At higher microwave frequencies, such elements are difficult to implement.
- ❑ Distributed elements such as stubs are often used to approximate the lumped elements.
- ❑ Conversion of lumped element to equivalent transmission line sections can be done using Richards's transformation.
- ❑ Moreover, at microwave frequencies the spacing between the filter elements are also to be considered.
- ❑ Kuroda's identities are used to physically separate filter elements by various transmission line sections.

Filter Implementation Tutorial

Steps to design a Microstrip Filter

1. Select the normalized filter parameters to meet the design criteria.
2. Replace the inductances and capacitances by equivalent $\lambda_0/8$ transmission lines.
3. Convert series stub lines to shunt stubs through Kudora's identities.
4. De-normalize and select equivalent microstrip lines (length, width, and dielectric constant).

Design Goal

Design a low-pass filter whose input and output are matched to a 50Ω impedance and that meets the following specifications: cut-off frequency of 3 GHz; equi-ripple of 0.5 dB; and rejection of at least 40 dB at approximately twice the cut-off frequency. Assume a dielectric material that results in a phase velocity of 60% of the speed of light.

1. Find the order of Filter (N)

Step 1
coefficients

the filter has to be of order $N = 5$, with

$$g_1 = 1.7058 = g_5, g_2 = 1.2296 = g_4, g_3 = 2.5408, g_6 = 1.0$$

The normalized low-pass filter is given in Figure

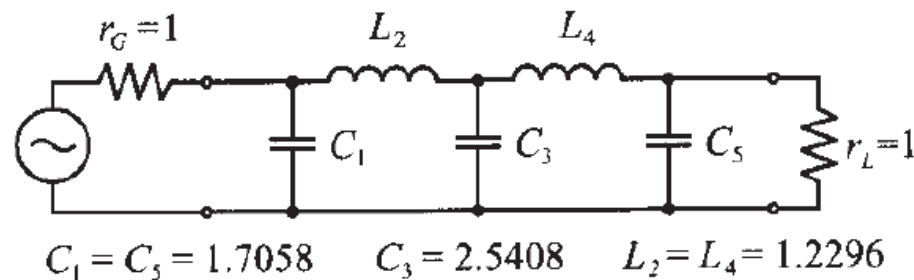


Figure 1 Normalized low-pass filter of order $N = 5$.

2. Apply Richard's Transformation

Step 2 The inductances and capacitances in Figure 1 are replaced by open and short circuit series and shunt stubs as shown in Figure 2. This is a direct consequence of applying Richards transformation. The characteristic line impedances and admittances are

$$Y_1 = Y_5 = g_1, Y_3 = g_3, Z_2 = Z_4 = g_4$$

2. Apply Richard's Transformation

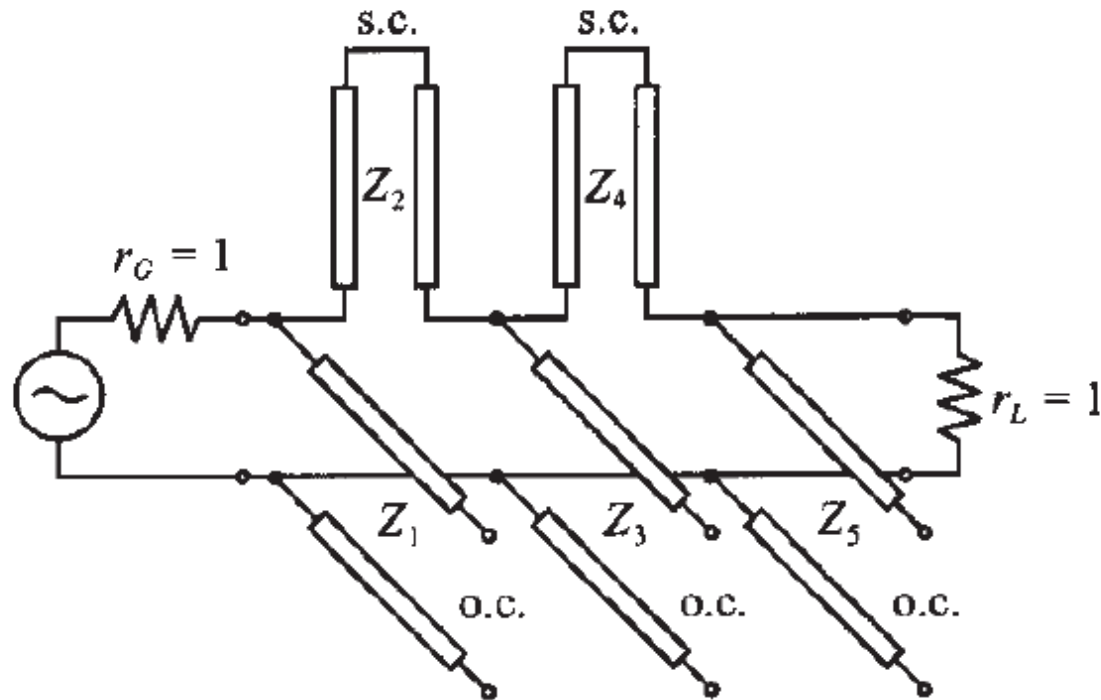


Figure 2 Replacing inductors and capacitors by series and shunt stubs (o.c. = open circuit line, s.c. = short circuit line).

3. Apply Kuroda's Identities

Step 3 To match source and load sides, and to make the filter realizable, unit elements are introduced with the intent to apply the first and second of Kuroda's identities to convert all series stubs into shunt stubs. Since we have a fifth-order filter we must deploy a total of four unit elements to convert all series connected short-circuited stubs into shunt connected open-circuit stubs. To clarify this process we divide this step into several substeps.

3. Apply Kuroda's Identities

First, we introduce two unit elements on the input and output ends of the filter, as shown in Figure 3.

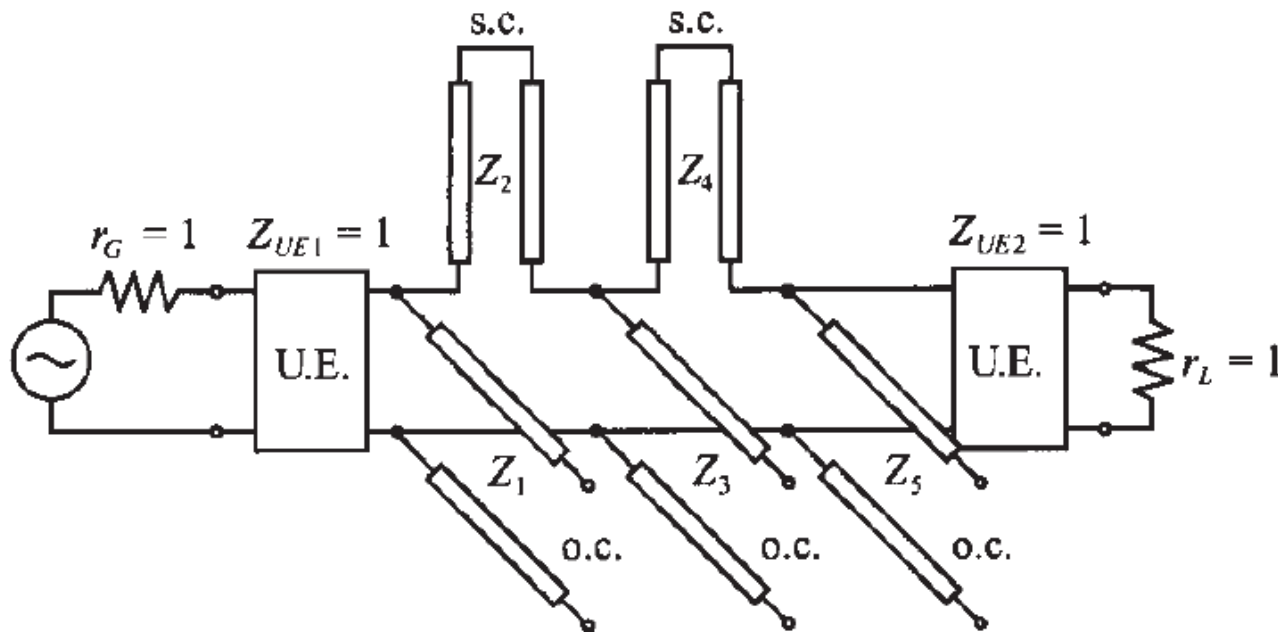


Figure 3 Deployment of the first set of unit elements (U.E. = unit element).

3. Apply Kuroda's Identities

The introduction of unit elements does not affect the filter performance since they are matched to source and load impedances. The result of applying Kuroda's identities to the first and last shunt stubs is shown in Figure 4.

3. Apply Kuroda's Identities

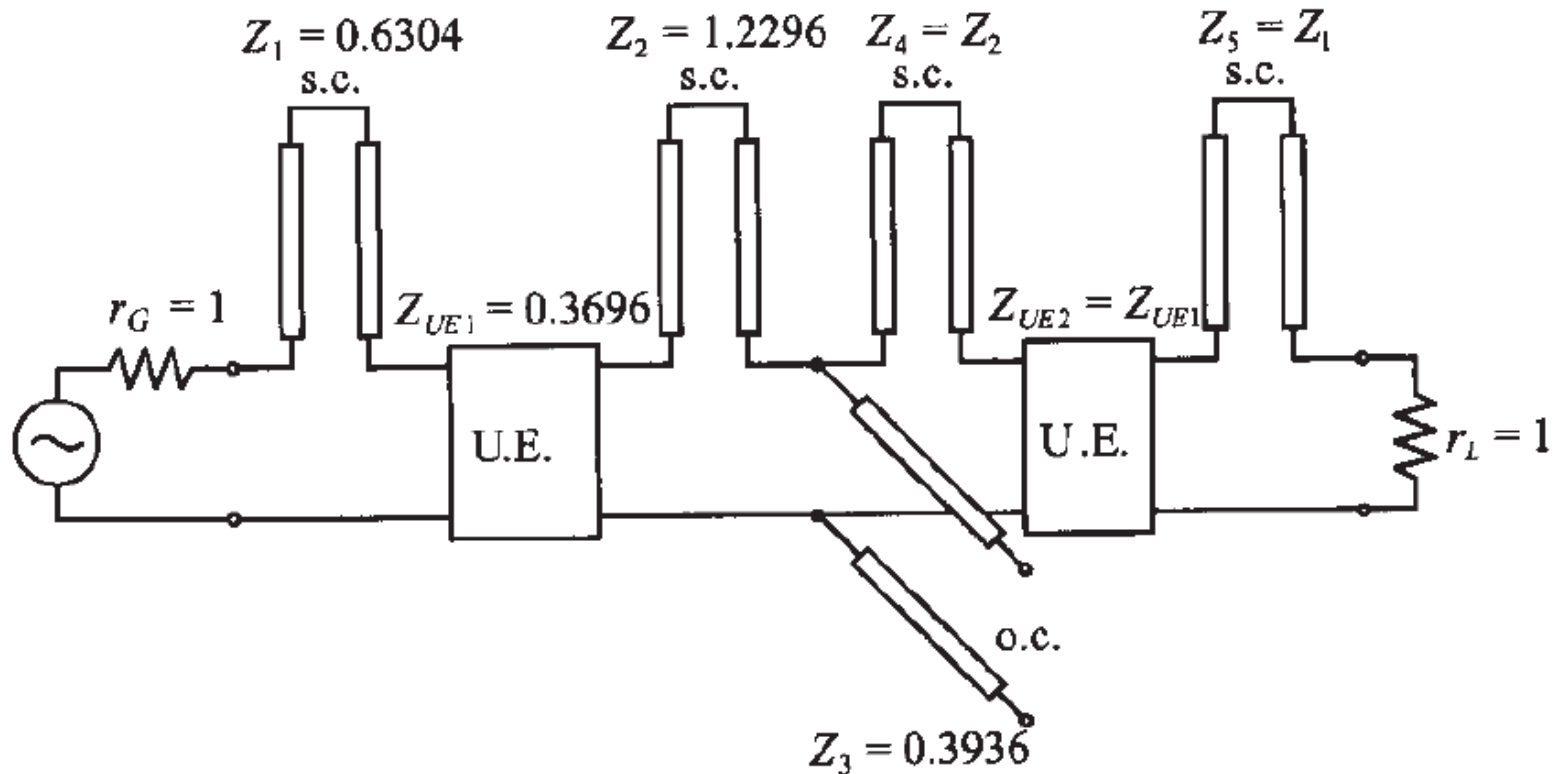


Figure 4 Converting shunt stubs to series stubs.

3. Apply Kuroda's Identities

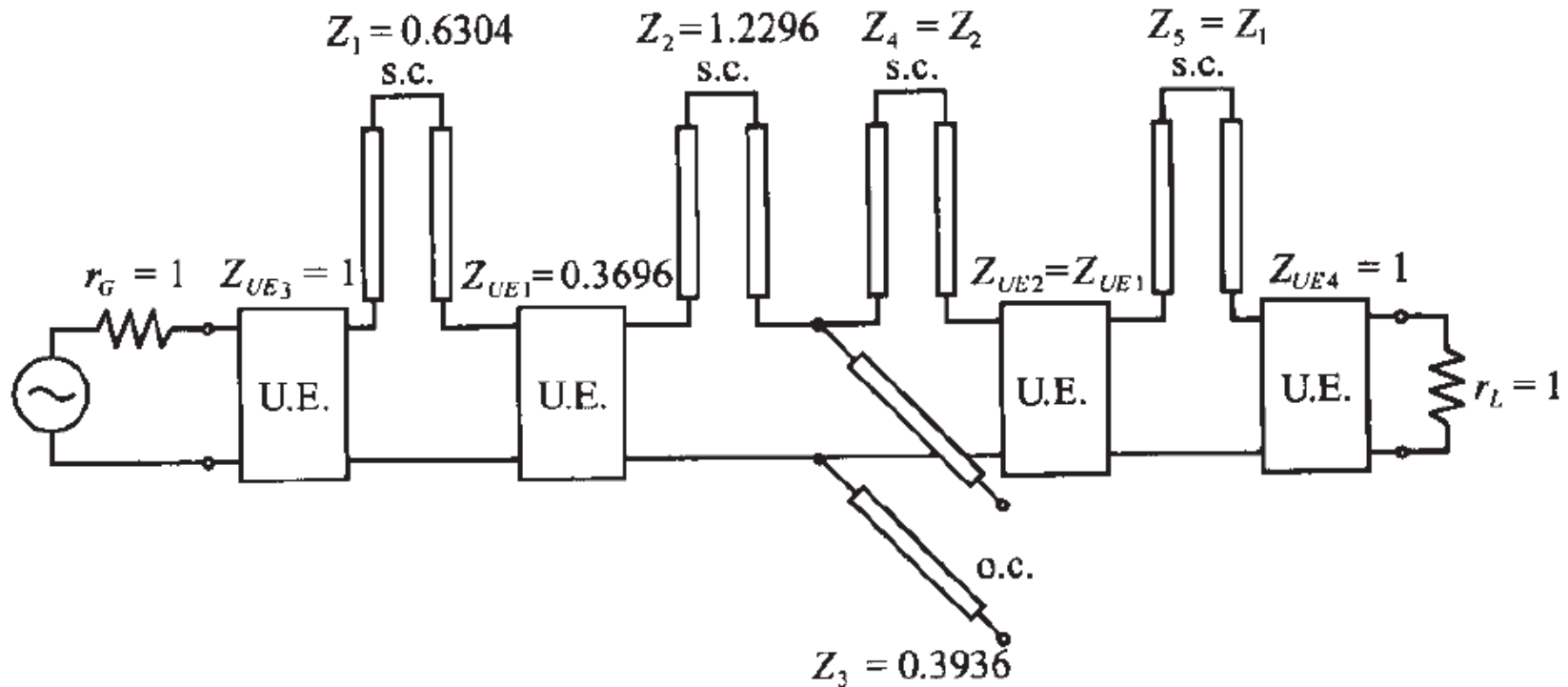


Figure 5 Deployment of the second set of unit elements to the fifth-order filter.

3. Apply Kuroda's Identities

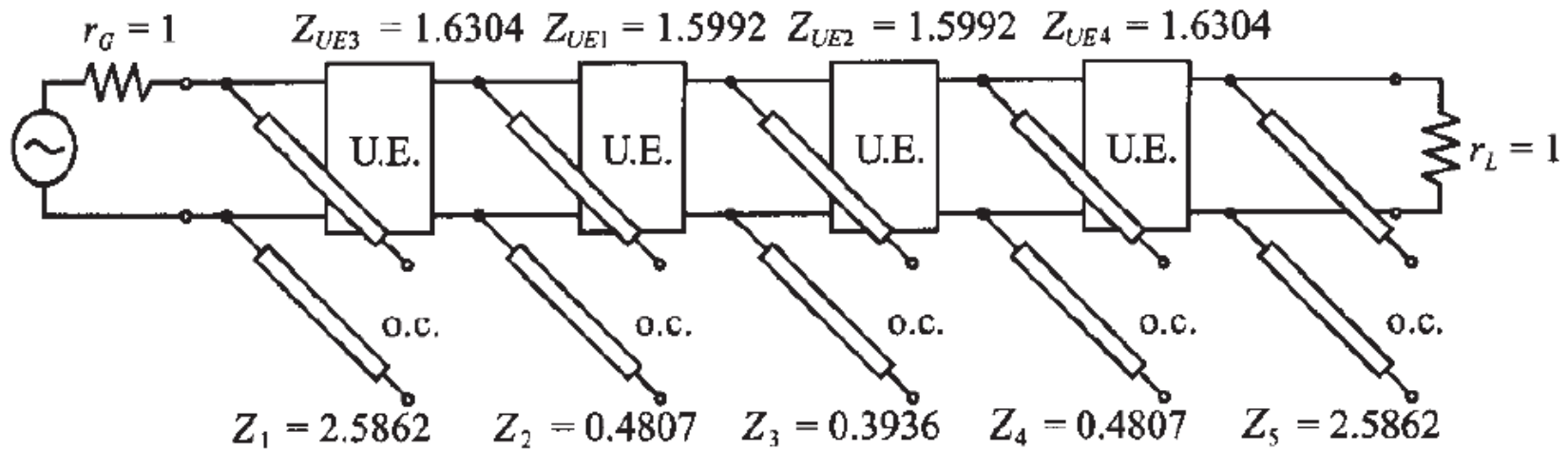
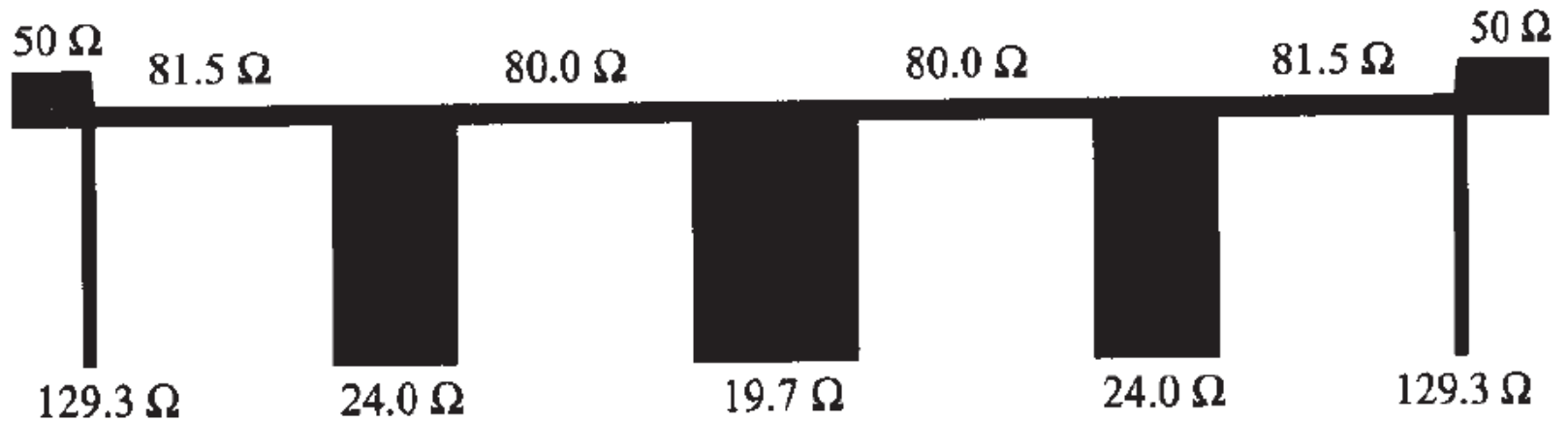


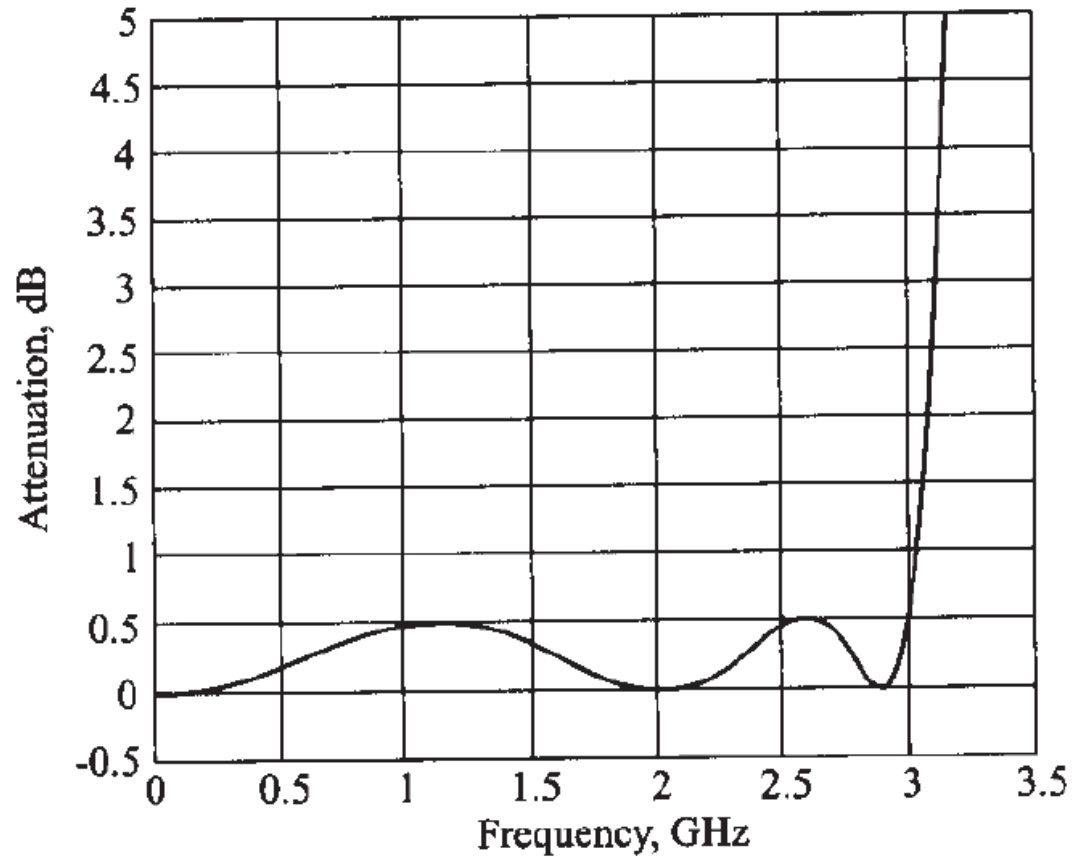
Figure 6 Realizable filter circuit obtained by converting series and shunt stubs using Kuroda's identities.

4. Denormalization to 50 ohms



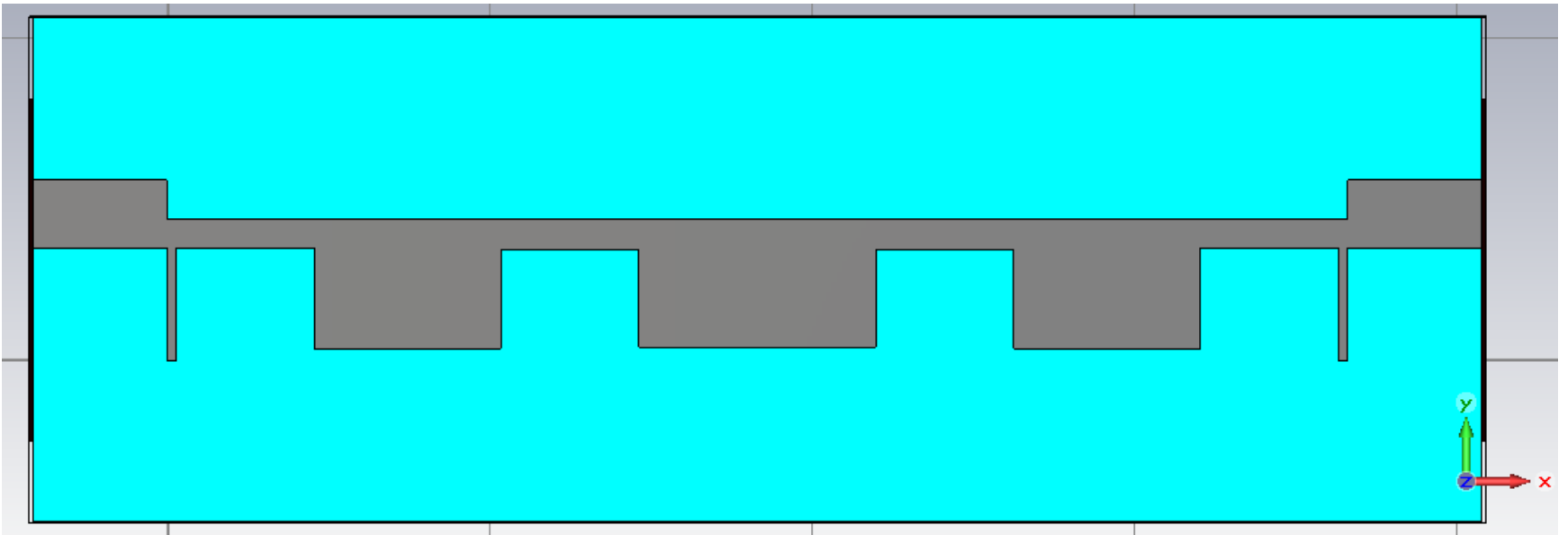
(a) Microstrip line low-pass filter implementation

Attenuation Profile



(b) Attenuation versus frequency response

CST MWS Simulation Front View

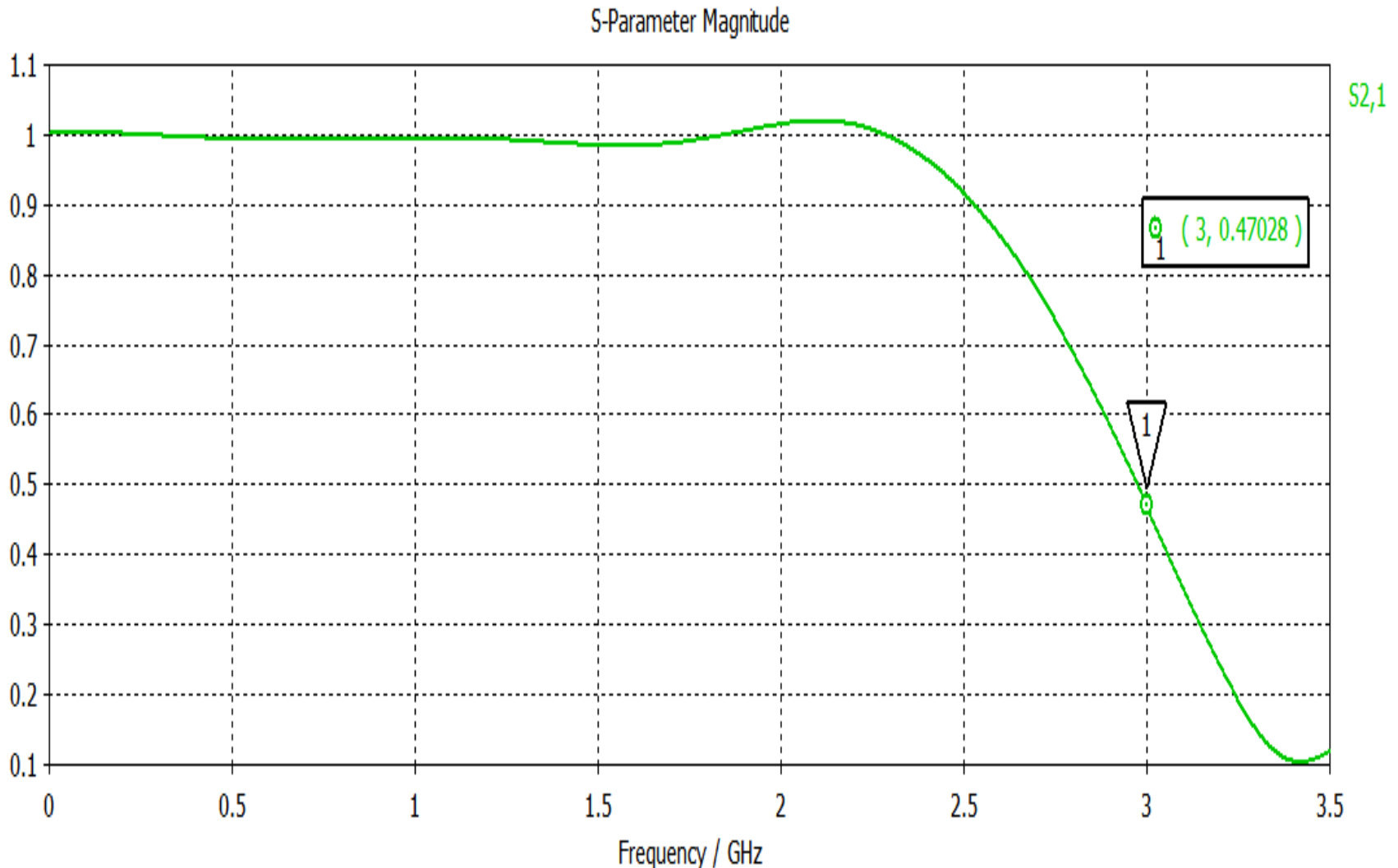


CST MWS Simulation Back View



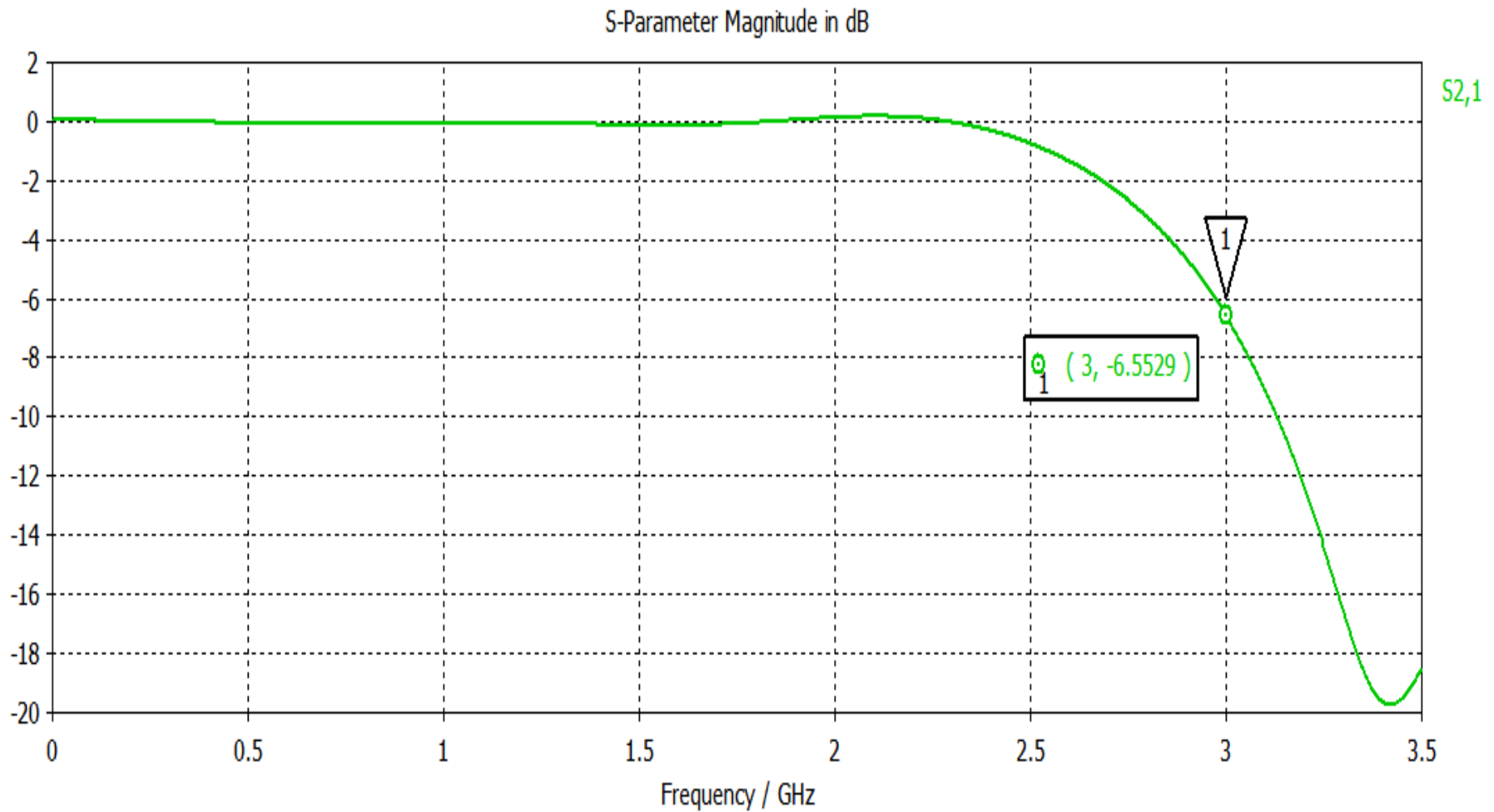
CST MWS Simulation

Transmission Co-efficient $S(2,1)$ (Linear)



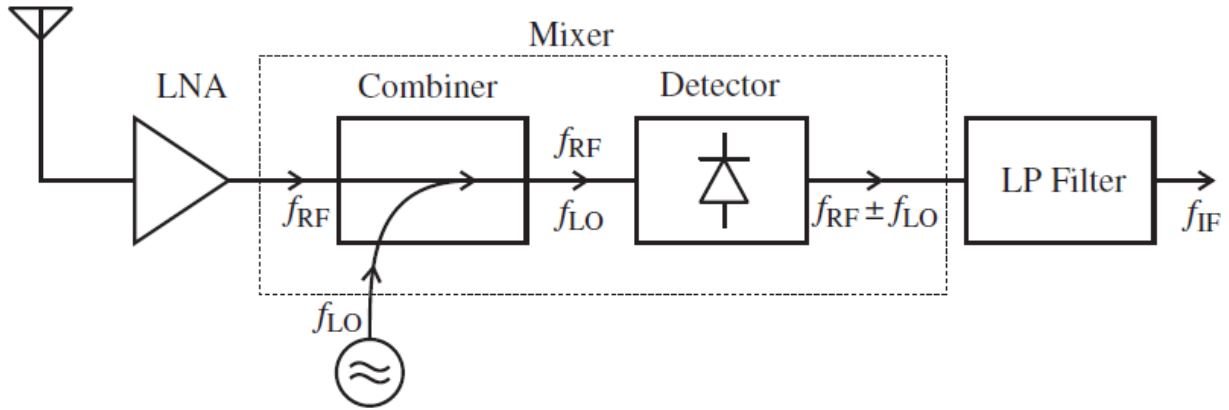
CST MWS Simulation

Transmission Co-efficient S(2,1) dB



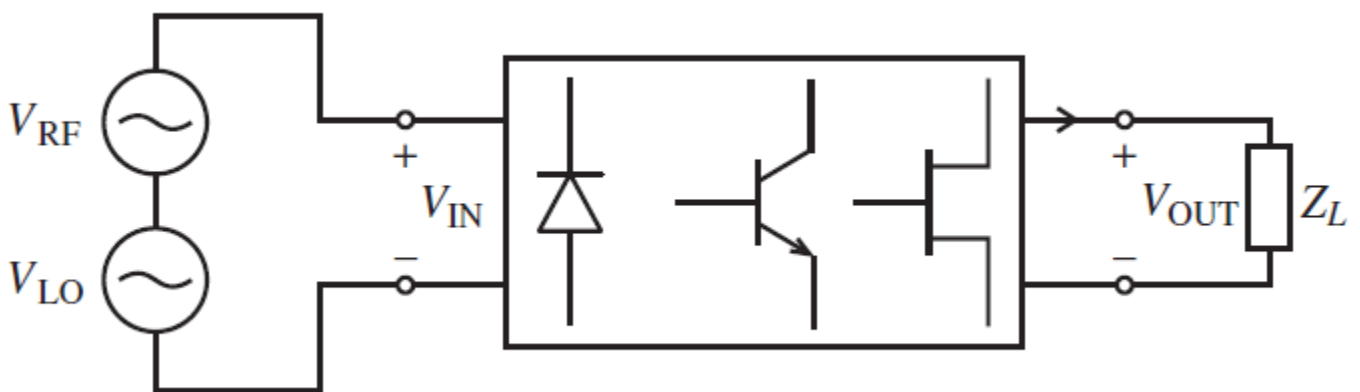
BASIC CHARACTERISTICS OF MIXERS

Mixers are commonly used to multiply signals of different frequencies in an effort to achieve frequency translation. The motivation for this translation stems from the fact that filtering out a particular RF signal channel centered among many densely populated, narrowly spaced neighboring channels would require extremely high Q-filters. This task, however, becomes much more manageable if the Rf signal carrier frequency can be reduced or down-converted within the communication system. Perhaps one of the best known systems is the down-conversion in a heterodyne receiver, schematically depicted in the following figure.



Heterodyne receiver system incorporating a mixer

Here the received RF signal is, after preamplification in a low-noise amplifier (LNA), supplied to a mixer whose task is to multiply the input signal of center frequency f_{RF} with a local oscillator (LO) frequency f_{LO} . The signal obtained after the mixer contains the frequencies $f_{RF} \pm f_{LO}$, of which after low-pass(LP) filtering, the low frequency component $f_{RF} - f_{LO}$, known as the intermediate frequency (IF), is selected for further processing. The two key ingredients constituting a mixer are the combiner and detector. The combiner can be implemented through the use of 90° or 180° directional coupler. The detector may employ non-linear devices like diode or BJT or MESFET.



Basic mixer concept: two input frequencies are used to create new frequencies at the output of the system

Above figure depicts the basic system arrangement of a mixer connected to an RF signal, $V_{RF}(t)$, and local oscillator signal, $V_{LO}(t)$, which is also known as the **pump signal**. It is seen that the RF input voltage signal is combined with the LO signal and supplied to a semiconductor device with a nonlinear characteristic at its output side driving a current into the load.

Both diode and BJT have an exponential transfer characteristic, as expressed for instance by the Shockley's diode equation as follows:

$$I = I_0 \left(e^{V/V_T} - 1 \right)$$

Alternatively, for a MESFET we have approximately a square behavior:

$$I(V) = I_{DSS} \left(1 - V/V_{T0} \right)^2$$

where the subscripts denoting drain current and gate-source voltage are omitted for simplicity. The input voltage is represented as sum of RF signal $v_{RF} = V_{RF} \cos(\omega_{RF}t)$ and the LO signal $v_{LO} = V_{LO} \cos(\omega_{LO}t)$ and a bias V_Q ; that is

$$V = V_Q + V_{RF} \cos(\omega_{RF}t) + V_{LO} \cos(\omega_{LO}t)$$

This voltage is applied to the nonlinear device whose current output characteristic can be found via a Taylor series expansion around the Q-point:

$$f(x) = f(a) + \frac{f'(a)}{1!} (x - a) + \frac{f''(a)}{2!} (x - a)^2 + \frac{f'''(a)}{3!} (x - a)^3 + \dots$$

$$I(V) = I_Q + V \left(\frac{dI}{dV} \right) \Big|_{V_Q} + \frac{1}{2} V^2 \left(\frac{d^2I}{dV^2} \right) \Big|_{V_Q} + \dots = I_Q + VA + V^2B + \dots$$

Where the constants A and B refer to $\left(\frac{dI}{dV} \right) \Big|_{V_Q}$ and $\left(\frac{d^2I}{dV^2} \right) \Big|_{V_Q}$, respectively. Neglecting the constant bias V_Q and I_Q ,

$$I(V) = VA + V^2B + \dots$$

$$I(V) = A \{V_{RF} \cos(\omega_{RF}t) + V_{LO} \cos(\omega_{LO}t)\} + B \{V_{RF} \cos(\omega_{RF}t) + V_{LO} \cos(\omega_{LO}t)\}^2 + \dots$$

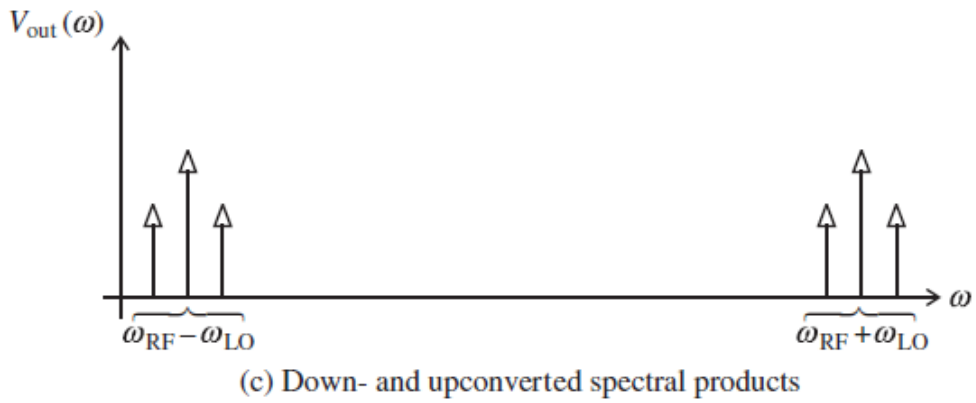
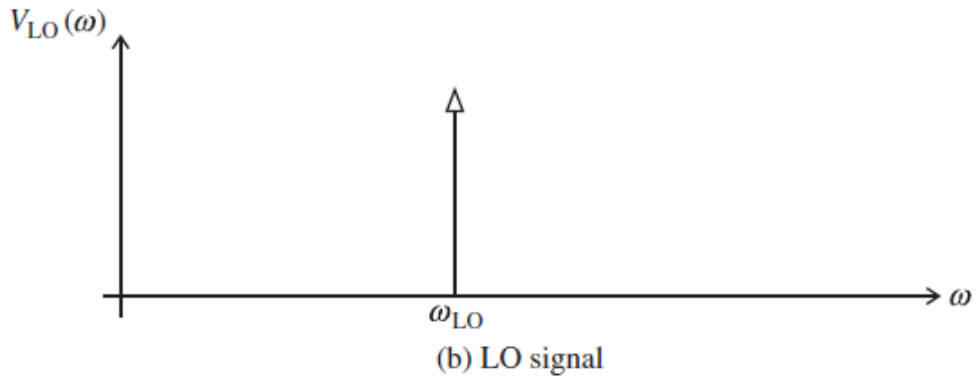
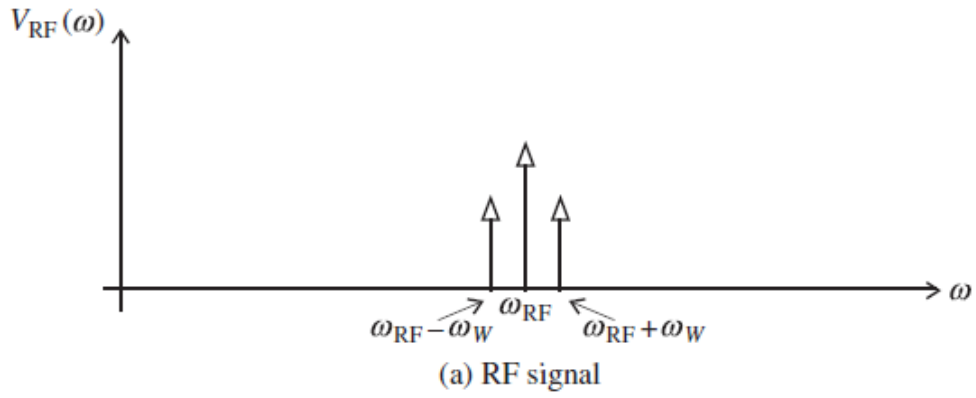
$$I(V) = A \{V_{RF} \cos(\omega_{RF}t) + V_{LO} \cos(\omega_{LO}t)\} + B \{V_{RF}^2 \cos^2(\omega_{RF}t) + V_{LO}^2 \cos^2(\omega_{LO}t) + 2 V_{RF} \cos(\omega_{RF}t) V_{LO} \cos(\omega_{LO}t)\}$$

The key lies in the last term

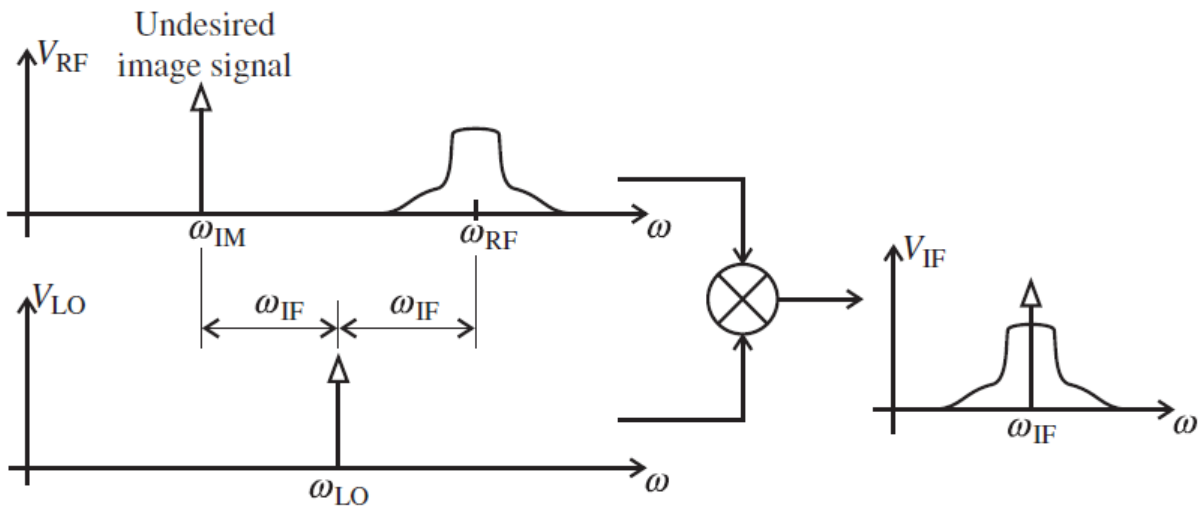
$$I(V) = \dots + 2B V_{RF} V_{LO} \cos(\omega_{RF}t) \cos(\omega_{LO}t)$$

$$I(V) = \dots + B V_{RF} V_{LO} \{ \cos[(\omega_{RF} + \omega_{LO})t] + \cos[(\omega_{RF} - \omega_{LO})t] \}$$

This expression makes clear that the non-linear action of the diode or transistor can generate new frequency components of the form $\omega_{RF} \pm \omega_{LO}$. It is also noted that the amplitudes are multiplied by $V_{RF} V_{LO}$, and B is a device dependent factor.



Spectral representation of mixing process

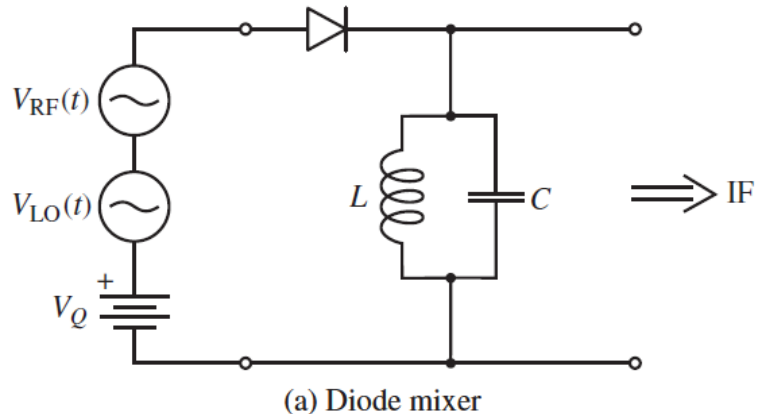


Problem of image frequency mapping

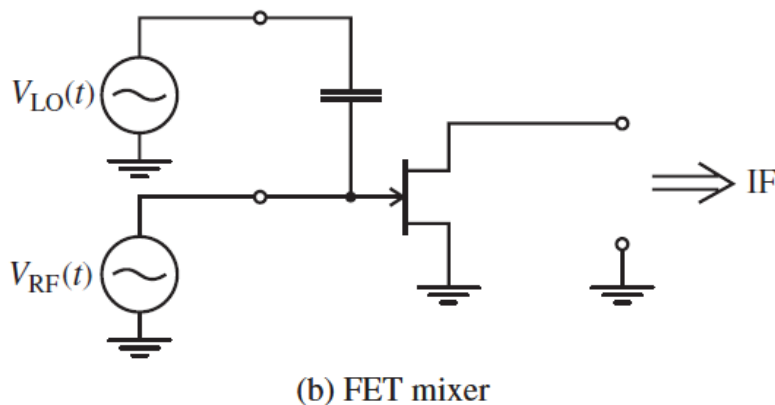
TYPES OF MIXERS

Single-Ended Mixer

- RF and LO sources are supplied to an appropriately biased diode followed by a resonator circuit tuned to the desired IF.



- Since LO and RF signals are not electrically isolated. There is a potential danger that the LO signal can interfere with the RF reception, possibly even reradiating portions of the LO energy through the receiving antenna.
- Following figure shows an improved design involving a FET, which, unlike the diode, is able to provide gain to the incoming RF and LO signals.



- FET realization allows not only allows for LO and RF isolation but also provides a signal gain and thus minimizes conversion loss
- Conversion Loss (CL)** of a mixer is generally defined in dB as the ratio of supplied input power P_{RF} over the obtained IF power P_{IF} . When dealing with BJTs and FETs, it is preferable to specify a **conversion gain (CG)** defined as the inverse of the power ratio.

$$CL = 10 \log \left(\frac{P_{RF}}{P_{IF}} \right)$$

- Noise Figure(F)** of a mixer is defined as

$$F = \frac{P_{n_{out}}}{CG P_{n_{in}}}$$

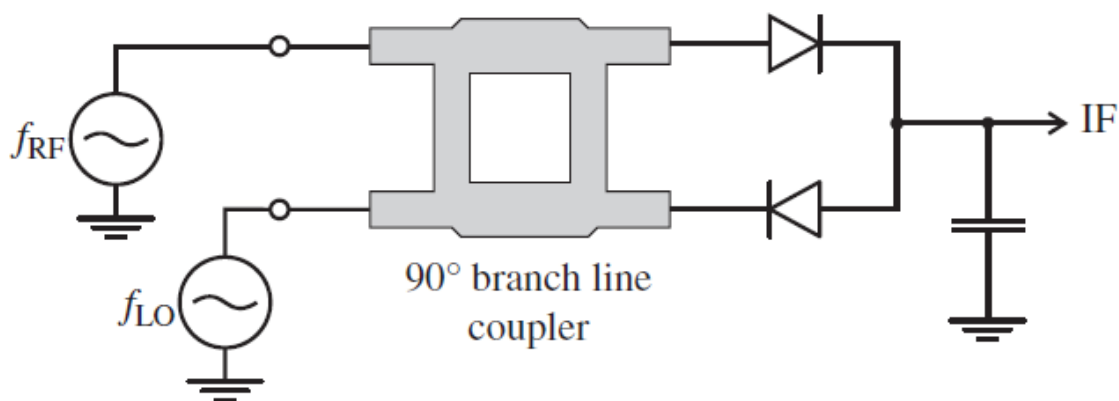
Where CG being the conversion gain and $P_{n_{out}}$, $P_{n_{in}}$ the noise power at the output due to the RF signal input (at RF) and the total noise power at the output (at IF).

- FET generally has a lower noise figure than a BJT, and because of a nearly quadratic transfer characteristic, the influence of higher-order nonlinear terms is minimized.

- BJT finds application when high conversion gain and low voltage bias conditions are needed.

Single-Balanced Mixer

- Single-ended mixers are rather easy to construct circuits and the main disadvantage of these designs is the difficulty associated with providing LO energy while maintaining separation between LO, RF, and IF signals for broadband applications
- Balanced dual-diode or dual transistor mixer in conjunction with a hybrid coupler offers the ability to conduct such broadband operations. Moreover, it provides further advantages related to noise suppression and spurious mode rejection
- Spurs arise in oscillators and amplifiers due to parasitic resonances and non-linearities and are only suppressed by the front-end. Thermal noise can critically raise the noise floor in the receiver.
- Following figure shows the basic mixer design featuring a quadratic coupler and a dual-diode detector followed by a capacitor acting as summation point.

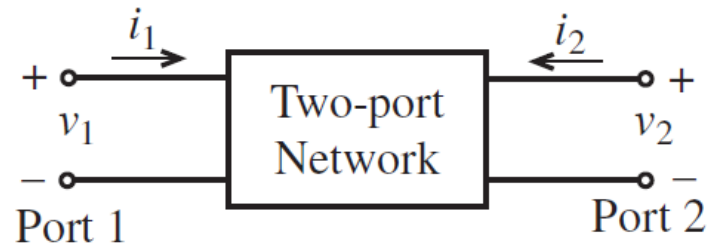


- This design provides excellent VSWR and is capable of suppressing a considerable amount of noise because the opposite diode arrangement in conjunction with the 90° phase shift provides a good degree of noise cancellation.

MICROWAVE NETWORKS

A microwave network is formed when several microwave devices and components such as sources, attenuators, resonators, filters, amplifiers, etc., are coupled by transmission lines or waveguides for the desired transmission of a microwave signal. The point of interconnection of two or more devices is called a junction.

For a low-frequency network, a port is a pair of terminals whereas for a microwave network, a port is a reference plane transverse to the length of the microwave transmission line or waveguide. At low frequencies, the physical length of the network is much smaller than the wavelength of the signal transmitted. Therefore, the measurable input and output variables are voltage and current which can be related in terms of the impedance Z-parameters, or admittance Y-parameters, or hybrid h-parameters, or ABCD parameters. For a two-port network as shown schematically in following figure, these relationships are given by



Basic voltage and current definitions for a two-port network

$$\begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix}$$

$$\begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

$$\begin{bmatrix} v_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} i_1 \\ v_2 \end{bmatrix}$$

$$\begin{bmatrix} v_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} v_2 \\ -i_2 \end{bmatrix}$$

At microwave frequencies, the physical length of the component or line is comparable to or much larger than the wavelength. Furthermore, the voltage and current cannot be uniquely defined at a given point in a single conductor waveguide.

Besides this constraint, measurement of Z, Y, h and ABCD parameters is difficult at microwave frequencies due to the following reasons:

1. Non-availability of terminal voltage and current measuring equipment even in the cases of TEM lines (coaxial, strip and microstrip lines) where voltage and current can be uniquely defined.
2. Short-circuit and open-circuit conditions are not easily achieved over a wide range of frequencies.
3. Presence of active devices makes the circuit unstable or open and short-circuit.

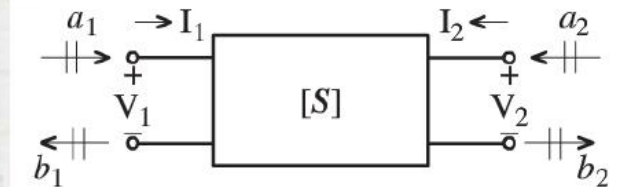
Therefore, microwave circuits are analysed using **Scattering or S-parameters** which linearly relate the amplitudes of scattered (reflected or transmitted) waves with those of incident waves. However, many of

the circuit-analysis techniques and circuit properties that are valid at low frequencies are also valid for microwave circuits.

S-Parameters

Consider a two-port with incident and reflected power waves as shown in the figure.

The power waves a_n and b_n are defined as follows



$$a_n = \frac{1}{2\sqrt{Z_0}}(V_n + Z_0 I_n)$$

Incident normalized power at port n

$$b_n = \frac{1}{2\sqrt{Z_0}}(V_n - Z_0 I_n)$$

Reflected normalized power at port n

These definitions are somewhat strange. For example, what are the units of a_n and b_n ? However, the purpose of the definition will become clear. Inverting the definitions we find that

$$I_n = \sqrt{Z_0}(a_n - b_n)$$

Net current at port n

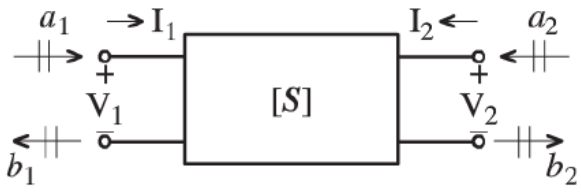
$$V_n = \sqrt{Z_0}(a_n + b_n)$$

Net voltage at port n

The power at port n is

$$P_n = \frac{1}{2} \operatorname{Re}(V_n I_n^*) = \frac{1}{2} (|a_n|^2 - |b_n|^2)$$

S-Parameters



$$a_n = \frac{1}{2\sqrt{Z_0}}(V_n + Z_0 I_n)$$

Incident normalized power at port n

$$b_n = \frac{1}{2\sqrt{Z_0}}(V_n - Z_0 I_n)$$

Reflected normalized power at port n

$$V_n = \sqrt{Z_0}(a_n + b_n)$$

$$I_n = \sqrt{Z_0}(a_n - b_n)$$

$$P_n = \frac{1}{2} \text{Re}(V_n I_n^*)$$

$$= \frac{1}{2} (|a_n|^2 - |b_n|^2)$$

One can now define 4 parameters that relate the reflected and transmitted powers at each port.

$$S_{11} = \left. \frac{b_1}{a_1} \right|_{a_2=0} = \frac{\text{reflected power at port 1}}{\text{incident power at port 1}}$$

$$S_{21} = \left. \frac{b_2}{a_1} \right|_{a_2=0} = \frac{\text{transmitted power at port 2}}{\text{incident power at port 1}}$$

$$S_{12} = \left. \frac{b_1}{a_2} \right|_{a_1=0} = \frac{\text{transmitted power at port 1}}{\text{incident power at port 2}}$$

$$S_{22} = \left. \frac{b_2}{a_2} \right|_{a_1=0} = \frac{\text{reflected power at port 2}}{\text{incident power at port 2}}$$

Further:
$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix}$$

Consider, for example the first row
$$b_1 = S_{11}a_{11} + S_{12}a_2$$

Power from port 1 to generator

Reflected power from port 1

Transmitted power from port 2

Measurement of S-Parameters

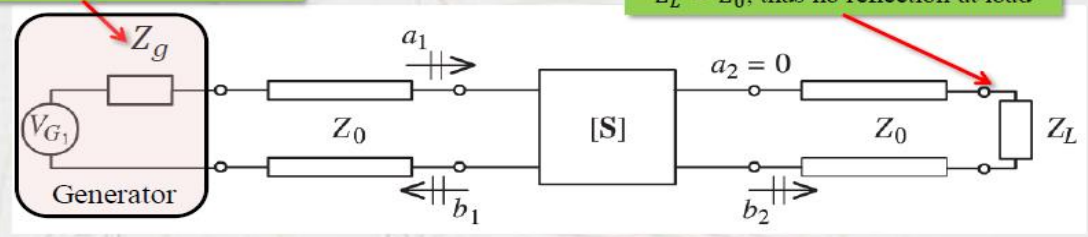
Setup for measuring which parameters?

$$S_{11} = \left. \frac{b_1}{a_1} \right|_{a_2=0}$$

$$S_{21} = \left. \frac{b_2}{a_1} \right|_{a_2=0}$$

$Z_g = Z_0$, thus no reflection generator

$Z_L = Z_0$, thus no reflection at load



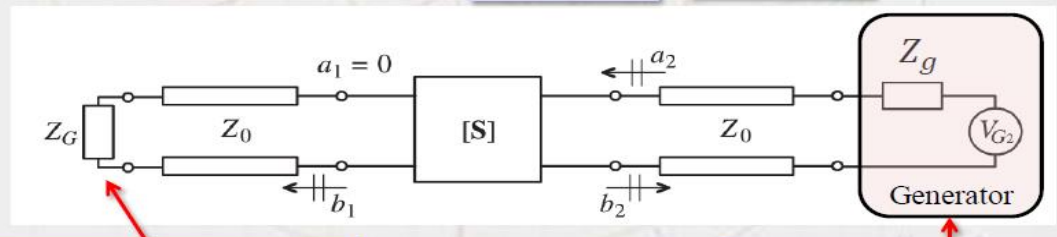
Setup for measuring which parameters?

$$S_{22} = \left. \frac{b_1}{a_2} \right|_{a_1=0}$$

$$S_{12} = \left. \frac{b_1}{a_2} \right|_{a_1=0}$$

$Z_L = Z_0$, thus no reflection at load

$Z_g = Z_0$, thus no reflection generator

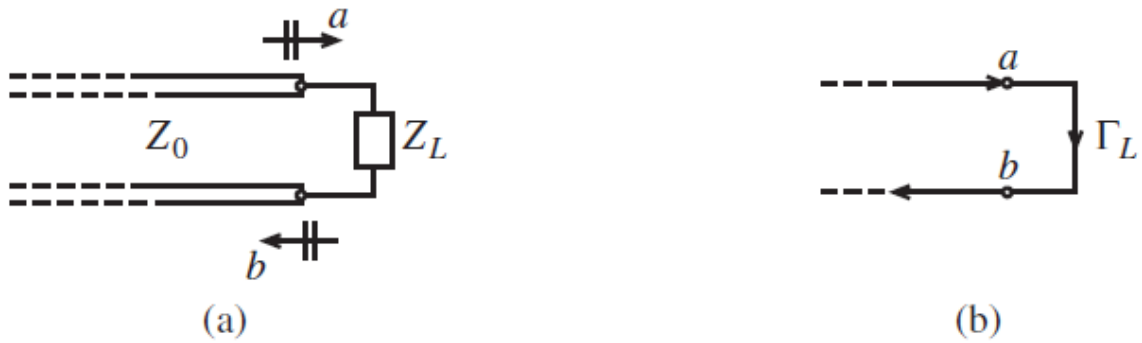


Meaning of S-parameters

S_{11} represents Input Reflection Coefficient; S_{21} indicates Forward Voltage Gain; S_{12} represents Reverse Voltage Gain; S_{22} indicates Output Reflection Coefficient

SIGNAL FLOW CHART MODELING

The analysis of RF networks and their overall interconnection is greatly facilitated through signal flow charts as commonly used in system and control theory. Even complicated networks are easily reduced to input output relations in which the reflection and transmission coefficients play integral parts.



Terminated transmission line segment with incident and reflected power wave description. (a) Conventional form, and (b) Signal flow form



(a) Source node a , which launches wave.

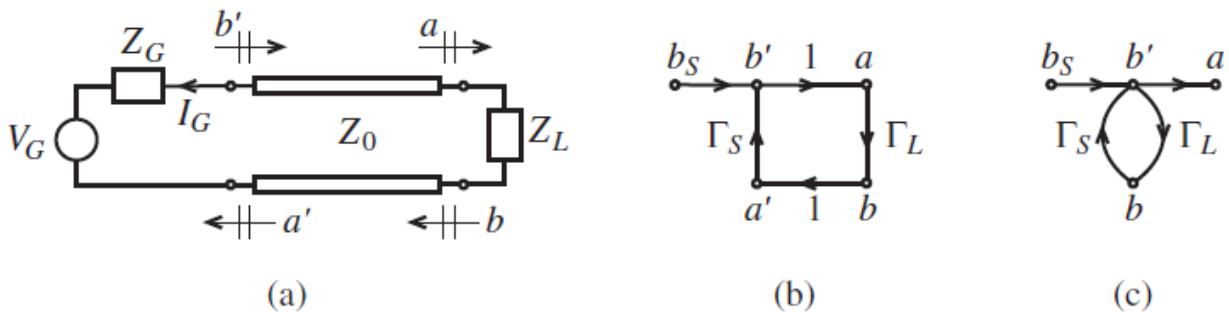


(b) Sink node b , which receives wave.

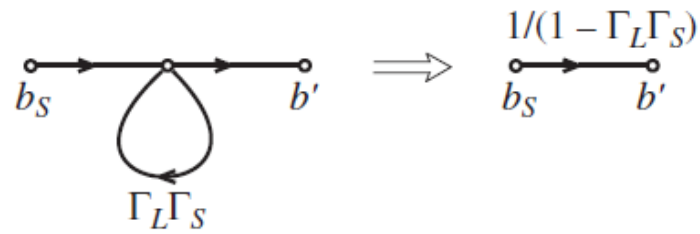


(c) Branch connecting source and sink.

Generic source node (a), receiver node (b), and the associated branch connection (c).



Terminated transmission line with source. (a) conventional form, (b) signal flow form, and (c) simplified signal flow form



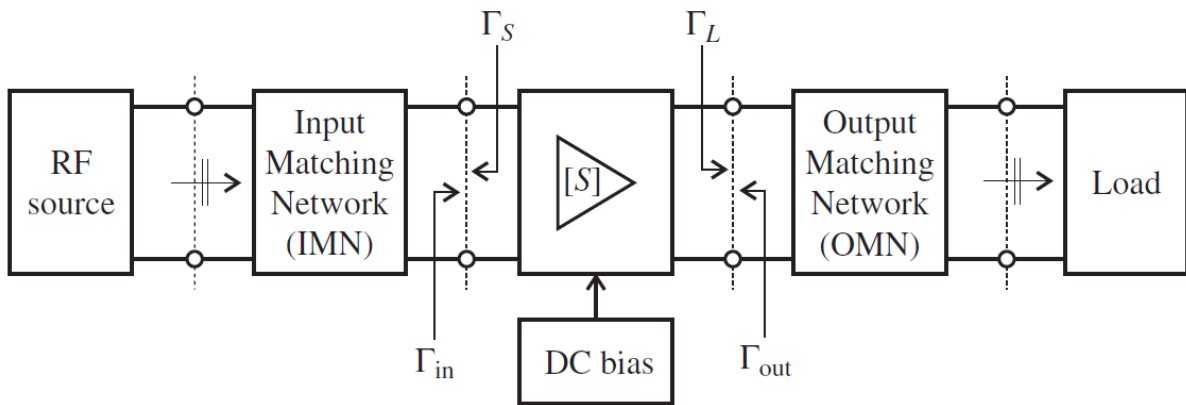
A self-loop that collapses to a single branch

Table: Signal Flowgraph Building Blocks

Description	Graphical Representation
Nodal Assignment	
Branch	
Series Connection	
Parallel Connection	
Splitting of Branches	
Self-Loop	

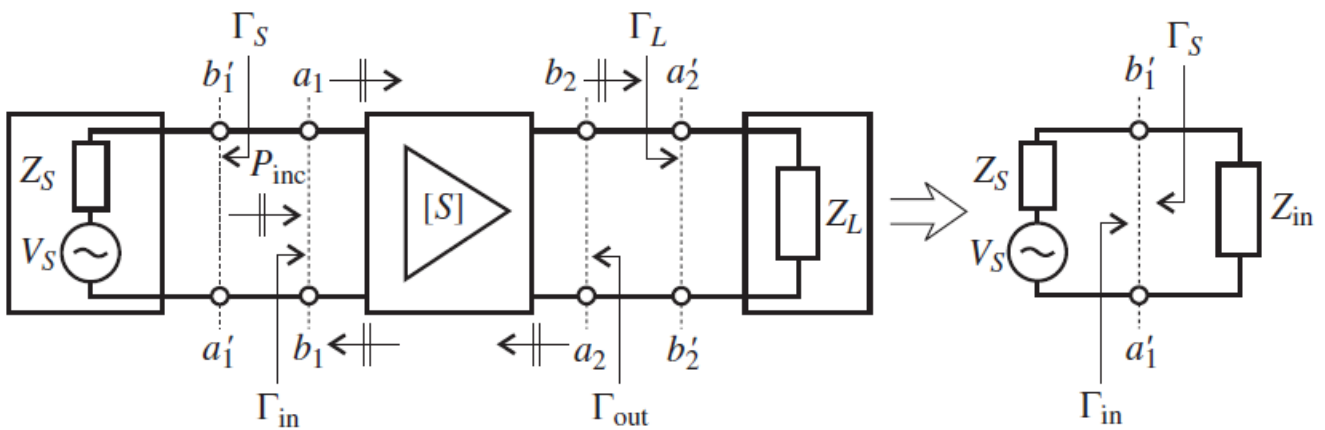
AMPLIFIER POWER RELATIONS

RF SOURCE

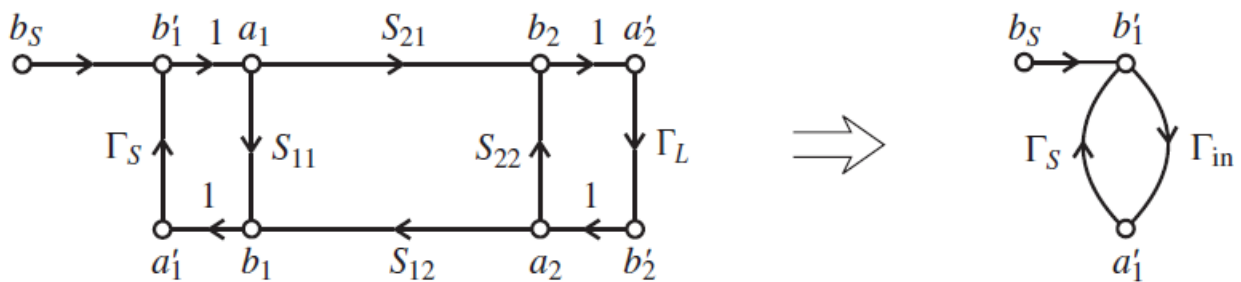


Generic amplifier system

Let us examine the above figure in terms of its power flow relations under the assumptions that the two matching networks are included in the source and load impedances. This simplifies our system to the configuration shown below:



(a) Simplified schematics of a single-stage amplifier



(b) Signal flow graph

Source and load connected to a single-stage amplifier network

The starting point of our power analysis is the RF source connected to the amplifier network. The source voltage is written as

$$b_S = b'_1(1 - \Gamma_{in}\Gamma_S) \quad (1)$$

The incident power wave associated with b'_1 is given as

$$P_{inc} = \frac{|b'_1|^2}{2} = \frac{1}{2} \frac{|b_S|^2}{|1 - \Gamma_{in}\Gamma_S|^2} \quad (2)$$

which is the power launched toward the amplifier. The actual input power P_{in} observed at the input terminal of the amplifier is composed of the incident and reflected power waves. With the aid of the input reflection coefficient Γ_{in} we can therefore write:

$$P_{in} = P_{inc}(1 - |\Gamma_{in}|^2) = \frac{1}{2} \frac{|b_S|^2}{|1 - \Gamma_{in}\Gamma_S|^2} (1 - |\Gamma_{in}|^2) \quad (3)$$

The **maximum power transfer** from source to the amplifier is achieved if the input impedance is complex conjugate matched $Z_{in} = Z_S^*$ or in terms of the reflection coefficients, if $\Gamma_{in} = \Gamma_S^*$. Under maximum power transfer condition, we define **available power** P_A as

$$P_A = P_{in}|_{\Gamma_{in}=\Gamma_S^*} = \frac{1}{2} \frac{|b_S|^2}{|1 - \Gamma_{in}\Gamma_S|^2} (1 - |\Gamma_{in}|^2) \Big|_{\Gamma_{in}=\Gamma_S^*} = \frac{1}{2} \frac{|b_S|^2}{1 - |\Gamma_S|^2} \quad (4)$$

Transducer Power Gain

The Transducer Power Gain quantifies the gain of the amplifier placed between source and load.

$$G_T = \frac{\text{Power delivered to the load}}{\text{Available power from the source}} = \frac{P_L}{P_A}$$

Or with $P_L = \frac{1}{2} |b_2|^2(1 - |\Gamma_L|^2)$ we obtain

$$G_T = \frac{P_L}{P_A} = \frac{|b_2|^2}{|b_S|^2} (1 - |\Gamma_L|^2) (1 - |\Gamma_S|^2) \quad (5)$$

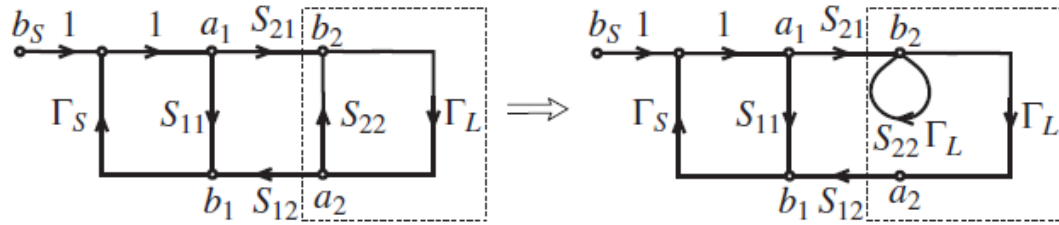
In this expression, the ratio b_2/b_S has to be determined. With the help of our signal flow graph discussion in above section and based on the following figure, we establish

$$b_2 = \frac{S_{21}a_1}{1 - S_{22}\Gamma_L} \quad (6a)$$

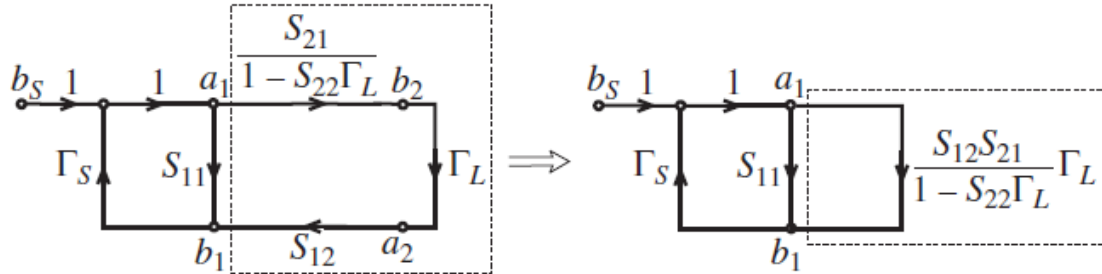
$$b_S = \left[1 - \left(S_{11} + \frac{S_{21}S_{12}\Gamma_L}{1 - S_{22}\Gamma_L} \right) \Gamma_S \right] a_1 \quad (6b)$$

The required ratio is therefore given by

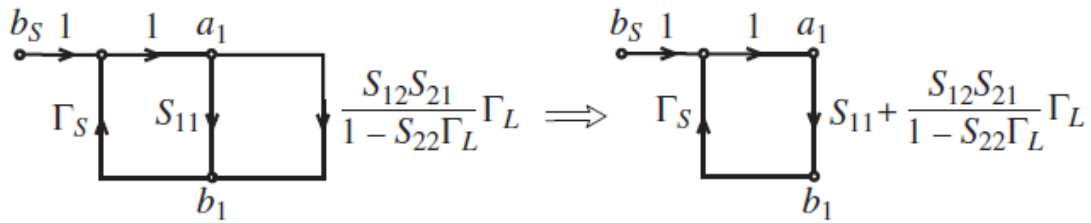
$$\frac{b_2}{b_S} = \frac{S_{21}}{(1 - S_{22}\Gamma_L)(1 - S_{11}\Gamma_S) - S_{21}S_{12}\Gamma_L\Gamma_S} \quad (7)$$



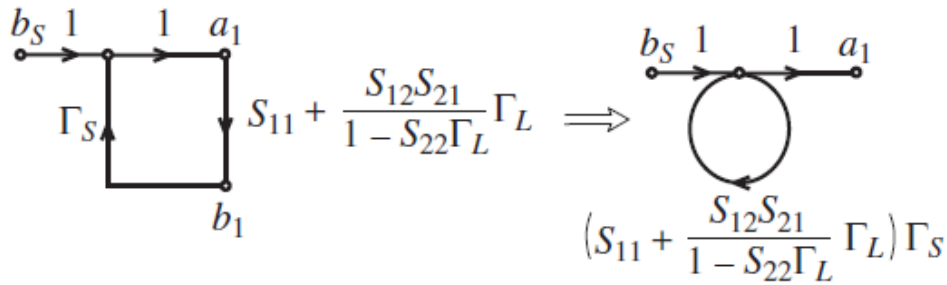
Step 1



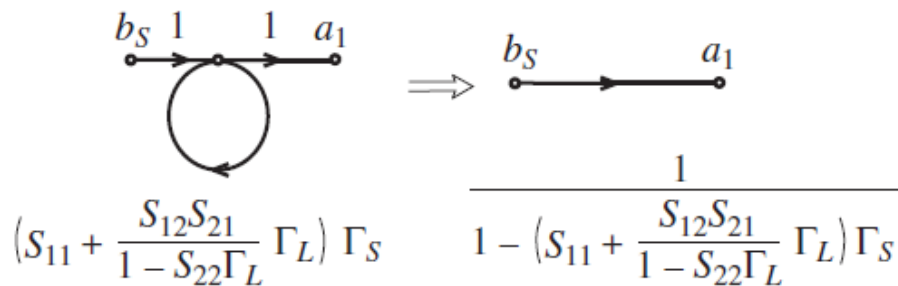
Step 2



Step 3



Step 4



Step 5

Step-by-step simplification to determine the ratio $\frac{b_2}{b_s}$

Inserting equation (7) into equation (5) results in

$$G_T = \frac{(1 - |\Gamma_L|^2) |S_{21}|^2 (1 - |\Gamma_S|^2)}{|(1 - S_{22}\Gamma_L)(1 - S_{11}\Gamma_S) - S_{21}S_{12}\Gamma_L\Gamma_S|^2} \quad (8)$$

Which can be rearranged by defining the input and output reflection coefficients

$$\Gamma_{in} = S_{11} + \frac{S_{21}S_{12}\Gamma_L}{1 - S_{22}\Gamma_L} \quad (9a)$$

$$\Gamma_{out} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1 - S_{11}\Gamma_S} \quad (9b)$$

With these two definitions, two more transducer power gain expressions can be derived. First, by incorporating equation (9a) into equation (8), it is seen that

$$G_T = \frac{(1 - |\Gamma_L|^2) |S_{21}|^2 (1 - |\Gamma_S|^2)}{|1 - \Gamma_S\Gamma_{in}|^2 |1 - S_{22}\Gamma_L|^2} \quad (10)$$

Second, using equation(9b) into equation (8), it is seen that

$$G_T = \frac{(1 - |\Gamma_L|^2) |S_{21}|^2 (1 - |\Gamma_S|^2)}{|1 - \Gamma_L\Gamma_{out}|^2 |1 - S_{11}\Gamma_S|^2} \quad (11)$$

An often employed approximation for the transducer power gain is the so-called **unilateral power gain**, G_{TU} , which neglects the feedback effect of the amplifier($S_{21} = 0$). This simplifies equation (11) to

$$G_{TU} = \frac{(1 - |\Gamma_L|^2) |S_{21}|^2 (1 - |\Gamma_S|^2)}{|1 - \Gamma_L S_{22}|^2 |1 - S_{11}\Gamma_S|^2} \quad (12)$$

Additional Power Relations

Available power gain for load side matching is defined as

$$G_A = G_T|_{\Gamma_L=\Gamma_{out}^*} = \frac{\text{power available from the amplifier}}{\text{power available from the source}}$$

With the aid of equation (11),

$$G_A = \frac{|S_{21}|^2 (1 - |\Gamma_S|^2)}{(1 - |\Gamma_{out}|^2) |1 - S_{11}\Gamma_S|^2} \quad (13)$$

Power gain (operating power gain) is defined as

$$G = \frac{\text{power delivered to the load}}{\text{power supplied to the amplifier}} = \frac{P_L}{P_{in}} = \frac{P_L}{P_A} \cdot \frac{P_A}{P_{in}} = G_T \cdot \frac{P_A}{P_{in}}$$

Using equations (3), (4) and (10)

$$G = \frac{(1 - |\Gamma_L|^2) |S_{21}|^2}{(1 - |\Gamma_{in}|^2) |1 - S_{22}\Gamma_L|^2} \quad (14)$$

Example: Power relations or an RF amplifier

An RF Amplifier has the following S-parameters:

$S_{11} = 0.3\angle -70^\circ$, $S_{12} = 0.2\angle -10^\circ$, $S_{21} = 3.5\angle 85^\circ$ and $S_{22} = 0.4\angle -45^\circ$. Furthermore, the input side of the amplifier is connected to a voltage source with $V_S = 5V\angle 0^\circ$ and source impedance $Z_S = 40\Omega$. The output is utilized to drive an antenna which has an impedance of $Z_L = 73\Omega$. Assuming that the S-parameters of the amplifier are measured with reference to a $Z_0 = 50\Omega$ characteristic impedance, find the following quantities:

- Transducer gain G_T , Unilateral transducer gain G_{TU} , available gain G_A , operating power gain G and
- Power delivered to the load P_L , available power P_A and incident power to the amplifier P_{inc}

Solution:

$$\Gamma_S = \frac{Z_S - Z_0}{Z_S + Z_0} = -0.111$$

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0} = 0.187$$

$$\Gamma_{in} = S_{11} + \frac{S_{21}S_{12}\Gamma_L}{1 - S_{22}\Gamma_L} = 0.146 - j0.151$$

$$\Gamma_{out} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1 - S_{11}\Gamma_S} = 0.265 - j0.358$$

$$G_T = \frac{(1 - |\Gamma_L|^2) |S_{21}|^2 (1 - |\Gamma_S|^2)}{|1 - \Gamma_L\Gamma_{out}|^2 |1 - S_{11}\Gamma_S|^2} = 12.56 \text{ or } 10.99 \text{ dB}$$

$$G_{TU} = \frac{(1 - |\Gamma_L|^2) |S_{21}|^2 (1 - |\Gamma_S|^2)}{|1 - \Gamma_L S_{22}|^2 |1 - S_{11}\Gamma_S|^2} = 12.67 \text{ or } 11.03 \text{ dB}$$

$$G_A = \frac{|S_{21}|^2 (1 - |\Gamma_S|^2)}{(1 - |\Gamma_{out}|^2) |1 - S_{11}\Gamma_S|^2} = 14.74 \text{ or } 11.68 \text{ dB}$$

$$G = \frac{(1 - |\Gamma_L|^2) |S_{21}|^2}{(1 - |\Gamma_{in}|^2) |1 - S_{22}\Gamma_L|^2} = 13.74 \text{ or } 11.38 \text{ dB}$$

$$P_{inc} = \frac{1}{2} \frac{|b_S|^2}{|1 - \Gamma_{in}\Gamma_S|^2} = \frac{1}{2} \frac{Z_0}{(Z_S + Z_0)^2} \frac{|V_S|^2}{|1 - \Gamma_{in}\Gamma_S|^2} = 74.7 \text{ mW}$$

$$P_{inc}(\text{dBm}) = 10 \log \left[\frac{P_{inc}}{1 \text{ mW}} \right] = 18.73 \text{ dBm}$$

$$P_A = \frac{1}{2} \frac{|b_S|^2}{1 - |\Gamma_S|^2} = \frac{1}{2} \frac{Z_0}{(Z_S + Z_0)^2} \frac{|V_S|^2}{1 - |\Gamma_S|^2} = 78.1 \text{ mW or } 18.93 \text{ dBm}$$

$$P_L(\text{dBm}) = P_A(\text{dBm}) + G_T(\text{dBm}) = 29.92 \text{ dBm}$$

STABILITY of Amplifier

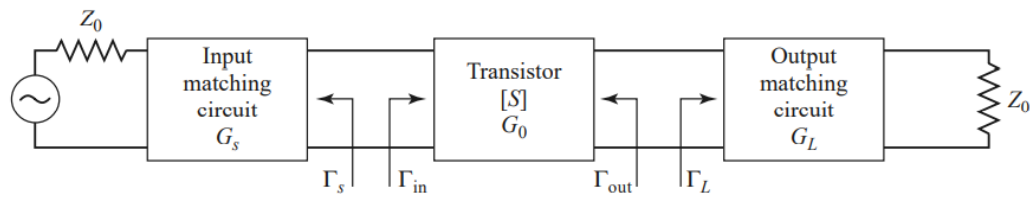


FIGURE The general transistor amplifier circuit.

A single-stage microwave transistor amplifier can be modelled by the circuit shown in Figure, where matching networks are used on both sides of the transistor to transform the input and output impedance Z_0 to the source and load impedances Z_S and Z_L .

In the circuit shown in Figure, oscillation is possible if either the input or output port impedance has a negative real part; this would then imply that $|\Gamma_{in}| > 1$ or $|\Gamma_{out}| > 1$. Because Γ_{in} and Γ_{out} depend on the source and load matching networks, the stability of the amplifier depends on Γ_S and Γ_L as presented by the matching networks.

Two types of stability:

(i) Unconditional stability:

The network is unconditionally stable if $|\Gamma_{in}| < 1$ and $|\Gamma_{out}| < 1$ for all passive source and load impedances (i.e., $|\Gamma_S| < 1$ or $|\Gamma_L| < 1$).

(ii) Conditional stability:

The network is conditionally stable if $|\Gamma_{in}| < 1$ and $|\Gamma_{out}| < 1$ only for a certain range of passive source and load impedances. This case is also referred to as potentially unstable.

Note that the stability condition of an amplifier circuit is usually frequency dependent since the input and output matching networks generally depend on frequency. It is therefore possible for an amplifier to be stable at its design frequency but unstable at other frequencies. Careful amplifier design should consider this possibility.

Requirements for unconditional stability

The following conditions must be satisfied by Γ_S and Γ_L , if the amplifier is to be unconditionally stable:

$$|\Gamma_{in}| = \left| S_{11} + \frac{S_{21}S_{12}\Gamma_L}{1-S_{22}\Gamma_L} \right| < 1 \quad \text{----- (1)}$$

$$|\Gamma_{out}| = \left| S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1-S_{11}\Gamma_S} \right| < 1 \quad \text{----- (2)}$$

If the device is unilateral ($S_{12} = 0$), these conditions reduce to the simple results that $|S_{11}| < 1$ and $|S_{22}| < 1$ are sufficient for unconditional stability. Otherwise, the inequalities define a range of values for Γ_S and Γ_L where the amplifier will be stable. Finding this range for Γ_S and Γ_L can be facilitated by using the Smith chart and plotting the input and output stability circles.

The stability circles are defined as the loci in the Γ_L (or Γ_S) plane for which $|\Gamma_{in}| = 1$ (or $|\Gamma_{out}| = 1$). The stability circles then define the boundaries between stable and potentially

unstable regions of Γ_S and Γ_L . Γ_S and Γ_L must lie on the Smith chart ($|\Gamma_S| < 1$, $|\Gamma_L| < 1$ for passive matching networks).

We can derive the equation for the output stability circle as follows:

Express the condition that $|\Gamma_{in}| = 1$ using equation (1)

$$|\Gamma_{in}| = \left| S_{11} + \frac{S_{21}S_{12}\Gamma_L}{1-S_{22}\Gamma_L} \right| = 1 \quad \text{----- (3)}$$

$$|S_{11}(1 - S_{22}\Gamma_L) + S_{21}S_{12}\Gamma_L| = |1 - S_{22}\Gamma_L| \quad \text{----- (4)}$$

Now define Δ as the determinant of the scattering matrix:

$$\Delta = S_{11} S_{22} - S_{12} S_{21}$$

We can write the above equation as,

$$|S_{11} - \Delta\Gamma_L| = |1 - S_{22}\Gamma_L| \quad \text{----- (5)}$$

Now square both sides and simplify to obtain

$$\begin{aligned} |S_{11}|^2 + |\Delta|^2|\Gamma_L|^2 - (\Delta\Gamma_L S_{11}^* + \Delta^* \Gamma_L^* S_{11}) &= 1 + |S_{22}|^2|\Gamma_L|^2 - (S_{22}^* \Gamma_L^* + S_{22}\Gamma_L) \\ (|S_{22}|^2 - |\Delta|^2)\Gamma_L \Gamma_L^* - (S_{22} - \Delta S_{11}^*)\Gamma_L - (S_{22}^* - \Delta^* S_{11})\Gamma_L^* &= |S_{11}|^2 - 1 \\ \Gamma_L \Gamma_L^* - \frac{(S_{22} - \Delta S_{11}^*)\Gamma_L + (S_{22}^* - \Delta^* S_{11})\Gamma_L^*}{|S_{22}|^2 - |\Delta|^2} &= \frac{|S_{11}|^2 - 1}{|S_{22}|^2 - |\Delta|^2}. \end{aligned}$$

Next, complete the square by adding $|S_{22} - \Delta S_{11}^*|^2 / (|S_{22}|^2 - |\Delta|^2)^2$ to both sides:

$$\left| \Gamma_L - \frac{(S_{22} - \Delta S_{11}^*)^*}{|S_{22}|^2 - |\Delta|^2} \right|^2 = \frac{|S_{11}|^2 - 1}{|S_{22}|^2 - |\Delta|^2} + \frac{|S_{22} - \Delta S_{11}^*|^2}{(|S_{22}|^2 - |\Delta|^2)^2},$$

or

$$\left| \Gamma_L - \frac{(S_{22} - \Delta S_{11}^*)^*}{|S_{22}|^2 - |\Delta|^2} \right| = \left| \frac{S_{12}S_{21}}{|S_{22}|^2 - |\Delta|^2} \right|.$$

In the complex plane, an equation of the form $|\Gamma - C| = R$ represents a circle having a center at C (a complex number) and a radius R (a real number). The above resultant equation defines the output stability circle with a center C_L and radius R_L where,

$$C_L = \frac{(S_{22} - \Delta S_{11}^*)^*}{|S_{22}|^2 - |\Delta|^2} \quad (\text{center}) \quad \text{----- (6a)}$$

$$R_L = \left| \frac{S_{12}S_{21}}{|S_{22}|^2 - |\Delta|^2} \right| \quad (\text{radius}) \quad \text{----- (6b)}$$

Similar results can be obtained for the input stability circle by interchanging S_{11} and S_{22} .

$$C_S = \frac{(S_{11} - \Delta S_{22}^*)^*}{|S_{11}|^2 - |\Delta|^2} \quad (\text{center}) \quad \text{----- (7a)}$$

$$R_S = \left| \frac{S_{12}S_{21}}{|S_{11}|^2 - |\Delta|^2} \right| \quad (\text{radius}) \quad \text{----- (7b)}$$

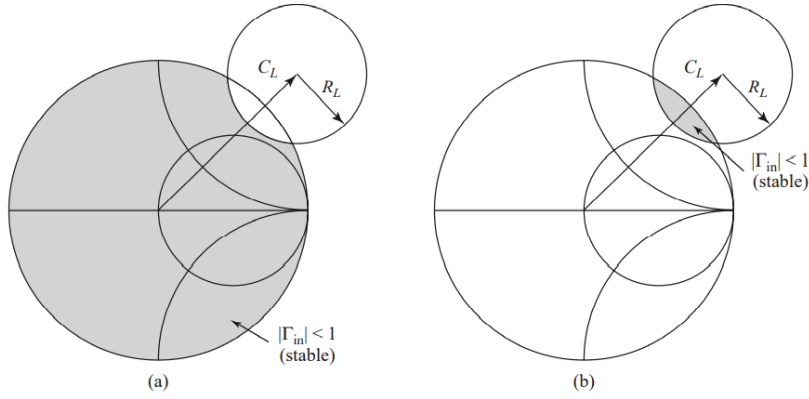


FIGURE 1 Output stability circles for a conditionally stable device. (a) $|S_{11}| < 1$. (b) $|S_{11}| > 1$.

Given the scattering parameters of the transistor, we can plot the input and output stability circles to define where $|\Gamma_{in}| = 1$ and $|\Gamma_{out}| = 1$. On one side of the input stability circle we will have $|\Gamma_{out}| < 1$, while on the other side we will have $|\Gamma_{out}| > 1$.

Similarly, we will have $|\Gamma_{in}| < 1$ on one side of the output stability circle, and $|\Gamma_{in}| > 1$ on the other side. We need to determine which areas on the Smith chart represent the stable region, for which $|\Gamma_{in}| < 1$ and $|\Gamma_{out}| < 1$.

Consider the output stability circles plotted in the Γ_L plane for $|S_{11}| < 1$ and $|S_{11}| > 1$, as shown in Figure. If we set $Z_L = Z_0$, then $\Gamma_L = 0$, and (1) shows that $|\Gamma_{in}| = |S_{11}|$. Now if $|S_{11}| < 1$, then $|\Gamma_{in}| < 1$, so $\Gamma_L = 0$ must be in a stable region. This means that the center of the Smith chart ($\Gamma_L = 0$) is in the stable region, so all of the Smith chart ($|\Gamma_L| < 1$) that is exterior to the stability circle defines the stable range for Γ_L . This region is shaded in Figure (a). Alternatively, if we set $Z_L = Z_0$ but have $|S_{11}| > 1$, then $|\Gamma_{in}| > 1$ for $\Gamma_L = 0$, and the center of the Smith chart must be in an unstable region. In this case the stable region is the inside region of the stability circle that intersects the Smith chart, as illustrated in Figure (b). Similar results apply to the input stability circle.

If the device is unconditionally stable, the stability circles must be completely outside (or totally enclose) the Smith chart. We can state this mathematically as,

$$|C_L - R_L| > 1 \quad \text{for} \quad |S_{11}| < 1 \quad \text{----- (8a)}$$

$$|C_S - R_S| > 1 \quad \text{for} \quad |S_{22}| < 1 \quad \text{----- (8a)}$$

If $|S_{11}| > 1$ or $|S_{22}| > 1$, the amplifier cannot be unconditionally stable because we can always have a source or load impedance of Z_0 leading to $\Gamma_S = 0$ or $\Gamma_L = 0$, thus causing $|\Gamma_{in}| > 1$ or $|\Gamma_{out}| > 1$. If the device is only conditionally stable, operating points for Γ_S and Γ_L must be chosen in stable regions, and it is good practice to check stability at several frequencies over the range where the device operates. Also note that the scattering parameters of a transistor depend on the bias conditions, and so stability will also depend on bias conditions.

Tests for Unconditional Stability

The stability circles discussed above can be used to determine regions for Γ_S and Γ_L where the amplifier circuit will be conditionally stable, but simpler tests can be used to determine unconditional stability. One of these is the $K - \Delta$ test, where it can be shown that a device will be unconditionally stable if **Rollet's condition**, defined as,

$$K = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2|S_{12}S_{21}|} > 1 \quad \text{----- (9)}$$

Along with the auxiliary condition that

$$|\Delta| = |S_{11}S_{22} - S_{12}S_{21}| < 1 \quad \text{-----(10)}$$

are simultaneously satisfied. These two conditions are necessary and sufficient for unconditional stability, and are easily evaluated. If the device scattering parameters do not satisfy the $K - \Delta$ test, the device is not unconditionally stable, and stability circles must be used to determine if there are values of Γ_S and Γ_L for which the device will be conditionally stable. Also recall that we must have $|S_{11}| < 1$ and $|S_{22}| < 1$ if the device is to be unconditionally stable. While the $K - \Delta$ test is a mathematically rigorous condition for unconditional stability, it cannot be used to compare the relative stability of two or more devices because it involves constraints on two separate parameters.

μ -test

combines the scattering parameters in a test involving only a single parameter, μ .

μ is defined as,

$$\mu = \frac{1 - |S_{11}|^2}{|S_{22} - \Delta S_{11}^*| + |S_{12}S_{21}|} > 1 \quad \text{----- (11)}$$

If $\mu > 1$, the device is unconditionally stable. In addition, the larger values of μ imply greater stability.

Development of μ -test

$$\Gamma_{out} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1 - S_{11}\Gamma_S} = \frac{S_{22} - \Delta\Gamma_S}{1 - S_{11}\Gamma_S} \quad \text{-----(12)}$$

Where, Δ is the determinant of the scattering matrix. Unconditional stability implies that $|\Gamma_{out}| < 1$ for any passive source termination, Γ_S . The reflection coefficient for a passive source impedance must lie within the unit circle on a Smith chart, and the outer boundary of this circle can be written as $\Gamma_S = e^{j\phi}$. The expression given in (12) maps this circle into another circle in the Γ_{out} plane.

Substituting $\Gamma_S = e^{j\phi}$ into (12) and solving for $e^{j\phi}$:

$$e^{j\phi} = \frac{S_{22} - \Gamma_{out}}{\Delta - S_{11}\Gamma_{out}}$$

Taking the magnitude of both sides gives

$$\left| \frac{S_{22} - \Gamma_{out}}{\Delta - S_{11}\Gamma_{out}} \right| = 1.$$

Squaring both sides and expanding gives

$$|\Gamma_{out}|^2(1 - |S_{11}|^2) + \Gamma_{out}(\Delta^* S_{11} - S_{22}^*) + \Gamma_{out}^*(\Delta S_{11}^* - S_{22}) = |\Delta|^2 - |S_{22}|^2.$$

Now divide by $1 - |S_{11}|^2$ to obtain

$$|\Gamma_{out}|^2 + \frac{(\Delta^* S_{11} - S_{22}^*)\Gamma_{out} + (\Delta S_{11}^* - S_{22})\Gamma_{out}^*}{1 - |S_{11}|^2} = \frac{|\Delta|^2 - |S_{22}|^2}{1 - |S_{11}|^2}.$$

Complete the square by adding $\frac{|\Delta^* S_{11} - S_{22}^*|^2}{(1 - |S_{11}|^2)^2}$ to both sides:

$$\left| \Gamma_{out} + \frac{\Delta S_{11}^* - S_{22}}{1 - |S_{11}|^2} \right|^2 = \frac{|\Delta|^2 - |S_{22}|^2}{1 - |S_{11}|^2} + \frac{|\Delta^* S_{11} - S_{22}^*|^2}{(1 - |S_{11}|^2)^2} = \frac{|S_{12}S_{21}|^2}{(1 - |S_{11}|^2)^2}.$$

This equation is of the form $|\Gamma_{out} - C| = R$, representing a circle with center C and radius R in the Γ_{out} plane. Thus the center and radius of the mapped $|\Gamma_S| = 1$ circle are given by,

$$C = \frac{S_{22} - \Delta S_{11}^*}{1 - |S_{11}|^2} \quad \text{----- (13a)}$$

$$R = \frac{|S_{12}S_{21}|}{1 - |S_{11}|^2} \quad \text{----- (13b)}$$

If points within this circular region are to satisfy $|\Gamma_{out}| < 1$, then we must have that ,

$$|C| + R < 1 \quad \text{----- (14)}$$

Substituting equation (13) in equation (14) gives,

$$|S_{22} - \Delta S_{11}^*| + |S_{12}S_{21}| < 1 - |S_{11}|^2$$

Rearranging the above equation yields the μ - test.

$$\frac{1 - |S_{11}|^2}{|S_{22} - \Delta S_{11}^*| + |S_{12}S_{21}|} > 1$$

Development of K – Δ test

The K – Δ test can be derived more simply from the μ -test.

$$\frac{1 - |S_{11}|^2}{|S_{22} - \Delta S_{11}^*| + |S_{12}S_{21}|} > 1$$

$$|S_{22} - \Delta S_{11}^*| < 1 - |S_{11}|^2 - |S_{12}S_{21}| \quad \text{----- (15)}$$

Rearranging and squaring gives,

$$|S_{22} - \Delta S_{11}^*|^2 < (1 - |S_{11}|^2 - |S_{12}S_{21}|)^2.$$

It can be verified by direct expansion that

$$|S_{22} - \Delta S_{11}^*|^2 = |S_{12}S_{21}|^2 + (1 - |S_{11}|^2)(|S_{22}|^2 - |\Delta|^2),$$

so

$$|S_{12}S_{21}|^2 + (1 - |S_{11}|^2)(|S_{22}|^2 - |\Delta|^2) < (1 - |S_{11}|^2)(1 - |S_{11}|^2 - 2|S_{12}S_{21}|) + |S_{12}S_{21}|^2.$$

Simplifying gives

$$|S_{22}|^2 - |\Delta|^2 < 1 - |S_{11}|^2 - 2|S_{12}S_{21}|,$$

Which yields the Rollet's condition

$$K = \frac{1 - |S_{11}|^2 - |S_{22}|^2 + |\Delta|^2}{2|S_{12}S_{21}|} > 1$$

The squaring of equation (15) introduces an ambiguity in the sign of the right hand side, thus requiring an additional condition. The right hand side of equation (15) should be positive before squaring. Thus,

$$|S_{12}S_{21}| < 1 - |S_{11}|^2.$$

Because similar conditions can be derived for the input side of the circuit, we can interchange S_{11} and S_{22} to obtain the analogous condition that

$$|S_{12}S_{21}| < 1 - |S_{22}|^2.$$

Adding these two inequalities gives

$$2|S_{12}S_{21}| < 2 - |S_{11}|^2 - |S_{22}|^2.$$

From the triangle inequality we know that

$$|\Delta| = |S_{11}S_{22} - S_{12}S_{21}| \leq |S_{11}S_{22}| + |S_{12}S_{21}|,$$

so we have that

$$|\Delta| < |S_{11}||S_{22}| + 1 - \frac{1}{2}|S_{11}|^2 - \frac{1}{2}|S_{22}|^2 < 1 - \frac{1}{2}(|S_{11}|^2 - |S_{22}|^2) < 1,$$

Which is the required additional condition.

SINGLE-STAGE TRANSISTOR AMPLIFIER DESIGN for maximum gain

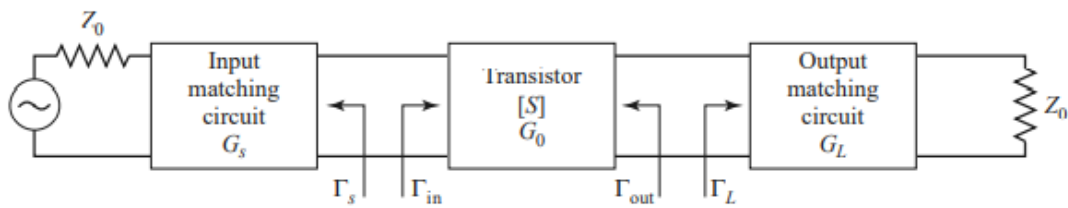


FIGURE . . . The general transistor amplifier circuit.

A single-stage microwave transistor amplifier can be modelled by the circuit shown in Figure, where matching networks are used on both sides of the transistor to transform the input and output impedance Z_0 to the source and load impedances Z_S and Z_L . The most useful gain definition for amplifier design is the transducer power gain, which accounts for both source and load mismatch. We can define separate effective gain factors for the input (source) matching network, the transistor itself, and the output (load) matching network as follows:

$$G_S = \frac{1-|\Gamma_S|^2}{|1-\Gamma_{in}\Gamma_S|^2} \quad \text{----- (1a)}$$

$$G_0 = |S_{21}|^2 \quad \text{-----(1b)}$$

$$G_L = \frac{1-|\Gamma_L|^2}{|1-S_{22}\Gamma_L|^2} \quad \text{-----(1c)}$$

The overall transducer gain is then $G_T = G_S G_0 G_L$. The effective gains G_S and G_L of the matching networks may be greater than unity. This is because the unmatched transistor would incur power loss due to reflections at the input and output of the transistor, and the matching sections can reduce these losses.

Design for Maximum Gain (Conjugate Matching)

After the stability of the transistor has been determined and the stable regions for Γ_S and Γ_L have been located on the Smith chart, the input and output matching sections can be designed. Since G_0 is fixed for a given transistor, the overall transducer gain of the amplifier will be controlled by the gains, G_S and G_L , of the matching sections.

Maximum gain will be realized when these sections provide a conjugate match between the amplifier source or load impedance and the transistor.

We know that maximum power transfer from the input matching network to the transistor will occur when

$$\Gamma_{in} = \Gamma_S^* \quad \text{-----(2a)}$$

and that maximum power transfer from the transistor to the output matching network will occur when

$$\Gamma_{out} = \Gamma_L^* \quad \text{----- (2b)}$$

With the assumption of lossless matching sections, these conditions will maximize the overall transducer gain.

The maximum gain is given by,

$$G_{T_{max}} = \frac{1}{1-|\Gamma_S|^2} |S_{21}|^2 \frac{1-|\Gamma_L|^2}{|1-S_{22}\Gamma_L|^2} \quad \text{----- (3)}$$

In the general case with a bilateral ($S_{12} = 0$) transistor, Γ_{in} is affected by Γ_{out} and vice versa, so the input and output sections must be matched simultaneously. The necessary equations are:

$$\Gamma_{in} = S_{11} + \frac{S_{21}S_{12}\Gamma_L}{1-S_{22}\Gamma_L} = \Gamma_S^* \text{-----(4)}$$

$$\Gamma_{out} = S_{22} + \frac{S_{12}S_{21}\Gamma_S}{1-S_{11}\Gamma_S} = \Gamma_L^* \text{-----(5)}$$

We can solve for Γ_S by first rewriting these equations as follows:

$$\Gamma_S = S_{11}^* + \frac{S_{12}^*S_{21}^*}{1/\Gamma_L^* - S_{22}^*},$$

$$\Gamma_L^* = \frac{S_{22} - \Delta\Gamma_S}{1 - S_{11}\Gamma_S},$$

where $\Delta = S_{11}S_{22} - S_{12}S_{21}$. Substituting the expression for Γ_L^* into the expression for Γ_S and expanding gives

$$\Gamma_S(1 - |S_{22}|^2) + \Gamma_S^2(\Delta S_{22}^* - S_{11}) = \Gamma_S(\Delta S_{11}^*S_{22}^* - |S_{11}|^2 - \Delta S_{12}^*S_{21}^*)$$

$$+ S_{11}^*(1 - |S_{22}|^2) + S_{12}^*S_{21}^*S_{22}.$$

Using the result that $\Delta(S_{11}^*S_{22}^* - S_{12}^*S_{21}^*) = |\Delta|^2$ allows this to be rewritten as a quadratic equation for Γ_S :

$$(S_{11} - \Delta S_{22}^*)\Gamma_S^2 + (|\Delta|^2 - |S_{11}|^2 + |S_{22}|^2 - 1)\Gamma_S + (S_{11}^* - \Delta^*S_{22}) = 0.$$

The solution is

$$\Gamma_S = \frac{B_1 \pm \sqrt{B_1^2 - 4|C_1|^2}}{2C_1}.$$

Similarly, the solution for Γ_L can be written as

$$\Gamma_L = \frac{B_2 \pm \sqrt{B_2^2 - 4|C_2|^2}}{2C_2}.$$

The variables B_1, C_1, B_2, C_2 are defined as

$$B_1 = 1 + |S_{11}|^2 - |S_{22}|^2 - |\Delta|^2,$$

$$B_2 = 1 + |S_{22}|^2 - |S_{11}|^2 - |\Delta|^2,$$

$$C_1 = S_{11} - \Delta S_{22}^*,$$

$$C_2 = S_{22} - \Delta S_{11}^*.$$

Solutions to Γ_S and Γ_L are only possible if the quantity within the square root is positive, and it can be shown that this is equivalent to requiring $K > 1$. Thus, unconditionally stable devices can always be conjugately matched for maximum gain, and potentially unstable devices can be conjugately matched if $K > 1$ and $|\Delta| < 1$.

The results are much simpler for the unilateral case. When $S_{12} = 0$, $\Gamma_S = S_{11}^*$ and $\Gamma_L = S_{22}^*$, and the maximum unilateral transducer gain is

$$G_{TU_{max}} = \frac{1}{1 - |S_{11}|^2} |S_{21}|^2 \frac{1}{1 - |S_{22}|^2}$$

The maximum transducer power gain occurs when the source and load are conjugately matched to the transistor. If the transistor is unconditionally stable, so that $K > 1$, the maximum transducer power gain can be simply rewritten as follows:

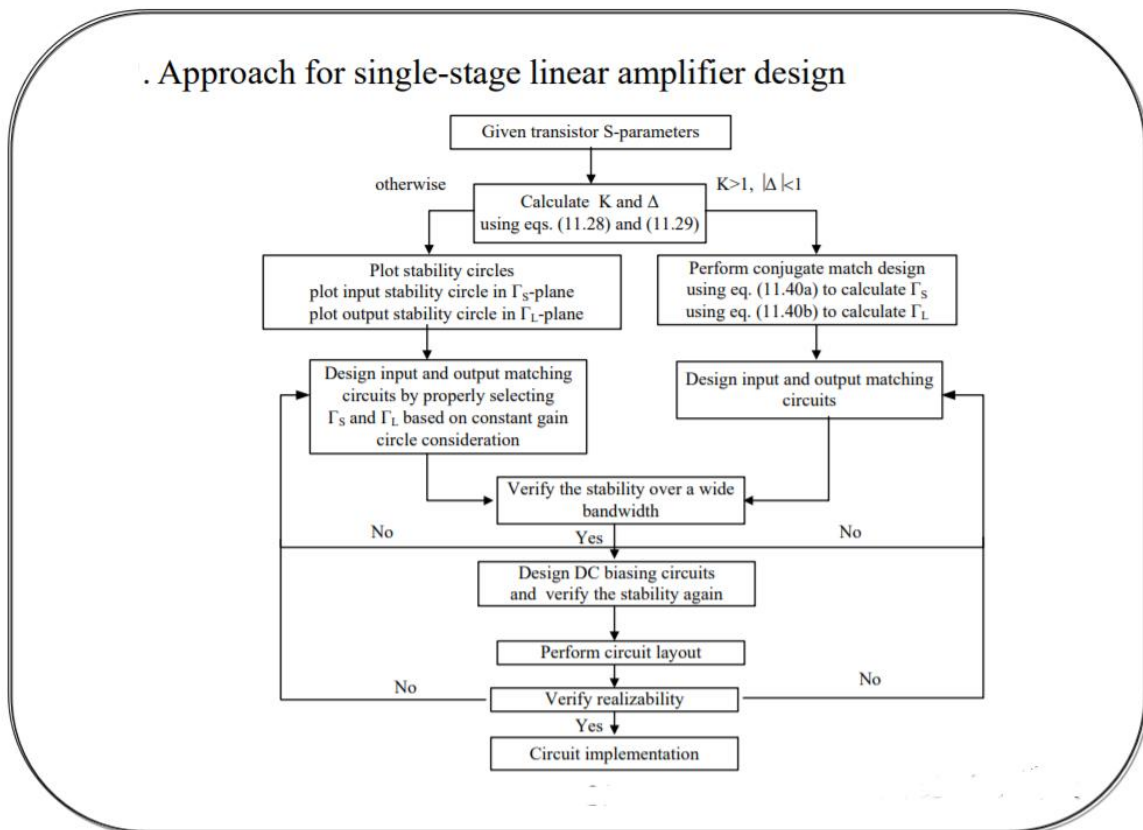
$$G_{T_{max}} = \frac{|S_{21}|}{|S_{12}|} (K - \sqrt{K^2 - 1})$$

The maximum transducer power gain is also sometimes referred to as the matched gain. The maximum gain does not provide a meaningful result if the device is only conditionally stable since simultaneous conjugate matching of the source and load is not possible if $K < 1$. In this case a useful figure of merit is the maximum stable gain, defined as the maximum transducer power gain with $K = 1$.

Thus, $G_{msg} = \frac{|S_{21}|}{|S_{12}|}$

The maximum stable gain is easy to compute and offers a convenient way to compare the gain of various devices under stable operating conditions.

Single stage amplifier design



Low-Noise Amplifier Design

Besides stability and gain, another important design consideration for a microwave amplifier is its noise figure. In receiver applications especially it is often required to have a preamplifier with as low a noise figure as possible since, the first stage of a receiver front end has the dominant effect on the noise performance of the overall system.

Generally it is not possible to obtain both minimum noise figure and maximum gain for an amplifier, so some sort of compromise must be made. This can be done by using constant-gain circles and circles of constant noise figure to select a usable trade-off between noise figure and gain. Here we will derive the equations for constant-noise figure circles and show how they are used in transistor amplifier design.

The noise figure of a two-port amplifier can be expressed as

$$F = F_{min} + \frac{R_N}{G_S} |Y_S - Y_{opt}|^2 \quad \text{----- (1)}$$

where the following definitions apply:

$Y_S = G_S + j B_S$ = source admittance presented to transistor.

Y_{opt} = optimum source admittance that results in minimum noise figure.

F_{min} = minimum noise figure of transistor, attained when $Y_S = Y_{opt}$.

R_N = equivalent noise resistance of transistor.

G_S = real part of source admittance.

Instead of the admittance Y_S and Y_{opt} , we can use the reflection coefficients Γ_S and Γ_{opt} ,

$$Y_S = \frac{1}{Z_0} \frac{1-\Gamma_S}{1+\Gamma_S} \quad \text{---- (2a)}$$

$$Y_{opt} = \frac{1}{Z_0} \frac{1-\Gamma_{opt}}{1+\Gamma_{opt}} \quad \text{---- (2b)}$$

Γ_S is the source reflection coefficient. The quantities F_{min} , Γ_{opt} and R_N are characteristics of the particular transistor being used, and are called the noise parameters of the device; they may be given by the manufacturer or measured.

Using equations for Y_S and Y_{opt} , we can express the quantity $|Y_S - Y_{opt}|^2$ in terms of Γ_S and Γ_{opt} as,

$$|Y_S - Y_{opt}|^2 = \frac{4}{Z_0^2} \frac{|\Gamma_S - \Gamma_{opt}|^2}{|1+\Gamma_S|^2 |1+\Gamma_{opt}|^2} \quad \text{-----(3)}$$

In addition,

$$G_S = \text{Re}\{Y_S\} = \frac{1}{2Z_0} \left(\frac{1-\Gamma_S}{1+\Gamma_S} + \frac{1-\Gamma_S^*}{1+\Gamma_S^*} \right) = \frac{1}{Z_0} \frac{1-|\Gamma_S|^2}{|1+\Gamma_S|^2} \quad \text{----(4)}$$

Using these results, Noise figure F is given as,

$$F = F_{min} + \frac{R_N}{G_S} |Y_S - Y_{opt}|^2$$

$$F = F_{min} + \frac{4R_N}{Z_0 (1-|\Gamma_S|^2) |1+\Gamma_{opt}|^2} \quad \text{-----(5)}$$

For a fixed Noise figure F, this result defines a circle in the Γ_S plane.

Noise figure parameter N is defined as,

$$N = \frac{|\Gamma_S - \Gamma_{opt}|^2}{(1 - |\Gamma_S|^2)} \quad \text{----- (6)}$$

Using equation (5), N can be written as,

$$N = \frac{F - F_{min}}{4R_N/Z_0} |1 + \Gamma_{opt}|^2 \quad \text{----- (7)}$$

N is a constant for a given noise figure and set of noise parameters. Equation (6) can be rewritten as,

$$\begin{aligned} (\Gamma_S - \Gamma_{opt})(\Gamma_S^* - \Gamma_{opt}^*) &= N(1 - |\Gamma_S|^2), \\ \Gamma_S \Gamma_S^* - (\Gamma_S \Gamma_{opt}^* + \Gamma_S^* \Gamma_{opt}) + \Gamma_{opt} \Gamma_{opt}^* &= N - N|\Gamma_S|^2, \\ \Gamma_S \Gamma_S^* - \frac{(\Gamma_S \Gamma_{opt}^* + \Gamma_S^* \Gamma_{opt})}{N + 1} &= \frac{N - |\Gamma_{opt}|^2}{N + 1}. \end{aligned}$$

Add $|\Gamma_{opt}|^2/(N + 1)^2$ to both sides to complete the square to obtain

$$\left| \Gamma_S - \frac{\Gamma_{opt}}{N + 1} \right| = \frac{\sqrt{N(N + 1 - |\Gamma_{opt}|^2)}}{(N + 1)}.$$

This result defines circles of constant noise figure with centers at

$$C_F = \frac{\Gamma_{opt}}{N + 1},$$

and radii of

$$R_F = \frac{\sqrt{N(N + 1 - |\Gamma_{opt}|^2)}}{N + 1}.$$

Generally it is not possible to obtain both minimum noise figure and maximum gain for an amplifier, so some sort of compromise must be made. This can be done by using constant-gain circles and circles of constant noise figure to select a usable trade-off between noise figure and gain.

Limitations of classical noise optimization:-

$$F_{min} = 1 + 2 R_n [G_{opt} + G_c]$$

$$F = F_{min} + \frac{R_n}{G_{S}} \left[(G_S - G_{opt})^2 + (B_S - B_{opt})^2 \right]$$

$$= F_{min} + \frac{R_n}{G_S} | Y_S - Y_{opt} |^2$$

* classical theory gives the idea of optimum source admittance (Y_{opt}) that will yield minimum noise figure for a given device.

What is the short-coming?

- ① In discrete RF design, we can deal with devices of fixed characteristics. In classical theory, there are no specific guidelines about what device size will minimize noise (in IC's).
- ② power consumption is not considered. (No relation connecting maximum power transfer & Minimum noise figure)

LNA DESIGN

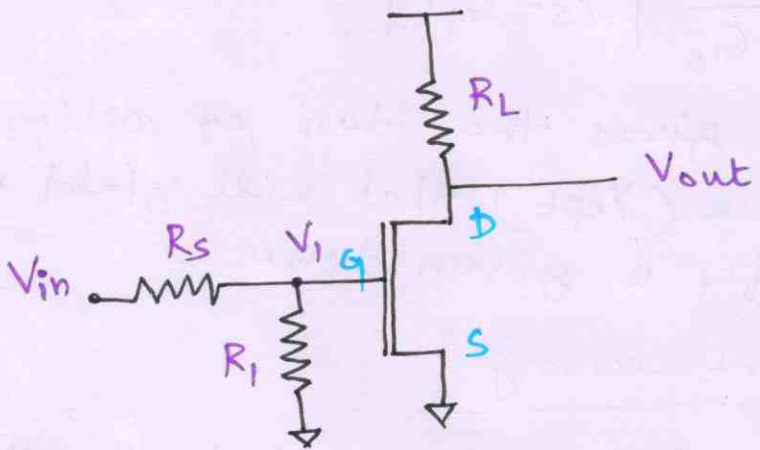
- * LNA: Low Noise Amplifier
- * First stage in most of the receivers.
- * Function
 - To provide Gain to overcome the noise of subsequent stages (Mixer, filter, etc.)
 - Added noise should be as low as possible.

LNA Topologies

(2)

* LNA's should have $50\ \Omega$ input impedance to achieve matching with other devices.

① Common-source Amplifier with shunt input Resistor :-



* To provide broad-band $50\ \Omega$ termination, put $50\ \Omega$ termination across the input terminals of CS Amplifier.

* R_1 - adds thermal noise. ①

* R_1 attenuates signal before amplifier (②) ($V_1 = V_{in}/2$)

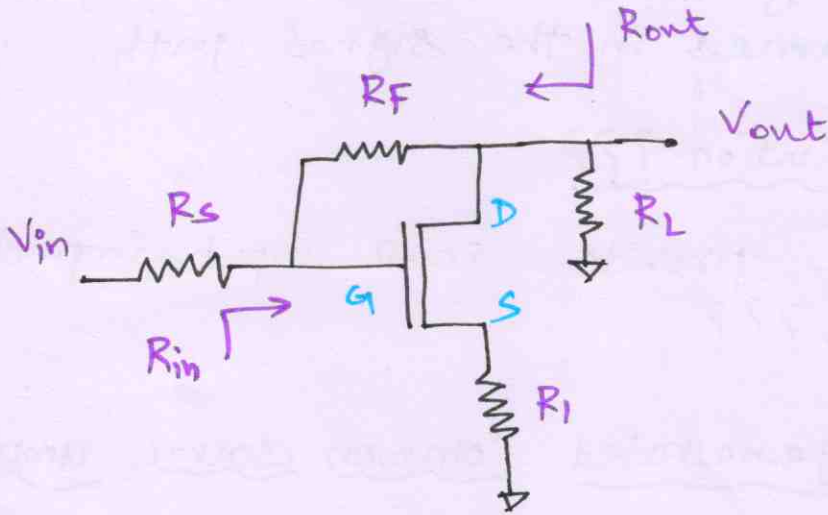
* ① & ② - results in high noise figures.

* When $R_s = R_1 = R$, the lower bound on noise figure is

$$F \geq 2 + \frac{4R^2}{d} \cdot \frac{1}{g_m R}$$

(Ignoring Gate current noise).

② Shunt-series Amplifier :-

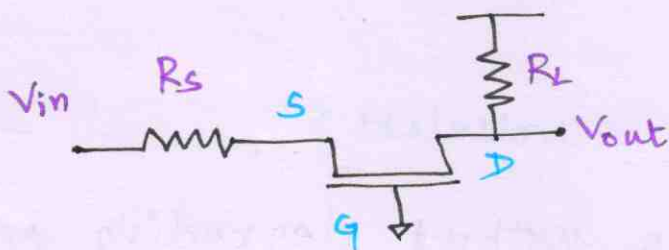


- * provides a broadband real input impedance.
- * No noisy attenuator before amplification
 ↳ noise figure better than previous circuit is expected.

* problems :-

- R_f generates thermal noise
- F is still larger than F_{min} .

③ Common Gate Amplifier :-



- * $R_{in} = 1/g_m$
- * proper selection of device size & bias current can provide the desired 50Ω resistance.

* noise figure lower bound: $F \geq 1 + \delta/\alpha$
 (Neglecting gate current noise).

NF = 2.2 dB → Long-channel
 NF = 4.8 dB → short devices.

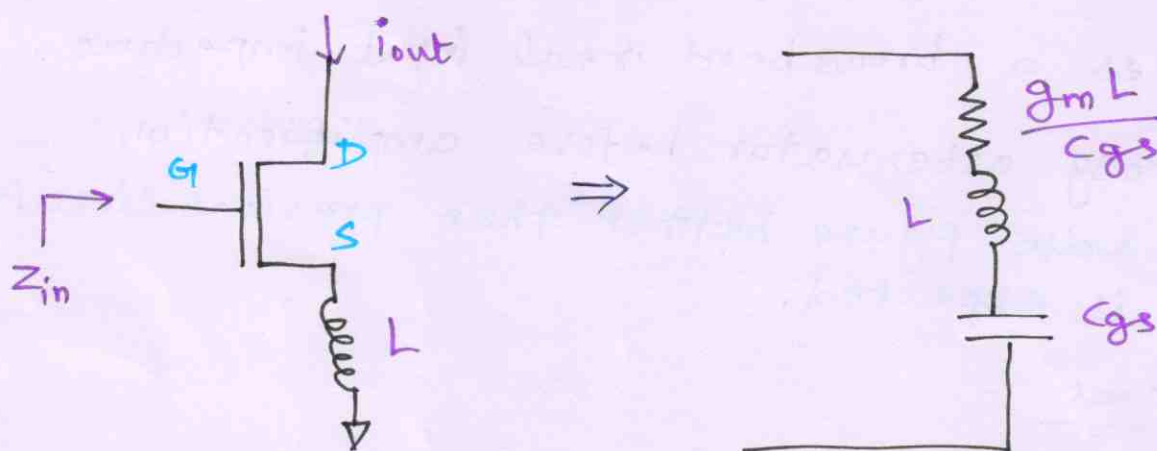
problem with previous topologies:-

- * Noise figure degradation due to the presence of **noisy resistances** in the signal path.

Think about solution???

Why can't we provide 50Ω input impedance without resistors???

④ Inductively degenerated common source amplifier:-



$$* Z_{in} = sL + \frac{1}{sC_{gs}} + \frac{g_m L}{C_{gs}} \approx sL + \frac{1}{sC_{gs}} + \omega_T L$$

* Z_{in} resembles series RLC network

* Resistive term $\propto L$

* No thermal noise ($L \rightarrow$ noiseless)

* Can provide desired Z_{in} without degrading NF.

* Can provide only narrow-band impedance matching.

Power Constrained Noise Optimization

* As per two-port classical noise theory,

$$F = F_{\min} + \frac{R_n}{G_s} \left[(G_s - G_{\text{opt}})^2 + (B_s - B_{\text{opt}})^2 \right]$$

* Goal: Re-formulate above expression in terms of **power consumption**.

* Minimize [Expression connecting F & P_d] with respect to the constraint of **fixed power**.

* For a fixed power constraint, solve for the **width of the transistor** that corresponds to this optimum condition.

$$W_{\text{opt},P} \approx \frac{1}{3\omega L C_{\text{ox}} R_s}$$

* For a device of width $W_{\text{opt},P}$, the noise figure obtained within the power constraint is

$$F_{\min,P} \approx 1 + 2.4 \frac{\gamma}{\alpha} \left[\frac{\omega}{\omega_T} \right]$$

* Excellent match - guaranteed by **inductive source degeneration technique**

→ Minimum possible NF

→ Maximum Gain (Reasonably Good)

$$g_m \propto W_T$$

$$F_{\min,P} \propto 1/W_T$$

$W_T \uparrow, F_{\min,P} \downarrow, g_m \uparrow$ (Gain \uparrow).

Power Amplifiers

* Narrowband vs. Broadband

* Linear vs. Constant Envelope operation

↓
AM etc.

↓
PM, FM etc.

(usually switching PAs)

* Tradeoffs

→ Power gain

→ Linearity

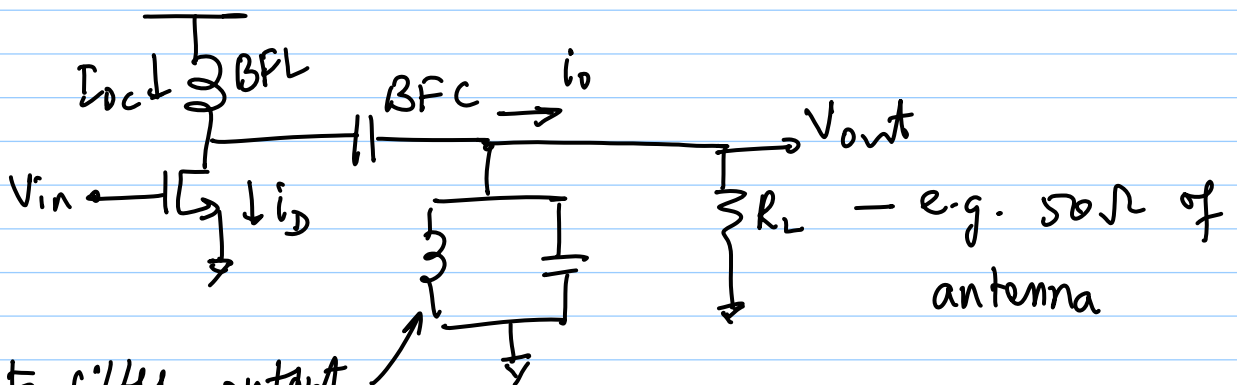
→ Output Power

→ Efficiency (drain eff. & power added eff.)

Classical PAs (linear)

→ class A, AB, B, C

→ classified based on bias conditions



tank to filter output
(tuned to ω_0)

* BFC prevents DC power diss. in R_L

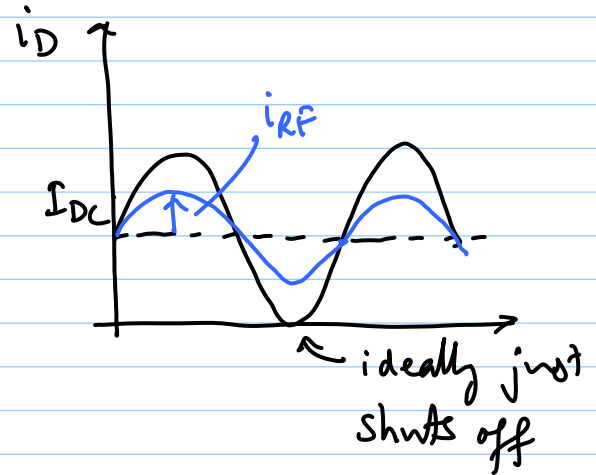
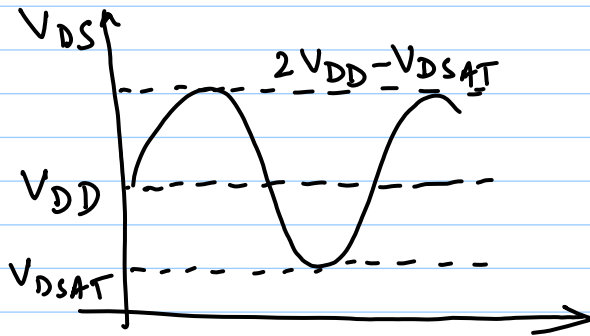
* BFL provides approximately constant current

* tank ckt with high Q provides linear output

I Class A

* 360° conduction angle

* $V_{in_{min.}} = V_T$



* high linearity

* poor efficiency

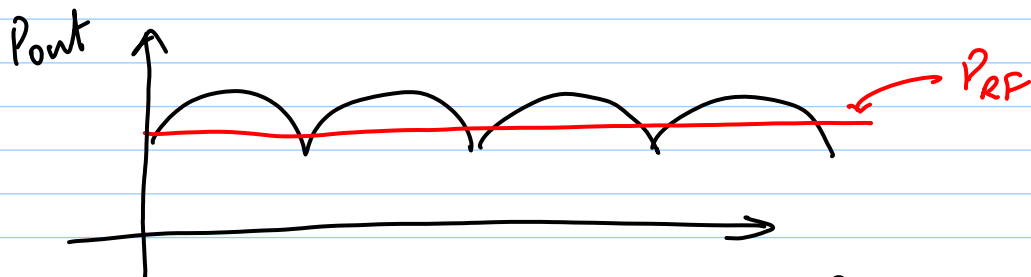
$$i_D = I_{DC} + i_{RF} \sin \omega t$$

$$i_o = I_{DC} - i_D = -i_{RF} \sin \omega t$$

$$V_{out} = i_o \cdot R_L = -i_{RF} R_L \sin \omega t$$

$$V_{DS} = V_{DD} + i_o \cdot R_L = V_{DD} - i_{RF} \cdot R_L \sin \omega t$$

$$P_{out} = i_{out} \cdot V_{out} = i_{RF}^2 R_L \sin^2 \omega t$$



$$P_{RF} = (i_{RF, rms})^2 \cdot R_L = \frac{i_{RF}^2 R_L}{2}$$

$$P_{DC} = \text{DC power from } V_{DD}$$

$$= V_{DD} \cdot I_{DC} = V_{DD} \cdot i_{RF} \text{ (assume } M_1 \text{ just wts off at lower extreme)}$$

$$\eta = \text{drain circuit efficiency}$$

$$= \frac{P_{out}}{P_{DC}} = \frac{\frac{1}{2} i_{RF}^2 R_L}{i_{RF} \cdot V_{DD}} = \frac{1}{2} \frac{i_{RF} \cdot R_L}{V_{DD}}$$

max. value of $i_{RF} \cdot R_L = V_{DD}$ (max. swing neglecting V_{DSAT})

$$\Rightarrow \eta_{\max} = \frac{1}{2} \text{ or } 50\%$$

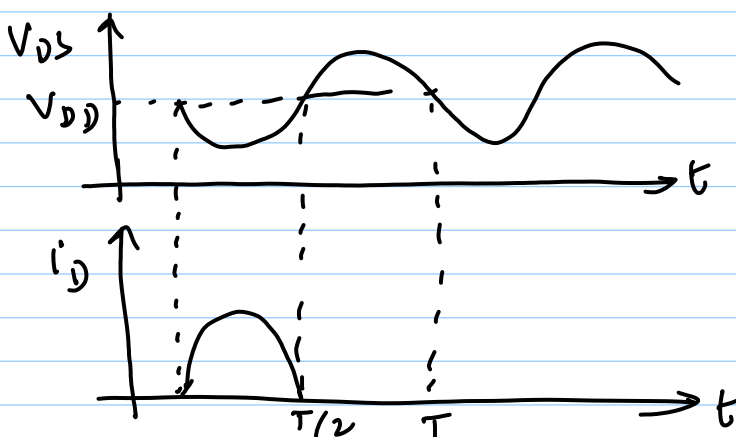
practical $\eta \sim 30-35\%$

Normalised power output capability $\equiv P_N$

$$P_N = \frac{P_{rf}}{V_{DSpk} \cdot i_{Dpk}} = \frac{V_{DD}^2 / 2R_L}{(2V_{DD}) \cdot \left(\frac{2V_{DD}}{R_L}\right)}$$

$= 1/8 \quad // \quad \text{high device stress}$

II Class B PA



* 180° conduction angle

* current flow only when V_{DS} is small \rightarrow low P_{dis} .

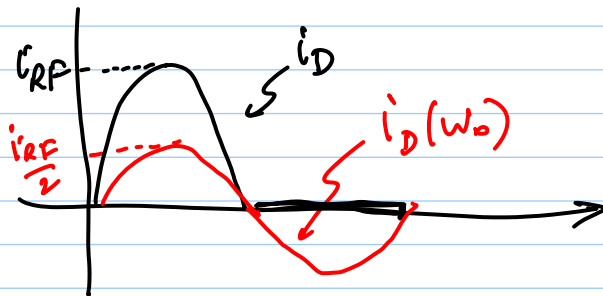
$$i_D = i_{RF} \sin \omega t \text{ for } 0 - T/2$$

* Tank filters out harmonics of i_D , leaving a sinusoidal voltage across R_L

* fundamental current:

$$i_D(\omega_0) = \frac{2}{T} \int_0^{T/2} i_{RF} \sin \omega t \cdot \sin \omega t dt$$

$$= \frac{i_{RF}}{2}$$



$$V_o = \frac{i_{RF}}{2} R_L \sin \omega t$$

$$V_o(\text{max}) \approx V_{DD} \Rightarrow i_{RF}(\text{max.}) = \frac{2V_{DD}}{R_L}$$

$$P_o(\text{max.}) = \frac{V_{DD}^2}{2R_L}$$

$$i_{DC} = \frac{1}{T} \int_0^{T/2} \frac{2V_{DD}}{R_L} \sin \omega t dt = \frac{2V_{DD}}{\pi R_L} //$$

$$\therefore P_{DC} = i_{DC} \cdot V_{DD}$$

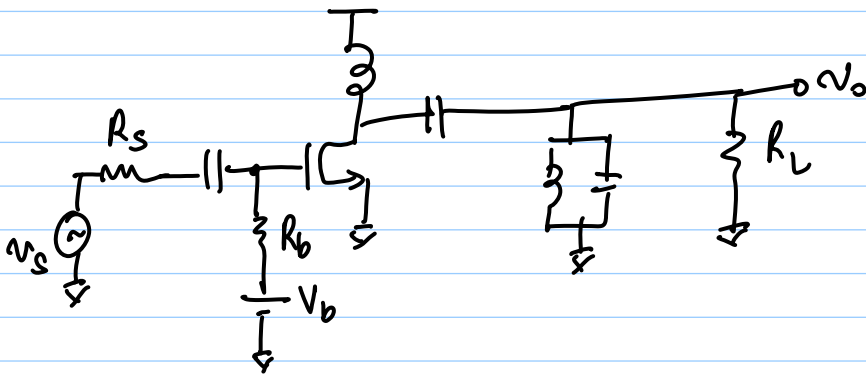
$$= \frac{2V_{DD}^2}{\pi R_L}$$

$$\eta = \frac{P_{out}}{P_{DC}} = \frac{V_{DD}^2 / 2R_L}{2V_{DD}^2 / \pi R_L} = \frac{\pi}{4} = 78.5\%$$

$$P_N = \frac{P_{RF}}{V_{DS(max)} i_D(max)}$$

$$= \frac{V_{DD}^2 / 2R_L}{2V_{DD} \cdot \frac{2V_{DD}}{R_L}} = \frac{1}{8} \quad \text{High stress}$$

With biasing:



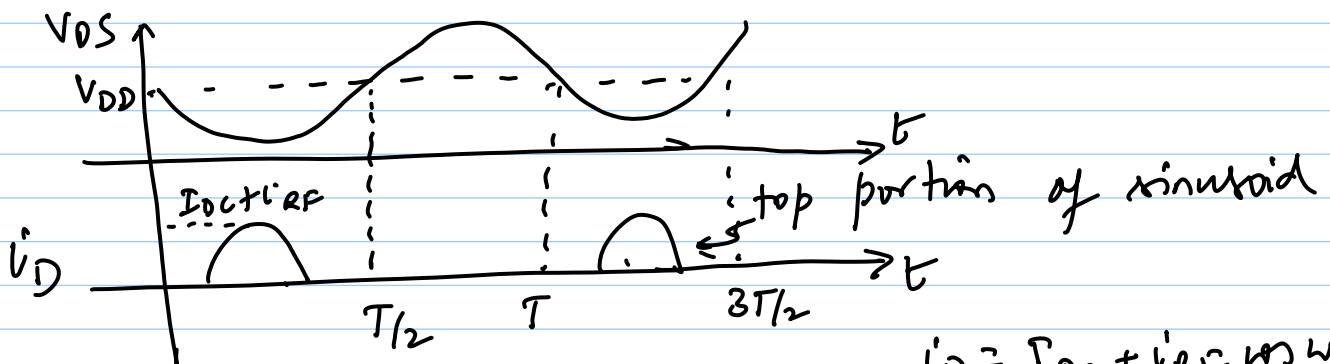
$v_s = V_{RF} \sin \omega t$
class-A

$$V_b > (V_T + V_{RF})$$

class-B

$$V_b = V_T$$

III class-C PA $\Rightarrow V_b < V_T$
 conducts for $< 180^\circ$



$$i_D = I_{DC} + i_{RF} \sin \omega t \quad \text{for } i_D > 0$$

$2\Phi \equiv$ conduction angle

$$\Phi = \cos^{-1} \left(\frac{-I_{DC}}{i_{RF}} \right)$$

"Bias current"

$$I_{DC} = -i_{RF} \cos \Phi$$

{ offset current i_D }
 negative

average current

$$\begin{aligned}\bar{i}_D &= \frac{1}{2\pi} \int_{-\Phi}^{\Phi} (I_{DC} + i_{RF} \cos \varphi) d\varphi \\ &= \frac{1}{2\pi} 2\Phi I_{DC} + \frac{1}{2\pi} (i_{RF} \sin \varphi) \Big|_{-\Phi}^{\Phi} \\ &= \frac{i_{RF}}{\pi} [\sin \Phi - \Phi \cos \Phi]\end{aligned}$$

fundamental:

$$\begin{aligned}i_{fund} &= \frac{2}{T} \int_0^T i_D \cos \omega_0 t dt \\ &= \frac{1}{2\pi} (4 I_{DC} \sin \Phi + 2 i_{RF} \Phi + i_{RF} \sin 2\Phi)\end{aligned}$$

$$= \frac{i_{RF}}{2\pi} (2\Phi - \sin 2\Phi)$$

max. swing = V_{DD}

$$\Rightarrow V_{DD} = i_{RF} \frac{R_L}{2\pi} (2\Phi - \sin 2\Phi)$$

$$\Rightarrow i_{RF_{max}} = \frac{2\pi V_{DD}}{R_L [2\Phi - \sin 2\Phi]}$$

$$\Rightarrow i_{Dpk.} = i_{RF_{max}} + I_{DC}$$

$$= \frac{2\pi V_{DD}}{R_L (2\Phi - \sin 2\Phi)} \left[1 + \frac{\sin \Phi - \Phi \cos \Phi}{\pi} \right]$$

$$\Rightarrow \eta_{max.} = \frac{2\Phi - \sin 2\Phi}{4 (\sin \Phi - \Phi \cos \Phi)}$$

as $\Phi \rightarrow 0, \eta \rightarrow 100\%$

but gain & $P_{out} \rightarrow 0$

* We can obtain high efficiency at the expense of linearity, gain & P_{out}

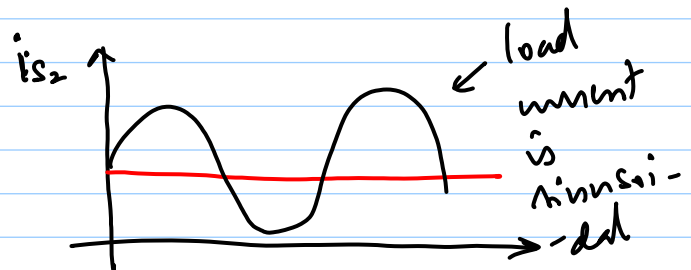
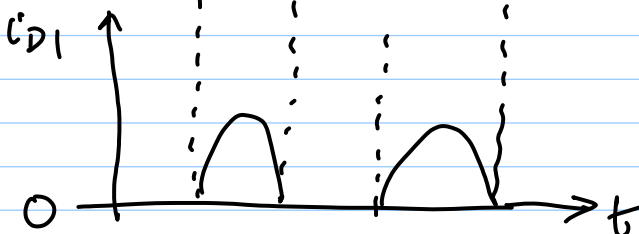
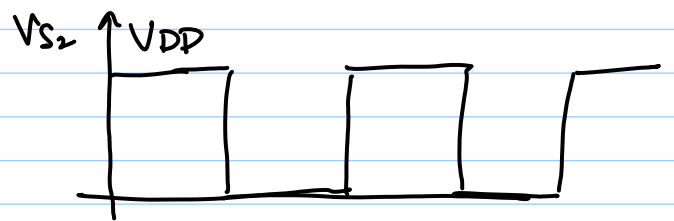
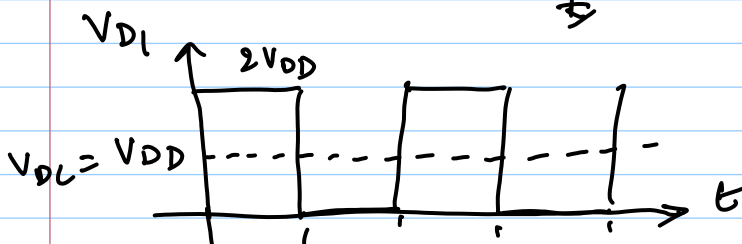
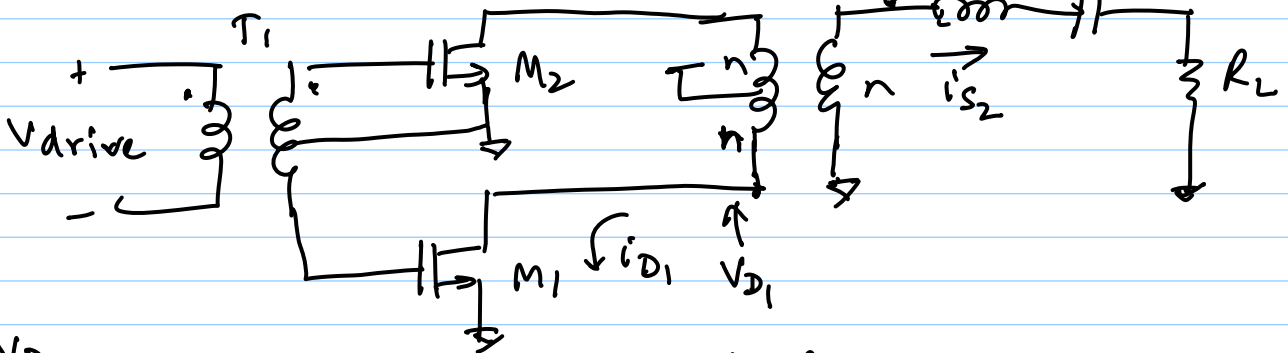
Switching PAs

Basic Principle: Use MOSFET as a switch rather than as a controlled current source in the case of linear PAs

ideal switch ON $\Rightarrow V=0, I>0$ Power $\rightarrow 0$
 OFF $\Rightarrow V>0, I=0$ Power $=0$

no loss in switch $\Rightarrow 100\%$ efficiency

Class-D PA



You can show that:

normalised power handling capability

$$P_N = \frac{P_{out}}{V_{DSpk} \cdot i_{Dpk}} = \frac{1}{\pi} \quad \leftarrow \text{much lower stress than linear PAs}$$

ideal $\eta = 100\%$

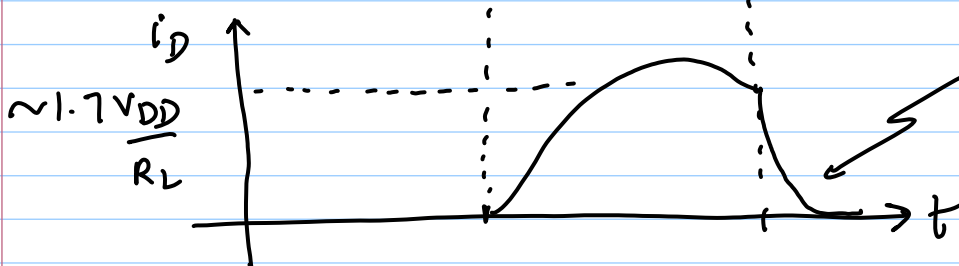
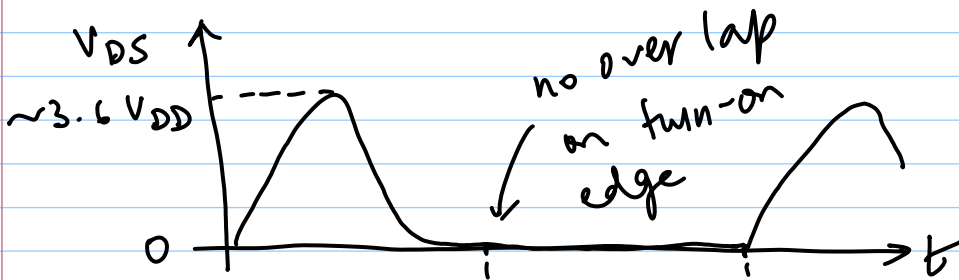
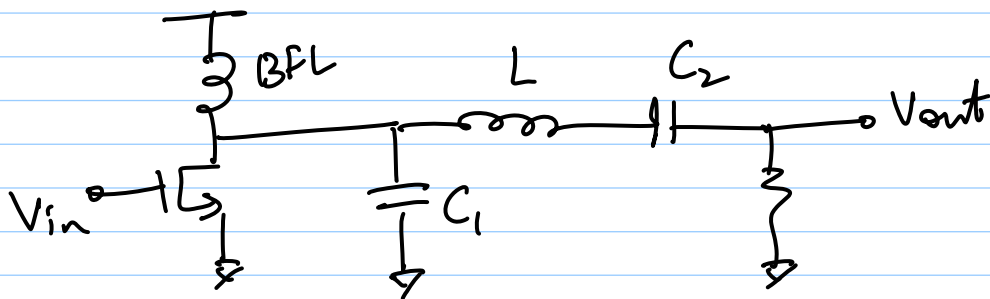
Practical: switches must be very fast relative to

ω_0 , otherwise $\eta < 100\%$.

V Class-E PAs

key ideas: * switch voltage ≈ 0 before current flows

* use higher order filter to shape the pulses



turn-off transient may not be as good (esp. BJT PA)

Ref: Sokal & Sokal, JSSC June 1975

Design Equations

$$L = \frac{QR_L}{\omega}$$

$$C_1 = \frac{1}{\omega R_L \left(\frac{\pi^2}{4} + 1 \right) \left(\frac{\pi}{2} \right)} \approx \frac{1}{5.447 \omega R_L}$$

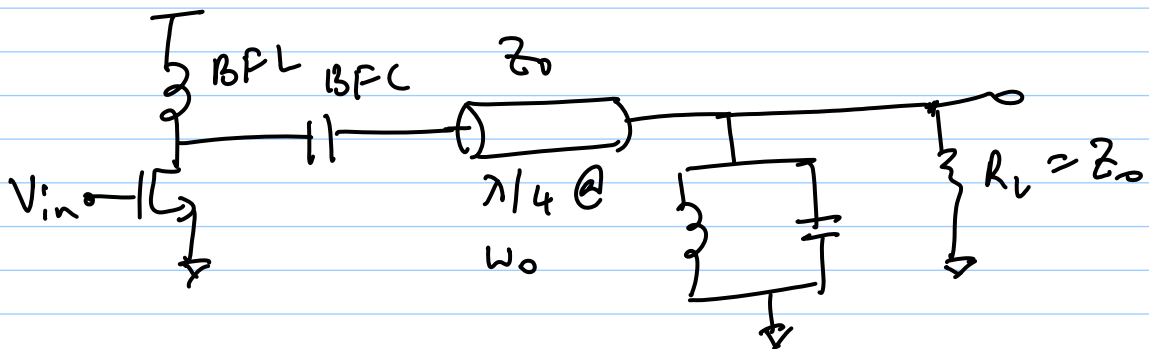
$$C_2 \approx C_1 \left(\frac{5.447}{Q} \right) \left(1 + \frac{1.42}{Q - 2.08} \right)$$

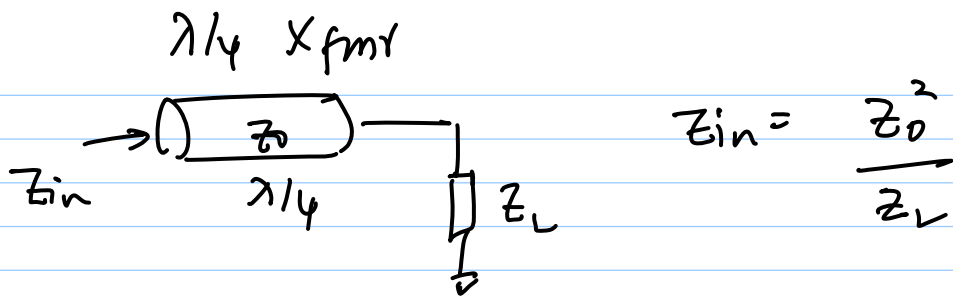
$$P_{out} (max.) = \frac{2}{1 + \pi^2/4} \cdot \frac{V_{DD}^2}{R_L} \approx 0.577 \frac{V_{DD}^2}{R_L}$$

$$P_N \approx \frac{P_o}{V_{DSpk} \cdot i_{Dpk}} \approx 0.098 \leftarrow \text{high stress}$$

note that $V_{DSpk} = 3.6 V_{DD}$
 $i_{Dpk} = 1.7 \frac{V_{DD}}{R_L}$) very high values

VI Class-F PAs





here $Z_{in} = R_L @ \omega_0$

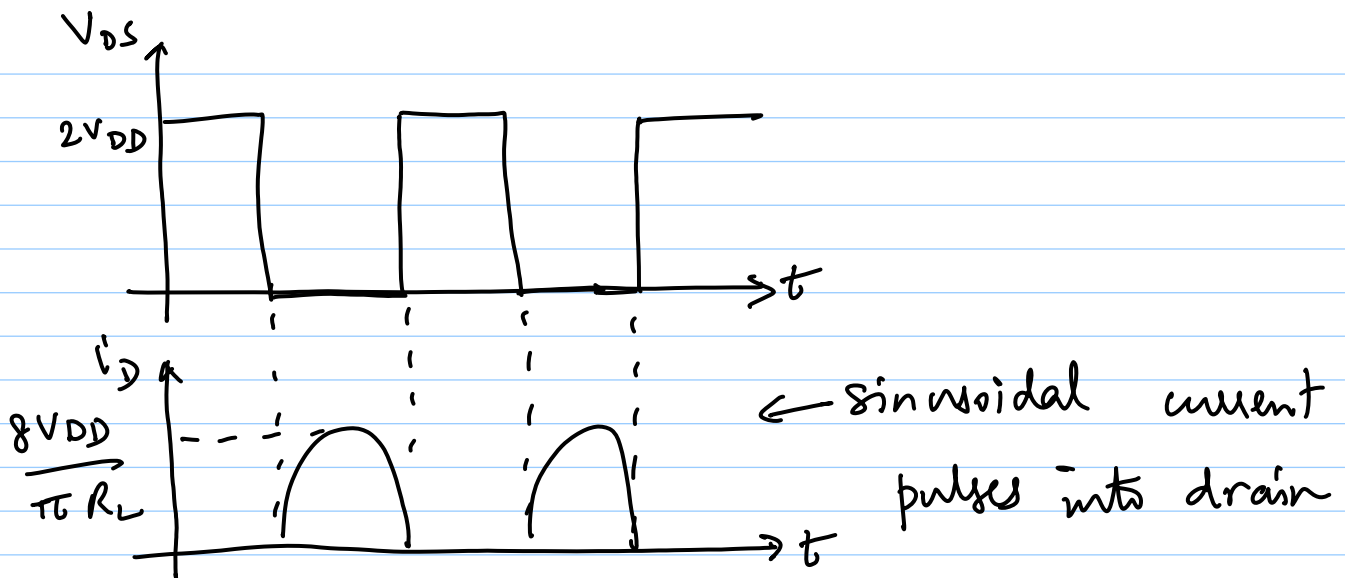
LC tank $\Rightarrow Z(\omega) = 0$ for $\omega \neq \omega_0$ (short ckt)

@ $\omega = 2n\omega_0$, T-line $l = 2n \frac{\lambda}{4} = n \frac{\lambda}{2}$

\Rightarrow short ckt - @ Drain

@ $\omega = (2n+1)\omega_0$, T-line $l = (2n+1)\lambda/4$

LC - tank short ckt \Rightarrow open ckt - @ Drain



$$V_{fund.} = \frac{4}{\pi} (V_{DD})$$

$$P_o = \left[\frac{4}{\pi} \left(\frac{V_{DD}}{\sqrt{2}} \right) \right]^2 \cdot \frac{1}{R_L} = \frac{8V_{DD}^2}{\pi^2 R_L}$$

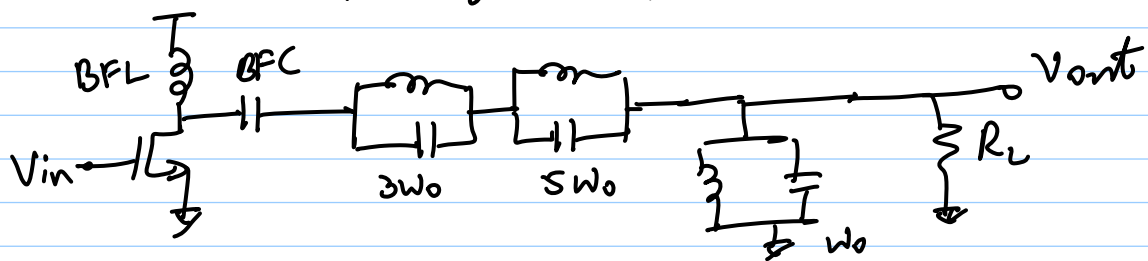
$$\eta_{ideal} = 100\%$$

in practice $\eta \gg \eta_{class-E}$

$$P_N = \frac{P_o}{V_{DSpk} \cdot I_{Dpk}} = \frac{8V_{DD}^2 / \pi^2 R_L}{2V_{DD} \cdot \frac{8V_{DD}}{\pi R_L}}$$

$$= \frac{1}{2\pi} \approx 0.16 \quad (\text{better than class-E})$$

alternative topology: replace T-lines with L-C



* Note: Switching PAs are constant envelope PAs

$$V_{out} = f(V_{DD}), \text{ \& not } f(V_{in})$$

Other design considerations:

1) Power-added Efficiency:

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}}$$

$$\text{obviously } PAE < \eta$$

→ takes power gain into account

2) Stability: * g_d is very important (layout)

* stability-gain trade off

3) Breakdown

* output swings upto $2V_{DD}$

* BV reduces as tech. scales

→ DB & SB diode breakdown (A)

→ D-S punchthrough (B)

→ Time-dependent dielectric breakdown (TDDB) (C)

→ gate oxide rupture (D)

(A) : Diode BV \sim few V ($2-3 \times V_{DD}$)

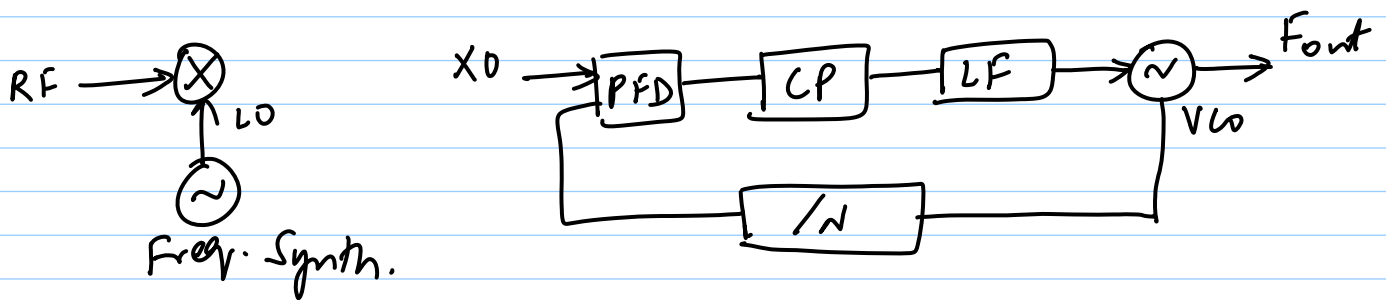
(B) : If V_D is large, depletion region extends to source, eliminating the channel

(C) : Gate oxide damage due to energetic carriers - @ high fields, high energy e's create oxide traps; charge trapped here shift device V_T (cumulative)

(D) : Gate oxide rupture occurs @ high gate fields

4) Large-signal impedance matching

VCOs - I



* Autonomous ckt (i.e. needs to be locked with a PLL)

* Applications: → Freq. Synthesis

→ Clock & Data Recovery ckt
(also optical comm.)

→ μ -processor clk gen.

Types of Oscillators (RF)

* Relaxation oscillators (typically C-based)

→ poor spectral purity

→ high phase noise

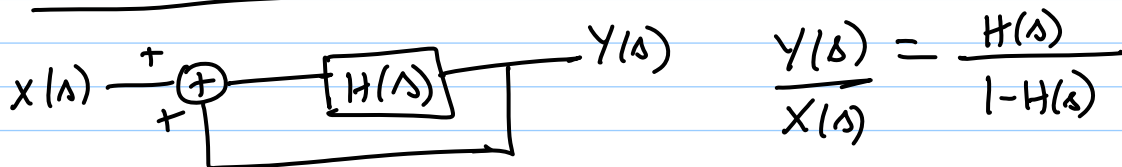
* Harmonic oscillators

→ LC VCOs, Crystal oscillators

→ good phase noise & spectral purity

In this course, we will study these

Feed back model (2-port)



$H(s_0) = 1 \Rightarrow$ self-sustaining oscillations

Amplitude is constant if $H(s_0 = j\omega_0) = 1$

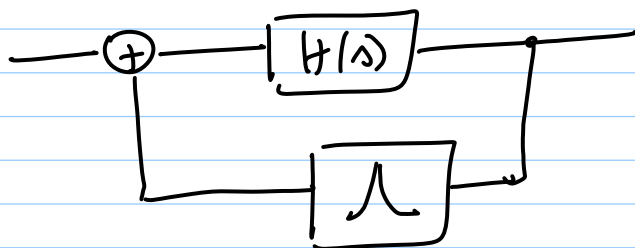
Barkhausen Criterion

(1) $|H(j\omega_0)| = 1$

(2) $\angle H(j\omega_0) = 360^\circ$ (loop phase shift is +ve)

\rightarrow if there is DC negative feedback, only 180° excess phase shift is required

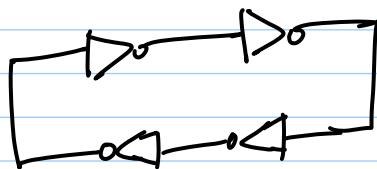
i.e. $\angle H(j\omega_0) = 180^\circ$



RLC tank = sets frequency ω_0

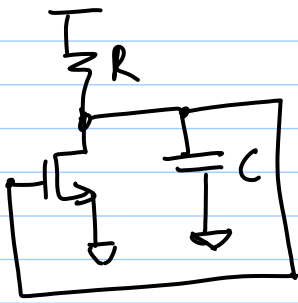
* Barkhausen criteria are necessary but not sufficient conditions

e.g. #1



\leftarrow does not oscillate

①

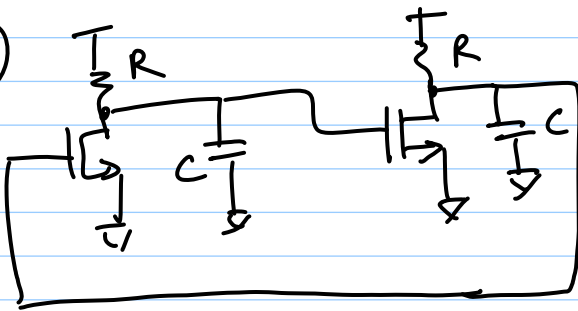


* Does not oscillate

only 1 pole; 270° phase shift

@ $\omega = \infty$

②



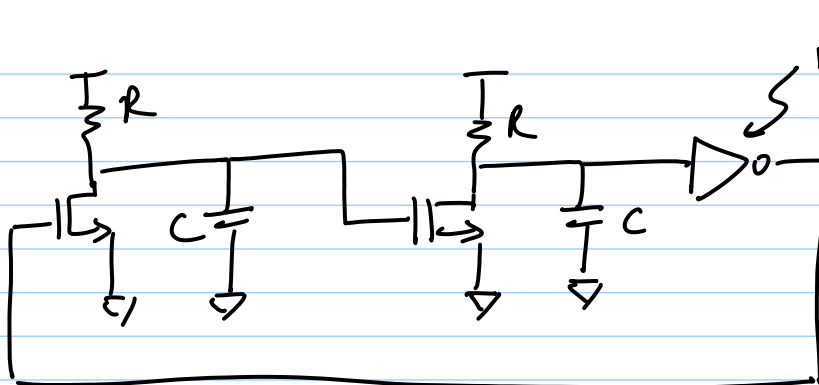
* Does not oscillate

2-poles = 180° @ $\omega = \infty$

but, DC f.b. = 360°

\Rightarrow latch-up

③



ideal inv.

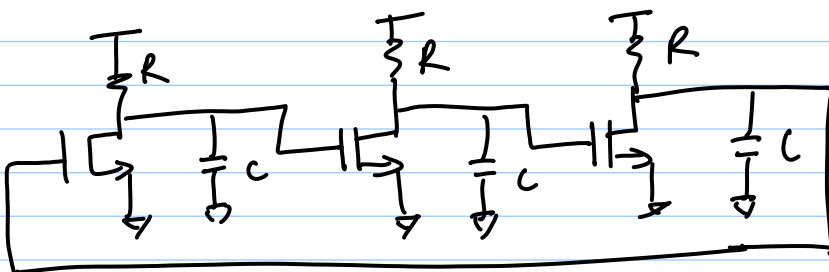
* Does NOT oscillate

DC f.b. = -ve (180°)

excess phase required = 180°

2-poles $\Rightarrow 180^\circ$ reached only @ $\omega = \infty$

④

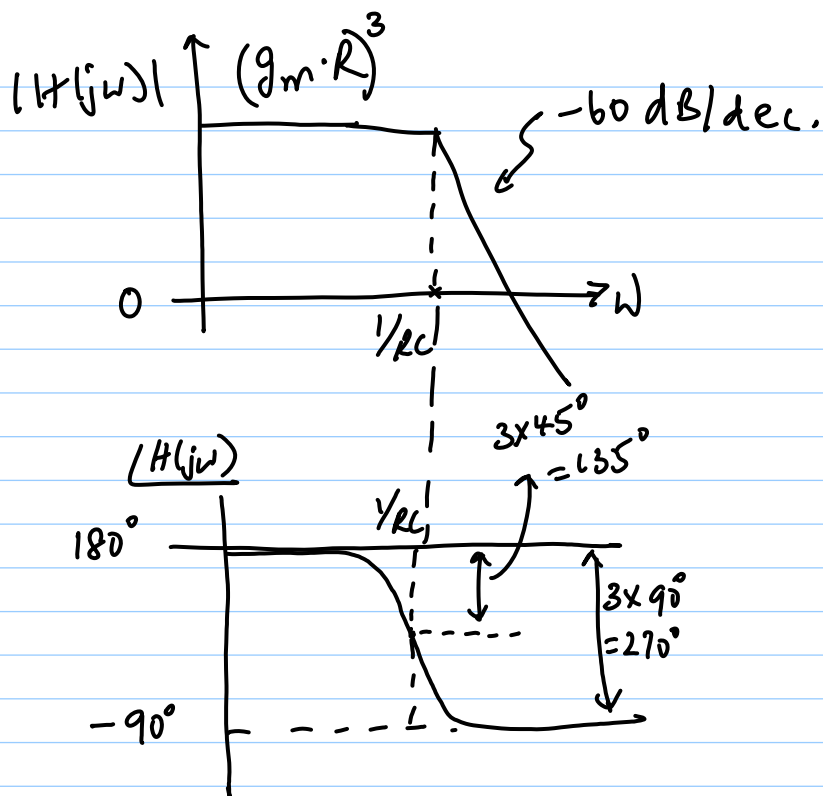
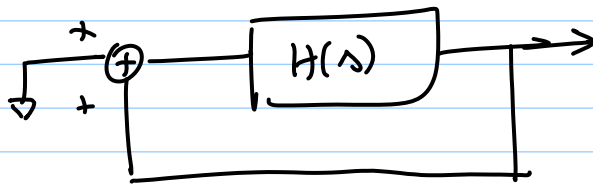


Does oscillate

at DC $\Rightarrow 180^\circ$ phase shift
for each stage,

$$\frac{out}{in} = \frac{-A_0}{1 + s/\omega_p}$$

$$\Rightarrow H(s) = - \left(\frac{A_0}{1 + s/\omega_p} \right)^3 ; \omega_p = \frac{1}{RC}$$



say oscillation happens @ ω_0

$|H(j\omega)| = 360^\circ/0^\circ$ has to be satisfied

$$\Rightarrow \pi - 3 \tan^{-1} \left(\frac{\omega_0}{\omega_p} \right) = 0^\circ$$

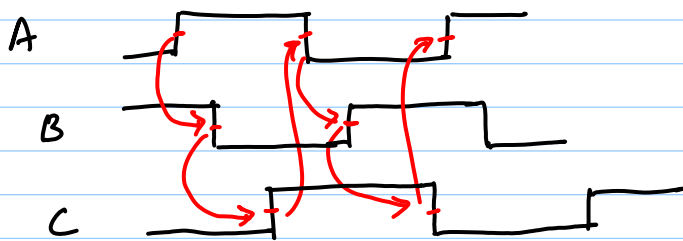
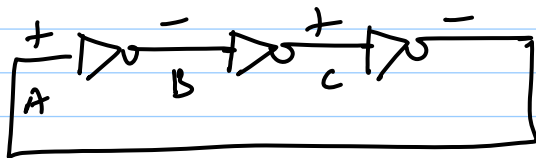
$$\tan^{-1} \left(\frac{\omega_0}{\omega_p} \right) = \frac{\pi}{3}$$

$$\Rightarrow \boxed{\omega_0 = \sqrt{3} \omega_p}$$

$$|H(j\omega_0)| = \frac{A_0^3}{\left[1 + \left(\frac{\omega_0}{\omega_p} \right)^2 \right]^{3/2}} = \frac{A_0^3}{8}$$

$|H(j\omega_0)| = 1$ has to be satisfied $\Rightarrow \boxed{A_0 = 2}$

* What happens if $A_0 \neq 2$?

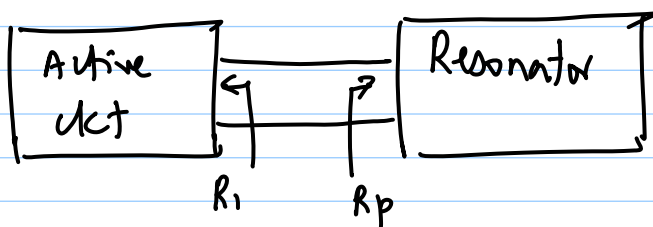


$$T = 6 \tau$$

$$\tau = \frac{t_r + t_f}{2}$$

In general $T = 2n\tau$ for n inv.

One-port View (useful for LC oscillators)



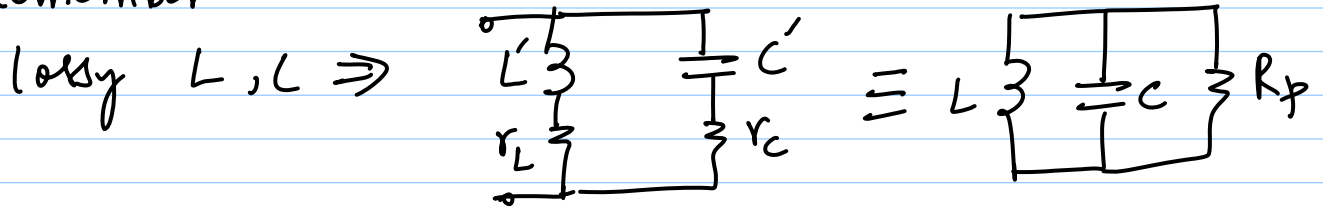
we want

$$R_i = -R_p$$

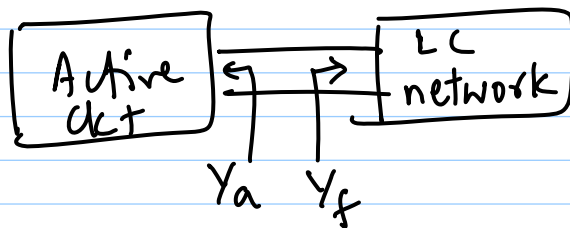
* Energy is lost in R_p every cycle

* This is replenished by active ckt

Remember



* Freq. determined by L, C



$$Y_a = G_a + jB_a$$

$$Y_f = G_f + jB_f$$

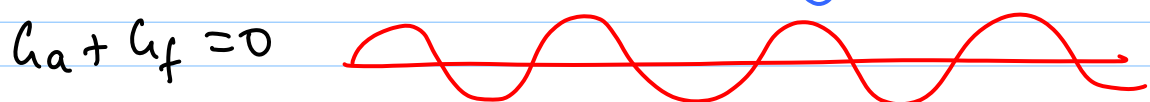
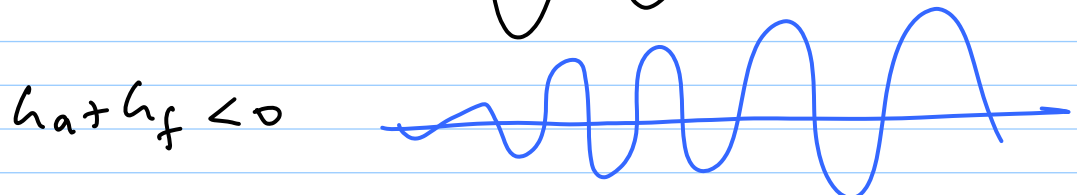
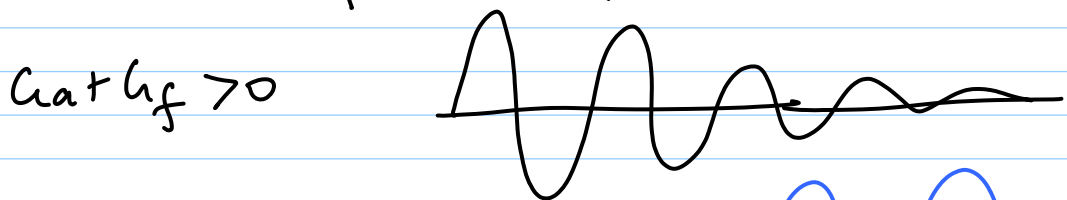
$$\Rightarrow G_a = -G_f \quad \text{and} \quad B_a = -B_f$$

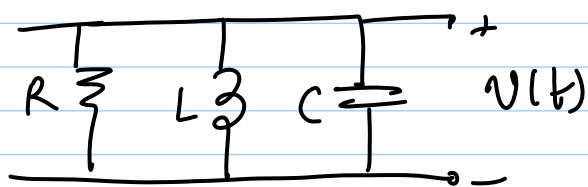
* How about amplitude?

\rightarrow usually require "startup" loop gain > 1 (e.g. 3)

\rightarrow at steady state, amp inside active block saturates with low gain @ peaks so that average $Lh = 1$

choose $G_a + G_f < 0$ for startup



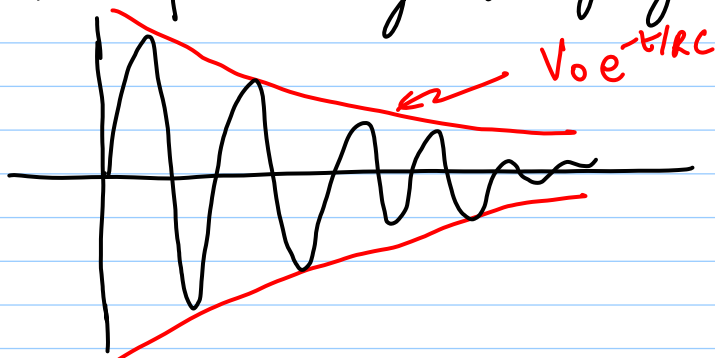


$v(0) = V_0 = \text{initial voltage across cap.}$

If $Q = \frac{R}{\omega_0 L} = \omega_0 RC = \sqrt{\frac{L}{C}} > \frac{1}{2}$,

the 2nd-order system will have imaginary roots

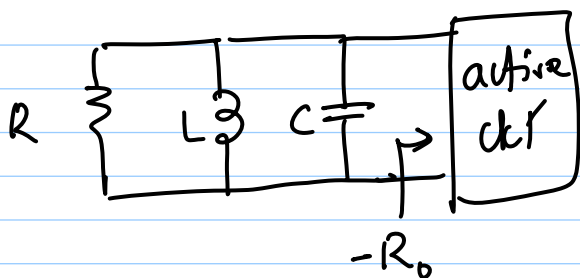
⇒ exponentially decaying sinusoid



$\tau = 1/RC$

as $R \rightarrow \infty$, $\tau \rightarrow 0$

⇒ amplitude does not decay (ideal LC)



$R_{eq.} = \frac{R R_0}{R_0 - R}$

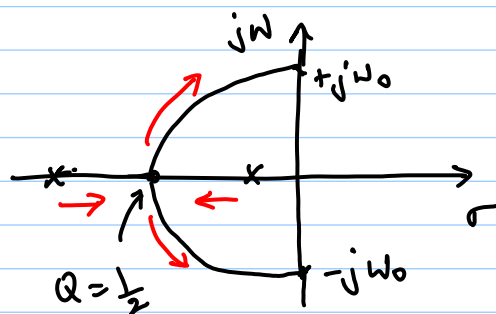
decaying exponential envelope = $V_0 \exp\left[-\frac{t}{R_{eq.} C}\right]$

$R = R_0 \Rightarrow$ amplitude never decays

$R < R_0 \Rightarrow$ decaying exp.

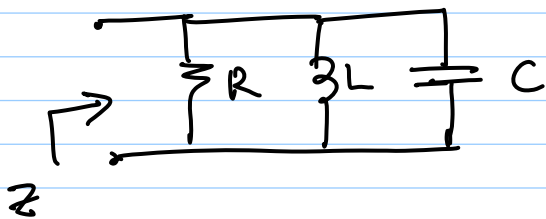
$R > R_0 \Rightarrow$ growing exp.

@ $Q = \frac{1}{2}$, poles split into complex conjugate pair

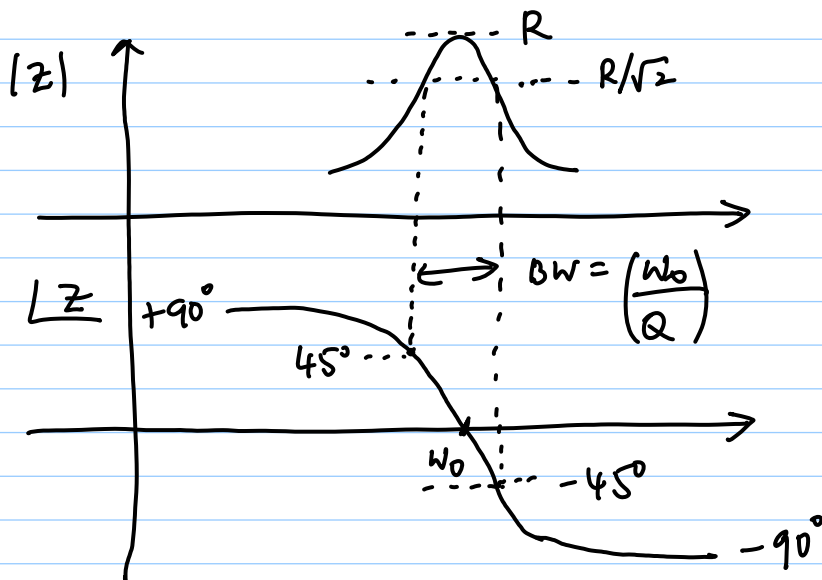


Linear oscillators

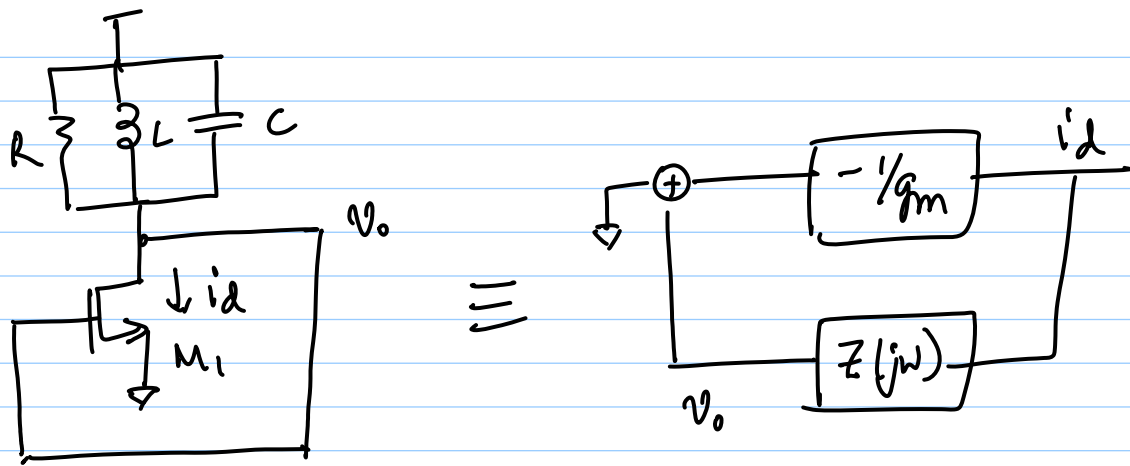
- 1) $G_{in} + G_p = 0 \Rightarrow$ oscillations are sustained, but startup is compromised
- 2) $G_{in} + G_p < 0 \Rightarrow$ startup possible, but amplitude not stable (\uparrow)
 \Rightarrow oscillator needs to be fundamentally non-linear



$$Z(j\omega) = \frac{sL}{LCs^2 + \frac{L}{R}s + 1}$$



$$\omega_0 = \frac{1}{\sqrt{LC}}$$

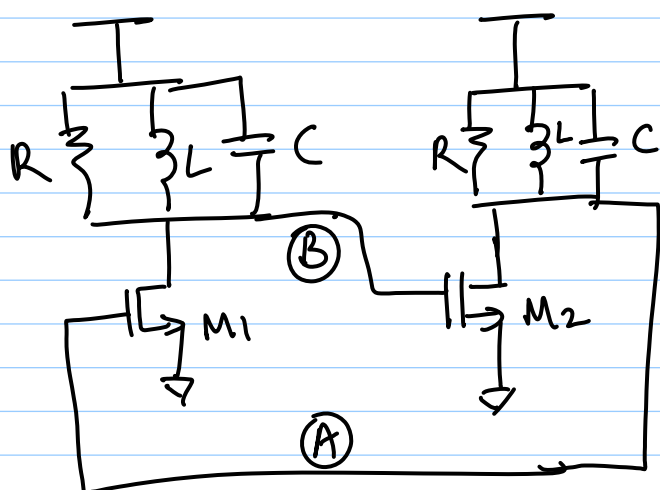


(a) resonance, RLC tank phase shift = 0

M_1 phase shift = 180°

$\Rightarrow |H(s)| = 180^\circ \Rightarrow$ no oscillations

Let us cascade two of these stages.



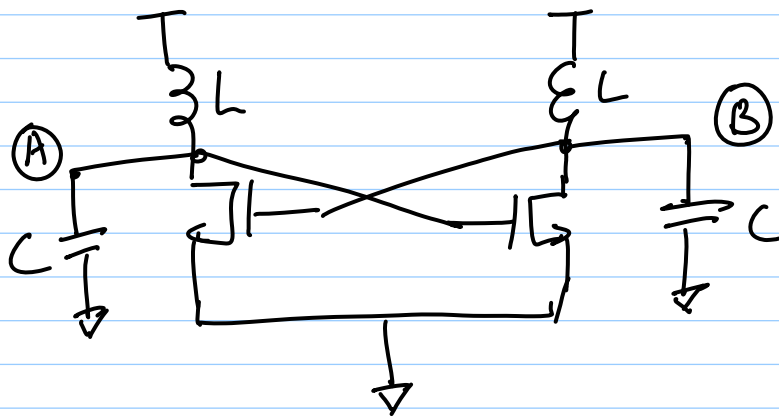
* at $\omega_0 = \frac{1}{\sqrt{LC}}$,
each stage contributes
 180° phase shift

* $|H(s)| = 360^\circ$

* oscillations are possible if $|H(j\omega_0)| \geq 1$

$\Rightarrow (g_m R)^2 \geq 1$ $\left\{ \begin{array}{l} > \Rightarrow \text{startup} \\ = \Rightarrow \text{steady-state} \end{array} \right\}$

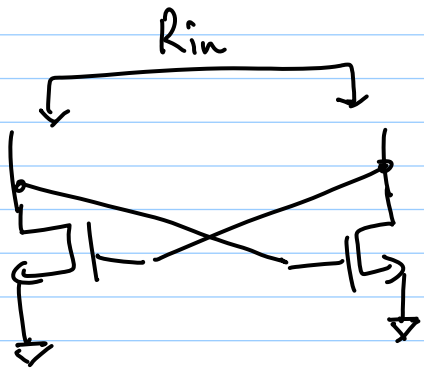
Redraw :



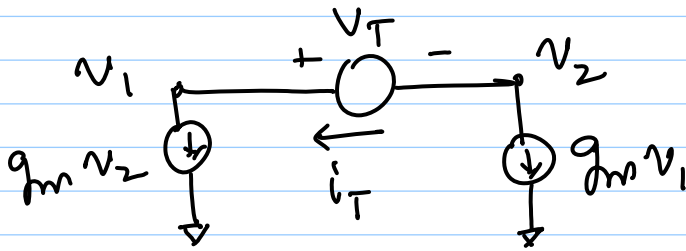
Modern on-chip
LC oscillator

A & B are 180° out of phase
 \Rightarrow differential outputs

VCOs - II



$R_{in} = ?$

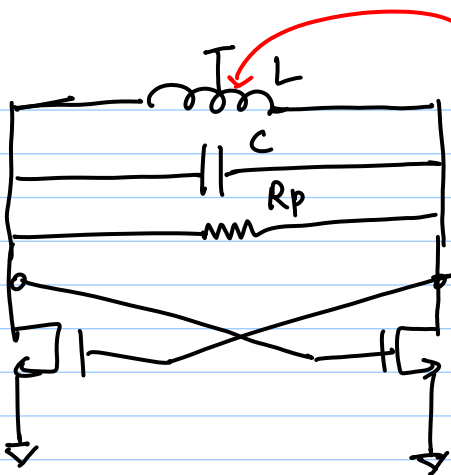


$$i_T = g_m v_2 = -g_m v_1$$

$$2i_T = g_m (v_2 - v_1)$$

$$= -g_m v_T$$

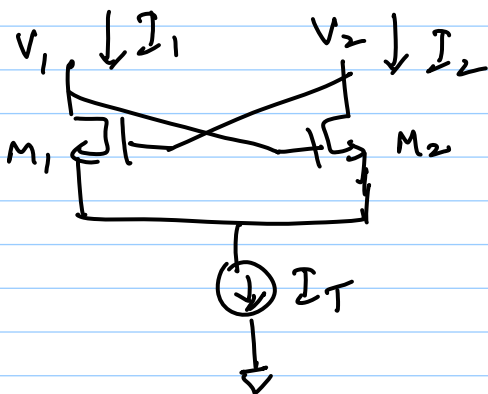
$$\Rightarrow R_{in} = \frac{v_T}{i_T} = -\frac{2}{g_m}$$



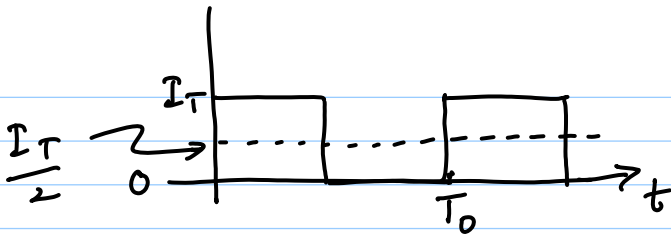
centre-tap to V_{DD}

for oscillations,

$$|R_p| = \left| -\frac{2}{g_m} \right| \Rightarrow \underline{\underline{g_m R_p = 2}}$$



positive feedback; assume M_1 & M_2 switch quickly
 $I_1 - I_2 = I_d$



$$I_1 = I_T \left[\frac{1}{2} + \frac{2}{\pi} \left\{ \sin(\omega_0 t) + \frac{1}{3} \sin(3\omega_0 t) + \dots \right\} \right]$$

LC tank filters out DC & harmonics of I

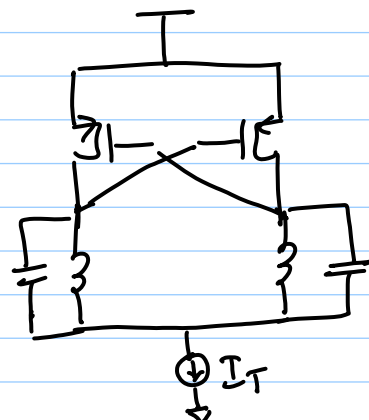
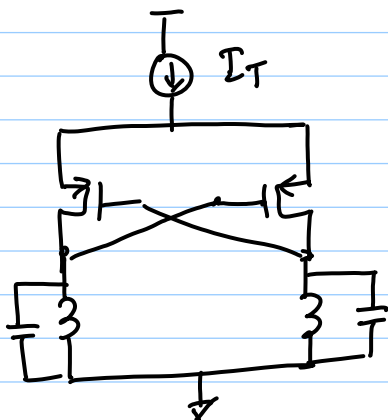
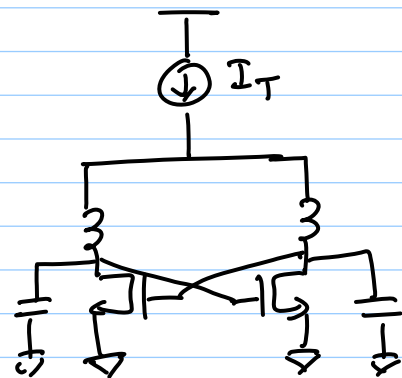
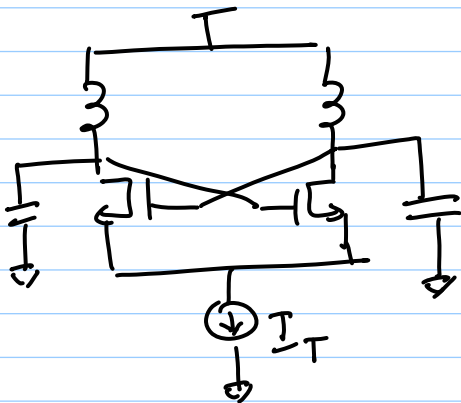
$$\Rightarrow V_1 \{ = -V_2 \} = I_1(\omega_0) \cdot Z(j\omega_0) = \frac{2}{\pi} \cdot I_T \cdot \frac{R_p}{2}$$

$$= \frac{1}{\pi} I_T \sin(\omega_0 t) \cdot R_p$$

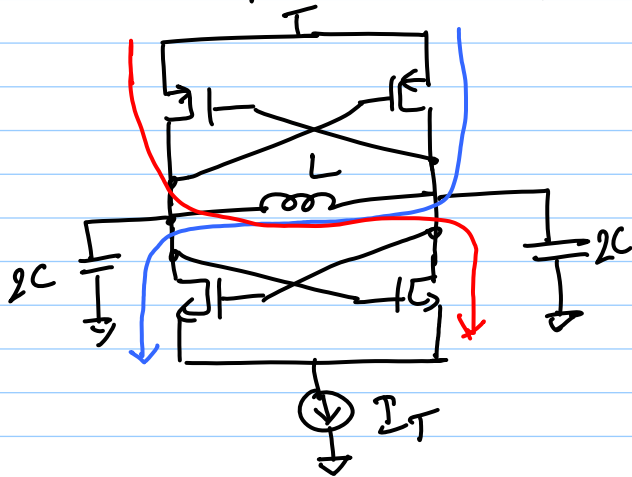
$$|V_{od}| = \frac{2}{\pi} I_T R_p \leftarrow \text{output amplitude}$$

* to a first order, oscillation ampl. is independent of device size!

Other flavours of CC VCO:

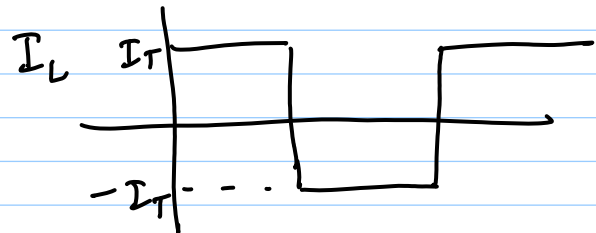


Another popular topology?



* current sense

$$\Rightarrow R_{in} = \frac{-2}{g_{m_n} + g_{m_p}}$$

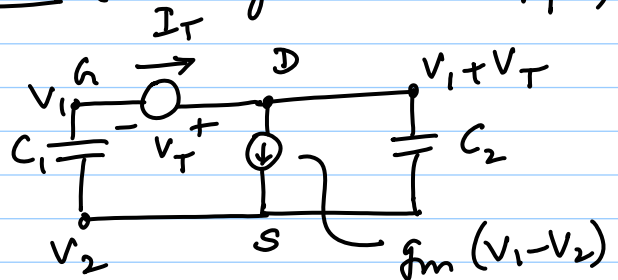
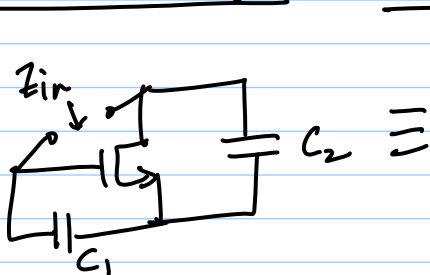


$$\Rightarrow I_L = \frac{4I_T}{\pi} \left[\sin \omega t + \frac{1}{3} \sin 3\omega t + \dots \right]$$

$$\Rightarrow \boxed{\text{Amplitude} = \frac{4}{\pi} I_T R_p}$$

← double the amplitude compared to nmos-c-c. VCO

Single-transistor oscillators (usually discrete apps)



$$V_1 = V_2 - I_T \cdot \frac{1}{sC_1}$$

$$\Rightarrow (V_1 - V_2) = -\frac{I_T}{sC_1}$$

KCL @ D:

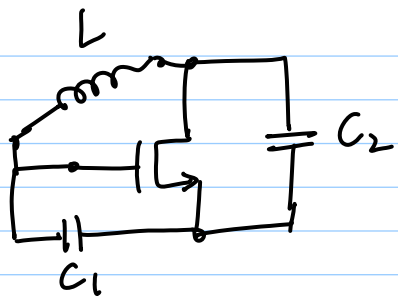
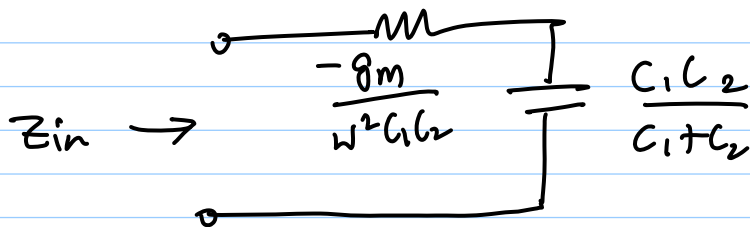
$$\begin{aligned} I_T &= g_m (V_1 - V_2) + \left\{ (V_1 - V_2) + V_T \right\} \cdot sC_2 \\ &= (g_m + sC_2) \cdot (V_1 - V_2) + V_T \cdot sC_2 \\ &= -\frac{(g_m + sC_2)}{sC_1} \cdot I_T + V_T \cdot sC_2 \end{aligned}$$

$$\Rightarrow Z_{in} = \frac{V_T}{I_T} = \frac{g_m + s(C_1 + C_2)}{s^2 C_1 C_2}$$

$$= \frac{g_m}{s^2 C_1 C_2} + \frac{1}{s C_{eq}} \quad \text{where } C_{eq} = \frac{C_1 C_2}{C_1 + C_2}$$

$$Z_{in}(j\omega) = -\frac{g_m}{\omega^2 C_1 C_2} + \frac{1}{j\omega C_{eq}}$$

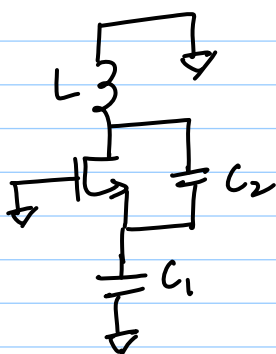
i.e. equivalent circuit is



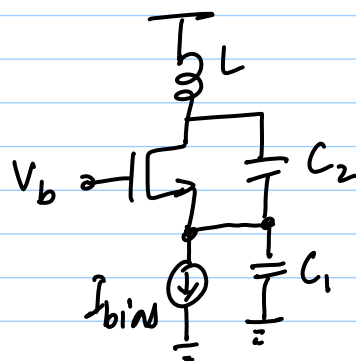
* For oscillations, we require

$$\frac{g_m}{\omega^2 C_1 C_2} \leq R_p$$

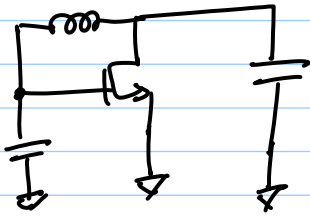
* Note that we can get 3 oscillator topologies by defining an AC ground
Ground gate \Rightarrow Colpitts oscillator



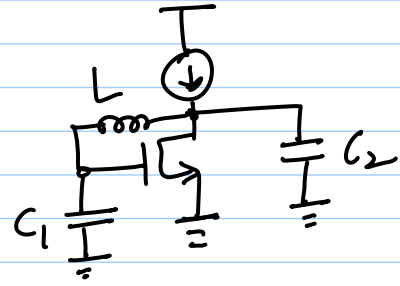
\Rightarrow
with
biasing



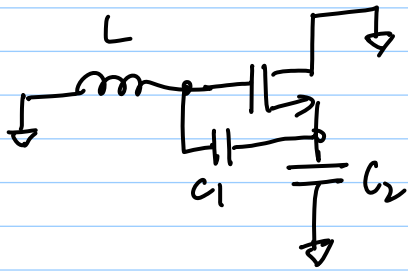
Ground source



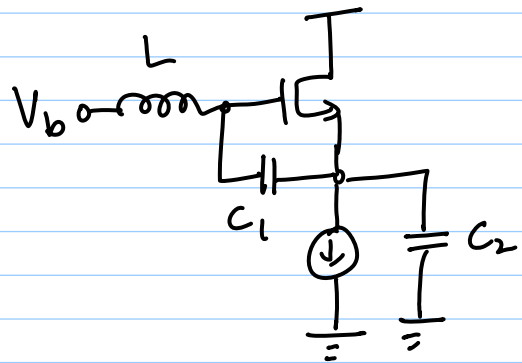
\Rightarrow
with
biasing



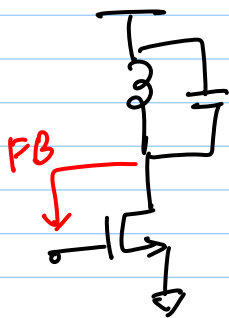
Ground drain



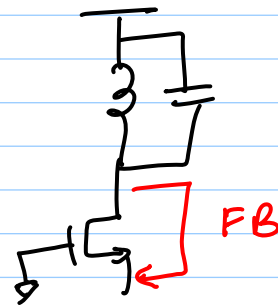
\Rightarrow
with
biasing



* back to FB model :



X



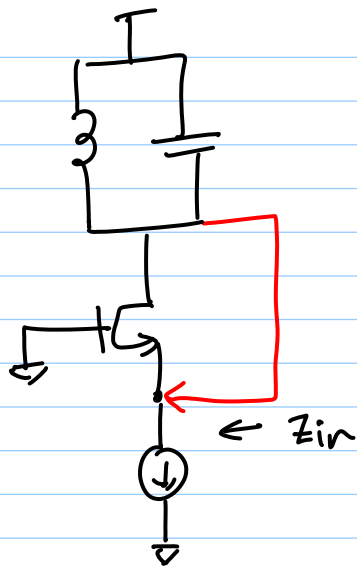
✓

\rightarrow (a) resonance, $z(\text{tank}) = R_p$

$\Rightarrow V$ & I are in-phase

total phase = 0 \Rightarrow f.b. signal goes to emitter

(without any additional phase in f.b.)

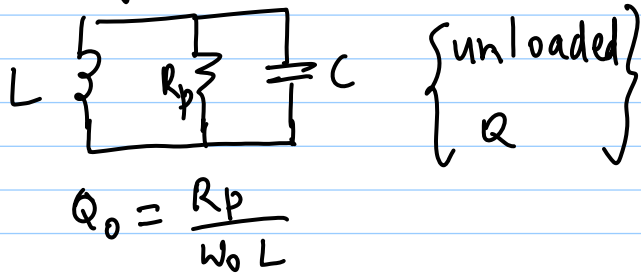


$$Z_{in} = \frac{1}{g_m + g_{mb}} + \frac{R_p}{(g_m + g_{mb})r_o}$$

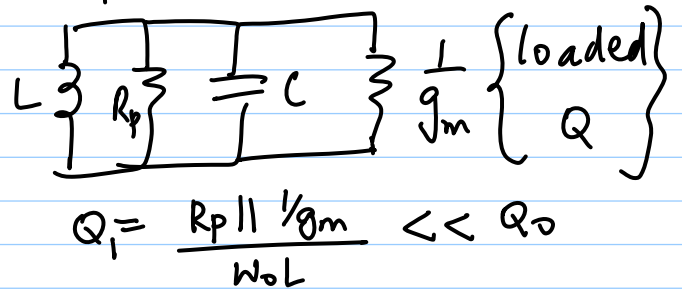
neglect this

= low impedance

original tank

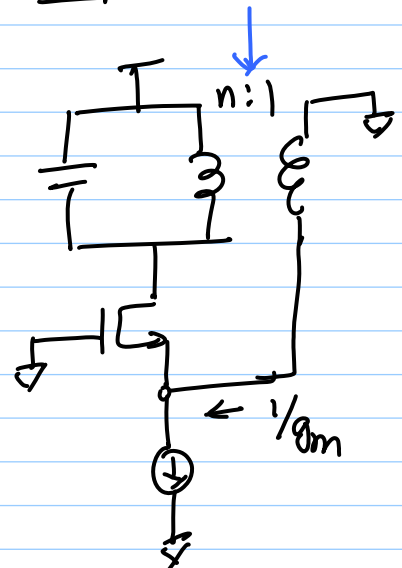
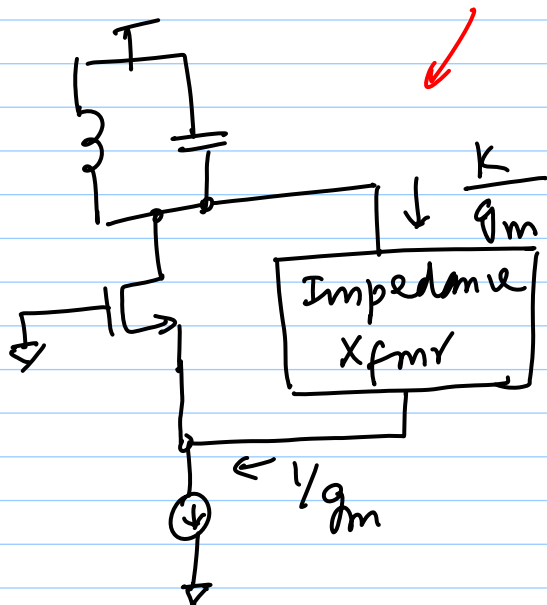


after f.b.



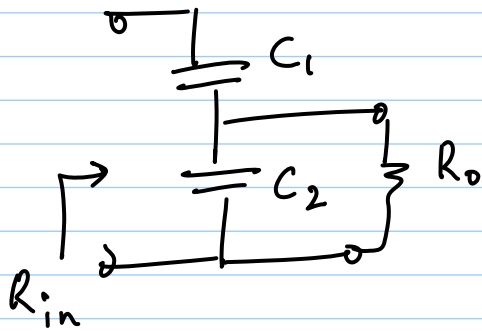
⇒ LC can become < 1 ⇒ no oscillations
 → Use an Impedance Xfmr

Explicit Xfmr

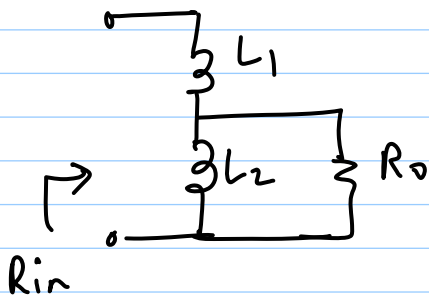


$\frac{n^2}{g_m}$ is seen by tank

Remember from impedance matching:



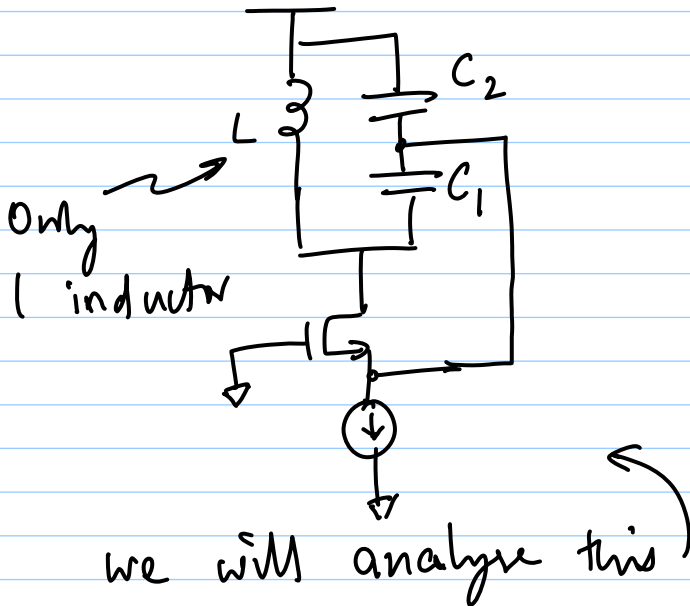
$$R_{in} = \left(1 + \frac{C_2}{C_1} \right)^2 R_0$$
$$= n^2 R_0$$



$$R_{in} = \left(1 + \frac{L_1}{L_2} \right)^2 R_0$$
$$= n^2 R_0$$

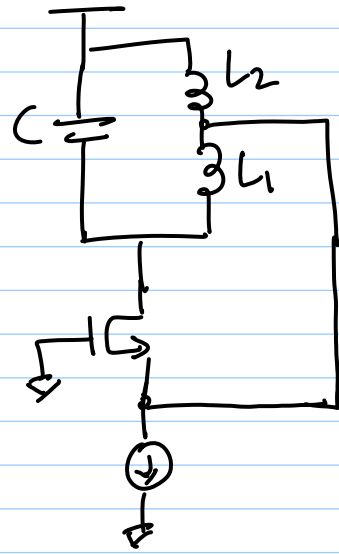
Colpitts Oscillator ; Quadrature Signal Generation

Colpitts oscillator

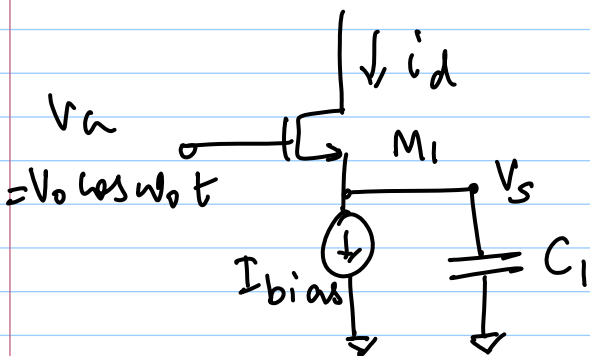


$$\omega_0 = \frac{1}{\sqrt{L - \frac{C_1 C_2}{C_1 + C_2}}}$$

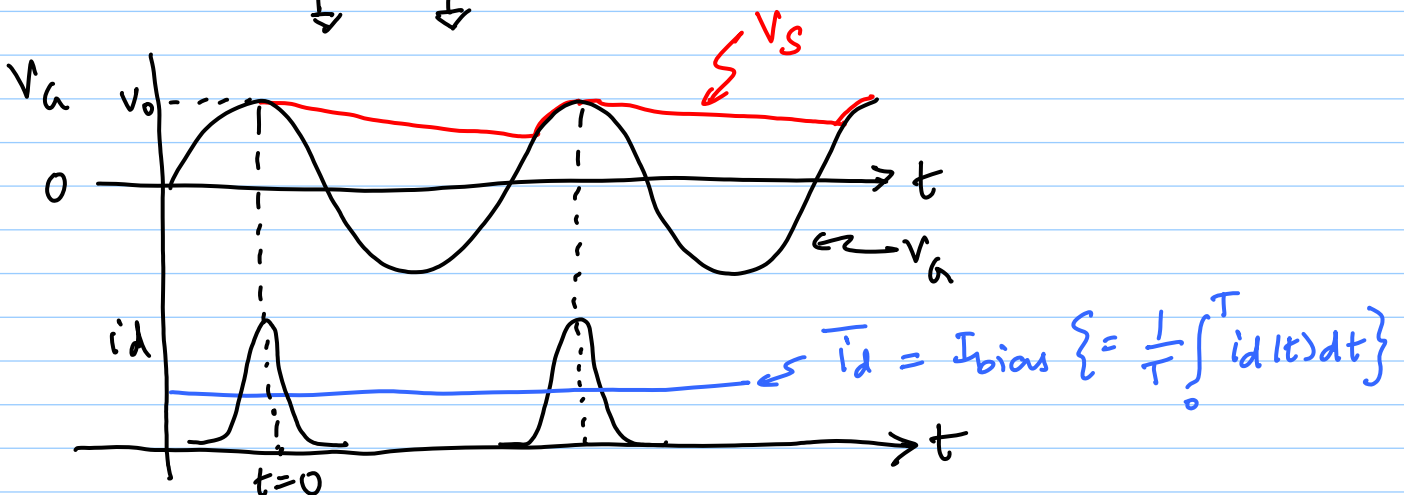
Hartley oscillator



Colpitts oscillator



- * Assume large amplitude
- * Assume $\omega_0 =$ high freq.
- * Assume C_1 is short @ ω_0



* M_1 conducts only when V_a is large ($\sim V_0$)
 \rightarrow when $V_a \ll V_0$, M_1 cuts off, I_{bias} discharges C_1

* $i_d =$ periodic pulses of current @ ω_0

Fourier series:

$$i_d(t) = I_0 + \sum_{n=1}^{\infty} \hat{I}_n \cos n\omega_0 t$$

$\rightarrow t=0$ reference = peak of $V_0 \cos \omega_0 t$

$$I_{DC}(C_1) = 0 \Rightarrow I_0 = I_{bias}$$

\rightarrow fundamental component of i_d is

$$\hat{I}_1 = \frac{2}{T} \int_0^T i_d(t) \cos \omega_0 t dt$$

$$\approx \frac{2}{T} \int_0^T i_d(t) dt \quad \left\{ \begin{array}{l} \text{current exists} \\ \text{only @ peak } V_a \end{array} \right\}$$

$$\hat{I}_1 = 2I_{bias}$$

* If i_d flows through a tank tuned to ω_0 ,
 only \hat{I}_1 creates a voltage

$$G_m \approx \frac{\hat{I}_1}{V_0} = \frac{2I_{bias}}{V_0}$$

\leftarrow applicable to
any device

\leftarrow assume V_S is

Recall that

$$g_m (\text{long channel}) = \frac{2I_{bias}}{V_{DSAT}}$$

constant (due
to C_1)

short-channel:

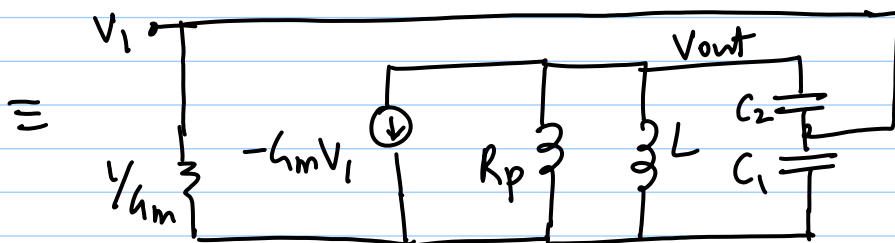
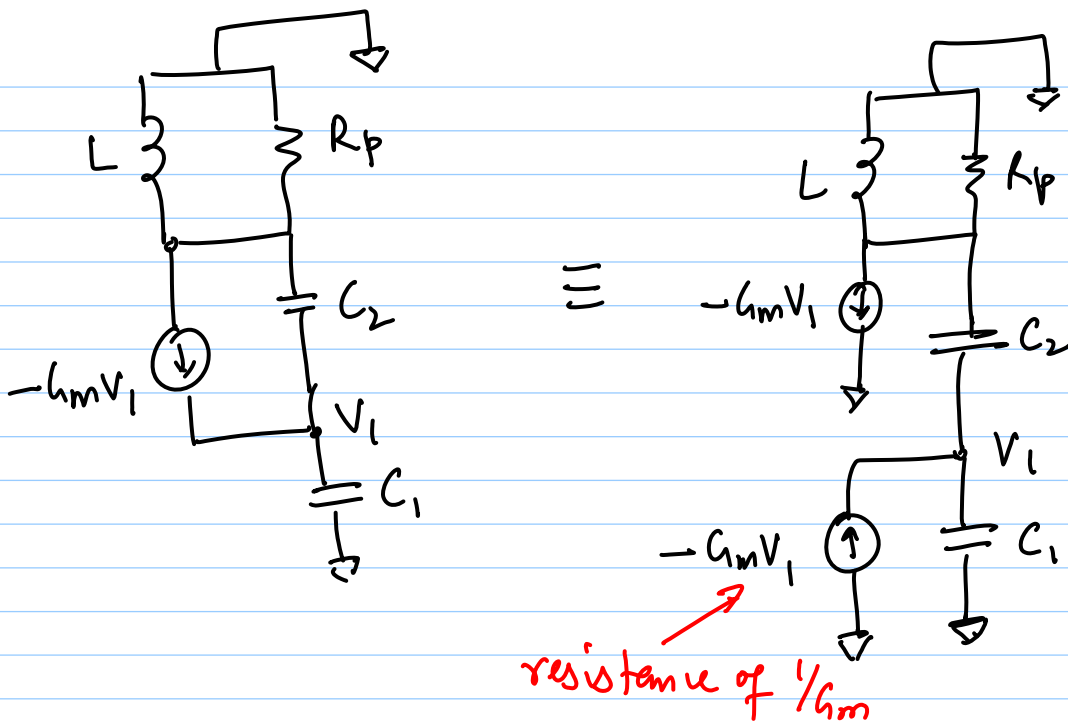
$$I_D = \mu_n \frac{C_{ox}}{2} \left(\frac{W}{L}\right) (V_{GS} - V_T) \cdot L E_c \quad (\text{vel. sat.})$$

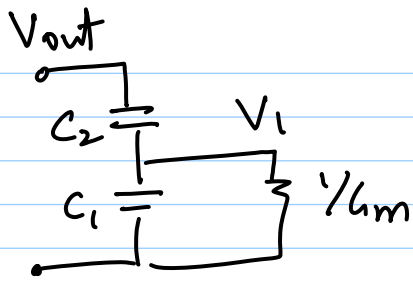
$$\Rightarrow g_m = \frac{\partial I_D}{\partial V_{GS}} = \frac{I_{bias}}{V_{DSAT}}$$

\Rightarrow in general,

$$\frac{I_{bias}}{V_{DSAT}} \leq g_m \leq \frac{2 I_{bias}}{V_{DSAT}}$$

$$\Rightarrow \boxed{\frac{V_{DSAT}}{V_0} \leq \frac{G_m}{g_m} \leq \frac{2 V_{DSAT}}{V_0}}$$



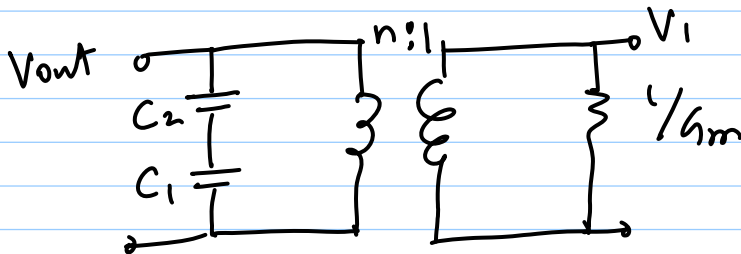


assume $1/G_m$ does not load
the cap divider
(i.e. $Q \gg 1$)

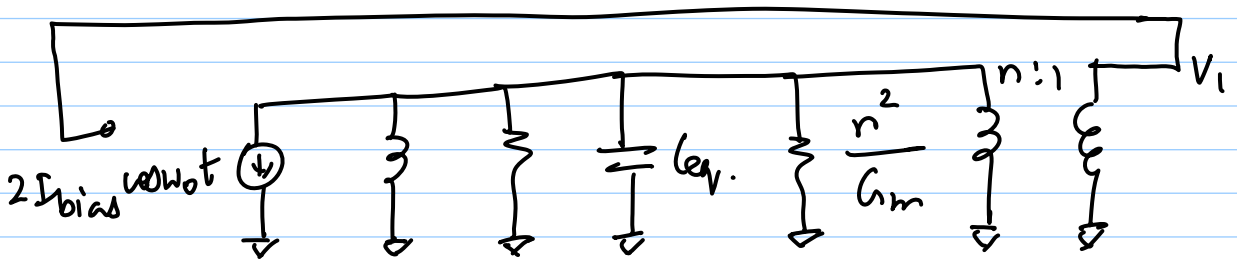
$$V_1 = \frac{C_2}{C_1 + C_2} \cdot V_{out} = \frac{1}{n} \cdot V_{out}$$

$$n = \frac{C_1 + C_2}{C_2} = 1 + \frac{C_1}{C_2} \quad \leftarrow \text{equivalent trans ratio}$$

equivalent ckt:



overall oscillator ckt becomes:



$$C_{eq} = \frac{C_1 C_2}{C_1 + C_2}$$

resonance: set $\omega_0 = \frac{1}{\sqrt{L \cdot C_{eq}}}$

$$V_{out} = 2 I_{bias} \cdot R_p \parallel \frac{n^2}{G_m}$$

$$= 2 I_{bias} \cdot \frac{R_p \cdot n^2 / G_m}{R_p + n^2 / G_m}$$

$$= \frac{2 I_{bias} R_p}{1 + \frac{R_p G_m}{n^2}}$$

$$\Rightarrow V_{out} \left[1 + \frac{R_p G_m}{n^2} \right] = 2 I_{bias} R_p$$

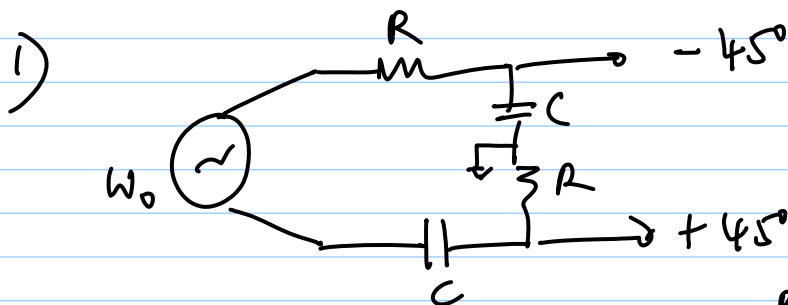
$$V_{out} \left[1 + \frac{R_p}{n^2} \cdot \frac{2 I_{bias}}{V_{out}/n} \right] = 2 I_{bias} R_p$$

$$\Rightarrow V_{out} + \frac{2 I_{bias} R_p}{n} = 2 I_{bias} R_p$$

$$V_{out} = 2 I_{bias} R_p \left(1 - \frac{1}{n} \right)$$

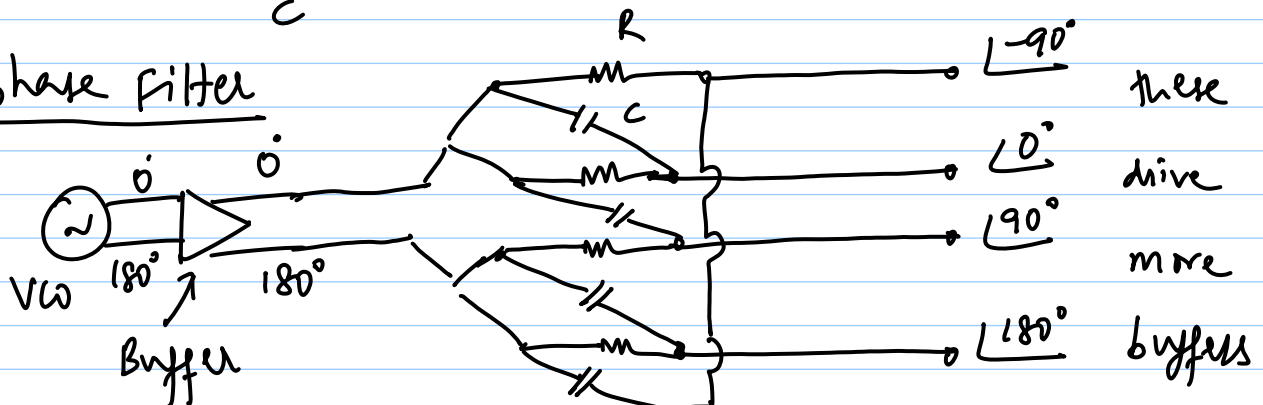
* For startup, use small-signal g_m

Quadrature Signal Generation



← we discussed this before

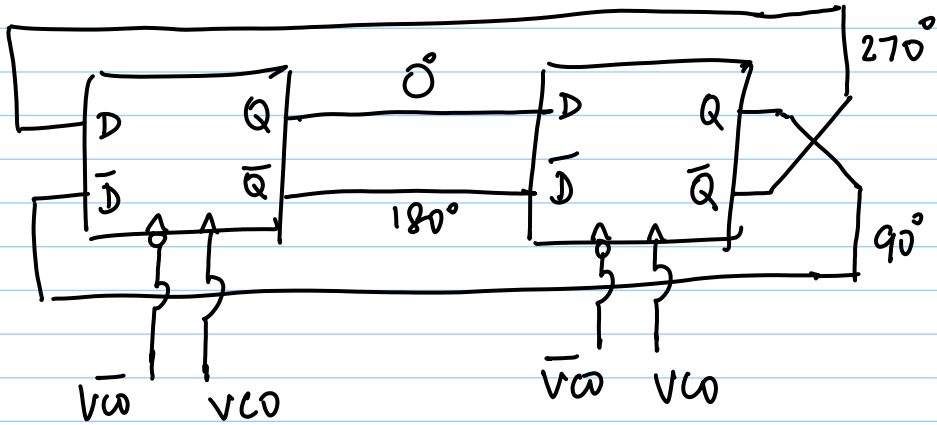
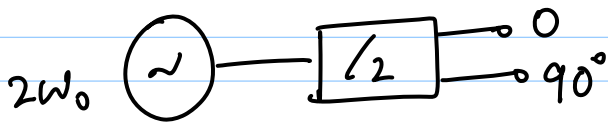
Polyphase Filter



* Buffers consume extra power

* R, C mismatches affect phase error

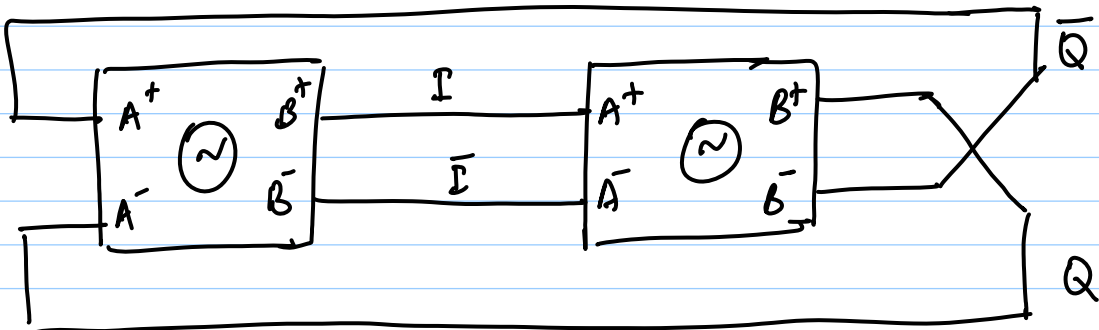
2)



- * Basically a synchronous counter (that counts to 2)
- * power consumption @ high frequencies \uparrow

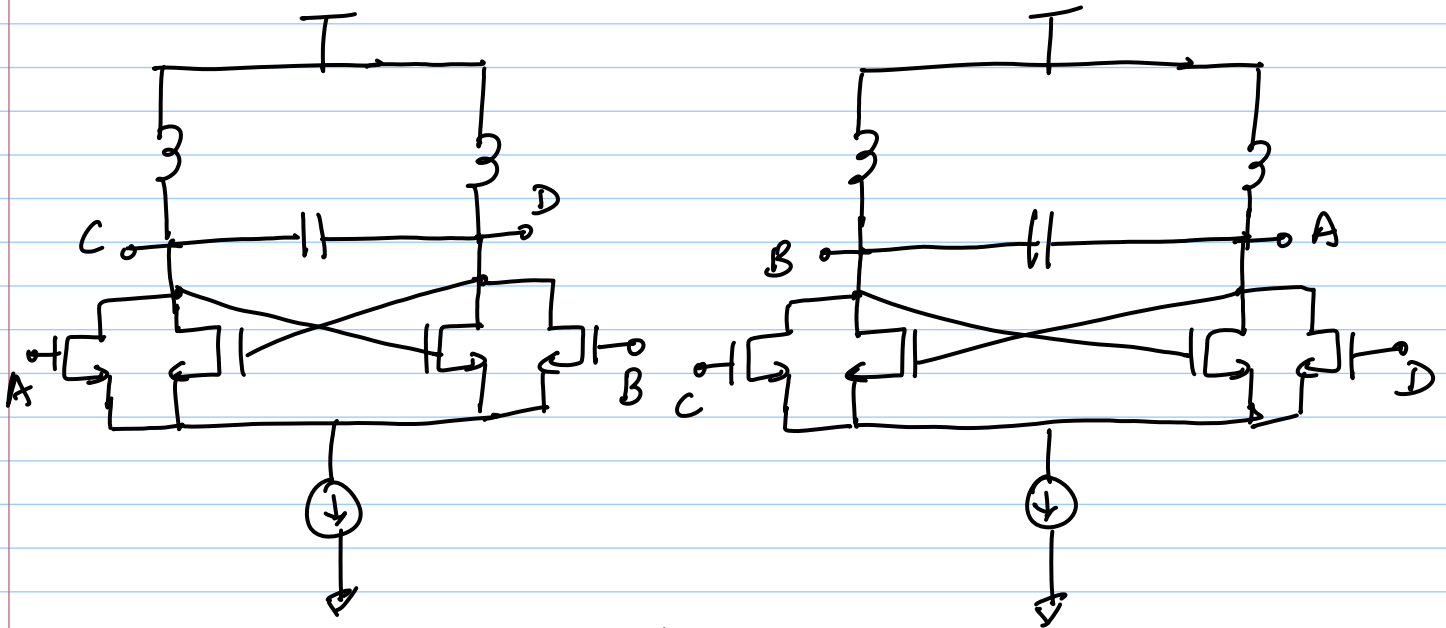
3) Quadrature VCOs

- * Couple 2 identical oscillators in quadrature



- * A - inputs (coupling)
- * B - oscillator output

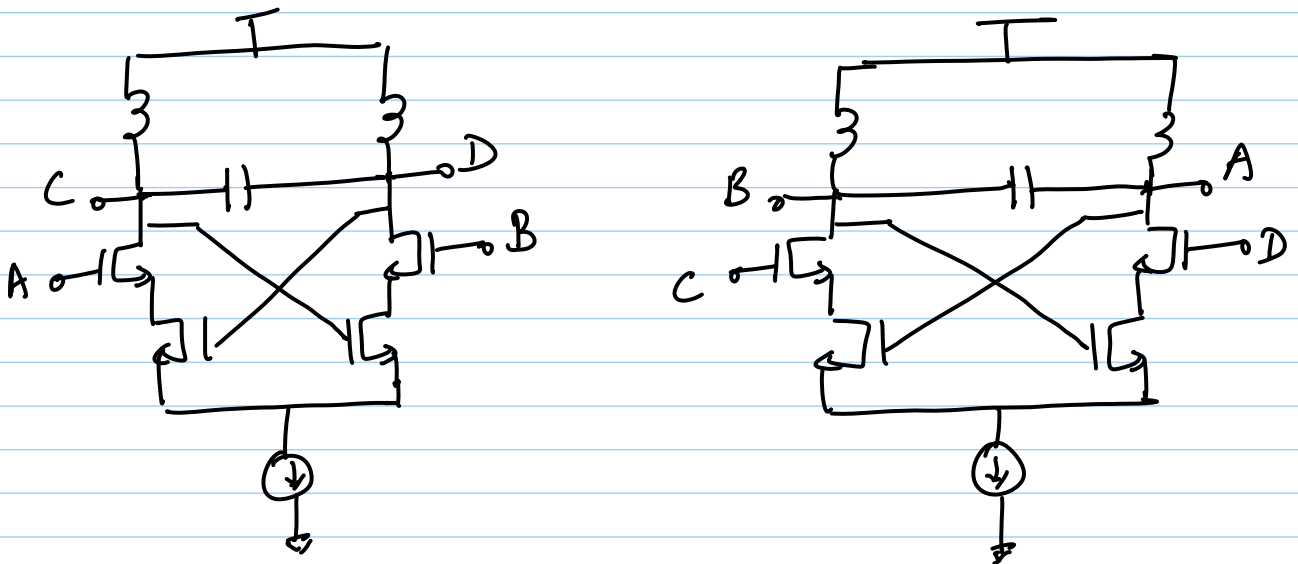
a) parallel coupling



* Exercise: Analyse the det with waveforms
→ assume outputs @ 0° & 180° phases and

see what happens (both cases will result in no oscillations - i.e. any $0^\circ/180^\circ$ components will die out)

b) Series coupling



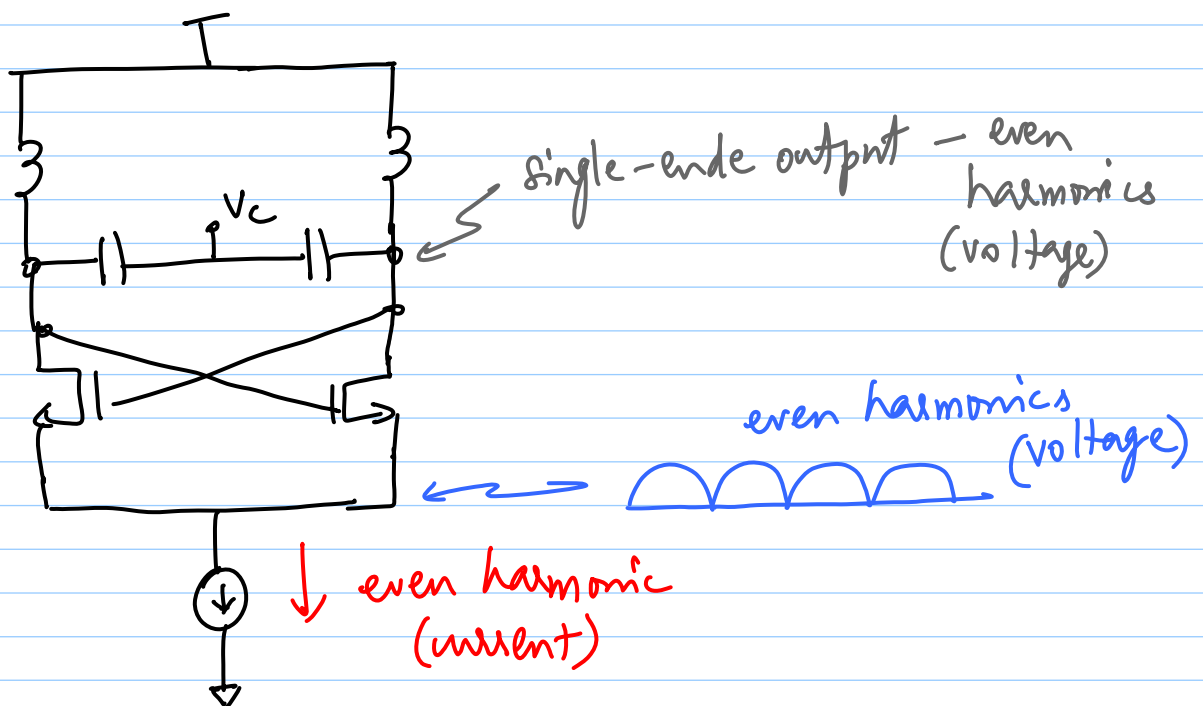
* (a) & (b) : extra parasitic cap on tank (\downarrow tuning range)

* (b) : quadrature coupling devices have to be much larger than uncoupled devices (to increase headroom for CC devices)

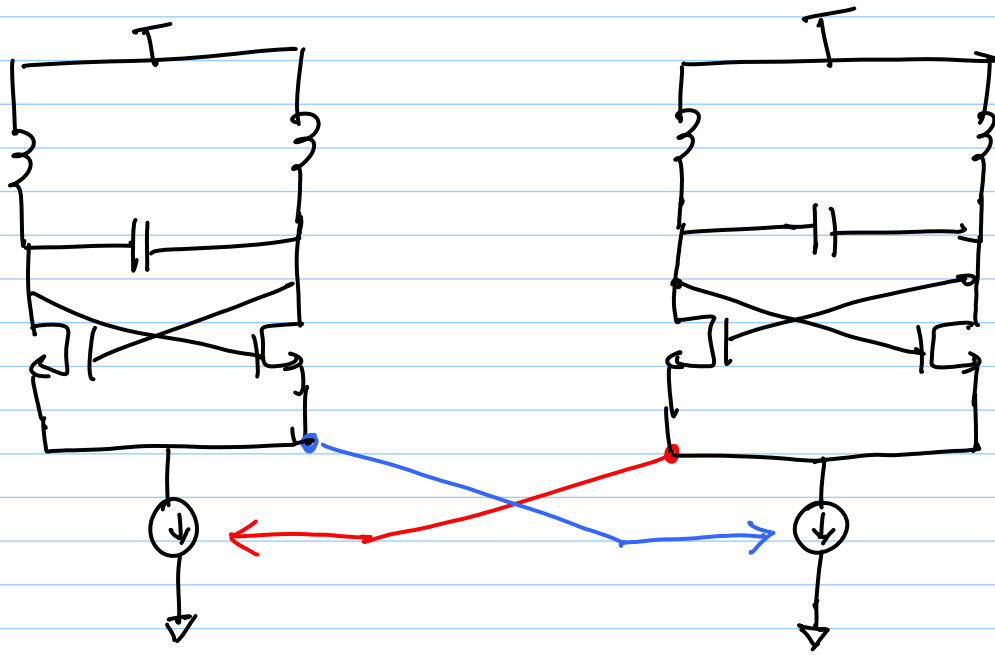
(c) Harmonic injection

* Single-ended outputs
Common-mode nodes
Common-mode paths } even harmonics exist in voltages and/or currents

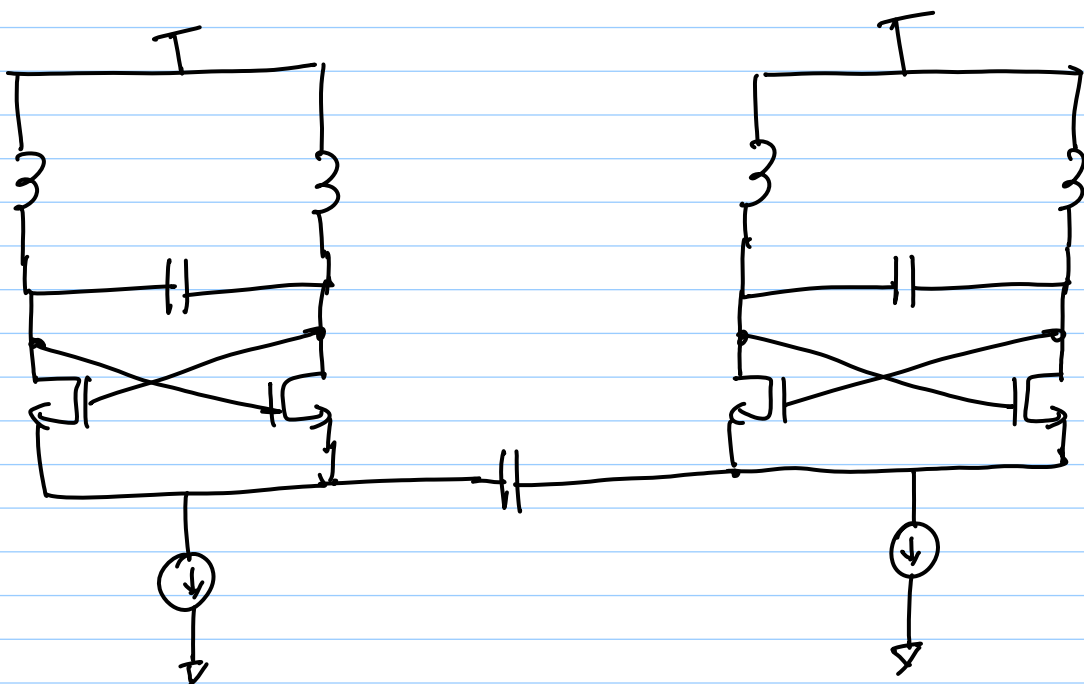
* Harmonics have specific phase relationship with fundamental freq.



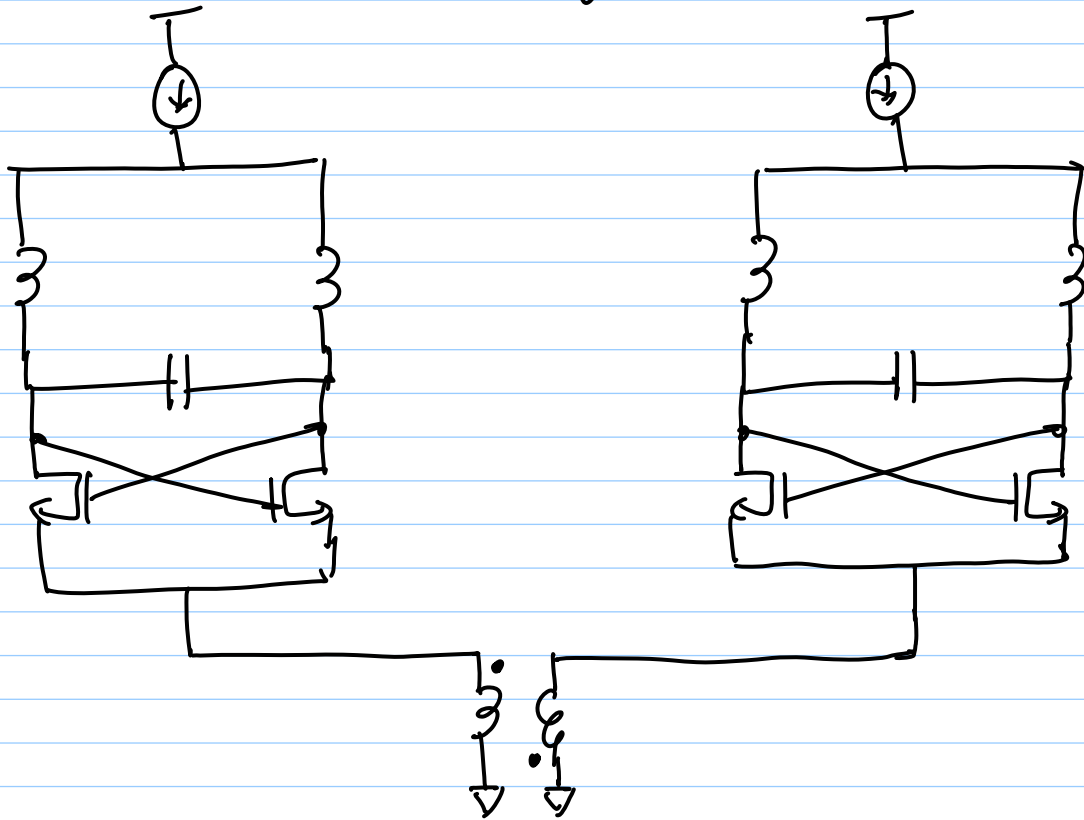
(i) C-S. coupling



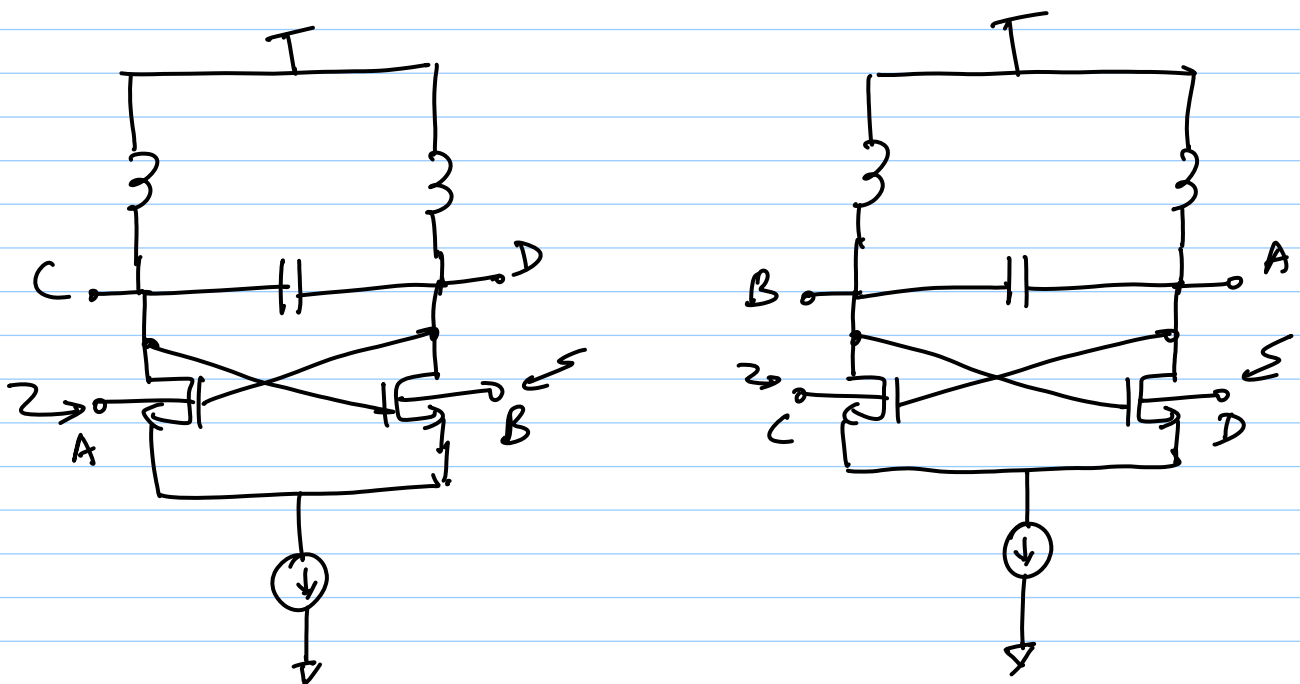
(ii) Capacitive coupling



(iii) Transformer coupling




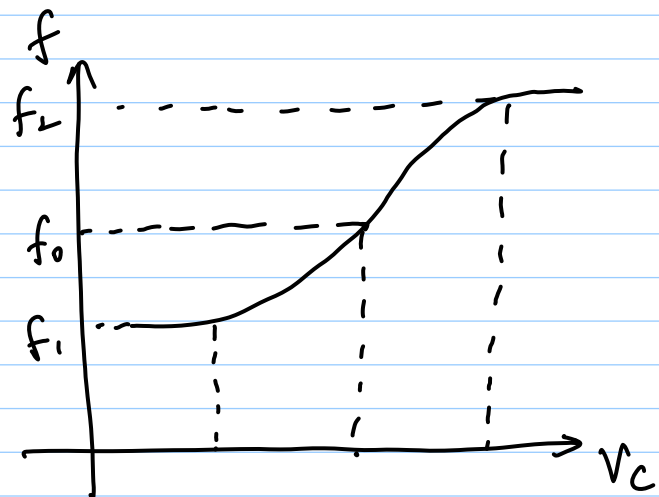
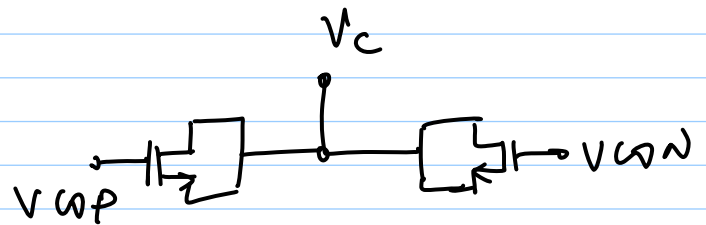
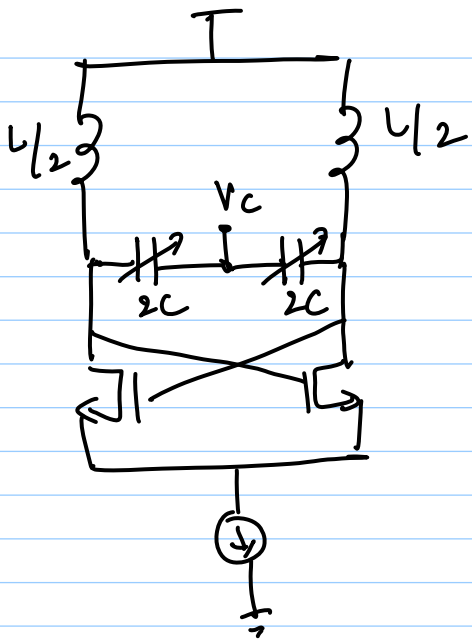
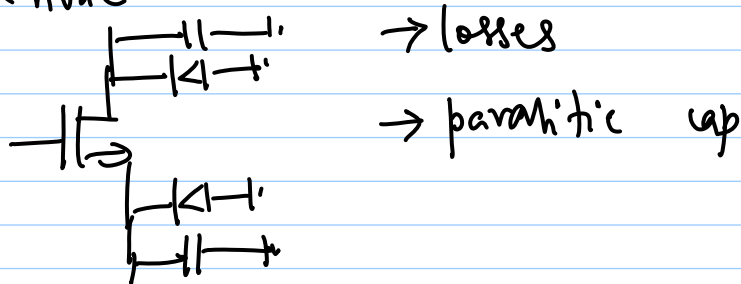
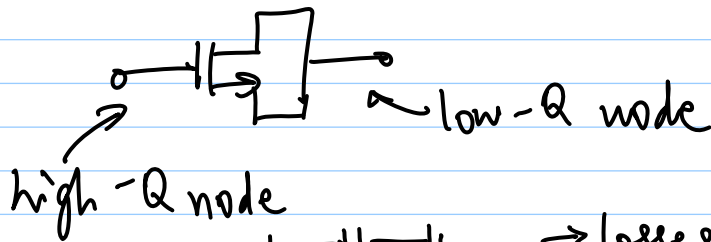
(iv) Backgate coupling



Other features of the VCO

1) tuning & tuning range

varactor is a 2-terminal device

$$\left. \begin{array}{l} \text{tuning} \\ \text{range} \end{array} \right\} = \frac{f_2 - f_1}{f_0}$$

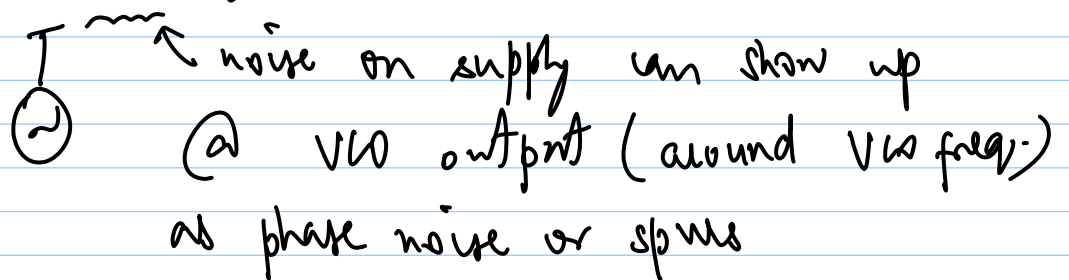
2) centre frequency $f_0 = \frac{f_1 + f_2}{2}$

3) VCO gain

$$K_{VCO} = \frac{f_2 - f_1}{\Delta V_c} \quad \left\{ \text{typically MHz/V} \right\}$$

4) Power consumption

5) Supply pushing/rejection



6) Phase noise