Chapter (1)

Introduction

1.1. The ionosphere:

Historically, the first long distance communication via electromagnetic waves is credited to G. Marconi. Marconi's experiments were started in 1894. Six years later, he successfully received a signal sent from England in St. John's, Newfoundland, Canada. Marconi believed the EM wave followed the curvature of the Earth as was also explained by Sommerfeld in 1909. There were other scientists who didnot believe that EM waves can really follow the earth's curvature for that long distance, and there must be some other reason for the success of Marconi's experiment.

In the year 1927, Edward Appleton proved the existence of the so called "Kenelly-Heaviside" layer. This is a layer in the upper atmosphere which is full of ionized gas. This layer was called later the "Ionosphere". It became widely accepted that early success of Marconi's experiment was due to the reflection of the EM waves on the Ionosphere.

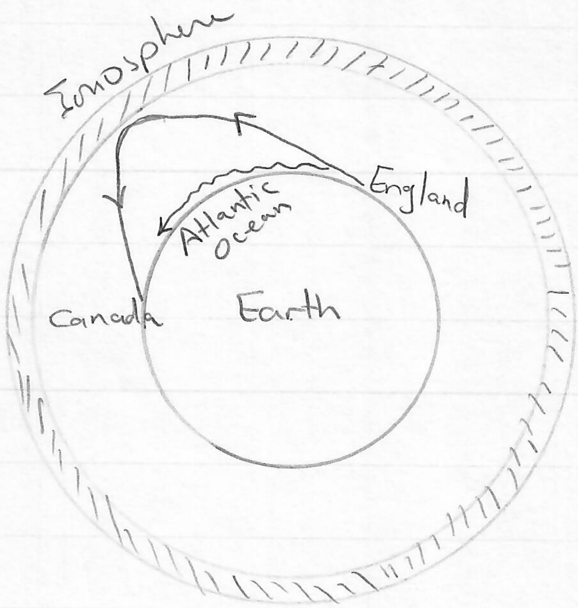


Figure: Showing the two possible explanations of Marconi's experiment.

Not to scale.

The ionosphere as we know it today extends roughly from ~80 km to ~300 km above sealevel.

The detailed characteristiss of this layer changes from day to night and from season to season.

The ionosphere reflects EM waves whose frequency is below a certain threshold called $f_{critical}$.

$$f_{critical} = 9\sqrt{N_{max}} \quad , \text{ where}$$

N_{max} is the maximum electron density of the ionosphere measured in electrons/m^3.

A known figure for this maximum density is roughly

$$N_{max} = 10^6 \text{ electron/cm}^3 = 10^{12} \text{ electron/m}^3$$

This gives a rough estimate for $f_{critical}$ as

$$f_{critical} = 9\sqrt{10^6 \times 10^6} = 9 \text{ MHz}.$$

$f_{critical}$ is calculated for EM waves incident perpendicularly on the ionosphere. Varying the angle of incidence increases $f_{critical}$ to 3~4 times the estimated figure of 9 MHz. This brings the maximum frequency to be reflected on the ionosphere to roughly 30 MHz. This is infact the upper limit of Short-wave "SW" band which we find labeled on commercial radios.

_ We learn from this short note that any EM wave to be transmitted from Earth to the outer-space past the atmosphere must have a frequency well above this mentioned limit of 30 MHz.

1.2. Frequency bands suitable for Satellite Systems:

As explained in the previous section, EM waves must have a frequency higher than the critical frequency of the ionosphere to be able to penetrate into the outer space. The question now is how high above this limit the frequency should be.

In the 1920's & till the start of world war II, Radio and Television applications were well developped and occupied the Medium Frequency (MF), high frequency (HF), Very high frequency (VHF) and the ultra high frequency (UHF) bands covering from hundreds of kilohertz up to several hundreds of megahertz.

After Wolld War II, and in the 1950's, satellite Communications started to be developed. The frequency had to be chosed at bands higher than those occupied by radio and TV applications.

There is also another reason for choosing frequencies much higher than the ionosphere critical frequency, which is the need for high directivity. Directive antennas must be large compared to the wavelength. Thus, the smaller the wavelength the more reasonable the size of antennas gets. Limited by the available technologies, the freq. of satellite systems stayed for decades between 1 and 12 GHz, that is the L, S, C, X bands.

1.3 Waveguide Technology:

In the range of frequencies higher than 1 GHz, the use of two conductor transmission lines becomes impractical due to its high conductor loss.

Single conductor transmission lines become more suitable at this range of frequencies due to its relatively low loss as compared to other transmission lines.

Furthermore, hollow (air-filled) waveguides can handle much higher power levels as compared to caxial cables and printed transmission lines.

Therefore, Waveguides are very important for satellite communications.

1.3.1 Plane wave solutions of Maxwell's Eqns in free space:

Recall: Maxwell's equations are given by:

$$(1)\ \nabla \times \overline{E} = -\frac{\partial \overline{B}}{\partial t} \qquad (3)\ \nabla \cdot \overline{D} = \rho_v$$

$$(2)\ \nabla \times \overline{H} = \frac{\partial \overline{D}}{\partial t} \qquad (4)\ \nabla \cdot \overline{B} = 0$$

Eqn. (1) & (2) give the wave equation

$$\nabla^2 \overline{E} + k^2 \overline{E} = 0, \quad k^2 = \omega^2 \mu_0 \epsilon_0 \quad \text{--- (5)}$$

One possible solution for the above equation is given by a plane wave propagating in say the z-axis direction:

$$\overline{E} = \hat{a}_x E_0 e^{j(\omega t - \beta z)} \quad \text{--- (6)}$$

where $\beta = k = \omega\sqrt{\mu_0 \epsilon_0} = \frac{\omega}{c}$ --- (7)

Note two things;

(i) Equation (6) is only one possible solution, and there are actually infinite other possibilities.

(ii) Equation (7) is called the dispersion relation or the ω-β relation.

The dispersion relation is very important because it shows the characteristics of the medium in which the wave is propagating.

For free space, it is a simple straight line whose slope gives the speed of light (EM wave) in free space

$$C = \frac{1}{\sqrt{\mu_0 \epsilon_0}}$$

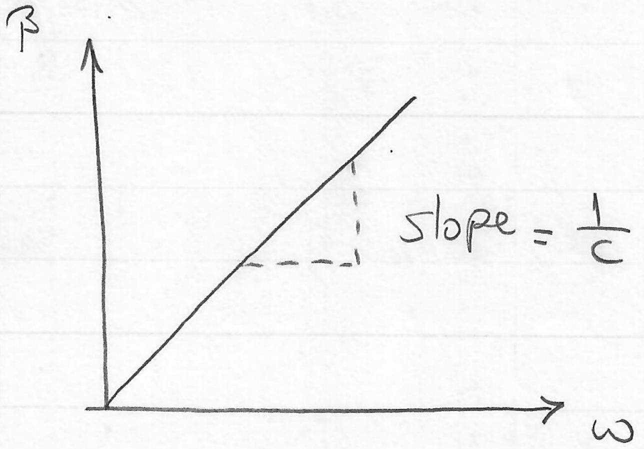
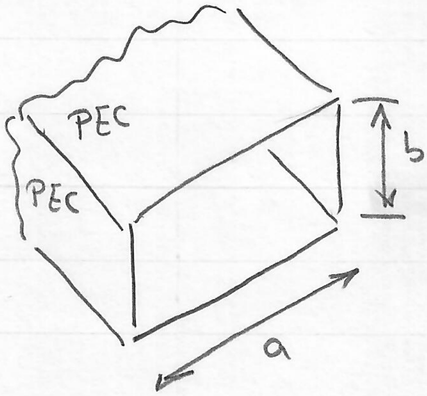
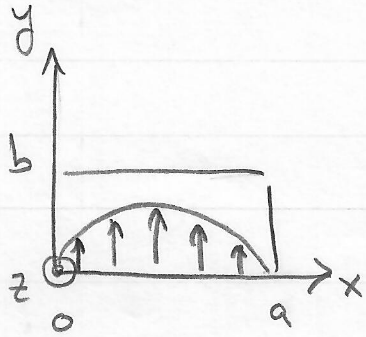


Fig.: Dispersion relation in free space

1.3.2 Rectangular Waveguide :





The solution of Maxwell's equations inside a hollow rectangular waveguide gives :

$$\nabla^2 \bar{E} + k^2 \bar{E} = 0 \quad \text{(wave eqn.)}$$

$$\Longrightarrow \quad \bar{E} = \hat{a}_y \, E_o \sin\left(\frac{\pi}{a} x\right) e^{j(\omega t - \beta z)} \quad \text{---(8)}$$

where $$k^2 = \left(\frac{\pi}{a}\right)^2 + \beta^2 = \omega^2 \mu_o \epsilon_o \quad \text{---(9)}$$

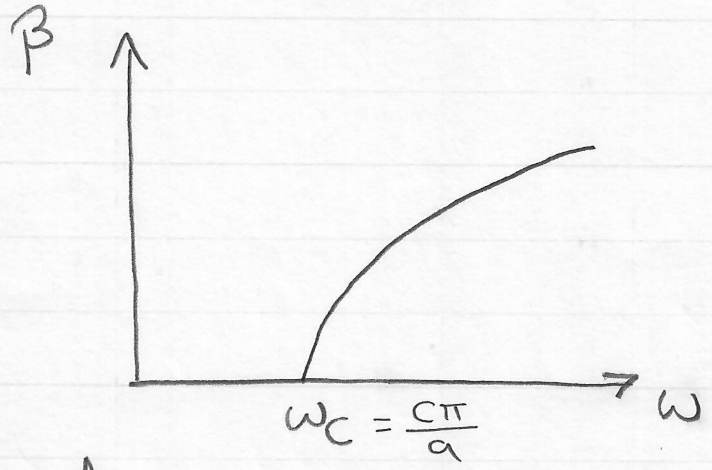
Eqn.(9) give the ω-β relation (dispersion) as:

$$\beta = \sqrt{\omega^2 \mu_o \epsilon_o - \left(\frac{\pi}{a}\right)^2} \quad \text{---(10)}$$

Note two things:

(i) Eqn(8) is only one possible solution.

(ii) The dispersion relation in (10) is not a straight line any more. It actually looks like this



Beta is real only for $\omega > \omega_C = 2\pi f_C = \frac{C\pi}{a}$

where f_C is called the cutoff frequency.

End of note.

Another possible solution is given by:

$$\bar{E} = \hat{a}_x E_0 \sin\left(\frac{\pi}{b} y\right) e^{j(\omega t - \beta z)} \quad --- (11)$$

where $$K^2 = \left(\frac{\pi}{b}\right)^2 + \beta^2 = \omega^2 \mu_0 \epsilon_0 \quad --- (12)$$

& the ω-β relation is $$\beta = \sqrt{\omega^2 \mu_0 \epsilon_0 - \left(\frac{\pi}{b}\right)^2} \quad --- (13)$$

We can also note here that Beta is real only for $\omega > \omega_c = 2\pi f_c = \frac{c\pi}{b}$

In fact, the general equation for the cutoff frequency of all possible solutions (modes) is given by:

$$f_{c_{mn}} = \frac{c}{2}\sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2} \quad ---(14)$$

where $m = 0, 1, 2, ---$

$n = 0, 1, 2, ---$

m & n cannot be simultaneously zero.

The mode with the lowest cutoff frequency is called the dominant mode

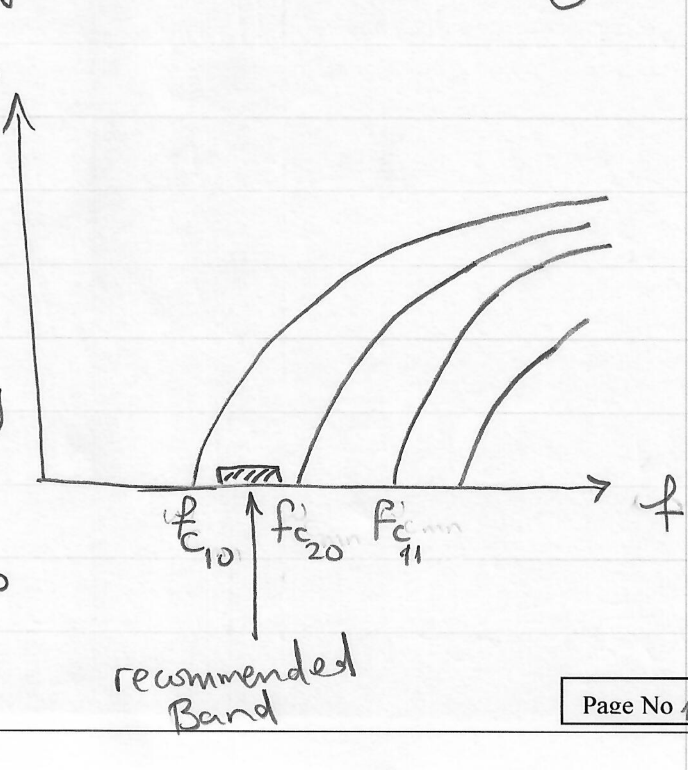
This is the first possible mode for EM wave to take inside the waveguide with $f_c = f_{c_{10}} = \frac{c}{2a}$ for $a > b$

In the industry, there is a standard figure for rectangular waveguides where $b \leq \frac{a}{2}$

This makes the second possible mode is the one with cutoff frequency $f_c = f_{c_{20}} = \frac{c}{a}$

It is always required to design waveguides which have only one possible mode in the band of operation. This is to avoid what is called intermodal distortion.

Thus if a rect. waveguide has the following dispersion relation

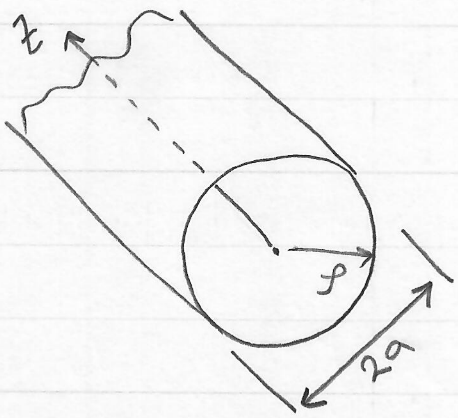


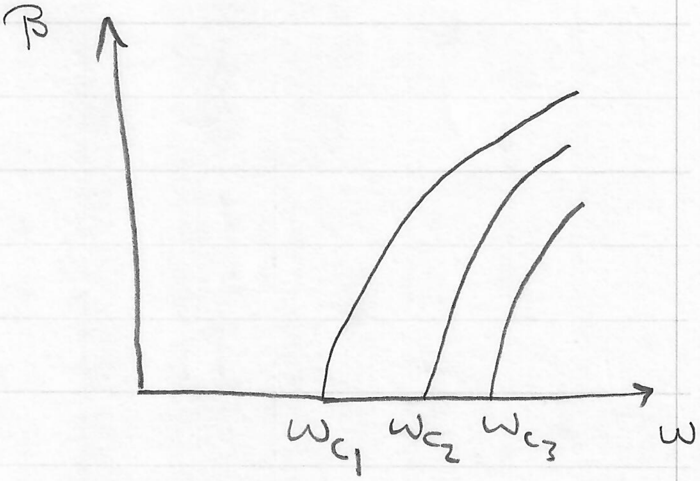
Then the recommended band of operation is roughly

$1.2\, f_{c_{10}} < f_{op} < 0.8\, f_{c_{20}}$

1.3.3. Circular Waveguides:

The waveguides discussed in the previous section can be of any cross-sectional shape as long as this shape is uniform along the length of the guide.





The circular waveguide shown in the figure above has a dispersion relation similar to the one shown to the right, where

$$\left.\begin{aligned} \omega_{c_1} &= 2\pi f_{c_1}, \quad f_{c_1} = \frac{c}{2a}\left(\frac{1.841}{\pi}\right) \\ \omega_{c_2} &= 2\pi f_{c_2}, \quad f_{c_2} = \frac{c}{2a}\left(\frac{2.405}{\pi}\right) \end{aligned}\right\} \quad (15)$$

The numbers in (15) come from Bessel functions & its derivatives

Notes:

(i) The bandwidth for single mode operation of circular waveguides is less than that of a rectangular waveguide with $a \geq 2b$.

(ii) Waveguides with homogeneous dielectric (ϵ_r) filling have lower cutoff frequencies compared to their air filled counterparts. To calculate those frequencies, we replace c by $\frac{c}{\sqrt{\epsilon_r}}$ in (14) & (15)

(iii) Modes inside metallic waveguides, made of one conductor as discussed here, are not TEM modes. In fact, they are either TE or TM modes. The dominant modes in rectangular waveguide and also in the circular ones are TE modes.

Recall: If the length of the waveguide is in the z-direction, then

* TEM modes are those who have $E_z = H_z = 0$.

those modes exist in two-conductors transmission lines but not in waveguides.

* TE modes are those who have $E_z = 0$.

The dominant mode of rect. w.g. is TE_{10} and the dominant mode of circular w.g. is TE_{11}

* TM modes are those who have $H_z = 0$.

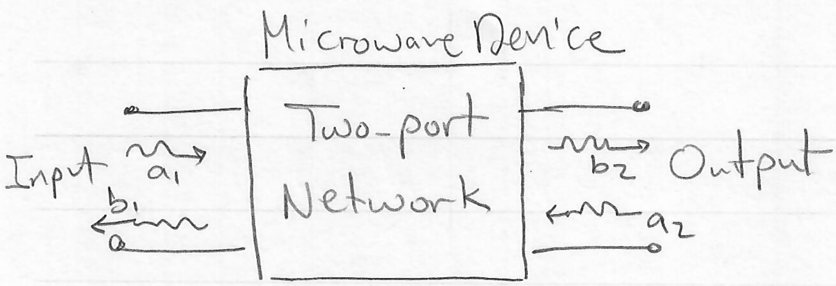
Those are the higher order modes which are usually not utilized in waveguides.

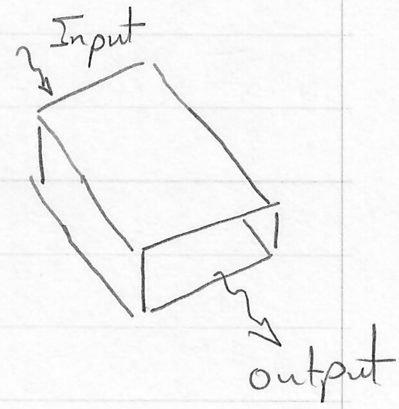
1.3.4. S-parameters:

Scattering parameters are the main tool used to analyse microwave devices.

It is clear that in waveguides, we cannot define potential as the structure has only one conductor. Instead, we define waves.

Waveguides can be considered as two-port networks





a_i & b_i are defined as the square root of the incident and reflected waves power.

$$S_{ij} = \left.\frac{b_i}{a_j}\right|_{a_i = 0} \qquad \text{--- (16)}$$

measured usually in dB

Equation (16) can be put in a matrix form as:

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} \quad --- (17)$$

The ideal loss-less waveguide has an S-matrix given by:

$$S = \begin{bmatrix} 0 & 1e^{-j\beta l} \\ 1e^{j\beta l} & 0 \end{bmatrix} \quad ---- (18)$$

$$S_{dB} = 20\log_{10}|S| = \begin{bmatrix} -\infty & 0 \\ 0 & -\infty \end{bmatrix} \quad -- (19)$$

β in (19) is given by the dispersion relation such as that given in (10).

Practically, a value of -10 dB is considered low, however this depends on the application.

Note: (i) For a reciprocal system,

$$S_{ij} = S_{ji}$$

i.e. the S-matrix should be symmetric.

This is the case for all passive devices such as the waveguide.

(ii) For loss-less systems:

$$\left.\begin{aligned} |S_{21}|^2 + |S_{11}|^2 &= 1 \\ \& \quad |S_{12}|^2 + |S_{22}|^2 &= 1 \end{aligned}\right\}$$ [due to conservation of energy]

(iii) Vector network analyzers VNAs are used in practice to measure the magnitude and phase of S-parameters of microwave devices. VNAs are usually much more expensive than oscilloscopes and spectrum analyzers.