ELECTROMAGNETIC WAVES AND ANTENNA THEORY

Nosa Bello

November 28, 2023

Contents

Preface vii

I TRANSMISSION LINE AND THE WAVE EQUATION 1

1 Introduction to Electromagnetic Waves 3
1.1 OBJECTIVES 3
1.2 Electromagnetic waves 3
1.3 Application of electromagnetic waves 4
1.3.1 Transmission Line 4
1.3.2 Waveguides 5
1.3.3 Antennas 5
1.3.4 Satellite communication 6
1.3.5 Fiber optics communication 6
1.3.6 Wireless communication 6
1.3.7 Cellular communication 6
1.3.8 Radar and remote sensing 6
1.3.9 Radio astronomy 9
1.3.10 EMI/EMC 9

2 Transmission lines 11
2.1 Objectives 11
2.2 Structures of transmission lines 11
2.2.1 Co-axial cable 11
2.2.2 Balanced Parallel wire transmission lines 11
2.2.3 Unbalanced Parallel wire transmission lines 11
2.2.4 Microstrip Line 12
2.2.5 Stripline 12
2.3 Concept of transit time effect 12
2.4 Concept of Distributed Elements in a transmission line 13
2.5 Telegrapher's Equations 14
2.6 Let's Understand the Telegrapher's Equations 14
2.7 The complex quantity γ 17
2.7.1 The Phase constant 17
2.7.2 The Attenuation constant 17
2.8 The Characteristic Impedance 20
2.9 Boundary Conditions 20
2.10 The Reflection Coefficient 21

3 MAXIMUM POWER TRANSFER AND MATCHED CONDITION 23
3.1 Objectives 23
3.2 Matched Condition of the Transmission Line 23
3.3 Impedance Transformation Relationship 24
3.4 Loss-Less Transmission Line 26
3.4.1 Propagation Constant of Loss-Less Transmission Line 26
3.4.2 Characterisitic Impedance Loss-Less Transmission Line 26
3.5 Low-loss Transmission Line 27
3.5.1 Propagation Constant of Low-loss Transmission Line 27
3.5.2 Characterisitic Impedance of Low-loss Transmission Line 27

4 VOLTAGE AND CURRENT VARIATION IN LOW-LOSS TRANSMISSION LINE **29**
4.1 Objectives . . . 29
4.2 Voltage And Current Variation on a Lossless Transmission Line . . . 30
4.3 Concept of Voltage Standing Wave Ratio (VSWR) . . . 32
4.4 Condition for Full Reflection at Load End . . . 33
4.4.1 Conditions for which $|\Gamma_L|= 1$ . . . 33
4.5 Maximum and Minimum Resistance . . . 34
4.6 Impedance Transformation on a Lossless Transmission Line . . . 36
4.6.1 Normalized impedance value repeat every $\frac{\lambda}{2}$ distance . . . 36
4.6.2 Normalized impedance inverts at every $\frac{\lambda}{4}$ distance . . . 37
4.6.3 The matching condition characteristics . . . 37
4.7 Power Transfer on Transmission Line . . . 37
4.8 Evaluation of $\mathbf{V}^+$ . . . 40
4.9 Introduction to Smith chart . . . 42
4.9.1 **Objectives** . . . 42
4.9.2 Plane Transformation . . . 42
4.9.3 Transformation analysis for complex plane . . . 43
4.9.4 Circle of Constant Resistance . . . 44
4.9.5 Circle of Constant Reactance . . . 45
4.10 . . . 46
4.11 Concept of constant VSWR circle in lossless transmission line analysis . . . 48
4.12 Transmission Line Analysis using Admittance . . . 48
4.13 Use of Smith chart for Transmission Line Calculation . . . 50
4.13.1 . . . 51
4.13.2 Voltage Standing Wave Ratio (VSWR) . . . 52
4.13.3 Location of . . . 52
4.13.4 Identifying from Standing Wave Pattern . . . 56
4.14 Applications of transmission lines . . . 59
4.14.1 Measurement of . . . 59
4.14.2 As a Circuit Element . . . 61
4.14.3 As a Resonant circuit . . . 64
4.14.4 As a Step Up Transformer . . . 69
4.14.5 Matching Of Impedances . . . 71
4.15 Lossy Transmission Line . . . 77
4.15.1 Very Lossy Transmission Line . . . 77
4.15.2 Moderately Lossy Transmission Line . . . 77
4.15.3 Characteristic Impedance and Propagation Constant measurement . . . 78
4.16 Various Types of Transmission Lines . . . 82
4.16.1 Coaxial line . . . 82
4.16.2 Parallel wire transmission line (balanced) . . . 82
4.16.3 Microstrip line structure . . . 83
4.17 Conclusion . . . 84
4.18 Exercises . . . 85

5 Coordinate System and Vectors **87**
5.1 Objectives . . . 87
5.2 Coordinate systems . . . 87
5.2.1 Cartesian coordinate system . . . 87
5.2.2 Cylindrical coordinate system . . . 88
5.2.3 Spherical coordinate system . . . 88
5.3 Vector analysis . . . 89
5.3.1 Basic operation of vectors . . . 89
5.3.2 Scalar product or dot product . . . 90
5.3.3 Vector or cross product . . . 90
5.4 Exercises . . . 91
5.5 Divergence and Stokes Theorem . . . 92
5.5.1 Divergence or Curl Producing Vector Fields . . . 92
5.5.2 Laplacian operator . . . 92
5.5.3 Integral Operators . . . 93
5.5.4 Line Integral . . . 93
5.5.5 Surface Integral for Infinitesimally small Area . . . 93
5.5.6 Volume Integral . . . 94
5.5.7 Divergence Theorem . . . 94

5.5.8 Stokes Theorem . . . 94
5.5.9 Summary . . . 95
5.6 EXERCISES . . . 96

6 The Basic Laws To Maxwell's Equation (1) 97
6.1 OBJECTIVES . . . 97
6.2 INTRODUCTION . . . 97
6.3 Magnetic Flux Density $\bar{B}$ . . . 97
6.4 Ohm's Law . . . 98
6.5 The Physical Laws To Maxwell's Equation . . . 98
6.6 Gauss Law . . . 99
6.7 Amperes Law . . . 100
6.8 Faraday's Law of Electromagnetic Induction . . . 101
6.9 Summary . . . 102
6.10 EXERCISE . . . 103

7 The Basic Laws To Maxwell's Equation (2) 105
7.1 OBJECTIVES . . . 105
7.2 INTRODUCTION . . . 105
7.3 Ampere Circuit Law Modification and Maxwell's Equations . . . 105
7.4 Concept of Surface Charges and Surface Current . . . 107
7.4.1 Surface Charge and Surface Charge Density . . . 107
7.4.2 Surface Current and Surface Current Density . . . 108
7.5 Problems . . . 108
7.6 EXERCISE . . . 110

8 Medium of Finite Conductivity 111
8.1 Objectives . . . 111
8.2 Maxwell's equation in relation to Ampere's Law . . . 111
8.3 Complex Dielectric Constant . . . 112
8.4 Propagation of Electromagnetic Waves . . . 113
8.4.1 Wave equation . . . 113
8.5 Loss Tangent . . . 114
8.6 ExerciseList . . . 115

9 Wave Propagation in a Medium 117
9.1 Objective . . . 117
9.2 **Effective Depth** . . . 118
9.3 **Phase Velocity** . . . 120
9.4 **The Case of a Pure Dielectric** . . . 120
9.5 ExerciseList . . . 121

10 Power Flow in an Electromagnetic Wave 123
10.1 Objective . . . 123
10.1.1 INTRODUCTION . . . 123
10.2 Poynting theorem and Poynting vector . . . 124
10.3 Uniform Plane wave . . . 126
10.4 ExerciseList . . . 127

11 Surface current and power loss in a conductor 129
11.1 Objective . . . 129
11.1.1 . . . 129
11.2 ExerciseList . . . 133

12 Uniform Plane Wave in an arbitrary direction 135
12.1 Objective . . . 135
12.2 . . . 135
12.3 Wave in arbitrary direction . . . 135
12.4 ExerciseList . . . 138

13 Wave propagating in a bound medium **139**
13.1 Learning Outcomes: 139
13.2 Introduction 139
13.3 Plane wave at media interface 139
13.4 Incident wave at a media interface 139

14 Total Internal Reflection **145**
14.1 Learning Outcomes: 145
14.2 Introduction 145
14.3 Snell's Law 145
14.4 Conditions For Total Internal Reflection 146
14.5 Reactive Fields 147
14.6 Wave Guides 148

15 Polarization across Dielectric Interface **151**
15.1 Learning Outcomes: 151
15.2 Introduction 151
15.3 Decomposition of Polarized wave 151
15.3.1 Incident wave 151
15.3.2 Reflected Wave 151
15.3.3 Transmitted Wave 151
15.3.4 Combination of $\mathbf{E}_i$, $\mathbf{E}_r$ and $\mathbf{E}_t$ 152
15.4 Linearly Polarized Wave 152
15.4.1 Normal Reflection 152
15.4.2 Total Internal Reflection 152
15.5 Circular Polarization 152
15.5.1 Normal Reflection 152
15.5.2 Total Internal Reflection 152
15.6 Brewster Angle 153
15.6.1 Parallel Polarization 153
15.6.2 Perpendicular Polarization 154

16 Reflection from a Conducting Boundary **157**
16.1 Learning Outcomes: 157
16.2 Introduction 157
16.3 Perpendicular Polarization 158
16.3.1 Electric Field 158
16.3.2 Magnetic Field 158
16.4 Electric fields in the dielectric medium 159
16.5 Magnetic fields in the dielectric medium 159

Preface

Electromagnetism and Antenna Theory is a book developed out of dedication and passion for comprehensive and curated concepts in transmission lines, wave equations and the foundations of antenna theory. The book is a compilation and review of lectures for the courses *Telecommunication Principle I* and *Microwave Engineering* taught by Dr. Nosa Bello at the University of Benin, Nigeria. The book is a work in progress and will be updated as the course progresses. The book is divided into two parts. Part one covers transmission lines and the wave equation while part two covers waveguides and antenna theory. The book is written in LaTeX and the source code is available on GitHub.

Part I

TRANSMISSION LINE AND THE WAVE EQUATION

Chapter 1

Introduction to Electromagnetic Waves

Figure 1.1: Overhead transmission line

1.1 OBJECTIVES

The objective of the chapter is to discuss the following

(i) Antennas and the different types of antennas their characteristics, such as radiation patterns, gain, and impedance.

(ii) Understand the principles and applications of transmission lines, including twisted pairs, coaxial cables, and waveguides.

(iii) Comprehend the advantages and disadvantages of each type of transmission line in terms of signal quality, bandwidth, and attenuation.

(iv) Gain knowledge about the practical considerations for selecting and installing antennas and transmission lines in various communication systems.

(v) Discuss Cellular communication and how it is related to transmission line.

(vi) Understand the principle and application of cellular communication, radar and remote sensing radio astronomy and EMI/EMC.

(vii) Develop practical skills on analyzing and trouble shooting these system.

(viii) Learn about Azimuth resolution and what it is defined by.

1.2 Electromagnetic waves

The concept of electromagnetic waves has fascinated man for so many years that man has asked so many questions varying from *"Why do stars twinkle and plant do not?"*, to *"Why do magnetic needles deflect?"*, and *"how does light travel from the sun when there is no medium between?"*

In modern-day, questions vary from *"how do we have tv reception?"* to *"how do we have radio stations operating"* to *"how does a mobile phone work?"* to *"why are certain things heated when they are kept inside a microwave?"*. All these phenomena revolve around the concept of electromagnetic waves.[1]

Electromagnetic waves can be divided into;

(i) low frequency and high power

(ii) high frequency and low power

In this part of the book, we shall concentrate on the high-frequency and low-power properties of electromagnetic waves. Devices and phenomena like electrical machines, electrical power generators, transformers, and distribution of electrical energy fall into the category of low-frequency high power. Whereas modern systems, like mobile communication, radars, satellite, and optical fibres, fall into the high-frequency low-power category. So we are mainly going to investigate what happens as frequency increases in electromagnetic waves[2] and how they can be transmitted from one position to another without loss.

Electromagnetic waves see applications in many areas namely[3]:

(i) Transmission lines and HF circuits.

(ii) Antennas.

(iii) Satellite communication.

(iv) Fibre-optic communication.

(v) Radars.

(vi) Radio astronomy.

(vii) Electromagnetic Interference/Compatibility (EMI/EMC)

We intend to investigate the behaviour of time-varying electric and magnetic fields especially when the frequency of operation is large. This investigation is going to be based on Maxwell's four equations of electromagnetic waves[4]. However, as we proceed certain approximations can be used to investigate

[1] The concept of electromagnetic waves is virtually applied in almost all advanced technology.

[2] When the frequency of a wave increases, the wavelength will decrease to compensate for this increment

[3] The applications of electromagnetic waves are not limited to this list but for this course, we streamline the application to these few.

[4] see chapter 6 for a detailed explanation of Maxwell's equations

the same phenomena in terms of voltage and current which are electrical circuits.

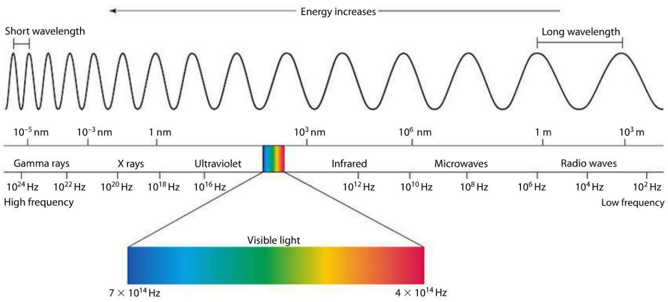


Figure 1.2: The electromagnetic spectrum

Electromagnetic spectrum: All electromagnetic radiations in the universe ranging from very low frequencies to very high frequencies can be referred to as the electromagnetic spectrum. This is shown in figure 1.2

1.3 Application of electromagnetic waves

Depending on the frequency of operation, there are different media used to transmit electromagnetic waves. When the frequency is between 30MHz to 300MHz, the coaxial cable is used for transmission of the wave, from 30GHz to 300GHz the waveguide structure is used and as the frequency goes higher, the media used is the optical fibre.

A question that comes to mind in electromagnetic wave transmission is *"Why do we have to increase frequency?"* To answer that we would have to think of the major application of high frequency which is in communication. Then for transmitting more information, we require large bandwidth. Since the frequency of operation is proportional to the bandwidth, by increasing the frequency of operation we increase the bandwidth and as such one can transmit more information on a given channel.

1.3.1 Transmission Line

Figure 1.3: Power transmission line

In transmission lines, our main concern is how voltage and current would flow in a two-conductor system called a *transmission line* and how losses occur during transmission These lines maintain the integrity and efficiency of the signal by managing impedance, minimizing reflections and controlling the transfer of power along a line. Looking into losses in transmission lines; they arise from resistance in conductors, dielectric losses in insulating materials, radiation losses, skin effect, and proximity effect. They reduce the efficiency of energy transfer in the line emphasizing the need for careful design and material selection to minimize these losses.[5].

There are different types of transmission media and the application of this media is based on the frequency range of the signal.

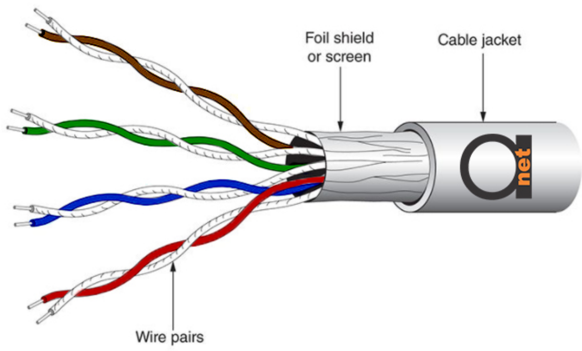


Figure 1.4: Twisted pair of wires

Twisted pairs

They are point-to-point transmission lines[6] (it is a balanced transmission line having voltages v^+ and v^- connected at its terminals). An example of the twisted pairs is the telephone wires; it is characterized by low data rate, high Electromagnetic Interference (EMI) and is lossy at radio frequencies.

Limitations of Twisted pairs

Limited bandwidths: suitable for moderate distances and bandwidths. However, at higher frequencies, they experience more signal degradation and are limited in the amount of data they can effectively transmit.
Susceptibility to interference: Since they lack strong shielding, they are more prone to electromagnetic interference from near by cables or other electrical devices
Attenuation: Higher frequencies suffer more attenuation in twisted pairs, limiting their use for high speed and long distance transmissions.

Coaxial cable

For the coaxial cable, we have an example in the LAN (Local Area Network) cable; it characterizes data rates of up to a few Mbps, low EMI and moderate loss.

Limitations of Coaxial cables
1. Expensive to install for longer distances because of its thickness and stiffness
2. The fault point is difficult to find and it is inconvenient to deal with the accident in time.

[5]When the frequency of a wave increases its wavelength decreases and when the wavelength of the wave is comparable to the length of the transmitting media (length of the wire) the losses along the wire become too significant to ignore.

[6]Connections between two nodes or endpoints

Figure 1.5: Coaxial cable

1.3.2 Waveguides

These are hollow circular or rectangular pipes and they are used when the frequency becomes high in other to reduce losses.

Inside the hollow metal conductor, an electromagnetic wave can propagate and here rigorous analysis of electromagnetic wave propagation is carried out to help find out what the field distribution would look like and how much energy loss will take place inside.

Limitations of waveguides
1. It is very bulky in size and weight
2. It is not very economical

Figure 1.6: Waveguide

1.3.3 Antennas

Figure 1.7: Parabolic dish antenna

An antenna is a device that can transmit electromagnetic energy into space and also can receive electromagnetic energy coming from space. An example is the parabolic dish antenna (see figure 1.7), signals coming from a point are like parallel rays. When it gets to the parabolic dish antenna the rays converge to a focal point called the feed from there they get processed into an electrical signal.

An antenna is a device that separately puts radiation in the desired direction. A simple antenna structure may not necessarily provide the desired characteristics of modern-day smart antenna systems where radiation characteristics can be automatically changed to maximize the reception of the signal.

Other more advanced antenna systems referred to as **smart antennas** can selectively steer the radiation in the desired direction. There are two types of smart antennas which are the **adaptive array** and **switched beam systems**.

Adaptive array

Adaptive arrays also known as adaptive antenna arrays, are a type of antenna system that can dynamically adjust their radiation pattern response to changing signal conditions. They use advanced signal processing techniques to optimize signal reception and transmission, improving the overall performance of wireless communication systems.
They steer the beam in the direction of the user/observer while simultaneously nulling interfering signals. It's like having a smart signal that can adapt to its surroundings.(see figure 1.8).

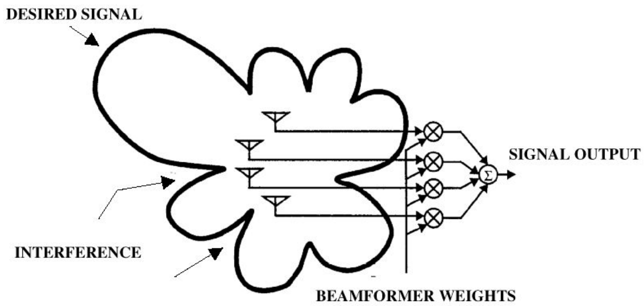


Figure 1.8: Adaptive array system

Limitations to adaptive array
1. Sensitivity to number of directional noise
2. Decorrelation of the jamming signal in the main channel with respect to the same signal in an auxiliary channel and multipath effects.

Switched beam antenna

It has multiple beams (see figure 1.9) which can be switched depending on the area an observer is situated so that a signal can be transmitted or received from that zone. Therefore, depending on the requirement of the system a beam is selected hence the name switched beam.

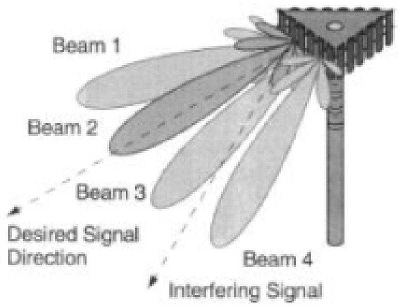


Figure 1.9: Illustration of a switched beam antenna

Limitations to switched bean antenna
1. They do not provide the precise beam steering capability of a phased array.

1.3.4 Satellite communication

A satellite is an object that is placed above the earth's surface. There is a satellite station on the earth's surface usually called the earth station which transmits signals from the earth to satellites and also receives signals from satellites to the earth. With satellite communication there are certain frequency bands are assigned. Satellite communication is a point-to-multi-point system.

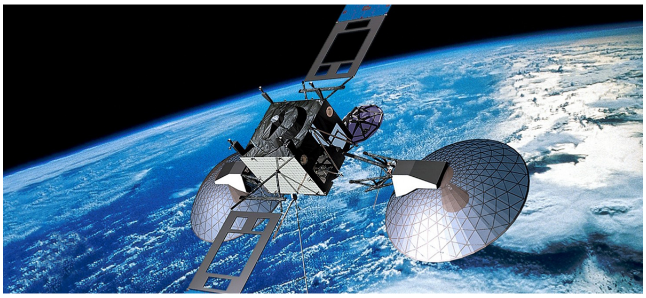


Figure 1.10: Space satellite

The whole propagation of the electromagnetic wave and proper placing of radiation in the direction towards the earth is achieved using the principles of electromagnetic waves. Satellite communication has a large time delay because of the turnaround trip time taken from the Earth station to the satellite and from the satellite back to the Earth station.

1.3.5 Fiber optics communication

Figure 1.11: Optical fiber

Knowledge of electromagnetic waves is required to investigate the propagation of light inside a fibre optic material. A fibre optic cable (see figure 1.11) is made up of very thin hollow glass strands in which light is reflected internally. As the light propagates inside the optical fibre, the signal gets distorted and hence the knowledge of electromagnetic waves is needed to know how the signal gets distorted.

1.3.6 Wireless communication

Wireless communication is used in cell phones and most modern systems like home entertainment systems using Bluetooth speakers, laptops using a wireless method to connect to routers and so on. All these work on the principle of electromagnetic waves.

Figure 1.12: Wireless connection of devices

1.3.7 Cellular communication

In cellular communication, we have the base station from where the signals are transmitted, all users located inside are called a cell (see figure 1.13). Any mobile call made goes from the handset to the base station in that sector and then goes to the desired handset within the same sector or in another sector.

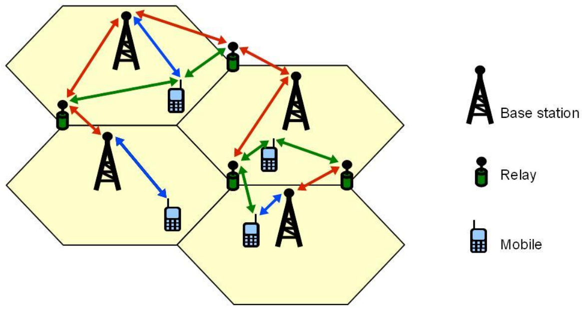


Figure 1.13: Cellular communication

It is observed that a cell phone user has signals from multiple paths all coming to his cell phone[7]. The mobile entity does not only get signals coming directly from the base station, but it also encounters signals due to reflection from associated objects. As a result of interference of these signals which can either be constructive or destructive, the final signal transmitted/received is altered (see figure 1.14).

When there is constructive interference, a strong signal strength is observed and when the interference is destructive, a weak signal strength is observed. The gradual weakening of a signal due to destructive interference is called *fading*. To understand fading phenomenon a good knowledge of electromagnetic waves is required. To avoid fading phenomena, antennas in mobile phones use *sectioned radiation patterns*. This implies that the cell phone gets its signal from the base station directly while minimizing signals from paths not in the direction of the base station of the antenna. A good knowledge of electromagnetic waves is required to design such antennas.

1.3.8 Radar and remote sensing

In radar altimetry, electromagnetic waves are used for finding the distance of an object from its position in space to the antenna. The antenna is excited with an electromagnetic pulse and the dish beams parallel waves downwards. This beam is

[7]Wireless relays are used to transmit and receive information between the base station and mobile when they are too far away to send the information to each other directly

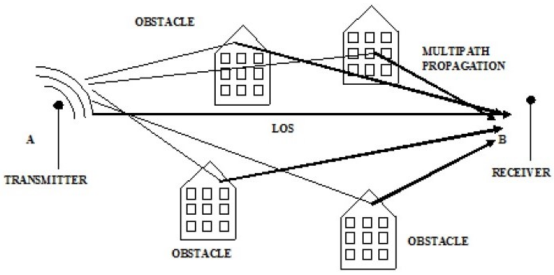


Figure 1.14: Illustration of signal inference

reflected from the earth's surface/target object back to the dish. The dish converges the received signal back to the antenna and processes it in the detector (see figure 1.15).

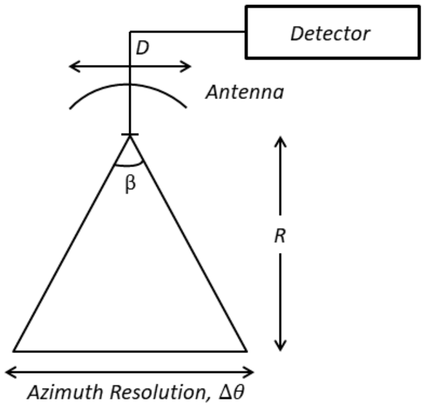


Figure 1.15: Illustration of Azimuth resolution

The *Azimuth (or bearing) resolution* of the radar is given as $\Delta\theta = \beta R = \frac{\lambda R}{D}$[8], where β is the beam width of the antenna given as $\dfrac{\lambda}{D}$, λ is the wavelength of the beam, R is the range of distance covered by the beam and D is the aperture of the antenna. It is the measure of the ability of an imaging radar to separate two closely spaced scatterers in the direction parallel to the motion of the beam (resolution at the same range but different bearings). Similarly, the measure of the ability of a radar to correctly resolve the position of two closely spaced objects in the range is called the *range resolution* which makes up the second category of resolution in radar systems. It is given as $\geq \dfrac{cT}{2}$ where c is the speed of light, and T is the transmission pulse width.

With radar systems, the distance travelled by the beam can be estimated and also if the object is moving in the radial direction, there would be a frequency change between the signal transmitted by the antenna and the signal reflected (Doppler shift) and from this result the velocity of the object can be estimated. The radar essentially uses the electromagnetic pulse to find the distance and the velocity of an object.

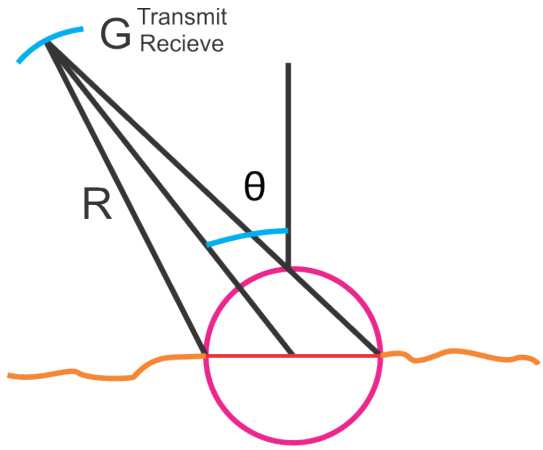


Figure 1.16: Illustration of the monostatic radar equation

The monostatic radar equation

The monostatic radar equation is given as

$$P_r = \frac{P_t G \sigma A}{(4\pi)^2 R^2}$$

P_r is the received power, P_t is the transmitted power, G is the antenna gain, σ is the radar cross section of the target, A is the effective area of the antenna and R is the range(see figure 1.3.8). The signal goes from the radar transmitting antenna to the object and is again received by the receiving antenna in the radar. The magnitude of the received signal can be calculated and this requires a good knowledge of the ability to model the propagating environment and a good modelling of the scatterer from which the energy is going to be reflected.

Figure 1.17: Radar locator

Figure 1.17 depicts a radar locator. We can see the various objects detected by the radar (represented by the white dots) and the green radial line represents the transmitting beam

Side looking airborne radar (SLAR)

It is a radar technique used for remote sensing. It is an imaging radar[9] mounted on a moving object like an aircraft, pointing perpendicular to the direction of flight (hence side-looking). A squinted (non-perpendicular) mode is possible also. SLAR can be fitted with a standard antenna (real aperture radar) or an antenna using synthetic aperture(this would be discussed in the next section). The platform of the radar moves in the direction of the x-axis. The radar looks with the looking angle θ (or so-called off-nadir angle).

[8]The implication of this equation lies in the size of the antenna. To get a high resolution the size of the antenna would be large

[9]Imaging radar provides its light to illuminate an area on the ground and takes the picture at radio wavelengths

The microwave beam is transmitted obliquely at right angles to the direction of flight illuminating a swath (see figure 1.18). *Swath width* refers to the strip of the earth's surface from which data are collected by a side-looking airborne radar. It is the width of the imaged scene in the range dimension. The longitudinal extent of the swath is defined by the motion of the aircraft with respect to the surface, whereas the swath width is measured perpendicularly to the longitudinal extent of the swath. Range refers to the across-track dimension perpendicular to the flight direction, while azimuth refers to the along-track dimension parallel to the flight direction[10].

To measure the azimuth resolution

The SLAR is primarily a real aperture radar. This requires a reasonably large antenna for adequate angular resolution. The azimuth resolution, Ra, is defined as

$$R_a = \frac{H\lambda}{Dcos\theta}$$

H is the height of the antenna(height of the airplane)
D is the geometric length of the antenna,
λ is the wavelength of the transmitted pulses, and
θ is the incidence angle

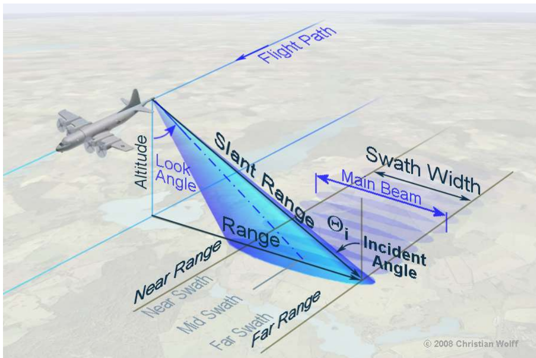


Figure 1.18: Diagram of the SLAR

The equation shows, that increasing altitude decreases the azimuthal resolution of SLAR. A very long antenna (i.e., large D)[11] would be required to achieve a good resolution from the aircraft. Synthetic Aperture Radar (SAR) is used to acquire higher resolution.

To measure the cross-track resolution

At all ranges, the radar antenna measures the radial line of sight distance between the radar and each target on the surface. This is the slant range distance. The ground range distance is the true horizontal distance along the ground corresponding to each point measured in the slant range. The cross-track resolution, R_r, is defined as:

$$R_r = \frac{c_0 t_p}{2sin\theta}$$

c_0 is the speed of light
t_p is the pulse duration of the transmitter and
θ is the incidence angle

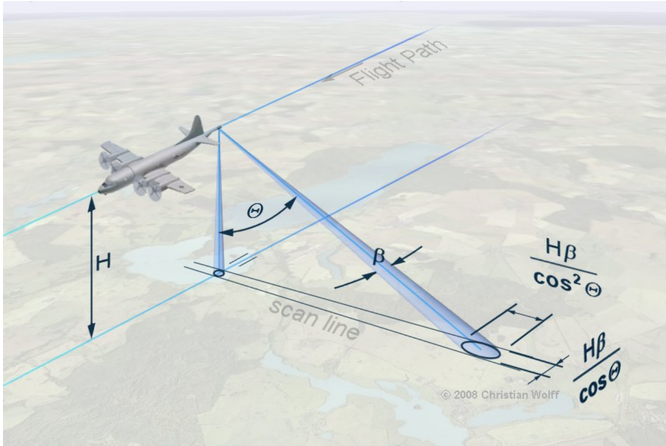


Figure 1.19: Diagram of the SLAR

Synthetic aperture radar(SAR)

To improve the resolution in remote sensing, a technique called the synthetic aperture dish is used. For an antenna like a parabolic dish, the angular resolution is given as the wavelength divided by the size[12] for the antenna, $\Delta\theta = \frac{\lambda}{D}$. To get a fine resolution for the image in remote sensing, a very large aperture D is required. A large D cannot be easily created especially in moving vehicles and aircraft.

Synthetic aperture radar (SAR)

To improve the resolution in remote sensing, a technique called the *synthetic aperture dish* is used. For an antenna like a parabolic dish, the angular resolution is given as the wavelength divided by the size (or aperture)[13] for the antenna $\Delta\theta = \frac{\lambda}{D}$. To get a fine resolution for the image in remote sensing, a very large aperture D is required. A large D cannot be easily created especially in moving vehicles and aircraft.

To mitigate this, there is a technique where the antenna is small but the vehicle moves and as the vehicle moves, the reflection information is stored after all the refection information is collected from different locations, then data processing can be done to get an angular resolution which will correspond to the total distance travelled by the vehicle. This technique is known as the *synthetic aperture radar* (see figure 1.20).

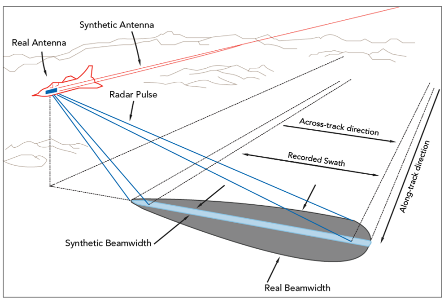


Figure 1.20: Diagram illustrating the SAR

[10]This technique is mostly used in aircraft
[11]For this book, D is the size of the antenna. hence the statement "large D" since the length (L) of an object is a function of the size (D)
[12]For a parabolic dish antenna the diameter 'd'is used instead of length 'L'. See footnote 11
[13]For a parabolic dish antenna the diameter "d" is used instead of length "L". See **??**

Features of the SAR

(i) It has a very high linear resolution independent of the range
(ii) It requires a source with higher coherence
(iii) The image is on the range-doppler coordinate grid
(iv) It requires a large data processing
(v) There are geometric and ratio metric distortions
(vi) There is speckle noise in the image

1.3.9 Radio astronomy

A typical radio telescope is shown in figure 1.21 with a passive receiver. In this case, no signal is transmitted but the signal is received.

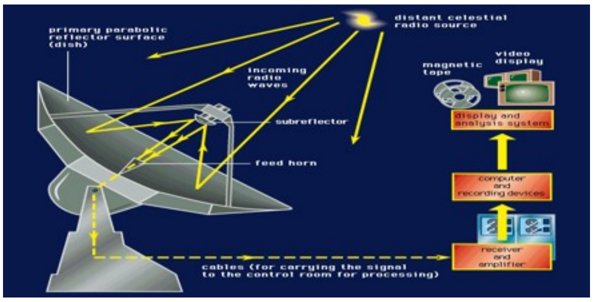


Figure 1.21: Radio telescope

The frequency of the signals is converted, detected and processed. With a radio telescope, we would like to get an image of the sky with as large a resolution as possible.

$\frac{\lambda}{D}$ comes into the picture and to get a very fine resolution of the image of the sky, a very large telescope would be needed.

Producing a very large telescope is difficult so we use the synthetic aperture technique as it was used on radars. In this method, we use a set-up of antennas, for example, if the dish in each array is of the order of $D = 25m$ and the total spread of the antennas of the order 21km. Therefore, we get an effective aperture through each antenna that has an aperture of only 25m.

1.3.10 EMI/EMC

EMI is *electromagnetic interference*. Firstly, let us investigate how a high-frequency device would create interfering signals and then what ways in which the interference can be reduced.

For example, you may get interference on our radios, whenever somebody starts a car or a motorcycle in the vicinity because starting a car or motorcycle involves sparking and because of that spark, you get electromagnetic interference which is picked up by the radio antenna and you get disturbance on your radios. It is essential to investigate the technique by which the interferer can be reduced or the mechanism by which the devices can be isolated. This technique is called ***shielding***.

EMC is electromagnetic compatibility. Today, whenever we design electromagnetic gadgets or electrical devices it is mandatory to make them electromagnetically compliant so it does not create additional electromagnetic interference which will affect other systems.

The figure below describes the whole concept of EMC, illustrating both causes(natural or man-made) and solutions (shielding and other techniques).

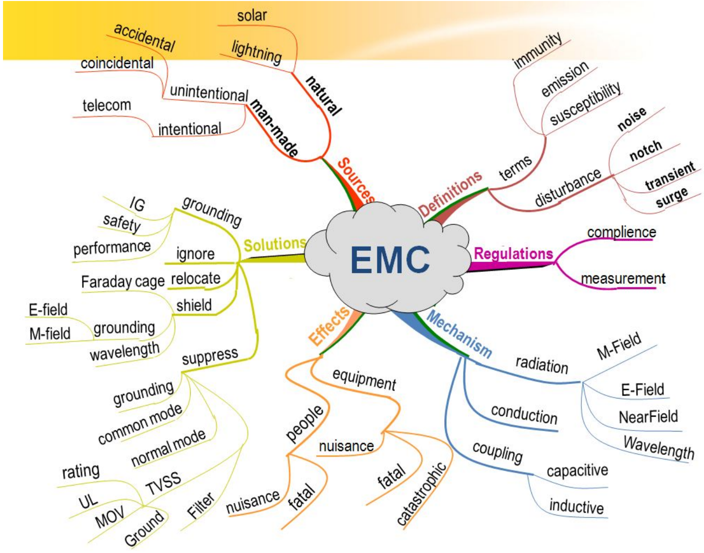


Figure 1.22: Electromagnetic Compatibility

ExerciseList

Ex. 1 — What is the difference between adaptive arrays and switched beam antennae?

Ex. 2 — What is Swath width in SLAR?

Ex. 3 — Explain the concept of fading in cellular communication.

Ex. 4 — If from 30MHz to 300MHz, the co-axial cable is used, and from 30GHz to 300GHz, the waveguide is used, what kind of transmission lines are used in the spectrum ranging from 300MHz to 30GHz?

Ex. 5 — For a SLAR with the following characteristics: $\lambda = 1cm$, $L = 3m$, $H = 6000m$, $\theta = 60\deg$, and pulse width = 100 ns. Find the resolutions.

Ex. 6 — Amongst coaxial cables, waveguides and twisted pairs, which suffers the most losses and why?

Ex. 7 — What is the basic working principle of the antenna?

Ex. 8 — How does the size of an antenna affect its performance?

Ex. 9 — what are the benefit of using adaptive arrays over switched beams in wireless communication?

Ex. 10 — In the realm of remote sensing, How can we effectively differentiate between man made structures and natural features using radar data?

Ex. 11 — When it comes to synthetic aperture radar (SAR) systems, what innovative techniques can we employ to mitigate the effect of speckle noise and enhance the quality of SAR imagery?

Ex. 12 — If a SLAR system operates at a frequency of 5.6GHz and has a wave length of approximately 5.36cm, what is the maximum achievable range resolution?

Ex. 13 — When it comes to interferometric synthetic aperture radar (SAR) techniques, what challenges do we face inaccurately measuring ground deformation and how can we over come them?

Ex. 14 — How does SLAR work to generate image of earth's surface?

Chapter 2

Transmission lines

Transmission lines are utilized for conveying special cases of electromagnetic waves generated by time-varying voltages and currents. In the previous courses, emphasis was placed on the concept of electromagnetic waves generated by electric and magnetic fields that vary with time.

2.1 Objectives

The objective of the chapter is to discuss the following

(i) The concept of voltage and current in high frequencies circuits using a distributed circuit analysis

(ii) The concept of maximum power transfer and matched condition.

(iii) Conditions for a lossless transmission line and the concept of Voltage Standing Wave Ratio (VSWR).

(iv) Impedance transformation in a lossless transmission line.

(v) Use of Smith chart for graphical analysis of transmission lines— how to develop the Voltage Standing Wave Ratio (VSWR) set of circles, relationship between impedance and admittance, how to determine reflection coefficient graphically, how to determine impedance transformation graphically, how to locate maximum and minimum voltage, and current points in the transmission line on the Smith chart.

(vi) Applications of transmission lines for impedance measurement, circuit elements, resonant circuits, and impedance matching and power flow in transmission lines.

(vii) Conditions of a lossy transmission line and its analysis as an approximation to a lossless transmission line.

(viii) The various types of transmission lines and their characteristic impedance.

2.2 Structures of transmission lines

Transmission lines can best be defined as a medium of power transfer from one point to another. This transmission is done through a variety of structures, some of which are explained below;

2.2.1 Co-axial cable

Figure 2.1 is a coaxial cable. In this particular setup, there's an inner and outer conductor and voltage is applied between the inner and outer conductor thus energy will be propagated along the length of the structure.

Figure 2.1: Co-axial cable transmission line

2.2.2 Balanced Parallel wire transmission lines

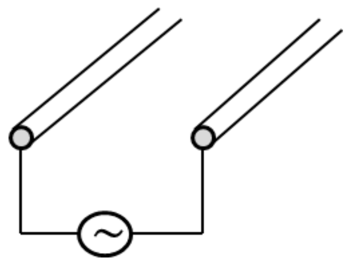


Figure 2.2: Balanced parallel wire transmission line

Figure 2.2 shows the parallel wire transmission line which has two separate conductors hence when voltage is applied at both ends, energy flows. The voltage on one wire is V^+ and the other V^-. The voltages are opposite in polarity hence the magnetic field generated by both wires cancels out.

2.2.3 Unbalanced Parallel wire transmission lines

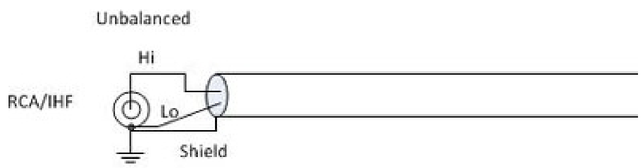


Figure 2.3: Unbalanced Line

In this structure shown in figure 2.3, we have a conducting rod above a ground surface and voltage is applied between the rod and the ground surface hence power is transmitted efficiently along the structure.

2.2.4 Microstrip Line

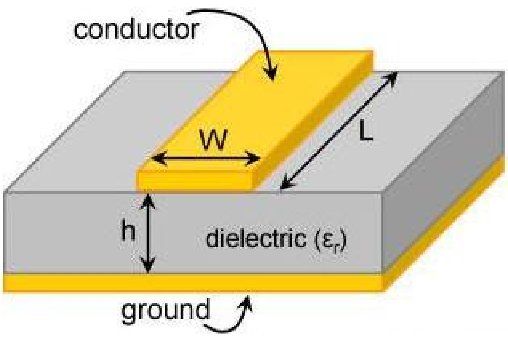


Figure 2.4: Microstrip line

The microstrip line is a transmission line geometry with a single conductor having conducting surfaces at the top and bottom that trace on one side of a dielectric substrate and a single ground plane on the other side (see figure 2.4). The single trace is carrying a voltage V with respect to the ground plane. There is no other trace close by carrying same voltage of opposite polarity to cancel out the magnetic field generated. Hence it is very noisy. The microstrip line is used extensively in printed circuit boards(PCBs).

2.2.5 Stripline

The structure is made up of a trace of conductor enmeshed in a dielectric with conducting surfaces at the top and bottom that are grounded.(see figure 2.5). The stripline is similar to the microstripline but with ground planes both above and below the trace. Striplines are mostly easily constructed on the inner layers of multi-layer printed circuit boards. Another form of a stripline known as a Differential-stripline is shown in figure 2.6.

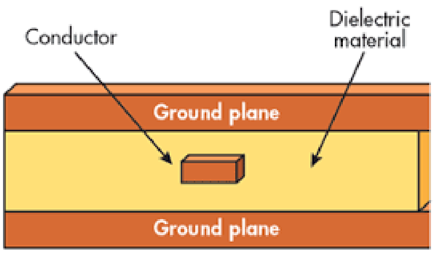


Figure 2.5: Stripline

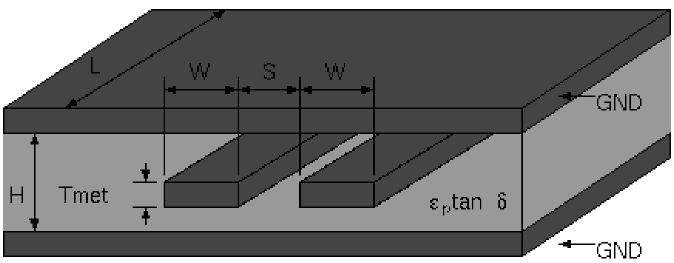


Figure 2.6: Differential Stripline

2.3 Concept of transit time effect

A similarity across the various structures is the presence of a two-conductor system, hence when voltage is applied between them because it is a time-varying voltage, current flows through the system. All the structures can be represented by a simple two-conductor system and this is also done for transmission line analysis where we ignore the structure and simply treat it as a two-conductor system, at one end we apply an energy source and at the other end, we apply a load and then carry out analysis from the source to the load.

Circuit laws like Kirchhoff's law, are limited to low-frequency circuits where we only take note of the value of the electrical components, i.e. for a circuit that comprises circuit elements like resistors, capacitors and inductors, only their values are needed for the circuit analysis. However, as frequency increases, the signal value changes significantly. Let us suppose a transmission line of length, L whose ends are connected to a voltage source and load at points XX'and YY' respectively as shown in figure 2.7.

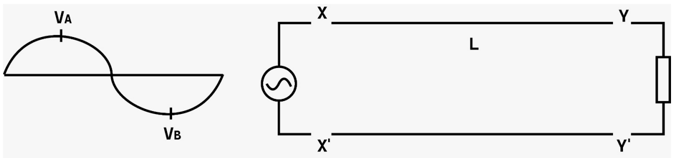


Figure 2.7: Circuit diagram of a transmission line

At some instant, V_A is applied at X, and when L is sufficiently small relative to the length (λ = wavelength), We assume that V_A appears at Y from X almost instantaneously.

In transit time effect one can realise that $L \gg \lambda$ and so it is very likely that $V_X \neq V_Y$, but when $\lambda \gg L$ (for low-frequency circuits) the sinusoidal wave changes almost instantaneously across all points on the line so that we can assume it is constant across the entire length, L. Mathematically, the transit time effect can be expressed in terms of the speed of the transmitted signal, v, from point XX' which shows the finite time it takes the voltage at X to appear at Y and it is called *transit time*.

$$t_r = \frac{L}{v} \tag{2.1}$$

So when a time-varying signal is applied at X, it requires a finite time to travel from X to Y. V_A at X is a time-varying sinusoidal signal such that it takes V_A to appear at Y in transit time, t_r. Thus, V_A would have changed to another value say V_B, so when a voltage at Y is V_B, the voltage at X is V_A. In other words, there is a potential difference between X and Y. This difference is related to the length of the cable, the more the length the more the difference, with L very small. V_x is close to V_y as the points A and B will be close to each other and the voltage difference will not be substantial. The important thing here is no matter how small the length we take, there will always be that voltage difference, $V_x - V_y \neq 0$ if $L > 0$.

Another observation is that as we increase the frequency, the role of the transit time effect becomes more and more important. So now to the big question; *When do you neglect transit time effect and when do you incorporate it in your design?*

If $V_x - V_y$ is small, the transit time effect can be neglected otherwise it should be taken into consideration. Generally, if the transit time is far less than the period of the signal, $t_r \ll T$ then the transit time effect can be neglected but if t_r is comparable to T, then we have to incorporate the effect of transit time.

In other words, when carrying out circuit analysis, the size of the structures starts to play an important role in the analysis.

Recall $T = \frac{1}{f}$, and $t_r = \frac{L}{v}$. Given that $\lambda = \frac{v}{f}$, then $T \gg t_r$ can be expressed as,

$$\frac{1}{f} \gg \frac{L}{v} \quad \text{or} \quad \lambda \gg L$$

Thus, $\lambda \gg L$ is the condition to neglect the transit time effect.

2.4 Concept of Distributed Elements in a transmission line

With the understanding of *transit time effect*, the transmission of signal along the transmission line causes a voltage difference across any arbitrary two points on the line. This is contrary to low-frequency circuit analysis (or the *lumped circuit analysis*), where the circuit idealizes the attributes such as resistance, capacitance and reactance to circuit elements joined by a network of perfecting conducting wires. So how can this observation be modelled? For an ideal conductor, the voltage drop is zero but when the angular frequency, ω is non-zero, there is reactive drop and the inductive reactance, ωL and capacitive susceptance, ωC becomes significant as frequency increases.

Thus, if an infinitesimally length of the transmission line is assumed to have the lumped element model where the transit time effect is neglible, then a transmission line can be modelled as as aggregation of all the lumped elements. Thus the lumped element is not situated in a particular location but distributed across the entire length of the transmission line. This approach, essentially, will break the **lumped element** model into a **distributed element** model. Again, the modelling of each subsection of the transmission line in space is done in such a way that its lenght is infinitesimally small, Δx so that the circuit laws at low frequency can apply within the subsection.

Transit time cannot be neglected for a two-conductor wire but when a small quantity in length is broken down into very Lineal elements of Δx with $\Delta x \to 0$, The transit time effect can be neglected as $L \gg \Delta x$ and so the circuit laws of low frequency can be applied to the arbitrary nodes separated by Δx length. In the low-frequency circuit, we should have resistance, capacitance and inductance stated in total value, instead, we have resistance, inductance and capacitance all per unit length.

The relationship between voltage and current is valid for high frequency like that of low frequency as $\Delta x \to 0$ (infinitesimally small). Hence, this solves for voltage and current of a transmission line in the presence of the transit time effect.

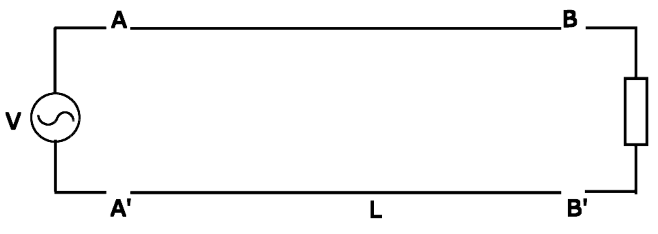


Figure 2.8: simple circuit diagram of a transmission line

In the circuit above, there is a voltage difference between both ends because of the transit time effect but we do not know why there is a voltage difference.

We all know that as frequency increases in the circuit, the two conductors have both electric and magnetic fields induced as shown below:

What would be the lumped element model for the subsection of the transmission line? As stated earlier, the voltage drop can be modelled as resistive and reactive drop across lumped elements. Put differently, We all know that as frequency increases in the circuit, the two conductors XY and X'Y' respectively in figure 2.7 have both electric and magnetic fields induced as shown in figure 2.9.

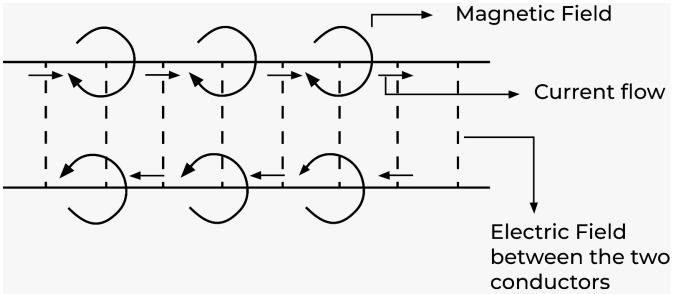


Figure 2.9: Distributed element model explained with generated fields across a transmission line

So when the magnetic field is linked with current, an inductance is produced while the electric field between the two conductors with air as dielectric produces capacitance as shown in figure 2.10. The inductance and capacitance produced are distributed along the length of the structure and these parameters are called the **distributed parameters of the line**.

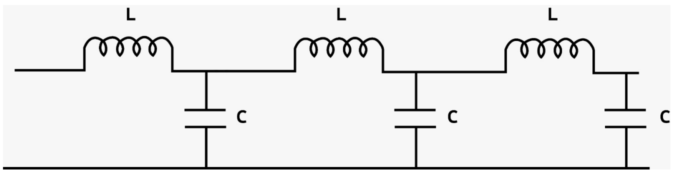


Figure 2.10: Distributed parameters of a lossless tranmission line

At low frequency, there is no voltage drop across the length of a transmission line but as frequency starts to increase, the inductive reactance, X_L, starts causing a voltage drop across the length of the transmission line that was not seen at low frequency. Conductors do not have zero resistance and the separation of the conductor, as a dielectric is not ideal, causes current flow between the conductors across the conductance of the dielectric. Hence, we have a more realistic transmission line modelled as shown in figure 2.11

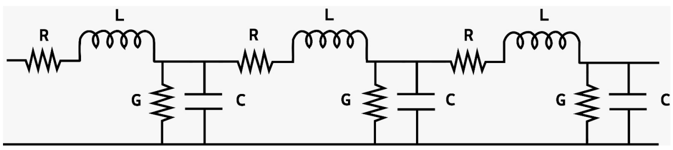


Figure 2.11: Distributed elements for lossy transmission

From figure 2.11, the resistance, capacitance, inductance and conductance are all per unit length values. We have defined the quantities which are called the *primary constants* of the transmission line which are:

(i) Resistance per unit length — Ohms/metre

(ii) Inductance per unit length — Henry/metre

(iii) Capacitance per metre — Farad/metre

(iv) Conductance per metre — Siemens/metre

After the primary constants have been determined then the analysis of the transmission line can be carried out by:

(i) Dividing the transmission line into small segments.

(ii) Writing down Kirchoff's voltage and current law for the segment.

(iii) Carrying out the analysis when the segment goes to zero to enable the analysis to be valid for any high frequency or any low wavelength.

(iv) Applying voltage to the lineal segment that causes current to flow into the circuit.

2.5 Telegrapher's Equations

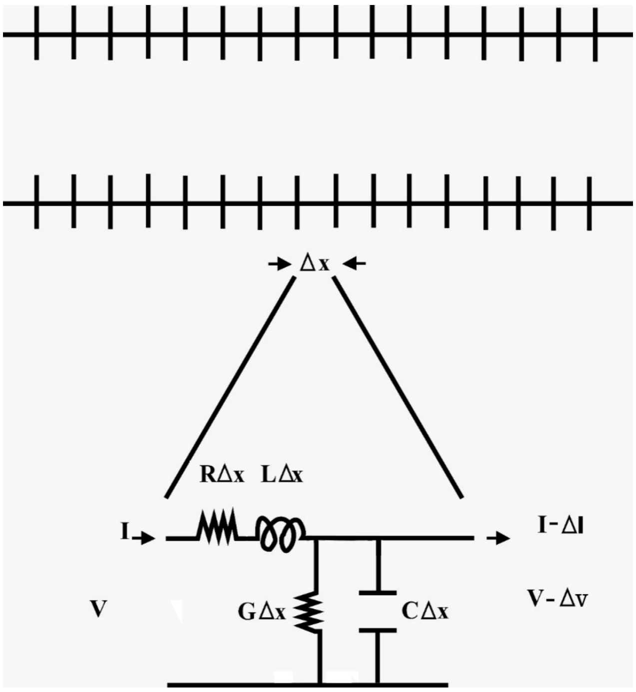


Figure 2.12: Circuit analysis of a small segment of the transmission line

Let us now analyse the small segment of the transmission line as shown in figure 2.12. Since the segment is very small and voltage has an angular frequency ω the voltage drop and consequent change in current (denoted by ΔI) will be contributed by the resistance and inductance, and conductance and capacitance of the lumped element model. These can be expressed as equations (2.2) and (2.3).

2.6 Let's Understand the Telegrapher's Equations

First, let's define some terms:

- V_R is the voltage across resistance
- V_G is the voltage across conductance
- V_C is the voltage across capacitance
- V_L is the voltage across inductance

Now, let's consider each term:

For V_R, we have $V = IR$, therefore $V_R = IR\Delta x$.
For V_L, we have $V_L = IX_L$, therefore $V_L = Ij\omega L\Delta x$.
For V_C, we have $V_C = I_C X_C$, therefore $V_C = \dfrac{I_C}{j\omega C\Delta x}$
hence $I_C = V_C j\omega C\Delta x$.
But $V_C = V - \Delta V$, so $I_C = (V - \Delta V)(j\omega C\Delta x)$.
For V_G, we have $V_G = I_G R_G$, therefore $V_G = \dfrac{I_G}{G\Delta x}$
hence $I_G = V_G G\Delta x$.
But $V_G = V_C = V - \Delta V$, so $I_G = (V - \Delta V)(G\Delta x)$.

Applying Kirchhoff's Voltage Law (KVL) to the small segment of the transmission line, we get
$V - V_R - V_L - (V - \Delta V) = 0$,
which implies $\Delta V = V_R + V_L$. Substituting V_R and V_L, we get $\Delta V = IR\Delta x + Ij\omega L\Delta x$.

$$V - IR\Delta x - jI\omega L\Delta x - (V - \Delta V) = 0$$

hence,

$$\Delta V = (R\Delta x + j\omega L\Delta x)I \tag{2.2}$$

Similarly, applying Kirchoff's current law(KCL), we would have;

$$I - G\Delta x(V - \Delta V) - j\omega C\Delta x(V - \Delta V) - (I - \Delta I) = 0$$

hence,

$$\Delta I = (G\Delta x + j\omega C\Delta x)(V - \Delta V) \tag{2.3}$$

Recall that for equations (2.2) and (2.3) to be valid at all frequencies, Δx must tend to zero.

From equation (2.2) we have,

$$\begin{aligned}\Delta V &= (R\Delta x + j\omega L\Delta x)I \\ &= (R + j\omega L)I\Delta x\end{aligned}$$

$$\begin{aligned}\frac{\Delta V}{\Delta x} &= \frac{(R + j\omega L)I\Delta x}{\Delta x} \\ &= (R + j\omega L)I\end{aligned}$$

$$\lim_{\Delta x \to 0} \frac{\Delta V}{\Delta x} = \frac{dV}{dx} = (R + j\omega L)I$$

$$\frac{dV}{dx} = (R + j\omega L)I \tag{2.4}$$

Also, from equation (2.3) we have,

$$\begin{aligned}\Delta I &= (G\Delta x + j\omega C\Delta x)(V - \Delta V) \\ &= (G + j\omega C)V\Delta x\end{aligned}$$

$$\begin{aligned}\frac{\Delta I}{\Delta x} &= \frac{(G + j\omega C)V\Delta x}{\Delta x} \\ &= (G + j\omega C)(V - \Delta V)\end{aligned}$$

$$\lim_{\Delta x \to 0} \frac{\Delta I}{\Delta x} = \frac{dI}{dx} = (G + j\omega C)V$$

$$\frac{dI}{dx} = (G + j\omega C)V \tag{2.5}$$

Equations (2.4) and (2.5) are said to be coupled equations since $\frac{dV}{dx}$ is related to I and $\frac{dI}{dx}$ is related to V.

Hence, we see that the voltage and current as $\Delta x \to 0$ are not related by algebraic equations but by a differential equation. By differentiating equations (2.4) and (2.5) again, we have;

$$\frac{d^2V}{dx^2} = (R + j\omega L)\frac{dI}{dx} \tag{2.6}$$

$$\frac{d^2I}{dx^2} = (G + j\omega C)\frac{dv}{dx} \tag{2.7}$$

Substituting equation (2.5) into (2.6) we have,

$$\begin{aligned}\frac{d^2V}{dx^2} &= (R + j\omega L) \times (G + j\omega C)V \\ &= (R + j\omega L)(G + j\omega C)V\end{aligned} \tag{2.8}$$

The coefficient of V in equation 2.8 must be a primary quantity of the transmission line since it depends on the primary constant of the transmission line and the frequency of operation ω. Let's call the quantity γ^2.

Therefore,

$$\frac{d^2V}{dx^2} = \gamma^2 V \tag{2.9}$$

$$\frac{d^2I}{dx^2} = \gamma^2 I \tag{2.10}$$

Equations 2.9 and 2.10 are referred to as the *telegrapher's equation*.

These are now the voltage and current expressions that governs a transmission line. The quantity γ is called the *propagation constant*. The quantity V and I are sinusoidal and so vary with $e^{j\omega t}$. The quantity $e^{j\omega t}$ helps take care of the harmonic time variation in V an I. Since γ is a constant for a given line and frequency,

$$\frac{d^2V}{dx^2} = \gamma^2 V$$

Therefore;

$$V(x) = V^+e^{-\gamma x} + V^-e^{\gamma x} \tag{2.11}$$

Equation (2.11) is the peak value and V^+ and V^- are the arbitrary constants which are evaluated by using appropriate boundary conditions. To find the instantaneous value of V, we multiply equation (2.11) by $e^{j\omega t}$. [1]So,

$$V(x,t) = V^+e^{-\gamma x}e^{j\omega t} + V^-e^{\gamma x}e^{j\omega t} \tag{2.12}$$

$V(x,t)$ is a complex quantity. The only way to show the amplitude and the initial phase is through a complex number. therefore, V^+ is the complex value of voltage at $x = 0$ location with a certain initial phase. Hence why V is always complex.

Now, what is the propagation constant, γ?. It is the quatity that defines the characteristics of the signal in the transmission line. It is expressed as $\gamma = \alpha + j\beta$ which is also a complex quantity. Substituting γ into $V^+e^{-\gamma x}e^{j\omega t}$ in equation (2.12) gives

$$\begin{aligned}V^+e^{-\gamma x}e^{j\omega t} &= V^+e^{-(\alpha+j\beta)x}e^{j\omega t} \\ &= V^+e^{-\alpha x}e^{j(\omega t-\beta x)}\end{aligned} \tag{2.13}$$

[1]V depends on x and t

$V^+e^{-\alpha x}$ is a voltage whose amplitude varies exponentially along the transmission line and the phase of the voltage is a combination of space and time i.e. $(\omega t - \beta x)$, ωt for time and βx for space. Hence $V^+e^{-\gamma x}e^{j\omega t} \equiv V^+e^{-\alpha x}e^{j(\omega t-\beta x)}$ is the first point of $V(x,t)$ and represent something whose amplitude varies along the length and the phase of which is a combination of space and time.

For understanding say, V^+ is real and positive and $\alpha = 0$ then

$$V^+e^{-\alpha x}e^{j(\omega t-\beta x)} = V^+e^{j(\omega t-\beta x)}$$

If we plot this as a function of space and time [2]

$$V^+e^{j(\omega t-\beta x)} = V^+cos(\omega t - \beta x) + V^+jsin(\omega t - \beta x)$$

Since we want real voltage we take only $V^+cos(\omega t - \beta x)$ and plot variation of V with respect to x and t. We have the graph shown in figure 2.13 for different t, as t increases, a point in the plot moves rightward. Hence with respect to space and time, the graph moves rightward as time increases. This is called a **travelling wave**.

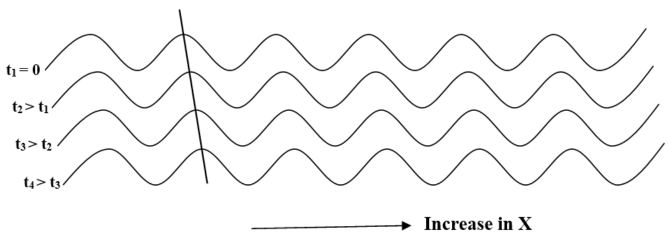


Figure 2.13: Voltage as a function of x for $V^+e^{-\lambda x}$

Hence $V^+e^{j(\omega t-\beta x)}$ is a positive or progressive travelling wave as it moves rightward with an increase in time. Similarly, when $V^-e^{j(\omega t+\beta x)}$ is plotted on the graph, the point starts moving leftward with an increase in time and it is the negative travelling wave(see figure 2.14).

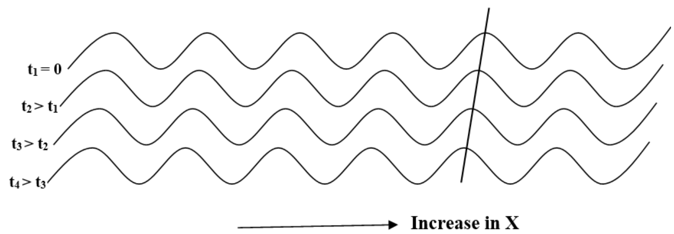


Figure 2.14: Voltage as a function of x for $V^+e^{-\lambda x}$

Hence with high-frequency circuit analysis, the voltages and current have to be visualized in the form of waves. In conclusion, we see that a departure from the lumped circuit analysis to the distributed circuit analysis radically changes the approach to circuit analysis to a voltage and current that exist in the form of waves on the electrical circuits.

Example 2.6.1 Transmission line Equation

Show that the voltage $V(x,t) = A\cos(\omega t + \theta)e^{j\beta x}$ *satisfies the transmission line equation* (2.8)*, for a uniform lossless line, if* $\beta = \omega\sqrt{LC}$.

[2]Recall from Euler's formula; $e^{j\theta} = cos\theta + jsin\theta$

Solution

In the coming sections, we will discuss the properties of a lossless transmission line in detail but as an introduction, for a lossless transmission line, $R = G = 0$, so equation (2.8) reduces to

$$\begin{aligned}\frac{\mathrm{d}^2 V}{\mathrm{d}x^2} &= -(0 + j\omega L) \times -(0 + j\omega C)V \\ &= -(\omega^2 LC)V\end{aligned}$$

For $V(x,t) = A\cos(\omega t + \theta)e^{j\beta x}$, the differential equation becomes

$$\frac{\mathrm{d}^2 A\cos(\omega t+\theta)e^{j\beta x}}{\mathrm{d}x^2} = -(\omega^2 LC) \times A\cos(\omega t+\theta)e^{j\beta x} \tag{2.14}$$

$$A \times -\cos(\omega t+\theta) \times \frac{\mathrm{d}^2 e^{j\beta x}}{\mathrm{d}x^2} = -(\omega^2 LC) \times A\cos(\omega t+\theta)e^{j\beta x} \tag{2.15}$$

$$-(A\cos(\omega t+\theta)) \times \beta^2 e^{j\beta x} = -(\omega^2 LC) \times A\cos(\omega t+\theta)e^{j\beta x} \tag{2.16}$$

$$\beta^2 = \omega^2 LC \tag{2.17}$$

This implies that $\beta = \omega\sqrt{LC}$ and proves that if $\beta = \omega\sqrt{LC}$, the voltage equation given for a lossless transmission line satifies the transmission line equation.

Exercises

Ex. 15 — Briefly state and Explain any 5 transmission line structures with appropriate diagrams.

Ex. 16 — What is the difference between a stripling and a microstripline in terms of better noise figure?

Ex. 17 — With appropriate diagram(s), explain the Transit time effect.

Ex. 18 — Contrast the lumped element model of a circuit with the distributed element model.

Ex. 19 — Why are circuit laws for low frequency sufficient for Lineal element created to analyze distributed circuit?

Ex. 20 — Briefly explain what we mean by the distributed parameters of a transmission line.

Ex. 21 — $\frac{dV}{dx} = -(R + j\omega L)I$ and $\frac{dI}{dx} = -(G + j\omega C)V$ are called coupled equations. Why?

Ex. 22 — From first principle and with appropriate diagram, derive the Telegrapher's equations for a distributed element of a transmission line.

Ex. 23 — Show that, the general wave equation of a transmission line given as $V(x,t) = V^+ e^{-\gamma x} e^{j\omega t} + V^- e^{\gamma x} e^{j\omega t}$ is made up of a forward traveling wave and a backward traveling wave.

Ex. 24 — Write short note on what we mean by primary constants of a transmission line.

Ex. 25 — Why is it not practical to use the lumped circuit element model in transmission line problems?

Ex. 26 — What are the possible sources of losses in a transmission line?

Ex. 27 — Show that

$$\alpha = -\left[\left(RG - \omega^2 LC\right)^2 + \omega^2 \left(LG + RC\right)^2\right]^{\frac{1}{4}} \sin\left(\frac{\tan^{-1}\left(\frac{\omega LG + \omega RC}{RG - \omega^2 LC}\right)}{2}\right) \tag{2.18}$$

if $RG - \omega^2 LC < 0$. Where α is the real part of the propagation constant (i.e attenuation constant) and R, G, L, C are the primary constants of the transmission line.

Ex. 28 — From the above question show that the attenuation constant $\alpha = 0$ for a lossless transmission line (i.e If $R = G = 0$).

Ex. 29 — Given the following primary constants and frequency of a transmission line. $R = 2$ ohms/m, $G = 3$ Siemens/m, $L = 7$ H/m, $C = 10$ f/m and $freq = 2000$ Hz. Find the propagation constant of the transmission line.

Ex. 30 — Suppose that the propagation constant of a lossless transmission line is $106629j$. Find the frequency of the transmission line if ohms/m, Siemens/m, $L = 6$ H/m and $C = 12$ f/m.

Now, let us try to understand the physical significance of this complex quantity γ.

2.7 The complex quantity γ

From equation (2.8), we have established that γ is given as

$$\begin{aligned}\gamma &= \sqrt{(R+\jmath\omega L)(G+\jmath\omega C)} \\ &= \alpha + \jmath\beta\end{aligned} \tag{2.19}$$

Where R is the resistance per unit length, L is the inductance per unit length, G is the conductance per unit length, and C is the capacitance per unit length.

For the forward travelling wave in equation (2.11), we can express it in phasor form as follows

$$V^+e^{-\gamma x} = \left|V^+\right| e^{-(\alpha+\jmath\beta)x}e^{\jmath\phi} \tag{2.20}$$

If we assume V^+ is real and have an initial phase $\phi = 0$, then;

$$\begin{aligned}V^+e^{-\gamma x} &= \left|V^+\right| e^{-(\alpha+\jmath\beta)x}e^{\jmath 0} \quad \text{But } e^{j0} = e^0 = 1 \\ &= \left|V^+\right| e^{-\alpha x}e^{-\jmath\beta x}\end{aligned} \tag{2.21}$$

Equation (2.21) is the phasor representation.

In phasor representation, the time-varying part is ignored so the complete expression for V^+ is

$$V^+ = \left|V^+\right| e^{j(\omega t+\phi)}e^{-\gamma x}$$

and the phasor form like equation (2.21) with initial phase ϕ is

$$V^+ = \left|V^+\right| e^{j\phi}e^{-\gamma x}$$

2.7.1 The Phase constant

From the equation (2.21), we see that as the wave propagates i.e as the value of x increases, the quantity $|V^+|\,e^{-\alpha x}$ is exponentially decreasing while $e^{-\jmath\beta x}$ is the sinusoidal part that oscillates because according to [3]Euler's formula $e^{-\jmath\beta x} = \cos\beta x - \jmath\sin\beta x$. Hence phase (spatial phase) is obtained from $e^{-\jmath\beta x}$.

We now see that the equation (2.19) has α that controls the amplitude of the wave as we move in the x-direction and β that controls the phase of the wave along the transmission line. Hence,

$$\text{Spatial phase} = -\jmath\beta \tag{2.22}$$

As we travel in the positive x-direction, the phase lags more and this linearly varies with x for a given value of β. Hence β represents the phase change per unit length.

$$\beta = \frac{\text{phase change}}{\text{unit length}} \quad \left(\frac{radian}{m}\right) \tag{2.23}$$

We know that a phase change of 2π corresponds to a wavelength. From

$$\phi = \beta x \quad \text{if, } \phi = 2\pi$$

thus,

$$2\pi = \beta\lambda \text{ or } \beta = \frac{2\pi}{\lambda}$$

where $x = \lambda$ is the distance travelled.

For most transmission line problems involving wave motion, λ is not given, instead, γ (propagation constant) is given in the complex form and β is analyzed from the value of γ given. The wavelength λ is then calculated from β. Since γ depends on the primary constants R, L, G and C at operating frequency ω, we then conclude that the quantity β called *phase constant* is also a function of frequency ω. In other words, we conclude that the wavelength of waves on a transmission line is a function of the line parameter, phase constant, and wavelength.

2.7.2 The Attenuation constant

Recall that amplitude varies as $|V^+|e^{-\alpha x}$ for the forward travelling wave as in equation (2.11) so that we have maximum amplitude at $x = 0$. As x increases, the amplitude decreases exponentially with αx.

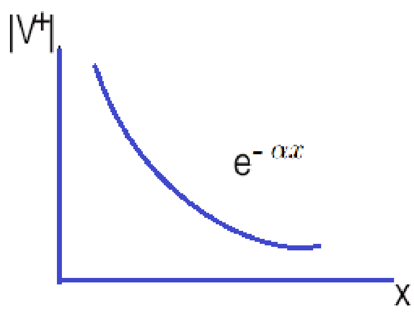


Figure 2.15: The amplitude versus distance plot

From figure 2.15, α is a parameter that measures how fast amplitude decay occurs in the transmission line wave. This quantity α is called the *attenuation constant*. The attenuation constant measures how the wave attenuates (reduces in its value) as it travels along the structure. So it is measured in Nepers/meter.

$$\text{Attenuation constant, } \alpha = \frac{Nepers}{meter}$$

[3] 

Leonhard Euler (15 April 1707 - 18 September 1783). He was an 18th-century Swiss mathematician and physicist who made significant contributions to various branches of mathematics and introduced numerous concepts that are widely used today. He was born on April 15, 1707, in Basel, Switzerland, and spent the majority of his career in St. Petersburg, Berlin, and Basel.

Euler's work spanned diverse areas of mathematics, including calculus, number theory, graph theory, and differential equations. He made groundbreaking contributions to each of these fields, revolutionizing the way mathematicians approached and solved problems. Euler's extraordinary output includes over 850 published papers, covering a vast range of mathematical topics.

One of Euler's most famous achievements is Euler's formula, also known as Euler's identity or Euler's equation. The formula relates five fundamental mathematical constants: e (the base of natural logarithms), π (pi, the ratio of a circle's circumference to its diameter), i (the imaginary unit, which satisfies $i^2 = -1$), 1 (the multiplicative identity), and 0 (the additive identity). The formula can be written as:

$$e^(i\pi) + 1 = 0.$$

Euler's formula connects exponential functions, complex numbers, and trigonometry in a profound way. It highlights the unexpected relationship between these seemingly unrelated mathematical concepts. The formula has been hailed as one of the most beautiful equations in mathematics due to its elegant simplicity and deep significance.

If α = 1(Nepers/meter)[4], then the voltage value will reduce from its initial value to $\frac{1}{e}$ for a distance x = 1 meter. So α *relates the distance over which the amplitude drops to* $\frac{1}{e}$ *of its initial value*. This length at which we get $\frac{1}{e}$ is called the *characteristic length*.

So, a distance $x = \frac{1}{\alpha}$ describes the effective travel distance in the transmission line beyond which the amplitude drops below $\frac{1}{e}$ of its initial value. Since the wave is reducing to $\frac{1}{e}$ of its initial value, the power of the wave also reduces. Taking the ratio of the initial amplitude and final amplitude after the effective travel distance we have,

$$\frac{|V^+|}{|V^+|e^{-\alpha x}}$$

Because the initial amplitude at $x = 0$ is $|V^+|$, with $x = \frac{1}{\alpha}$, the expression reduces to $\frac{1}{e}$.

We would now proceed to express the attenuation constant in decibels per metre $\left(\frac{\text{dB}}{\text{metre}}\right)$[5].

$$\begin{aligned} &dB \\ &= -20\log_{10}\left(\frac{1}{e^{\alpha x}}\right) \\ &= -20\log_{10}(e^{-\alpha x}) \quad \text{With } \alpha = 1(\text{Neper/metre}) \text{ and } x = 1m \\ &= -20\log_{10}(e^{-1}) \\ &= 8.68\ dB/m \end{aligned}$$

Therefore, 1 Neper/metre = 8.68 dB/m.

As in propagation constant γ, the attenuation constant α depends on the primary constant of the transmission line as well as the frequency of operation ω. In general, the propagation constant γ which is a combination of phase constant β and attenuation constant α is a function of primary line parameters and frequency of operation. Hence as ω increases, α increases. This is the reason some structures which were satisfactorily good conductors, at low frequencies become bad conductors at high frequencies (i.e. the conductor becomes a more lossy line at high frequencies).

Example 2.7.1 Measure the propagation constant

Let R = 0.5 Ω /m, L = 0.2μH /m, C = 100pF/m, G = 0.1 Ω /m, freq = 1GHz. Calculate the propagation constant, attenuation constant and phase constant for this line.

Solution

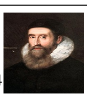
$$\gamma = \sqrt{(R + \jmath\omega L)(G + \jmath\omega C)}$$

But, $\omega = 2\pi f$ *and* $f = 1\text{GHz} = 1 \times 10^9\text{Hz}$

$$\omega = 2\pi \times 10^9 \text{rad/s}$$

Substituting into the γ *expression;*

$$\gamma = \sqrt{(0.5 + \jmath 400\pi)(0.1 + \jmath 0.2\pi)}$$



[4] The unit's name is derived from the name of John Napier. John Napier of Merchiston (1550 – 4 April 1617); also signed as Neper, Nepair; nicknamed Marvellous Merchiston, was a Scottish landowner known as a mathematician, physicist, and astronomer. He is best known as the discoverer of logarithms, he also invented the so-called "Napier's bones"

[5] This is the unit in which the attenuation constant is given in most data sheets

Expanding,

$$= \sqrt{0.5(0.1) + 0.5(\jmath 0.2\pi) + \jmath 400\pi(0.1) + \jmath 400\pi(\jmath 0.2\pi)}$$

Recall that, $\jmath \times \jmath = \sqrt{-1} \times \sqrt{-1} = (\sqrt{-1})^2 = -1$

$$\begin{aligned} &= \sqrt{0.05 + \jmath 0.31416 + \jmath 125.6637 - 789.568} \\ &= \sqrt{-789.518 + \jmath 125.97786} \end{aligned}$$

We are now faced with the challenge of finding the square root of a complex number.

To find the square root, we apply [6]DeMoivre's theorem. It states;

$$Z^{\frac{1}{n}} = |Z|^{\frac{1}{n}} \angle \frac{\theta}{n}$$

We first convert the complex number to polar form, that is,

$$-789.518 + \jmath 125.9778 = 799.5055\angle 170.934^o$$

Thus,

$$\begin{aligned} \gamma &= \sqrt{-789.518 + \jmath 125.9778} \\ &= \sqrt{799.5055}\angle\frac{170.934^o}{2} \\ &= 28.2\angle 85.467^o \end{aligned}$$

Converting back to the cartesian form, we have the propagation constant as $\gamma = 2.23 + \jmath 28.1$.

This give that attenuation constant, $\alpha = 2.23467$ *Nepers/m and phase constant,* $\beta = 28.1871$ *rad/m*

We now convert the attenuation constant to dB/m, thus,

$$1\textit{Nepers/metre} = 8.68 dB/m$$

$$\Rightarrow 2.23467\textit{Nepers/metre} = 19.3969 dB/m$$

Example 2.7.2 Determine the forward travelling wave

From previous example, say at x = 0, t = 0, V = 8.66V. Find the voltage at x = 1 and t = 100ns at point B as shown in figure 2.16 on the transmission line. Also, find the peak voltage at x = 1m. Assume the wave travels from left to right, and the initial phase $\phi = 30^o$.
If the wave travels from right to left, find the voltage at B.

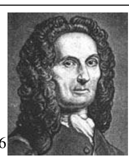
Solution

At x = 1, the voltage is maximum, at x = 0 i.e. A, V = 8.66v. Due to the direction of wave travel, we expect its amplitude to reduce to a smaller value at B. The expression

$$V(x,t) = \mathfrak{Re}\{\left|V^+\right| e^{-\alpha x}.e^{-\jmath\beta x + \jmath\omega t}.e^{+\jmath\phi}\} \qquad (2.24)$$

represents the forward travelling or progressive wave moving along the +x direction.
Extracting the real part and including the initial phase.

$$V(x,t) = |V^+|\cos(\phi + \omega t - \beta x)e^{-\alpha x}$$



[6] Named after Abraham de Moivre (26 May 1667 – 27 November 1754). He was a French mathematician known for de Moivre's formula, a formula that links complex numbers and trigonometry, and for his work on the normal distribution and probability theory. He was a friend of Isaac Newton, Edmond Halley, and James Stirling.

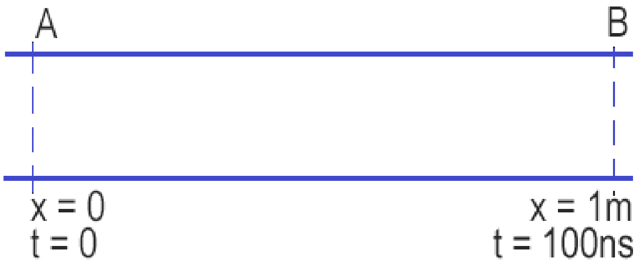


Figure 2.16: The transmission line showing points A and B

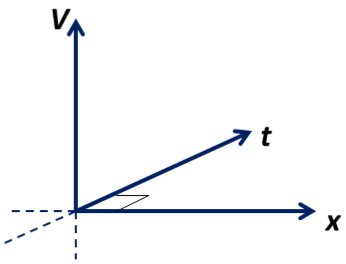


Figure 2.17: Voltage versus distance with time

$|V^+|e^{-\alpha x}$ gives the amplitude variation with distance x.
$\phi + \omega t - \beta x$ gives the phase (including the initial phase ϕ).
Substituting x = 0, t = 0, and $\phi = 30^o$ into equation 2.24,

$$V(x,t) = V(0,0) = 8.66V$$
$$V(x,t) = |V^+|\cos(\phi + \omega(0) - \beta(0))e^{-\alpha(0)}$$
$$8.66 = |V^+|\cos(\phi)$$
$$8.66 = |V^+|\cos(30)$$
$$|V^+| = 10V$$

Substituting x = 1m, t = 100ns, and $\phi = 30^o$ into equation 2.24,

$$V(x,t)$$
$$= 10\cos\Big(\frac{\pi}{6} + 2\pi \times 10^9 \times 100 \times 10^{-9} - 28.18 \times 1\Big)e^{-2.235\times 1}$$
$$= 10 \times -0.815 \times e^{-2.235}$$
$$= -0.872V$$

To find the peak voltage at x = 1,

$$V(1,t) = 10\cos\Big(\frac{\pi}{6} + 2\pi \times 10^9 t - 28.18\Big)e^{-2.235}$$

But the value is maximum when $\cos\big(\frac{\pi}{6} + 2\pi \times 10^9 t - 28.18\big) = 1$,

$$V_{\max} = 10e^{-2.235}$$
$$V_{\max} = 1.07528V$$

We observe from the solution that attenuation took place since the amplitude reduced from 8.66V to -0.88V.

If the wave travels from right to left

$$V(x,t) = \Re\{V^+e^{+\alpha x}.e^{+j\beta x + j\omega t}e^{+j\phi}\}$$
$$= |V^+|\cos(\phi + \omega t + \beta x)e^{+\alpha x}$$
$$= 10\cos\Big(\frac{\pi}{6} + 2\pi \times 10^9 \times 100 \times 10^{-9} + 28.18 \times 1\Big)e^{+2.235\times 1}$$
$$= 10 \times -0.9093 \times e^{+2.235}$$
$$= -84.99V$$

This is the amplitude of the voltage at point B that will undergo attenuation in moving backwards from B to A.

Example 2.7.3 Determine the backward travelling wave

Given that the wave in the previous example is travelling in the negative x direction and all other parameters are kept the same, what is the instantaneous voltage at the same time and at the same location?

Solution

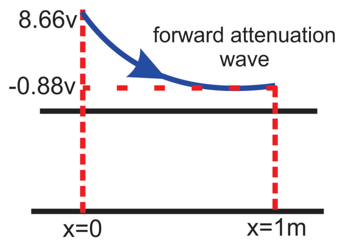


Figure 2.18: Solution for example 2.7.2

For the wave travelling in the negative direction, we have $V_{(x,t)} = \Re\left\{V^-e^{\gamma x}e^{j\omega t}\right\}$ for backward wave.

$$V(x,t) = \Re\left\{V^-e^{\alpha x}e^{j\beta x}e^{j\omega t}\right\}$$
$$= |V^-|e^{\alpha x}cos(\phi + \omega t + \beta x)$$

At $t = 0, x = 0, V = 8.66$, then

$$8.66 = |V^-|e^0 cos(\phi + 0 + 0)$$
$$= |V^-|cos\phi \quad \text{at } \phi = 30^\circ$$
$$= |V^-|cos30$$
$$= |V^-|0.866$$

$$\frac{8.66}{0.866} = |V^-|$$
$$= 10V$$

$$V(x,t) = 10e^{2.23x}\cos\Big(\frac{\pi}{6} + 2\pi \times 10^9 t + 28.2x\Big)$$

At $t = 100nsec$ and $x = 1m$

$$V(x,t) = 10e^{2.23}\cos\Big(\frac{\pi}{6} + 2\pi \times 10^9 \times 100 \times 10^{-9} + 28.2\Big)$$
$$= -83.7701V$$

So what we note here is, if we look at the transmission line with $x = 0$ and $x = 1m$. If the wave travels in the forward direction, then the wave will attenuate in the direction of propagation. The variation is from $8.66V$ to $-0.88V$. In the backward direction, the wave will attenuate in the reverse direction so it has to have a higher amplitude at $x = 1m$ and $t = 100nsec$ from where it attenuates from $-83.7701V$ to $8.66V$ in the other direction. The direction of propagation of the wave is very important.

2.8 The Characteristic Impedance

As derived in equations (2.4) and (2.5)

$$\frac{dV}{dx} = -(R + \jmath\omega L)I \tag{2.25}$$

$$\frac{dI}{dx} = -(G + \jmath\omega C)V \tag{2.26}$$

Also, the voltage wave solution derived in equation (2.11) is given and the same can be written for the current wave equation as follows

$$V = V^+e^{-\gamma x} + V^-e^{+\gamma x}$$
$$I = I^+e^{-\gamma x} + I^-e^{+\gamma x}$$

Substituting V and I into the differential equation (2.25),

$$\frac{d}{dx}(V^+e^{-\gamma x}+V^-e^{+\gamma x}) = -(R+\jmath\omega L)(I^+e^{-\gamma x}+I^-e^{+\gamma x})$$

Thus,

$$-\gamma V^+e^{-\gamma x}+\gamma V^-e^{+\gamma x} = -(R+\jmath\omega L)(I^+e^{-\gamma x}+I^-e^{+\gamma x})$$

(V^+, I^+) represents the forward travelling waves, while (V^-, I^-) represents the backward travelling waves for voltage and current respectively. The relationship between voltage and current has to be satisfied at every point along the transmission line. This will happen if and only if

$$-\gamma V^+e^{-\gamma x} = -(R + \jmath\omega L)I^+e^{-\gamma x}$$

So,

$$\frac{V^+}{I^+} = \frac{R + \jmath\omega L}{\gamma} \tag{2.27}$$

Similarly, for the backward travelling wave,

$$\gamma V^-e^{+\gamma x} = -(R + \jmath\omega L)I^-e^{+\gamma x}$$

So,

$$\frac{V^-}{I^-} = -\frac{R + \jmath\omega L}{\gamma} \tag{2.28}$$

The equations (2.27) and (2.28) show the relationship between forward and backward travelling waves for voltage and current.

Recall that $\gamma = \sqrt{(R + \jmath\omega L)(G + \jmath\omega C)}$, therefore from equation (2.27)

$$\begin{aligned}\frac{V^+}{I^+} &= \frac{R + \jmath\omega L}{\sqrt{(R + \jmath\omega L)(G + \jmath\omega C)}}\\ &= \frac{\sqrt{(R + \jmath\omega L)^2}}{\sqrt{(R + \jmath\omega L)(G + \jmath\omega C)}}\\ &= \sqrt{\frac{(R + \jmath\omega L)(R + \jmath\omega L)}{(R + \jmath\omega L)(G + \jmath\omega C)}}\\ &= \sqrt{\frac{R + \jmath\omega L}{G + \jmath\omega C}}\end{aligned} \tag{2.29}$$

Similarly for equation (2.28)

$$\begin{aligned}\frac{V^-}{I^-} &= \frac{-(R + \jmath\omega L)}{\sqrt{(R + \jmath\omega L)(G + \jmath\omega C)}}\\ &= -\sqrt{\frac{R + \jmath\omega L}{G + \jmath\omega C}}\end{aligned} \tag{2.30}$$

The expression $\sqrt{\frac{R+\jmath\omega L}{G+\jmath\omega C}}$ is another characteristic of the transmission line since it depends only on the primary constants and the frequency of operation. Also, this parameter is the ratio of voltage and current and as such has a definition of impedance. Hence the reason it is called the *characteristic impedance* of the line and it is denoted by Z_0. Thus,

$$Z_0 = \sqrt{\frac{R + \jmath\omega L}{G + \jmath\omega C}} \tag{2.31}$$

The characteristic impedance, Z_0 governs energy flow on the transmission line and will be discussed later in this chapter. So these two parameters, the propagation constant γ and the characteristic impedance Z_0 completely characterize the propagation of wave along a transmission line. Though R, L, G and C are the primary constants, they are hardly given in any transmission line problem. Instead, a transmission line is characterized by its propagation constant γ and characteristic impedance Z_0. These values are usually given in the datasheet of transmission lines and this is sufficient information to solve any transmission line problem.

From equations (2.29) and (2.30), substituting for equation (2.31) we can write

$$\frac{V^+}{I^+} = Z_0 \tag{2.32}$$

$$\frac{V^-}{I^-} = -Z_0 \tag{2.33}$$

These show that at any point on the transmission line, the ratio of voltage to current (forward or backward) is always constant i.e equal to characteristic impedance. Hence a forward travelling wave sees an impedance of Z_0 and the reverse travelling wave sees an impedance of $-Z_0$. If Z_0 is real, it means the forward travelling wave sees a positive resistance while the backward travelling wave sees a negative resistance.

But, *what does a negative resistance mean?* Ordinarily, energy flow is from the generator to the load which the positive resistance represents. The negative resistance means energy is being carried backwards i.e energy is flowing from the load into the generator.

In conclusion, irrespective of the boundary condition of the transmission line, the forward travelling wave always sees an impedance equal to the characteristic impedance while a backward travelling wave sees a negative of the characteristic impedance.

We have thus established that;

$$\frac{V^+}{I^+} = Z_0, \quad I^+ = \frac{V^+}{Z_0} \quad \text{and}$$
$$\frac{V^-}{I^-} = -Z_0, \quad I^- = -\frac{V^-}{Z_0}$$

Therefore, we can now rewrite the voltage and current wave equations as

$$V = V^+e^{-\gamma x} + V^-e^{+\gamma x} \tag{2.34}$$

$$I = \frac{V^+}{Z_0}e^{-\gamma x} - \frac{V^-}{Z_0}e^{+\gamma x} \tag{2.35}$$

2.9 Boundary Conditions

Up until now, we have not defined the boundary conditions of the transmission line, so let us discuss that in this section. We have two spatial locations on the transmission line, one at the generator and the other where an arbitrary load Z_L is connected as shown in figure 2.19.

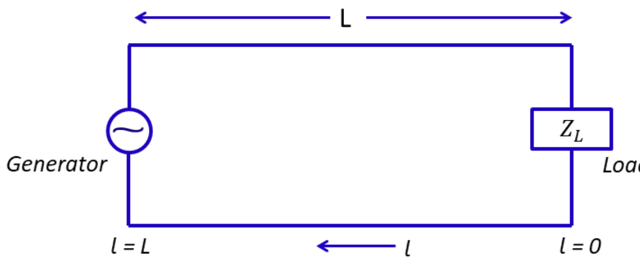


Figure 2.19: The transmission line showing generator and load

To define boundary conditions, we define the spatial distance from the load end, the origin, so that all distances move from the load end towards the generator. So we now have a parameter that moves towards the generator from the load side. At the load point $l = 0$ and at the generator, $l = L$. So $l = -x$ in our transmission line equations, therefore, substituting $l = -x$ in equations 2.34 and 2.35, we have

$$V = V^+e^{+\gamma l} + V^-e^{-\gamma l} \tag{2.36}$$

$$I = \frac{V^+}{Z_0}e^{+\gamma l} - \frac{V^-}{Z_0}e^{-\gamma l} \tag{2.37}$$

Beyond this point, all of our analysis will be done using the wave equations (2.36) and (2.37)

2.10 The Reflection Coefficient

At $l = 0$, $Z = Z_L$ since Z_L terminates the transmission line at this point. Substituting $l = 0$ into Equations (2.36) and (2.37),

$$V = V^+e^{+\gamma(0)} + V^-e^{-\gamma(0)} = V^+ + V^-$$
$$I = \frac{V^+}{Z_0}e^{+\gamma(0)} - \frac{V^-}{Z_0}e^{-\gamma(0)} = \frac{V^+}{Z_0} - \frac{V^-}{Z_0}$$
$$I = \frac{V^+ - V^-}{Z_0}$$

$$\frac{V}{I} = \left(\frac{V^+ + V^-}{1}\right) \times \left(\frac{Z_0}{V^+ - V^-}\right)$$

$$Z_L = \frac{V}{I}\bigg|_{l=0} = Z_0\left[\frac{V^+ + V^-}{V^+ - V^-}\right]$$

Thus,

$$Z_L = Z_0\left[\frac{V^+ + V^-}{V^+ - V^-}\right] \tag{2.38}$$

It is clear from the equation that the load impedance is related to the characteristic impedance and also related to the amplitude of the forward and backward waves. We can simplify further by dividing the top and bottom by V^+, then

$$\begin{aligned} Z_L &= Z_0\left[\frac{\frac{V^+ + V^-}{V^+}}{\frac{V^+ - V^-}{V^+}}\right] \\ &= Z_0\left[\frac{1 + \frac{V^-}{V^+}}{1 - \frac{V^-}{V^+}}\right] \end{aligned} \tag{2.39}$$

Here we see that the absolute values of V^+ and V^- do not matter, instead, the ratio $\frac{V^-}{V^+}$ is what's important. Hence we define a new parameter on which Z_L depends. This parameter is known as the *reflection coefficient*. It can be defined as the ratio of a backward travelling wave to a forward travelling wave on the transmission line.

The reflection coefficient is denoted by Γ[7].

$$\Gamma = \frac{\text{Backward travelling wave}}{\text{Forward travelling wave}}$$

$$\Gamma(l) = \frac{V^-e^{-\gamma l}}{V^+e^{+\gamma l}} \tag{2.40}$$

$\Gamma(l)$ is the reflection coefficient at any point on the line.

At $l = 0$,

$$\begin{aligned} \Gamma(0) &= \frac{V^-e^{-\gamma(0)}}{V^+e^{+\gamma(0)}} \\ &= \frac{V^-}{V^+} \\ &= \Gamma_L \end{aligned} \tag{2.41}$$

Γ_L is the reflection coefficient at the load point on the line and it will be used to represent the reflection coefficient at the load point from the point onwards.

Substituting equation (2.41) into equation (2.40),

$$\Gamma(l) = \Gamma_L\frac{e^{-\gamma l}}{e^{+\gamma l}}$$

$$\Gamma(l) = \Gamma_L e^{-2\gamma l} \tag{2.42}$$

Recall, equation 2.39 was derived at the load point. We would now derive the relationship at any point on the line.

Dividing Equation (2.36) by equation (2.37),

$$\begin{aligned} \frac{V}{I} &= \frac{V^+e^{+\gamma l} + V^-e^{-\gamma l}}{1} \times \frac{Z_0}{V^+e^{+\gamma l} - V^-e^{-\gamma l}} \\ &= Z_0\left(\frac{V^+e^{+\gamma l} + V^-e^{-\gamma l}}{V^+e^{+\gamma l} - V^-e^{-\gamma l}}\right) \end{aligned}$$

Dividing top and bottom by $V^+e^{+\gamma l}$

$$\begin{aligned} \frac{V}{I} &= Z_0\left(\frac{\frac{V^+e^{+\gamma l}}{V^+e^{+\gamma l}} + \frac{V^-e^{-\gamma l}}{V^+e^{+\gamma l}}}{\frac{V^+e^{+\gamma l}}{V^+e^{+\gamma l}} - \frac{V^-e^{-\gamma l}}{V^+e^{+\gamma l}}}\right) \\ &= Z_0\left(\frac{1 + \frac{V^-e^{-\gamma l}}{V^+e^{+\gamma l}}}{1 - \frac{V^-e^{-\gamma l}}{V^+e^{+\gamma l}}}\right) \quad \text{From Equation 2.40} \\ &= Z_0\left(\frac{1 + \Gamma(l)}{1 - \Gamma(l)}\right) \end{aligned}$$

Substituting for $\Gamma(l)$ from equation (2.42), we have the expression at any point on the line.

$$Z(l) = Z_0\left[\frac{1 + \Gamma_L e^{-2\gamma l}}{1 - \Gamma_L e^{-2\gamma l}}\right] \tag{2.43}$$

[7] Γ is capital Gamma, the third letter of the Greek alphabet.

At $l = 0$

$$
\begin{aligned}
Z_L &= Z_0 \left[\frac{1+\Gamma_L e^{-2\gamma(0)}}{1-\Gamma_L e^{-2\gamma(0)}}\right] \\
&= Z_0 \left[\frac{1+\Gamma_L}{1-\Gamma_L}\right]
\end{aligned}
\tag{2.44}
$$

$$
\begin{aligned}
Z_L(1-\Gamma_L) &= Z_0(1+\Gamma_L) \\
Z_L - Z_0 &= (Z_0 + Z_L)\Gamma_L
\end{aligned}
$$

$$
\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{2.45}
$$

The reflection coefficient in equation (2.45) decribes the ratio between reflected and incident voltage. *It is a measure of how much energy is reflected from the transmission line and is related to the terminating impedance of the line and the characteristic impedance.*

Exercises

Ex. 31 — With an appropriate equation of amplitude variation along a transmission line, explain what is meant by Characteristic length.

Ex. 32 — In what unit is the attenuation constant stated in most data sheets?

Ex. 33 — Starting from the differential equation building transmission lines, derive the characteristic impedance relation. What does negative impedance imply from the backward voltage to current ratio?

Ex. 34 — What are the physical significance of the propagation constant(γ), attenuation constant(α) and phase constant(β)?

Ex. 35 — From $V = V^+e^{+\lambda l} + V^-e^{-\lambda l}$ and $I = \frac{V^+}{Z_0}e^{+\lambda l} - \frac{V^-}{Z_0}e^{-\lambda l}$, show that $\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0}$.

Ex. 36 — Explain in details, how the characteristic impedance Z_0, Governs energy flow on the transmission line.

Ex. 37 — When Z_0 is real, A forward traveling wave sees a positive resistance and a backward traveling wave sees a negative resistance, differentiate therefore between“positive and negative resistance in this context”.

Ex. 38 — Show that the load impedance of a transmission line is related to the characteristic impedance and also related to the amplitude of the forward and backward wave.

Ex. 39 — Show mathematically that irrespective of the boundary condition of the transmission line, the forward traveling wave always sees an impedance equal to the characteristic impedance, while the backward traveling wave sees a negative of the characteristic impedance.

Ex. 40 — Explain the role the characteristics impedance of the transmission line play in the reflection coefficient.

Ex. 41 — Explain briefly, the influence of change in load impedance on the reflection coefficient.

Ex. 42 — What is the effect of the length of the transmission line on the reflection coefficient.

Ex. 43 — Define characteristic impedance.

Chapter 3

MAXIMUM POWER TRANSFER AND MATCHED CONDITION

3.1 Objectives

At the end of this chapter, you should be able;

(i) To find the condition necessary to deliver maximum power to the load i.e no reflection on the transmission line.

(ii) To understand the concept of lossless transmission line.

(iii) To understand the concept of low less transmission line.

3.2 Matched Condition of the Transmission Line

Equation (2.45) can be further expressed as

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{V^-}{V^+}$$

It can be observed that when $Z_L = Z_0$, then there is no reflection from the load and hence maximum power is delivered to the generator. That is

$$\Gamma_L = \frac{V^-}{V^+} = 0 \text{ (when } Z_L = Z_0.)$$

Although Z_0 cannot be located on the transmission line, it does govern **power flow** in the transmission line. This condition that $Z_L = Z_0$ is called the *matched condition* of the transmission line. The terminating impedance of the line is matched to the characteristic impedance. This condition is similar to the maximum power transfer theorem in a lumped circuit— where if the load impedance is equal to the conjugate of the generator impedance, maximum power is transferred from the generator to the load.

$$Z_L = Z_0 \quad \text{(Matched Load Condition)}$$

As the wave moves along the line, it sees Z_0. At Z_L, it suddenly sees an impedance discontinuity from Z_0 to Z_L which is like a **steep change** because of that part of the energy that tends to get reflected by the generator on the transmission line. So, for maximum power transfer, the terminating load impedance must be equal to the characteristics impedance otherwise, maximum power transfer will not take place and there will always be reflection.

The reflection coefficient at any point on the transmission line from equation (2.40) is given as:

$$\Gamma(l) = \frac{V^- e^{-\gamma l}}{V^+ e^{\gamma l}} = \frac{\text{Backward wave}}{\text{Forward wave}}$$

Therefore we can define the voltage and current equations at any point on the transmission line with respect to reflection coefficient $\Gamma(l)$. That is

$$V(l) = V^+ e^{\gamma l} + V^- e^{-\gamma l} \tag{3.1}$$

$$I(l) = \frac{V^+}{Z_0} e^{\gamma l} - \frac{V^-}{Z_0} e^{-\gamma l} \tag{3.2}$$

Dividing equation 3.1 by $V^+ e^{\gamma l}$

$$\begin{aligned} \frac{V(l)}{V^+ e^{\gamma l}} &= 1 + \frac{V^- e^{-\gamma l}}{V^+ e^{\gamma l}} \quad \text{But } \Gamma(l) = \frac{V^- e^{-\gamma l}}{V^+ e^{\gamma l}} \\ &= 1 + \Gamma(l) \end{aligned}$$

Therefore the voltage equation becomes

$$V(l) = V^+ e^{\gamma l}(1 + \Gamma(l)) \tag{3.3}$$

This gives the voltage at any point on the line. Similary, dividing equation 3.2 by $\frac{V^+}{Z_0} e^{\gamma l}$

$$\begin{aligned} \frac{I}{\frac{V^+}{Z_0} e^{\gamma l}} &= 1 - \frac{\frac{V^-}{Z_0} e^{-\gamma l}}{\frac{V^+}{Z_0} e^{\gamma l}} \\ &= 1 - [\frac{V^- e^{-\gamma l}}{Z_0}][\frac{Z_0}{V^+ e^{\gamma l}}] \\ &= 1 - \frac{V^- e^{-\gamma l}}{V^+ e^{\gamma l}} \\ &= 1 - \Gamma(l) \end{aligned}$$

Thus the current equation becomes

$$I(l) = \frac{V^+}{Z_0} e^{\gamma l}(1 - \Gamma(l)) \tag{3.4}$$

From lumped circuit analysis, we recall that $Z(l) = \frac{V(l)}{I(l)}$, thus, substituting equations (3.3) and (3.4) we have

$$\begin{aligned} Z(l) &= \frac{V(l)}{I(l)} \\ &= \frac{V^+ e^{\gamma l}(1 + \Gamma(l))}{\frac{V^+}{Z_0} e^{\gamma l}(1 - \Gamma(l))} \end{aligned}$$

$$= V^+ e^{\gamma l}(1+\Gamma(l)) \times \frac{Z_0}{V^+ e^{\gamma l}(1-\Gamma(l))}$$

$V^+ e^{\gamma l}$ cancels out to give,

$$Z(l) = Z_0 \left[\frac{1+\Gamma(l)}{1-\Gamma(l)}\right] \tag{3.5}$$

$Z(l)$ is the impedance measured at any location of the transmission line. It is related to the reflection coefficient at any point and the characteristic impedance. $\Gamma(l)$ is the reflection coefficient at any point on the transmission line. Hence $Z(l)$ and $\Gamma(l)$ have a one-to-one relationship at any point along the transmission line.

3.3 Impedance Transformation Relationship

Recall,

$$\begin{aligned} Z(l) &= \frac{V(l)}{I(l)} \quad \text{Substituting equations (3.1) and (3.2)} \\ &= \frac{V^+ e^{\gamma l} + V^- e^{-\gamma l}}{\frac{V^+}{Z_0} e^{\gamma l} - \frac{V^-}{Z_0} e^{-\gamma l}} \\ &= \frac{V^+ e^{\gamma l} + V^- e^{-\gamma l}}{\frac{V^+ e^{\gamma l} - V^- e^{-\gamma l}}{/Z_0}} \\ &= Z_0 \left[\frac{V^+ e^{\gamma l} + V^- e^{-\gamma l}}{V^+ e^{\gamma l} - V^- e^{-\gamma l}}\right] \end{aligned}$$

Dividing the numerator and denominator by $V^+ e^{\gamma l}$

$$Z(l) = Z_0 \left[\frac{1 + \frac{V^- e^{-\gamma l}}{V^+ e^{\gamma l}}}{1 - \frac{V^- e^{-\gamma}}{V^+ e^{\gamma}}}\right]$$

Where $\Gamma_L = \frac{V^-}{V^+}$ and $\frac{e^{-\gamma l}}{e^{\gamma l}} = e^{-2\gamma l}$

$$Z(l) = Z_0 \left[\frac{1+\Gamma_L e^{-2\gamma l}}{1-\Gamma_L e^{-2\gamma l}}\right] \tag{3.6}$$

But $\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0}$,

$$Z(l) = Z_0 \left[\frac{1 + \left(\frac{Z_L - Z_0}{Z_L + Z_0}\right) e^{-2\gamma l}}{1 - \left(\frac{Z_L - Z_0}{Z_L + Z_0}\right) e^{-2\gamma l}}\right]$$

Multiplying the numerator and denominator by $Z_L + Z_0$ gives

$$\begin{aligned} Z(l) &= Z_0 \left[\frac{(Z_L+Z_0) + \frac{(Z_L-Z_0)}{(Z_L+Z_0)} \times (Z_L+Z_0) \times e^{-2\gamma l}}{(Z_L+Z_0) - \frac{(Z_L-Z_0)}{(Z_L+Z_0)} \times (Z_L+Z_0) \times e^{-2\gamma l}}\right] \\ &= Z_0 \left[\frac{(Z_L+Z_0) + (Z_L-Z_0)e^{-2\gamma l}}{(Z_L+Z_0) - (Z_L-Z_0)e^{-2\gamma l}}\right] \end{aligned}$$

Multiplying the numerator and denominator by $e^{\gamma l}$ gives

$$\begin{aligned} &= Z_0 \left[\frac{(Z_L+Z_0)e^{\gamma l} + (Z_L-Z_0)e^{-2\gamma l} \times e^{\gamma l}}{(Z_L+Z_0)e^{\gamma l} - (Z_L-Z_0)e^{-2\gamma l} \times e^{\gamma l}}\right] \\ &= Z_0 \left[\frac{(Z_L+Z_0)e^{\gamma l} + (Z_L-Z_0)e^{-\gamma l}}{(Z_L+Z_0)e^{\gamma l} - (Z_L-Z_0)e^{-\gamma l}}\right] \\ &= Z_0 \left[\frac{Z_L e^{\gamma l} + Z_0 e^{\gamma l} + Z_L e^{-\gamma l} - Z_0 e^{-\gamma l}}{Z_L e^{\gamma l} + Z_0 e^{\gamma l} - Z_L e^{-\gamma l} + Z_0 e^{-\gamma l}}\right] \end{aligned}$$

Collect like terms

$$Z(l) = Z_0 \left[\frac{Z_L(e^{\gamma l} + e^{-\gamma l}) + Z_0(e^{\gamma l} - e^{-\gamma l})}{Z_L(e^{\gamma l} - e^{-\gamma l}) + Z_0(e^{\gamma l} + e^{-\gamma l})}\right]$$

Before we proceed let us recall that

$$\begin{aligned} \cosh(\gamma l) &= \frac{e^{\gamma l} + e^{-\gamma l}}{2} \\ 2\cosh(\gamma l) &= e^{\gamma l} + e^{-\gamma l} \end{aligned}$$

and

$$\begin{aligned} \sinh(\gamma l) &= \frac{e^{\gamma l} - e^{-\gamma l}}{2} \\ 2\sinh(\gamma l) &= e^{\gamma l} - e^{-\gamma l} \end{aligned}$$

Therefore,

$$\begin{aligned} Z(l) &= Z_0 \left[\frac{2(Z_L \cosh(\gamma l) + Z_0 \sinh(\gamma l))}{2(Z_L \sinh(\gamma l) + Z_0 \cosh(\gamma l))}\right] \\ &= Z_0 \left[\frac{Z_L \cosh(\gamma l) + Z_0 \sinh(\gamma l)}{Z_L \sinh(\gamma l) + Z_0 \cosh(\gamma l)}\right] \end{aligned} \tag{3.7}$$

Equation (3.7) shows that the impedance at any point l is related to the load impedance Z_L and the characteristics impedance Z_0. This enables us to move Z_L to any point along the transmission line. Hence at $l = 0$, we have Z_L but at $l = L$, the input impedance measured at the generator end will not be Z_L. It will depend on Z_L and also the length of the line. So, if the length of the line keeps varying, the impedance you measure at the input end of the line will keep varying.

Hence, if we design a circuit at high frequency and connect a load to the end, depending on the length of the transmission line, the input impedance will keep varying with length. From a circuit perspective, the input impedance is very important, so, at high frequencies, the length connecting the load to the circuit matters as the input impedance seen by the generator varies with length.

From Equation 3.7, we normalize with respect to Z_0 such that, $\frac{Z(l)}{Z_0} = \bar{Z}(l)$ and $\frac{Z_L}{Z_0} = \bar{Z}_L$ which are the normalized values.

Therefore, dividing both sides by Z_0

$$\frac{Z(l)}{Z_0} = \frac{Z_0}{Z_0}[\frac{Z_L \cosh(\gamma l) + Z_0 \sinh(\gamma l)}{Z_L \sinh(\gamma l) + Z_0 \cosh(\gamma l)}] \tag{3.8}$$

$$\bar{Z}(l) = \left[\frac{Z_L \cosh(\gamma l) + Z_0 \sinh(\gamma l)}{Z_0}\right] \times \left[\frac{Z_0}{Z_L \sinh(\gamma l) + Z_0 \cosh(\gamma l)}\right]$$

$$= \left[\frac{Z_L \cosh(\gamma l) + Z_0 \sinh(\gamma l)}{Z_0}\right] \div \left[\frac{Z_L \sinh(\gamma l) + Z_0 \cosh(\gamma l)}{Z_0}\right]$$

$$\begin{aligned}
&= \frac{\frac{Z_L \cosh(\gamma l)+Z_0 \sinh(\gamma l)}{Z_0}}{\frac{Z_L \sinh(\gamma l)+Z_0 \cosh(\gamma l)}{Z_0}} \\
&= \frac{\frac{Z_L}{Z_0}\cosh(\gamma l) + \frac{Z_0}{Z_0}\sinh(\gamma l)}{\frac{Z_L}{Z_0}\sinh(\gamma l) + \frac{Z_0}{Z_0}\cosh(\gamma l)} \\
&= \frac{\frac{Z_L}{Z_0}\cosh(\gamma l) + \sinh(\gamma l)}{\frac{Z_L}{Z_0}\sinh(\gamma l) + \cosh(\gamma l)} \\
&= \frac{\bar{Z}_L\cosh(\gamma l) + \sinh(\gamma l)}{\bar{Z}_L\sinh(\gamma l) + \cosh(\gamma l)} \text{ (normalized impedance)}
\end{aligned} \tag{3.9}$$

That is, $\bar{Z} = \frac{Z}{Z_0}$.

Once again we see that in transmission line calculation, the absolute impedance is not relevant, rather it is the normalized impedance that matters most. Hence, at $Z_0 = 50\Omega$ and $Z_L = 100\Omega$ gives same reflection as $Z_0 = 300\Omega$ and $Z_L = 600\Omega$. Hence the reason why before starting any transmission line calculation, we ask ourselves what the characteristic impedance is, every impedance we have is then normalized to the characteristic impedance. So every calculation done in the transmission line is always done with respect to **normalized impedances**.

We then see that characteristic impedance which is not located anywhere or seen anywhere is always governing the energy flow of the transmission line.

Now if $\bar{Z}_L = 1$, then equation (3.9) reduces to

$$\begin{aligned}
\bar{Z}(l) &= \frac{\bar{Z}_L\cosh(\gamma l) + \sinh(\gamma l)}{\bar{Z}_L\sinh(\gamma l) + \cosh(\gamma l)} \\
&= 1
\end{aligned}$$

Recall,

$$\frac{Z(l)}{Z_0} = \bar{Z}(l) \text{ and } \frac{Z_L}{Z_0} = \bar{Z}_L$$

So, $\frac{Z(l)}{Z_0} = 1$ and $\frac{Z_L}{Z_0} = 1$ means the impedance at every point equal to Z_0, if $Z_L = Z_0$ therefore $Z(l) = Z_0$. So, in this matched condition, the impedance $Z(l)$ measured at any point no longer depend on the length of the line. If a line is terminated by its characteristic impedance, there is no need to factor in the length of the line as every point along the line has an impedance equal to the characteristic impedance and this takes away all the worry about the **line length**.

The relationship of impedance transform will help us give a proper definition of characteristics impedance Z_0 which we know was related to the primary constants of the transmission line.

Characteristic Impedance *is defined as that impedance which if used to terminate the line, the impedance measured at all points along the line will be the same and equal to that terminating impedance*.

At matched conditions, there is no reflected wave. With infinite line length, there can be no reflection as the incident wave never gets to the end to get reflected. Hence, the reflection is zero. However, reflection is zero when the line is terminated with the characteristic impedance. Hence, *the characteristic impedance can also be defined as the input impedance measured with infinite line length*, since with infinite line length, we always have a forward wave and never a reflected wave.

Now we can generalize the impedance transformation relationship. Till now, we have transformed impedance Z_L which is at the load end to $Z(l)$. Nothing special about the load end other than we defining our origin to be at that point. So, the load point was used as our reference point. In general, given a transmission line as shown in figure 3.1 recall from equation (3.7)

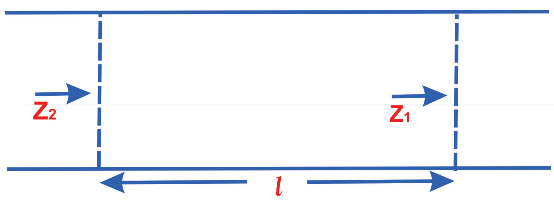


Figure 3.1: General representation of a transmission line

that

$$Z(l) = Z_0\left[\frac{Z_L\cosh(\gamma l) + Z_0\sinh(\gamma l)}{Z_L\sinh(\gamma l) + Z_0\cosh(\gamma l)}\right]$$

Let $Z_L = Z_1$ and $Z(l) = Z_2$.

$$Z_2 = Z_0\left[\frac{Z_1\cosh(\gamma l) + Z_0\sinh(\gamma l)}{Z_1\sinh(\gamma l) + Z_0\cosh(\gamma l)}\right] \tag{3.10}$$

Dividing both sides by Z_0

$$\frac{Z_2}{Z_0} = \frac{Z_0}{Z_0}\left[\frac{Z_1\cosh(\gamma l) + Z_0\sinh(\gamma l)}{Z_1\sinh(\gamma l) + Z_0\cosh(\gamma l)}\right]$$

$$\begin{aligned}
\bar{Z}(2) &= \left[\frac{Z_1\cosh(\gamma l) + Z_0\sinh(\gamma l)}{Z_0}\right] \times \\
&\quad \left[\frac{Z_0}{Z_1\sinh(\gamma l) + Z_0\cosh(\gamma l)}\right]
\end{aligned}$$

$$\begin{aligned}
&= \left[\frac{Z_1\cosh(\gamma l) + Z_0\sinh(\gamma l)}{Z_0}\right] \div \\
&\quad \left[\frac{Z_1\sinh(\gamma l) + Z_0\cosh(\gamma l)}{Z_0}\right]
\end{aligned}$$

$$\begin{aligned}
&= \frac{\frac{Z_1 \cosh(\gamma l)+Z_0 \sinh(\gamma l)}{Z_0}}{\frac{Z_1 \sinh(\gamma l)+Z_0 \cosh(\gamma l)}{Z_0}} \\
&= \frac{\frac{Z_1}{Z_0}\cosh(\gamma l) + \frac{Z_0}{Z_0}\sinh(\gamma l)}{\frac{Z_1}{Z_0}\sinh(\gamma l) + \frac{Z_0}{Z_0}\cosh(\gamma l)} \\
&= \frac{\frac{Z_1}{Z_0}\cosh(\gamma l) + \sinh(\gamma l)}{\frac{Z_1}{Z_0}\sinh(\gamma l) + \cosh(\gamma l)} \\
&= \frac{\bar{Z}_1\cosh(\gamma l) + \sinh(\gamma l)}{\bar{Z}_1\sinh(\gamma l) + \cosh(\gamma l)}
\end{aligned} \tag{3.11}$$

From equation (3.10) we can also invert the relationship to get Z_1 from Z_2 by cross multiplying.

$$Z_2 = Z_0\left[\frac{Z_1\cosh(\gamma l) + Z_0\sinh(\gamma l)}{Z_1\sinh(\gamma l) + Z_0\cosh(\gamma l)}\right]$$

$$\begin{aligned}
&Z_2\left[Z_1\sinh(\gamma l) + Z_0\cosh(\gamma l)\right] \\
&\quad = Z_0\left[Z_1\cosh(\gamma l) + Z_0\sinh(\gamma l)\right]
\end{aligned}$$

$$Z_1Z_2\sinh(\gamma l)+Z_0Z_2\cosh(\gamma l) = Z_0Z_1\cosh(\gamma l)+Z_0^2\sinh(\gamma l)$$

Collect like terms and factorise out Z_1 on the left and Z_0 on the right-hand side of the equation

$$\begin{aligned}
&Z_1Z_2\sinh(\gamma l) - Z_1Z_0\cosh(\gamma l) \\
&\quad = Z_0^2\sinh(\gamma l) - Z_0Z_2\cosh(\gamma l)
\end{aligned}$$

$$Z_1\left[Z_2\sinh(\gamma l)-Z_0\cosh(\gamma l)\right] = Z_0\left[Z_0\sinh(\gamma l)-Z_2\cosh(\gamma l)\right]$$

Make Z_1 the subject of the formula

$$Z_1 = Z_0\left[\frac{Z_0\sinh(\gamma l)-Z_2\cosh(\gamma l)}{Z_2\sinh(\gamma l)-Z_0\cosh(\gamma l)}\right] \tag{3.12}$$

Multiply the numerator and denominator of equation (3.12) by -1.

$$Z_1 = Z_0\left[\frac{Z_2\cosh(\gamma l)-Z_0\sinh(\gamma l)}{Z_0\cosh(\gamma l)-Z_2\sinh(\gamma l)}\right]^{1} \tag{3.13}$$

So, equation (3.13) can be used to transform impedance between Z_1 and Z_2 knowing the positive l direction is from load towards the generator.

Now we go to a special case of the transmission lines.

3.4 Loss-Less Transmission Line

In practice, we want to transfer maximum power from the generator to the load. However, some losses occur in real transmission lines due to **ohmic resistances**. Every effort is usually made to make sure maximum power is delivered to the load by minimizing these losses. Hence, a good transmission line loss should be very small at its operating frequency.

Once we have this condition in practice, then we can make some simplification to the transmission line problem and come up with an idea of what is called a *lossless transmission line* whose ideal loss is zero.

3.4.1 Propagation Constant of Loss-Less Transmission Line

The transmission line has four parameters R, L, G and C all in per unit length. R and G are ohmic values in the primary constants which is the resistance between the two ends of the conductors and the leakage current in the dielectric separating the two conductors respectively. Hence, power-losing elements exist because of the resistance and conductance. The inductance and capacitance exchange energy between themselves but do not result in any loss. Ideally, a line will be lossless if $R = G = 0$.

Substituting R and G as zero in the propagation constant from equation (2.19) given as follows simplifies to,

$$\begin{aligned}\gamma &= \sqrt{(R+\jmath\omega L)(G+\jmath\omega C)}\\ &= \sqrt{(\jmath\omega L)(\jmath\omega C)}\\ &= \jmath\omega\sqrt{LC}\end{aligned}$$

Recall that the propagation constant is a complex quantity represented as,

$$\gamma = \alpha + \jmath\beta$$

Therefore,

$$\alpha + j\beta = \jmath\omega\sqrt{LC}$$

so,

$$\alpha = 0 \quad \text{and} \quad \beta = \omega\sqrt{LC}$$

With no loss, there is no reason for the wave amplitude to reduce on the transmission line since the attenuation constant, $\alpha = 0$. Hence, there is a sustained propagation of electromagnetic waves along the transmission line. We can further simplify the phase constant as follows

$$\begin{aligned}\beta &= \omega\sqrt{LC} \quad \text{But } \beta = \frac{2\pi}{\lambda} \text{ and } \omega = 2\pi f\\ \frac{2\pi}{\lambda} &= 2\pi f\sqrt{LC}\\ \frac{1}{\lambda} &= f\sqrt{LC}^{2}\end{aligned}$$

Recall, $\lambda f = v$ which is the velocity of the wave in the transmission line. Thus

$$\lambda f = \frac{1}{\sqrt{LC}} = v$$

The velocity of the wave in the transmission line is related to the inductance and capacitance of the transmission line. Hence the velocity is fixed once the inductance and capacitance are given.

Can we vary load L and C to vary v? Not really, L and C are coupled. Varying C changes L and varying L changes C to make v a constant of the transmission line. The velocity is decided by the **field condition** in the transmission line and is fixed by the **boundary condition**. For instance, with varying separation between the two conductors, the **mutual inductance** will vary and at the same time, the separation between the two conductors vary thereby varying the capacitance. Hence, L and C of the transmission line are not independent quantities.

3.4.2 Characterisitic Impedance Loss-Less Transmission Line

We then calculate the characteristic impedance on a lossless transmission line which is given as,

$$\begin{aligned}Z_0 &= \sqrt{\frac{R+\jmath\omega L}{G+\jmath\omega C}}\\ &= \sqrt{\frac{\jmath\omega L}{\jmath\omega C}} \quad \text{Since } R = 0 \text{ and } G = 0\\ &= \sqrt{\frac{L}{C}} \quad \text{(real quantity)}\end{aligned}$$

Therefore, for a lossless line, it can be observed that Z_0 is real. We do not have any ohmic resistance or conductance in the line and yet it is at this point we have a characteristic impedance that is real (that is, pure ohmic loss). Hence, the wave which travels forward or backwards always sees a characteristic impedance that is a real impedance like resistance. This makes sense since we have a line carrying only forward waves which will go forever and no energy is reflected. That is, power somewhere is going to get dumped at the load end. So if we have a real quantity for Z_0, it means that power is completely transferred to the line. It does not mean that power is lost in the line. So for a lossless line, if the impedance on the line is measured, something more like resistance will be measured.

Now, we go to a case where little loss is accepted in the transmission line and then find the derivations for γ and Z_0.

[1] Recall, $\cosh -\gamma l = \cosh\gamma l$ and $\sinh -\gamma l = -\sinh -\gamma l$

[2] 2π will cancel out on both sides

3.5 Low-loss Transmission Line

$R \ll \omega L$, $G \ll \omega C$ are the two conditions for a low-loss transmission line. So we derive the propagation constant of this line.

3.5.1 Propagation Constant of Low-loss Transmission Line

Using equation (2.19), we have:

$$\begin{aligned}\gamma &= \sqrt{(R+\jmath\omega L)(G+\jmath\omega C)} \\ &= \sqrt{(\jmath\omega L)(\jmath\omega C)\left(1+\frac{R}{\jmath\omega L}\right)\left(1+\frac{G}{\jmath\omega C}\right)} \\ &= \sqrt{(\jmath\omega L)(\jmath\omega C)\left(1-j\frac{R}{\omega L}\right)\left(1-j\frac{G}{\omega C}\right)^3}\end{aligned} \tag{3.14}$$

The expressions $\left(1-j\frac{R}{\omega L}\right)^{\frac{1}{2}}$ and $\left(1-j\frac{G}{\omega C}\right)^{\frac{1}{2}}$ in equation (3.14) can be expressed as a power series. Recall from Binomial expansion the a power series with fraction power can be expressed as follows where x is the fractional part.

$$(1+x)^{\frac{1}{2}} = {}^{\frac{1}{2}}C_0(1)^{\frac{1}{2}}x^0 + {}^{\frac{1}{2}}C_1(1)^{\frac{1}{2}-1}x^1 + \ldots$$

Where ${}^{\frac{1}{2}}C_0(1)^{\frac{1}{2}}x^0 = 1$ and ${}^{\frac{1}{2}}C_1(1)^{\frac{1}{2}-1}x = {}^{\frac{1}{2}}C_1(1)x$.

Let's expand the combination ${}^{\frac{1}{2}}C_1 x$

$$\begin{aligned}{}^{\frac{1}{2}}C_1 x &= \binom{\frac{1}{2}}{1}x \\ &= \frac{\frac{1}{2}!\,x}{(\frac{1}{2}-1)!\,1!} \\ &= \frac{\frac{1}{2}(\frac{1}{2}-1)!\,x}{(\frac{1}{2}-1)!\,1!} \\ &= \frac{1}{2}x\end{aligned}$$

Thus,

$$(1+x)^{\frac{1}{2}} = 1 + \frac{1}{2}x + \ldots$$

and

$$(1-x)^{\frac{1}{2}} = 1 - \frac{1}{2}x + \ldots$$

Hence from the equation (3.14)

$$jw\sqrt{LC}\left(1-\jmath\frac{R}{wL}\right)^{\frac{1}{2}}\left(1-\jmath\frac{G}{wC}\right)^{\frac{1}{2}}$$

Let $x = \jmath\frac{R}{\omega L}$ and $x' = \jmath\frac{G}{\omega C}$

$$\left(1-\frac{\jmath R}{\omega L}\right)^{\frac{1}{2}} = 1 - \frac{1}{2}\frac{\jmath G}{\omega C} + \ldots$$

and

$$\left(1-\frac{jG}{\omega C}\right)^{\frac{1}{2}} = 1 - \frac{1}{2}\frac{\jmath R}{\omega L} + \ldots$$

[3] Recall $\frac{1}{\jmath} = -\jmath$

Thus, equation (3.14) becomes

$$\begin{aligned}\gamma &= jw\sqrt{LC}\left(1-\jmath\frac{R}{2\omega L}+\ldots\right)\left(1-j\frac{G}{2\omega C}+\ldots\right) \\ &= jw\sqrt{LC}\left[1-\left(\frac{\jmath R}{2\omega L}\right)-\left(\frac{\jmath G}{2\omega C}\right)+\left(\frac{\jmath R}{2\omega L}\right)\left(\frac{\jmath G}{2\omega C}\right)\right]\end{aligned}$$

Recall that two combination terms were used earlier because of the complex variable $\frac{jR}{wL}$ and $\frac{\jmath G}{\omega C}$, the product of $\frac{\jmath R}{2\omega C}$ and $\frac{\jmath G}{2\omega L}$ tends to zero since $R \ll wL$ and $G \ll \omega C$ for a low loss transmission line.

Thus,

$$\begin{aligned}\gamma &= \jmath\omega\sqrt{LC}\left[1-\frac{\jmath R}{2\omega L}-\frac{\jmath G}{2\omega C}\right] \\ &= \jmath\omega\sqrt{LC} - \jmath\frac{R}{2\omega L}\left(\jmath\omega\sqrt{LC}\right) - j\frac{G}{2\omega C}\left(\jmath\omega\sqrt{LC}\right)\end{aligned} \tag{3.15}$$

Expressing $L = \sqrt{L}\times\sqrt{L}$ and $C = \sqrt{C}\times\sqrt{C}$, we have

$$\begin{aligned}\gamma &= \jmath\omega\sqrt{LC} + \frac{R\sqrt{L}\sqrt{C}}{2\sqrt{L}\sqrt{L}} + \frac{G\sqrt{L}\sqrt{C}}{2\sqrt{C}\sqrt{C}} \\ &= \jmath\omega\sqrt{LC} + \frac{1}{2}R\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}}\end{aligned}$$

Recall that $\gamma = \alpha + j\beta$, therefore,

$$\jmath\omega\sqrt{LC} + \frac{1}{2}R\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}} = \alpha + \jmath\beta$$

So,

$$\alpha = \frac{1}{2}R\sqrt{\frac{C}{L}} + \frac{1}{2}G\sqrt{\frac{L}{C}} \tag{3.16}$$

$$\beta = \omega\sqrt{LC} \tag{3.17}$$

Compared to the lossless transmission line where $\alpha = 0$, there is a value for the attenuation constant. Also, we have that $\beta = \omega\sqrt{LC}$ for both cases of lossless and low-loss, which means that the phase constant does not change when we introduce low-loss to the transmission line. Hence, if we are interested in only the phase constant for a low-loss line, it can be treated as a lossless transmission line. However, for a low-loss transmission line $\alpha \neq 0$, there is that small loss so that as one travels with the wave, amplitude reduces slowly at $e^{-\alpha}$.

3.5.2 Characterisitic Impedance of Low-loss Transmission Line

Recall, $Z_0 = \sqrt{\frac{L}{C}}$ for a lossless line, substituting it in equation (3.16), we have,

$$\alpha = \frac{1}{2}R\sqrt{\frac{C}{L}} + \frac{1}{2}G\sqrt{\frac{L}{C}} = \frac{1}{2}\left(\frac{R}{Z_0}+GZ_0\right)$$

Therefore,

$$\gamma = \jmath\omega\sqrt{LC} + \frac{1}{2}\left(\frac{R}{Z_0}+GZ_0\right)$$

So if R and G are known and Z_0 for the lossless line has been determined, we can use $\alpha = \frac{1}{2}(\frac{R}{Z_0}+GZ_0)$ to calculate α for the low-loss line.

$$Z_0 = \sqrt{\frac{R+\jmath\omega L}{G+\jmath\omega C}} = \sqrt{\frac{\jmath\omega L(1-\jmath\frac{R}{\omega L})}{\jmath\omega C(1-\jmath\frac{G}{\omega C})}}$$

$j\omega$ will cancel out, leaving us with

$$Z_0 = \sqrt{\frac{L}{C}}\sqrt{\frac{1-j\frac{R}{\omega L}}{1-j\frac{G}{\omega C}}}$$
$$= \sqrt{\frac{L}{C}}\left(1-\jmath\frac{R}{\omega L}\right)^{\frac{1}{2}}\left(1-\jmath\frac{G}{\omega C}\right)^{-\frac{1}{2}}$$

From Binomial expansion, we have that:

$$Z_0 = \sqrt{\frac{L}{C}}\left(1-\jmath\frac{R}{2\omega L}+\ldots\right)\left(1-\jmath\frac{G}{2\omega C}+\ldots\right)$$

$$= \sqrt{\frac{L}{C}}\left(1-\jmath\frac{R}{2\omega L}-\jmath\frac{G}{2\omega C}+\ldots\textit{(smaller terms)}\right)$$

The characteristic impedance is no more real but complex. Its real value is the same as that of a lossless line. We have a small imaginary part which implies the presence of losses in the transmission line.

From now on, when given a transmission line problem, we assume it to be lossless unless we are told specifically that the line is lossy.

Exercises

Ex. 44 — Show that for a lossless transmission line, the velocity of the wave is a function of L and C i.e. $\left(v = \frac{1}{\sqrt{LC}}\right)$.

Ex. 45 — Given a transmission line with primary constants, $R = 0.00259\Omega/\text{m}$, $L = 2\mu H/\text{m}$, $G = 0$ and $C = 5.56pF/\text{m}$ operating at a frequency of 5kHz, find the characteristic impedance, propagation constant (attenuation and phase constant), velocity of propagation, and wavelength.

Ex. 46 — Prove that for both a lossless and a low-loss transmission line, the phase constant remains the same while the attenuation constant changes.

Ex. 47 — Consider a low-loss transmission line with a characteristic impedance (Z_0) of 50 ohms and a length (l) of 0.2 wavelengths at a frequency of 1 GHz. If the load impedance (Z_L) is 75 ohms, calculate the reflection coefficient (Γ) at the load and the voltage standing wave ratio (VSWR).

Ex. 48 — What is the required resistance, in ohms, for a transmission line to be considered a low-loss transmission line, with parameters $L = 0.4\mu H/m$, $C = 200pF/m$, $G = 0$, and operating at a frequency of 150MHz?

Ex. 49 — Consider a lossless transmission line with a characteristic impedance (Z_0) of 75 ohms and a length (l) of 0.4 wavelengths. The line is terminated with a load impedance (Z_L) of 100 ohms. The operating frequency is 2 GHz. Calculate the voltage and current traveling wave expressions along the transmission line.

Ex. 50 — Consider a lossless transmission line with a characteristic impedance of $Z_0 = 50\Omega$. The forward voltage wave is given by $V+(z,t) = 100\cos\big(2\pi\times10^9 t-\beta z\big)$, where z is the distance along the transmission line, t is time, and β is the phase constant. The phase constant is related to the wavelength λ by $\beta = \lambda 2\pi$. (a) Determine the maximum voltage on the transmission line. (b) Find the maximum current on the transmission line. (c) Calculate the minimum current on the transmission line. (d) Given that the load impedance (Z_L) is 75Ω, calculate the Voltage Standing Wave Ratio (VSWR) on the transmission line.

Ex. 51 — A transmission line with a characteristics impedance of 70ojms is linked to a parallel configuration comprising of a 150 ohms resistor and a 2nF capacitor (a) Determine the VSWR along the line when operating at a frequency of 3MHz (b) Calculate the maximum and minimum resistance observed on the line

Ex. 52 — Consider a lossless transmission line with a characteristic impedance (Z_0) of 50 ohms and a wavelength (λ) of 0.1 meters. If the frequency of the signal traveling on the transmission line is 1 GHz, calculate the phase velocity (v_p) of the signal on the transmission line.

Chapter 4

VOLTAGE AND CURRENT VARIATION IN LOW-LOSS TRANSMISSION LINE

4.1 Objectives

At the end of this chapter, you should be able;

(i) To know the conditions necessary for treating the transmission line as lowloss at a particular frequency.

(ii) To understand the variation of voltage and current along the transmission line.

(iii) To understand the concept of Voltage Standing Wave Ratio(VSWR).

(iv) To know the condition necessary for full reflection at the load end.

(v) To know the relationship between the maximum and minimum resistances and VSWR.

Again, the concept of a low-loss transmission line is a more practical transmission line where $R \ll \omega L$, $G \ll \omega C$.

Let us now study the conditions for treating the transmission line as low-loss at a particular frequency, the variation of voltage and current along a transmission line, the concept of voltage standing wave ratio (VSWR), the condition for full reflection at the load end, the relationship between $R_{\max}$ and VSWR and also the relationship between $R_{\min}$ and VSWR.

The propagation constant in equation (3.14) given as follows states the conditions defined for a low-loss transmission line in terms of the primary constants (R, L, C and G) of the line.

$$\gamma = \frac{1}{2}R\sqrt{\frac{C}{L}} + \frac{G}{2}\sqrt{\frac{L}{C}} + \jmath\omega\sqrt{LC} \tag{4.1}$$

To treat a line as a low-loss transmission line, one will have to express these conditions ($R \ll \omega L$ and $G \ll \omega C$) in terms of the secondary parameters (α and β) since these parameters are readily available on the datasheet. If we can establish a relationship between these parameters (α and β) for the low-loss nature of the line, then we can find out whether a particular line is low-loss at a particular frequency.

The attenuation constant as derived in equation (3.16) is

$$\alpha = \frac{1}{2}R\sqrt{\frac{C}{L}} + \frac{1}{2}G\sqrt{\frac{L}{C}} \tag{4.2}$$

And the phase constant remains the same as that of lossless transmission line

$$\beta = \omega\sqrt{LC} \tag{4.3}$$

Since we have expressed equation (4.1) in terms of primary constants, let us find out under what condition the transmission line can be treated as low-loss.

To do this we further simplify equation (4.2) such that,

$$\begin{aligned}
\alpha &= \frac{1}{2}R\sqrt{\frac{C}{L}} + \frac{1}{2}G\sqrt{\frac{L}{C}} \\
&= \frac{1}{2}R\sqrt{\frac{C}{L}}\sqrt{\frac{L}{L}} + \frac{1}{2}G\sqrt{\frac{L}{C}}\sqrt{\frac{C}{C}}\,\text{[1]} \\
&= \frac{1}{2}R\sqrt{\frac{CL}{L^2}} + \frac{1}{2}G\sqrt{\frac{LC}{C^2}} \\
&= \frac{1}{2}\frac{R}{L}\sqrt{LC} + \frac{1}{2}\frac{G}{C}\sqrt{LC} \\
&= \frac{1}{2}\frac{R}{\omega L}\omega\sqrt{LC} + \frac{1}{2}\frac{G}{\omega C}\omega\sqrt{LC}\,\text{[2]} \\
&= \frac{1}{2}\omega\sqrt{LC}\left(\frac{R}{\omega L} + \frac{G}{\omega C}\right)
\end{aligned}$$

But $\beta = \omega\sqrt{LC}$, so

$$\alpha = \beta\frac{1}{2}\left(\frac{R}{\omega L} + \frac{G}{\omega C}\right) \tag{4.4}$$

Recall that for low-loss, $R \ll \omega L$ and $G \ll \omega C$, thus

$$\frac{R}{\omega L} \ll 1 \text{ and } \frac{G}{\omega C} \ll 1$$

Therefore, $\frac{1}{2}\left(\frac{R}{\omega L} + \frac{G}{\omega C}\right) \ll 1$ and $\alpha = \beta \times$ something far smaller than 1.

Hence, for a low-loss transmission line, $\alpha \ll \beta$ where $\beta = \frac{2\pi}{\lambda}$. Considering a wave which travels a distance of one wavelength along the transmission line, its phase, β becomes 2π, that is, $\beta = 2\pi$, then the amplitude of the wave varies by

$$e^{-\alpha x} = e^{-\alpha\lambda}$$

Where $\lambda = \frac{2\pi}{\beta}$

$$e^{-\alpha x} = e^{-\alpha\dfrac{2\pi}{\beta}}$$

Since for a low-loss transmission line $\alpha \ll \beta$, therefore the quantity $e^{-\alpha\frac{2\pi}{\beta}} \approx 1$. Hence, amplitude reduction is very small as the original amplitude is close to the final amplitude. So, a line can be treated as a low-loss transmission line if the change

[2]The numerator and denominator of the first term is multiplied by $\sqrt{\frac{L}{L}}$ and the second term by $\sqrt{\frac{C}{C}}$,

[2]The numerator and denominator is multiplied by ω

in amplitude over one wavelength is negligibly small. Let negligibly small be 1 percent change from the original or starting amplitude. Then

$$\frac{\alpha 2\pi}{\beta} \approx \frac{1}{100}, e^{-0.01} = 0.99005 \text{ of the initial value}$$

The wave amplitude only reduces by $1 - 0.99005 = 0.00995$ or 1 per cent. That means the wave amplitude only reduces by 1 per cent after a distance of 1 wavelength.

A line can be treated as a low-loss transmission line if $\alpha \ll \beta$ depending on the given frequency but if the frequency changes such that $\alpha \geq \beta$, the line cannot be treated as a low-loss transmission line.

Example 4.1.1 Analysis of a Low-Loss transmission line

Let's say we have a transmission line with $L = 0.25\mu H/m$, $C = 100pF/m$, $G = 0$, what is the resistance of the transmission line so that the line can be treated as a low-loss transmission line given the frequency of operation as $100MHz$?

Solution

$$L = 0.25\mu H/m, C = 100pF/m, G = 0$$

$$\begin{aligned}\beta &= 2\pi f\sqrt{LC} \\ &= 2\pi \times 10^8\sqrt{(0.25 \times 10^{-6}) \times (100 \times 10^{-12})} \\ &= \pi rad/m\end{aligned}$$

For a low-loss transmission line, taking 1 percent of $\beta, \alpha = \frac{1}{100}\beta = \frac{\pi}{100}$

.

Given that, $G = 0$

$$\begin{aligned}\alpha &= \frac{1}{2}R\sqrt{\frac{C}{L}} + \frac{1}{2}G\sqrt{\frac{L}{C}} \\ &= \frac{1}{2}R\sqrt{\frac{C}{L}}\end{aligned}$$

$$\frac{\pi}{100} = \frac{1}{2}R\sqrt{\frac{100 \times 10^{(-12)}}{0.25 \times 10^{(-6)}}}$$

$$R = \pi\Omega/m.$$

If $R \leq \pi\Omega/m$, the line is low-loss at $f = 100MHz$. If the frequency changes, the line may not satisfy this low-loss condition, hence we have to check again.

Therefore, unless specifically told that a line is a lossy line, we are at liberty to treat the line as a lossless line since the phase constant as we have seen for lossy and lossless lines are the same. Also, the characteristics impedance of a low-loss line is almost real and the same as the characteristics impedance of a lossless line. Hence for all transmission line problems, we consider it lossless and thus,

$$Z_0 = \sqrt{\frac{\jmath\omega L}{\jmath\omega C}} = \sqrt{\frac{L}{C}} = \text{real}, \qquad \gamma = \jmath\beta = \jmath\omega\sqrt{LC}$$

4.2 Voltage And Current Variation on a Lossless Transmission Line

Revisiting the general voltage and current expression of the transmission line with origin at the load point, we have:

$$V = V^+e^{-\gamma x} + V^-e^{\gamma x} \tag{4.5}$$

$$I = \frac{V^+}{Z_0}e^{-\gamma x} - \frac{V^-}{Z_0}e^{\gamma x} \tag{4.6}$$

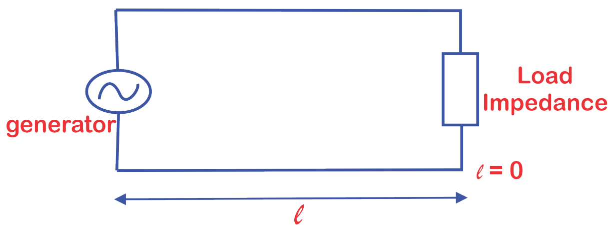


Figure 4.1: Wave moving towards the Generator

For a lossless transmission line at any point l, substituting $\gamma = \jmath\beta$ and $x = -l$ into equation (4.5), we have;

$$\begin{aligned}V(l) &= V^+e^{\jmath\beta l} + V^-e^{-\jmath\beta l} \\ \frac{V(l)}{V^+e^{\jmath\beta l}} &= \left(1 + \frac{V^-e^{-\jmath\beta l}}{V^+e^{\jmath\beta l}}\right)^3 \\ V(l) &= V^+e^{\jmath\beta l}\left(1 + \frac{V^-}{V^+}e^{-\jmath 2\beta l}\right)\end{aligned}$$

Thus,

$$V(l) = V^+e^{\jmath\beta l}(1 + \Gamma_L e^{-\jmath 2\beta l}) \quad \text{where } \Gamma_L = \frac{V^-}{V^+} \tag{4.7}$$

Similarly, substituting $\gamma = \jmath\beta$ and $x = -l$ into equation (4.6):

$$\begin{aligned}I(l) &= \frac{V^+}{Z_0}e^{\jmath\beta l} - \frac{V^-}{Z_0}e^{-\jmath\beta l} \\ \frac{I(l)}{\frac{V^+}{Z_0}e^{\jmath\beta l}} &= \left(1 - \frac{\frac{V^-}{Z_0}e^{-\jmath\beta l}}{\frac{V^+}{Z_0}e^{\jmath\beta l}}\right)^4 \\ \frac{I(l)}{\frac{V^+}{Z_0}e^{\jmath\beta l}} &= \left(1 - \frac{V^-}{V^+}e^{-\jmath 2\beta l}\right)\end{aligned}$$

Thus,

$$I(l) = \frac{V^+}{Z_0}e^{\jmath\beta l}\left(1 - \Gamma_L e^{-\jmath 2\beta l}\right) \quad \text{where } \Gamma_L = \frac{V^-}{V^+} \tag{4.8}$$

From equations (4.7) and (4.8), $V = \mathfrak{f}(V^+, V^-)$ and $I = \mathfrak{f}(I^+, I^-)$ means V and I are a superposition of forward and backward travelling wave, which is then a *standing wave*. Recall the reflection coefficient at the load end is given as

$$\Gamma_L = |\Gamma_L|e^{j\phi_L} \tag{4.9}$$

where ϕ_L is the phase of the reflection coefficient at the load end. Then, we can write down the $V(l)$ and $I(l)$ explicitly in terms of the magnitude of the reflection coefficient and the phase.

[3] The result is divided through by $V^+e^{\jmath\beta l}$

[4] The result is divided through by $\frac{V^+}{Z_0}e^{\jmath\beta l}$

Therefore substituting equation (4.9) into equations (4.7) and (4.8) respectively, we will get;

$$\begin{aligned} V(l) &= V^{+}e^{\jmath\beta l}(1+|\Gamma_L|e^{\jmath\phi_L}.e^{-\jmath 2\beta l}) \\ &= V^{+}e^{\jmath\beta l}(1+|\Gamma_L|e^{j(\phi_L-2\beta l)}) \end{aligned} \tag{4.10}$$

And

$$\begin{aligned} I(l) &= \frac{V^{+}}{Z_0}e^{j\beta l}(1-|\Gamma_L|e^{j\phi_L}.e^{-\jmath 2\beta l}) \\ &= \frac{V^{+}}{Z_0}e^{j\beta l}(1-|\Gamma_L|e^{j(\phi_L-2\beta l)}) \end{aligned} \tag{4.11}$$

As we move from load to generator, l increases positively, making $\phi_L - 2\beta l$ more and more negative. In the complex plane, when a phase gets more negative, it means we are moving in a clockwise direction as shown in figure 4.2.

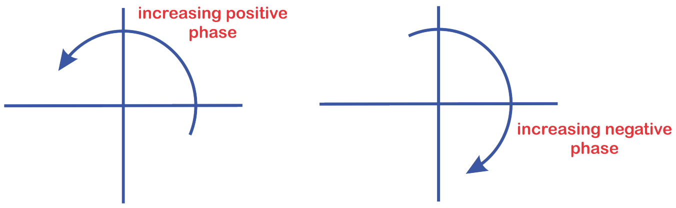


Figure 4.2: Clockwise and Anticlockwise movement of distance, l

So by moving towards the generator, the phase becomes more and more negative, while the amplitude of $|\Gamma_L|$ remains constant. The total voltage is scaled by the vector sum of the real term (1) and complex term $|\Gamma_L|e^{j(\phi_L-2\beta l)}$. Hence, we have a summation of a real vector, whose magnitude is 1, plus a complex term as shown in figure 4.3.

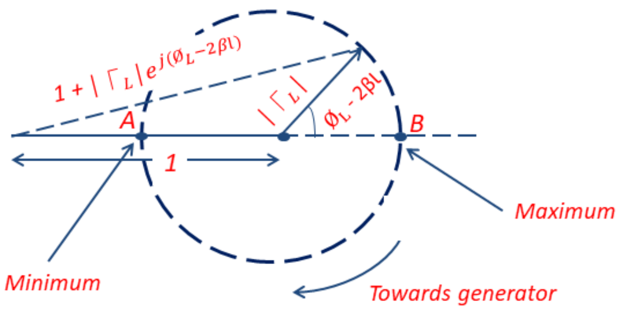


Figure 4.3: Variation of the $1 + |\Gamma_L|e^{j(\phi_L-2\beta l)}$ term in the complex plane (Argand diagram)

The magnitude $1 + |\Gamma_L|e^{j(\phi_L-2\beta l)}$ varies as we move along the transmission line.

From the [5]Argand diagram in figure 4.3, the dash line circle is the variation plot of the term $|\Gamma_L|e^{j(\phi_L-2\beta l)}$ whose magnitude is $|\Gamma_L|$ and phase $\phi_L - 2\beta l$. The term $|1+|\Gamma_L|e^{j(\phi_L-2\beta l)}|$ is the resolved vector of vector 1 and vector $|\Gamma_L|e^{j(\phi_L-2\beta l)}$. The same Argand diagram can be developed for the term $|1 - |\Gamma_L|e^{j(\phi_L-2\beta l)}|$ where the vector $|\Gamma_L|e^{j(\phi_L-2\beta l)}$ points in the negative direction.

For even multiples of π, that is, $2n\pi$ where n is a positive integer then $\phi_L - 2\beta l = 2n\pi$ and $e^{j(\phi_L-2\beta l)} = 1$. This is the maximum value of the resolved vector, that is, $1 + |\Gamma_L|$ and corresponds to point B in figure 4.3.

For odd multiples of π, that is, $(2n + 1)\pi$, then $\phi_L - 2\beta l = (2n+1)\pi$ and $e^{j(\phi_L-2\beta l)} = -1$. This is the minimum value of the resolved vector, that is, $1-|\Gamma_L|$ and corresponds to point A in figure 4.3. Similarly, $|1-|\Gamma_L|e^{j(\phi_L-2\beta l)}|$ resolves to a maximum at odd multiples of π and a minimum at even multiples of π.

Thus, $|1 + |\Gamma_L|e^{j(\phi_L-2\beta l)}|$ resolves to a maximum when $(\phi_L-2\beta l) = 0, 2\pi, 4\pi, \ldots$ and $|1+|\Gamma_L|e^{j(\phi_L-2\beta l)}|$ resolves to a minimum when $(\phi_L - 2\beta l) = \pi, 3\pi, 5\pi \ldots$. This is the voltage term and implies the voltage is maximum when the phase equals even multiples of π and minimum when the phase equals odd multiple of π.

This is opposite to the behaviour of the current relationship with the term $1 - |\Gamma_L|e^{j(\phi_L-2\beta l)}$. At odd multiples of π we have a maximum and at even multiples of π we have a minimum.

Thus we have established that when $e^{j(\phi_L-2\beta l)} = 1$, the voltage is maximum and the current is minimum and when $e^{j(\phi_L-2\beta l)} = -1$, the voltage is minimum and the current is maximum. So at the same location along the transmission line, maximum voltage corresponds to minimum current, and minimum voltage corresponds to maximum current. This is opposite to what we observed in a lumped circuit, where the maximum voltage point corresponds to maximum current and minimum voltage corresponds to minimum current.

The maximum voltage and current does not occur at the same point on the transmission line, they are staggered in space. Whenever there is a maximum voltage, there is a minimum current and vice versa. So the standing wave of voltage and current are shifted with respect to each other in space on the transmission line as shown in figure 4.4.

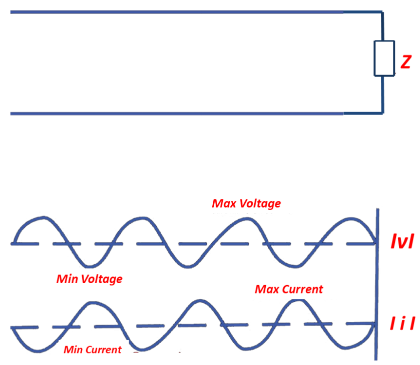


Figure 4.4: Voltage and current variations for lossless transmission line in the direction of increasing l

The plot of the voltage and current variations in figure 4.4 shown that maximum voltage corresponds to minimum current and minimum voltage corresponds to maximum current.

Example 4.2.1 Maximum and Minimum voltage and current

A lossless transmission line has 75Ω characteristic impedance. The line is terminated in a load impedance of $50 - j100\Omega$. *The*

[5]Argand Diagram refers to a geometric plot of complex numbers as points $z = x+iy$ using the x-axis as the real axis and the y-axis as the imaginary axis. Such plots are named after Jean-Robert Argand (1768-1822), although they were first described by Norwegian-Danish land surveyor and mathematician Casper Wessel (1745-1818.)

maximum voltage measured on the line is 100V.

(a) Find the maximum and minimum current, and the minimum voltage on the line.

(b) At what distance from the load are the voltage and the current maximum?

Solution

We are given $Z_0 = 75\Omega$, $Z_L = 50 - j100\Omega$, *and* $|V|_{\max}= 100V$

(a) Maximum voltage and maximum and minimum current

The first step is to calculate the reflection coefficient.

$$\begin{aligned}\Gamma_L &= \frac{Z_L - Z_0}{Z_L + Z_0} \\ &= \frac{50 - j100 - 75}{50 - j100 + 75} \\ &= 0.64\angle - 65.3^o\end{aligned}$$

Where $|\Gamma_L|= 0.64$ *and* $\phi_L = -65.3^o$. *Since we are told that the line is lossless, then the reflection coefficient will remain constant throughout the line.*

We know that the maximum voltage which we will see on the line $|V|_{\max}= |V^+|(1 + |\Gamma_L|)$, *and we are given* $|V|_{\max}$, *we also know now the reflection coefficient, so we can find out the magnitude of the incident wave* $|V^+|$.

So,

$$\begin{aligned}100 &= |V^+|(1 + 0.64) \\ |V^+| &= \frac{100}{1.64} \\ &= 60.8V\end{aligned}$$

Let us determine the maximum current and minimum voltage and current. Given the characteristic impedance and the maximum voltage, we can find the maximum current as

$$\begin{aligned}|I|_{\max} &= \frac{|V|_{\max}}{Z_0} \\ &= \frac{100}{75} \\ &= 1.33A\end{aligned}$$

The minimum current can be expressed in terms of minimum voltage as well so that

$$\begin{aligned}|I|_{\min} &= \frac{|V|_{\min}}{Z_0} \quad \textit{But } |V|_{\min} = |V^+|(1 - |\Gamma_L|) \\ &= \frac{|V^+|}{Z_0}(1 - |\Gamma_L|) \\ &= \frac{60.8}{75}(1 - 0.64) \\ &= 0.29A\end{aligned}$$

And we can determin the minimum voltage using either $|V|_{\min}= Z_0|I|_{\min}$ *or* $|V|_{\min}= |V^+|(1 - |\Gamma_L|)$. *Using the first equation, then the minimum voltage is*

$$|V|_{\min} = Z_0|I|_{\min} = 75 \times 0.29 = 21.88V$$

(b) Location of maximum voltage and current

Lastly, let us locate the distance from the load where the current and voltage are maximum. Let us note that when the two traveling waves have a constructive interference, we have a voltage maximum which corresponds to the current minimum. Also, when two traveling waves have a destructive interference, we have a voltage minimum which corresponds to the current maximum.

Recall that the phase difference between the forward and the backward wave is given by $\phi_L - 2\beta l$. *If* $\phi_L - 2\beta l$ *equals even multiples of* π, *the two waves will have a constructive interference and the voltage will be maximum. Similarly, when* $\phi_L - 2\beta l$ *equals odd multiples of* π, *the two waves will have a destructive interference and a voltage minimal will be observed. So mathematically, voltage maximum occurs at* $\phi_L - 2\beta l = \pm 2m\pi$, *where m is an integer quantity.*

$$\begin{aligned}\phi_L - 2\beta l_{\max} &= -2m\pi \\ \textit{Thus,} \quad 2\beta l_{\max} &= \phi_L + 2m\pi \\ \textit{But, } \beta &= \frac{2\pi}{\lambda} \\ 2 \times \frac{2\pi}{\lambda} \times l_{\max} &= \phi_L + 2m\pi \\ l_{\max} &= \frac{(\phi_L + 2m\pi)\lambda}{4\pi}\end{aligned}$$

We know $\phi_L = -65.3^o = -1.14rad$

$$l_{\max} = \frac{(-1.14 + 2m\pi)\lambda}{4\pi}$$

When $m = 1, 2, 3, \ldots \quad l_{\max} = 0.41\lambda, 0.91\lambda, 1.41\lambda$ *and so on.*

Similarly, to find the current minimum, let us recall that the current minimum occurs at a point where the voltage is minimum and vice versa. They are apart by a distance of $\frac{\lambda}{4}$, *so the current maximum is at*

$$l_{\max}(\textit{voltage}) \pm \frac{\lambda}{4}$$

Using the negative and positive sign will both give correct answers. Here, we will use the negative sign, that is, $l_{\max} - \frac{\lambda}{4}$. *When* $m = 1, 2, 3, \ldots$, *therefore* $|I|_{\min}$ *occurs at* $0.16\lambda, 0.66\lambda, 1.16\lambda$ *and so on.*

4.3 Concept of Voltage Standing Wave Ratio (VSWR)

Let us suppose we are interested in finding the impedance at an arbitrary distance l on a lossless transmission line. Then the ratio of voltage and current would be

$$\frac{V(l)}{I(l)} = \frac{V^+ e^{j\beta l}(1 + |\Gamma_L|e^{j(\phi_L - 2\beta l)})}{\frac{V^+}{Z_0} e^{j\beta l}(1 - |\Gamma_L|e^{j(\phi_L - 2\beta l)})}$$

Therefore,

$$\frac{V(l)}{I(l)} = Z_0 \frac{1 + |\Gamma_L|e^{j(\phi_L - 2\beta l)}}{1 - |\Gamma_L|e^{j(\phi_L - 2\beta l)}}$$

For maximum voltage, $\phi_L - 2\beta l =$ Even multiples of π that is $0, 2\pi, 4\pi, 6\pi, \ldots$. For minimum voltage, $\phi_L - 2\beta l =$

Odd multiples of π that is $\pi, 3\pi,\ 5\pi,\ \ldots$. Thus the euler formular

$$e^{j(\phi_L - 2\beta l)} = \cos(\phi_L - 2\beta l) + jsin(\phi_L - 2\beta l)$$

Since $\sin(\phi_L - 2\beta l) = 0$ for all multiples of π or $n\pi$, where n is positive integer, therefore the euler relation is real. So when a voltage is maximum or minimum, $\frac{V(l)}{I(l)} = Z_0\times$ (real quantity).

So irrespective of what the line is terminated with, at maximum voltage, the impedance at that point is always real. Even if the line is terminated with a complex impedance, the maximum voltage is always real. If we move to the point of maximum voltage, the impedance that will be measured at that point will always be real. Similarly, at the location where the voltage is minimum, the impedance will again be real.

In conclusion, on a transmission line, wherever there is a maximum voltage the impedance measured at that point will be real, same for the point of minimum voltage.

Mathematically, the maximum impedance is expressed as

$$\begin{aligned} Z_{\max} &= \frac{|V_{\max}|}{|I_{\min}|} \\ &= R_{\max} \\ &= Z_0 \left\{ \frac{1 + |\Gamma_L|}{1 - |\Gamma_L|} \right\} \end{aligned} \tag{4.12}$$

Similarly, at the location of minimum voltage, the minimum impedance is expressed as

$$\begin{aligned} Z_{\min} &= \frac{|V_{\min}|}{|I_{\max}|} \\ &= R_{\min} \\ &= Z_0 \left\{ \frac{1 - |\Gamma_L|}{1 + |\Gamma_L|} \right\} \end{aligned} \tag{4.13}$$

Once load impedance, characteristic impedance Z_0 as well as the reflection coefficient Γ_L are known, then we can calculate the maximum and minimum value of impedance that we see on the transmission line. As we move along the transmission line, the impedance will vary. However, there is a bound on the upper and lower value of this impedance; the lowest according to equation (4.12) and the highest according to equation (4.13).

At high frequencies, with the knowledge of the voltage standing wave on the transmission line, the measurement of phase is very difficult and complicated. One can measure the amplitude of a signal reliably but the measurement of phase is rather uncertain. So at high frequencies, we estimate the phase not in a direct manner but rather an indirect one by measuring only the magnitude quantities. As we have seen the phase of the signal in time gets translated into the space phase because the total phase we see on a wave is a combination of space and time relationship between the two waves, the forward and backward wave.

These waves are related to the time phase and the space phase, $\omega t + \beta x$, where ωt is the temporal phase and βx is the spatial phase. Since the total phase governs the location of maximum and minimum on the standing wave, noting the location of maximum and minimum on the transmission line, one can estimate the phase of the signal. Now we define a parameter for the standing wave which is a parameter of only amplitude variation of the transmission line. This quantity is called the **Voltage Standing Wave Ratio (VSWR)** *which is the measure of the relative contribution of the reflected wave with respect to the incident wave*. If the reflected wave is zero, there is no standing wave but only a travelling wave and if the reflected wave is the same as the transmitted wave, then we have a completely developed standing wave. So the interference of the two waves, the forward and backward wave is going to give the variation of $R_{\max}$ to $R_{\min}$ and $R_{\min}$ to $R_{\max}$. So we define a quantity relating $V_{\max}$ and $V_{\min}$ as the VSWR.

Then

$$\frac{V_{\max}}{V_{\min}} = VSWR, \rho$$

VSWR is a very important quantity because without carrying out phase measurements, we can measure this quantity on a transmission line. Recall that the reflection coefficient is a complex quantity, so a complete knowledge of the reflection coefficient is achieved when both its amplitude and phase are known. However, the quantity which we are defining is known as VSWR or ρ which is measured only by amplitude.

Therefore, by measuring the maximum and minimum values of the standing wave, we get;

$$\begin{aligned} VSWR, \rho &= \frac{|V|_{\max}}{|V|_{\min}} \\ &= \frac{|V^+|\{1 + |\Gamma_L|\}}{|V^+|\{1 - |\Gamma_L|\}} \\ VSWR &= \frac{1 + |\Gamma_L|}{1 - |\Gamma_L|} \end{aligned} \tag{4.14}$$

4.4 Condition for Full Reflection at Load End

Since the transmission line is lossless, the reflection coefficient at the load is;

$$|\Gamma_L| = \frac{Z_L - Z_0}{Z_L + Z_0}$$

Recall that Z_0 is real for a lossless transmission line and Z_L can be any complex impedance as the load impedance. So, $|\Gamma_L| = \frac{Z_L - Z_0}{Z_L + Z_0}$ is always less than 1 that is:

$$|\Gamma_L| \leq 1 \quad \text{(for a passive load)}$$

What this means is that $|\Gamma_L|$ is the relative amplitude of the reflected wave with respect to the incident wave and its value no more than 1 implies that we do not have any energy source at the load point so the transmitted energy can only be reflected. Therefore, the amplitude of the reflected wave has to always be less than or equal to the amplitude of the incident wave.

4.4.1 Conditions for which $|\Gamma_L| = 1$

Case 1: $Z_L = 0$ (the line is short-circuited at the load end)

$$\begin{aligned} \Gamma_L &= \frac{Z_L - Z_0}{Z_L + Z_0} \\ &= \frac{0 - Z_0}{0 + Z_0} = -1 \\ \Rightarrow & |\Gamma_L| = 1 \end{aligned}$$

Case 2: $Z_L = \infty$ (Open circuit line)

$$\begin{aligned} \Gamma_L &= \frac{Z_L - Z_0}{Z_L + Z_0} \\ &= \frac{1 - \frac{Z_0}{Z_L}}{1 + \frac{Z_0}{Z_L}} = \frac{1 - 0}{1 + 0} = 1 \\ \Rightarrow & |\Gamma_L| = 1 \end{aligned}$$

Case 3: $Z_L = \jmath X$ (Pure reactance)

$$\begin{aligned}\Gamma_L &= \frac{Z_L - Z_0}{Z_L + Z_0} \\ &= \frac{\jmath X - Z_0}{\jmath X + Z_0} = \text{complex} \\ \Rightarrow &|\Gamma_L| = 1\end{aligned}$$

So the three cases under which $|\Gamma_L| = 1$ are

(i) When the line is short-circuited.

(ii) When the line is open-circuited.

(iii) When the line is terminated in a pure reactance

Since there is no energy-absorbing circuit at the load end of the line, short circuits, open circuits and ideal reactance cannot absorb power. So, whatever power the generator takes to the load end, it has no option but to return all the power in the reflected waveform.

$$VSWR = \frac{1 + |\Gamma_L|}{1 - |\Gamma_L|}$$

The VSWR will always be greater than 1, the best case is $|\Gamma_L| = 0$ (zero reflection) which gives $\rho = 1$ whereas at $|\Gamma_L| = 1$ (full reflection), $\rho = \infty$.

So $1 \leq \rho \leq \infty$, where $\rho = 1$ represents no reflected wave on the transmission line that is, full power is transferred to the load and $\rho = \infty$ means so much reflection on the transmission line or less efficiency of power transfer to the load. So, every circuit design tries to make VSWR as close to 1 as possible. A higher value of VSWR indicates more mismatch on the transmission line or a higher value of the reflected wave on the transmission line.

So VSWR is one of the most important quantities at high frequencies. When designing a circuit, we try to make sure the VSWR on the transmission line is close to one as possible, making sure the circuit is efficiently transferring power to the load end of the line. Once ρ or VSWR is defined, we can relate it to the maximum and minimum impedance that can be seen on the transmission line.

4.5 Maximum and Minimum Resistance

From equations (4.12) and (4.13), we can express the maximum and minimum impedance in terms of ρ[6] such that

$$\begin{aligned}R_{\max} &= Z_0\left(\frac{1 + |\Gamma_L|}{1 - |\Gamma_L|}\right) \\ &= Z_0\rho\end{aligned} \tag{4.15}$$

$$\begin{aligned}R_{\min} &= Z_0\left(\frac{1 - |\Gamma_L|}{1 + |\Gamma_L|}\right) \\ &= \frac{Z_0}{\rho}\end{aligned} \tag{4.16}$$

This implies that for any transmission line terminated with an arbitrary impedance, if we move to a point where the voltage is maximum or minimum, then we know the value of the impedance at that location. This is because with $V_{\max}$ and $V_{\min}$ known, we can determine ρ and Z_0 is also known beforehand, hence $R_{\max} = Z_0\rho$ and $R_{\min} = \frac{Z_0}{\rho}$.

We know that the impedance known at any point in a transmission line can always be transformed to any other point using the impedance transformation relation. Hence knowing the location at voltage maximum or minimum, since the impedance value is known there, then we can transform the impedance value towards the load to get the load impedance. This essentially opens up a measurement technique for unknown impedance.

At high frequencies, if we have a complex impedance, its measurement is quite tedious because we cannot measure phase accurately. Now we have a mechanism for measuring the phase indirectly. As we have mentioned, the phase gets reflected into the standing wave pattern on the location of voltage maximum and minimum. So if we measure VSWR and the location of $V_{\max}$ and $V_{\min}$ we can always transform $R_{\max}$ or $R_{\min}$ to the location of the load which is nothing but the load impedance. So if we transform from the load impedance to the voltage maximum or minimum we would get $R_{\max}$ and $R_{\min}$. In summary, ig the location of $V_{\max}$ from the load is known and the value of impedance at that location, and the distance of the load from that point are also known, then the transformation of $R_{\max}$ or $R_{\min}$ to load end should give load impedance.

When we discuss the application of transmission lines we will learn that this technique is used for measuring complex impedances. Then we will explicitly derive the expression for the unknown impedance which is terminated to a transmission line.

Example 4.5.1 Maximum and Minimum Resistance

A 50Ω transmission line is connected to a parallel combination of a 100Ω resistance and a 1nF capacitance.

(a) Find the VSWR on the line at a frequency of 2MHz.

(b) Find the maximum and minimum resistance seen on the line.

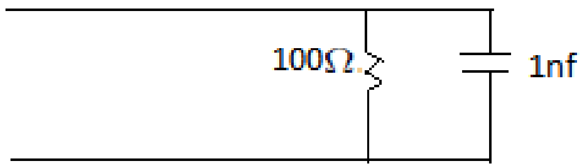


Figure 4.5: Circuit diagram

Solution

(i) First, we will find out the complete impedance for the parallel combination of the resistor and capacitor, given as $Z_L = R \parallel X_C$, that is

$$\begin{aligned}Z_L &= \frac{R \times X_c}{R + X_C} \quad \textit{Where } X_C = \frac{1}{j\omega C} \\ &= \frac{R \times \frac{1}{j\omega C}}{\frac{j\omega RC + 1}{j\omega C}} \\ &= \frac{R}{j\omega C} \times \frac{j\omega C}{j\omega RC + 1} \\ &= \frac{R}{j\omega RC + 1}\end{aligned}$$

[6]Recall from equation (4.14) that $\rho = \frac{1+|\Gamma_L|}{1-|\Gamma_L|}$

Given $f = 2MHz$

$$\begin{aligned}\omega &= 2\pi f\\ &= 2\pi \times 2 \times 10^6\\ &= 1.256 \times 10^7 rad/s\end{aligned}$$

Therefore the load impedance is

$$\begin{aligned}Z_L &= \frac{100}{1 + j1.2566}\\ &= 38.77 - j48.7\Omega\\ &= 62.3\angle - 51.5^o\Omega\end{aligned}$$

(ii) *Next, we find the refection coefficient.*
Reflection coefficient at load point,

$$\begin{aligned}\Gamma_L &= \frac{Z_L - Z_0}{Z_L + Z_0}\\ &= \frac{38.77 - j48.7 - 50}{38.77 - j48.7 + 50}\\ &= 0.13 - j0.47\\ &= 0.5\angle - 74.23^o\end{aligned}$$

We get $|\Gamma_L| = 0.5$ *and* $\phi_L = -74.23^o$

(iii) **(a) VSWR**

Now, we know the reflection coefficient, we can find out VSWR, ρ*, given as;*

$$\begin{aligned}\rho &= \frac{1 + |\Gamma_L|}{1 - |\Gamma_L|}\\ &= \frac{1 + 0.5}{1 - 0.5}\\ &= 3\end{aligned}$$

(iv) **(b) Maximum and Minimum Resistance**

$R_{\max}$ *and* $R_{\min}$ *can be calculated from equations* (4.15) *and* (4.16) *such that*

$$R_{\max} = 50 \times 3 = 150\Omega$$
$$R_{\min} = \frac{50}{3} = 16.67\Omega$$

Exercises

Ex. 53 — Measurements are made at 5kHz on a 0.5-mile long transmission line. The results show that the characteristic impedance is $94\angle - 23.2^o$, the total attenuation is 0.06Nepers/mi and the phase shift between the input and output is 8°. Find the R, L, G and C per mile.

Ex. 54 — A 600Ω transmission line is 150m long operates at 400kHz with $\alpha = 2.4 \times 10^{-3}$Nepers/m and $\beta = 0.0212$rad/m and supplies a load impedance $Z_L = 424.3\angle 45^o\Omega$. Find the length of line in wavelength, $\Gamma_L, \Gamma(L)$, and $Z(L)$. For a received voltage $V_L = 50\angle 0^o$V, find $V(L)$, the position on the line where the voltage is a maximum and the value of $|V|_{\max}$.

Ex. 55 — Prove that the attenuation constant is comparable to the phase constant (that is $\alpha \approx \beta$) for a low-loss transmission line.

Ex. 56 — In a lossless transmission line, how does the absence of resistance impact the power transfer capability?

Ex. 57 — Calculate the characteristic impedance of a lossless transmission line with inductance $L = 0.2$ mH/m and capacitance $C = 80$ nF/m.

Ex. 58 — What is the significance of the surge impedance loading in a lossless transmission line?

Ex. 59 — In a lossless transmission line, explain how the voltage and current vary along the line for a given signal frequency.

Ex. 60 — How does the absence of resistance in a lossless transmission line affect signal reflection at the line termination?

Ex. 61 — Calculate the wavelength of a signal with a frequency of 100 MHz propagating in a lossless transmission line with a velocity factor of 0.8.

Ex. 62 — Explain how the absence of resistive losses influences the phase velocity of signals in a lossless transmission line.

In the previous sections, we investigated the Standing Wave Pattern on a transmission line and also an important parameter which is the voltage standing wave ratio (VSWR) *which is the ratio of the maximum voltage seen on the transmission line to the minimum voltage seen on the transmission line*.

Let us now study impedance transformation in a lossless transmission line and then establish some of the very important characteristics of impedance transformation on a lossless transmission line. Thereafter, we will proceed to an important calculation of power transfer to the load and also derive the expression for V^+.

4.6 Impedance Transformation on a Lossless Transmission Line

Recall that, the impedance transformation relationship for any point is given as:

$$^{1}Z(l) = Z_0 \left\{ \frac{Z_L \cosh(\gamma l) + Z_0 \sinh(\gamma l)}{Z_0 \cosh(\gamma l) + Z_L \sinh(\gamma l)} \right\}$$

$$^{2}\bar{Z}(l) = \left\{ \frac{\bar{Z}_L \cosh(\gamma l) + \sinh(\gamma l)}{\bar{Z}_L \sinh(\gamma l) + \cosh(\gamma l)} \right\}$$

if $\bar{Z}_L = 1$ that is, $Z_L = Z_0$ and $\bar{Z}(l) = 1 \rightarrow Z(l) = Z_0$

For a lossless Transmission Line,$\gamma = j\beta$ where

$$\cosh \gamma l = \frac{e^{\gamma l} + e^{-\gamma l}}{2} \quad and \quad \sinh \gamma l = \frac{e^{\gamma l} - e^{-\gamma l}}{2}$$

so,

$$\cosh \gamma l = \cosh(j\beta l) = \frac{e^{j\beta l} + e^{-j\beta l}}{2} = \cos(\beta l)$$

And

$$\begin{aligned} \sinh \gamma l &= \sinh(j\beta l) \\ &= \frac{e^{j\beta l} - e^{-j\beta l}}{2} \\ &= j\left(\frac{e^{j\beta l} - e^{-j\beta l}}{2j}\right) \\ &= j\sin(\beta l) \end{aligned}$$

Therefore impedance transformation relationship for a lossless transmission line is given as;

$$\bar{Z}(l) = \left\{ \frac{\bar{Z}_L \cos(\beta l) + j\sin(\beta l)}{\cos(\beta l) + j\bar{Z}_L \sin(\beta l)} \right\} \tag{4.17}$$

Z_0 is a real quantity for a lossless transmission line same as:

$$\bar{Z}(l) = \left\{ \frac{\bar{Z}_L \cos(\beta l) + j\sin(\beta l)}{\cos(\beta l) + j\bar{Z}_L \sin(\beta l)} \right\} \tag{4.18}$$

With this impedance transformation relationship, we can establish some very important characteristics of the transmission line. When we move on a transmission line a distance of $\frac{\lambda}{2}$ (half wavelength), the voltage standing wave characteristics repeat themselves, so if we take a ratio of voltage and current at a certain location, we expect that this characteristic will repeat at $\frac{\lambda}{2}$.

[1] $Z(l)$ is the impedance at any point on the line
[2] $\bar{Z}(l)$ is the normalized impedance at any point on the line

Moving $\frac{\lambda}{4}$ from the point of $R_{\max}$ and $R_{\min}$, should bring something interesting. Similarly, if we terminate the line into its characteristics impedance, the impedance measured at any point equals the characteristics impedance. So we have three important points we can draw from this.

Let's show with proof these three characteristics.

4.6.1 Normalized impedance value repeat every $\frac{\lambda}{2}$ distance

Let's say at l we have maximum or minimum impedance, then moving $l = \frac{\lambda}{2}$ from that point we should get the same impedance again!

At location l, Impedance $= \bar{Z}(l)$,

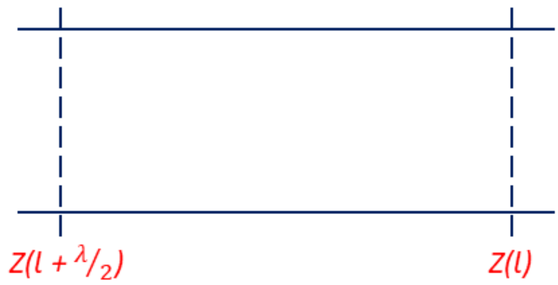


Figure 4.6: Z(l) over distance $\frac{\lambda}{2}$

At location $\left(l + \frac{\lambda}{2}\right)$, Impedance $= \bar{Z}\left(l + \frac{\lambda}{2}\right)$ substituting $\left(l + \frac{\lambda}{2}\right)$ for l in equation (4.17), we get:

$$\bar{Z}\left(l + \frac{\lambda}{2}\right) = \left\{ \frac{\bar{Z}_L \cos\left(\beta\left(l + \frac{\lambda}{2}\right)\right) + j\sin\left(\beta\left(l + \frac{\lambda}{2}\right)\right)}{\cos\left(\beta\left(l + \frac{\lambda}{2}\right)\right) + j\bar{Z}_L \sin\left(\beta\left(l + \frac{\lambda}{2}\right)\right)} \right\}$$

Since [7]$\beta = \frac{2\pi}{\lambda}$, therefore,

$$\begin{aligned} \beta\left(l + \frac{\lambda}{2}\right) &= \frac{2\pi}{\lambda}\left(l + \frac{\lambda}{2}\right) \\ &= \frac{2\pi}{\lambda} l + \pi \\ &= \beta l + \pi \end{aligned}$$

$$\bar{Z}\left(l + \frac{\lambda}{2}\right) = \left\{ \frac{\bar{Z}_L \cos(\beta l + \pi) + j\sin(\beta l + \pi)}{\cos(\beta l + \pi) + j\bar{Z}_L \sin(\beta l + \pi)} \right\}$$

But

$$\cos(\beta l + \pi) = -\cos(\beta l) \text{ and } \sin(\beta l + \pi) = -\sin(\beta l)$$

$$\begin{aligned} \bar{Z}\left(l + \frac{\lambda}{2}\right) &= \left\{ \frac{-\bar{Z}_L \cos(\beta l) - j\sin(\beta l)}{-\cos(\beta l) - j\bar{Z}_L \sin(\beta l)} \right\} \\ &= \left\{ \frac{\bar{Z}_L \cos(\beta l) + j\sin(\beta l)}{\cos(\beta l) + j\bar{Z}_L \sin(\beta l)} \right\} \\ &\equiv \bar{Z}(l) \end{aligned}$$

This is the same as the original formula obtained for $\bar{Z}(l)$.

Hence $\bar{Z}\left(l + \frac{\lambda}{2}\right) = \bar{Z}(l)$, which proves that impedance repeat itself at $\frac{\lambda}{2}$.

In other words, no matter the length of the transmission line, modulus $\frac{\lambda}{2}$ is special information that is available from the impedance relationship.

[7] β represents phase constant

4.6.2 Normalized impedance inverts at every $\frac{\lambda}{4}$ distance

If we move a distance of $\frac{\lambda}{4}$, we move from a maximum to minimum impedance and vice versa.

At location l, Impedance $= \bar{Z}(l)$.

Then at location $(l+\frac{\lambda}{4})$, Impedance $= \bar{Z}(l+\frac{\lambda}{4})$. Substituting $(l+\frac{\lambda}{4})$ for l in equation 4.17, we get:

$$\bar{Z}\left(l+\frac{\lambda}{4}\right) = \left\{\frac{\bar{Z}(l)\cos\left(\beta(l+\frac{\lambda}{4})\right) + jsin(\beta(l+\frac{\lambda}{4}))}{\cos\left(\beta(l+\frac{\lambda}{4})\right) + j\bar{Z}(l)\sin\left(\beta(l+\frac{\lambda}{4})\right)}\right\}$$

Where

$$\begin{aligned}\beta(l+\frac{\lambda}{4}) &= \beta l + \frac{2\pi}{\lambda}\cdot\frac{\lambda}{4}\\ &= \beta l + \frac{\pi}{2}\end{aligned}$$

$$\bar{Z}\left(l+\frac{\lambda}{4}\right) = \left\{\frac{\bar{Z}(l)\cos\left(\beta l+\frac{\pi}{2}\right) + j\sin\left(\beta l+\frac{\pi}{2}\right)}{\cos\left(\beta l+\frac{\pi}{2}\right) + j\bar{Z}(l)\sin\left(\beta l+\frac{\pi}{2}\right)}\right\}$$

But $\cos\left(\beta l+\frac{\pi}{2}\right) = -\sin(\beta l), \sin\left(\beta l+\frac{\pi}{2}\right) = \cos(\beta l)$

$$\bar{Z}\left(l+\frac{\lambda}{4}\right) = \left\{\frac{-\bar{Z}(l)\sin(\beta l) + j\cos(\beta l)}{-\sin(\beta l) + j\bar{Z}(l)\cos(\beta l)}\right\}$$

$$\begin{aligned}\bar{Z}\left(l+\frac{\lambda}{4}\right) &= \frac{j}{j}\left\{\frac{\cos(\beta l) + j\bar{Z}(l)\sin(\beta l)}{\bar{Z}(l)\cos(\beta l) + j\sin(\beta l)}\right\}\\ &= \frac{1}{\left\{\frac{\bar{Z}(l)\cos(\beta l)+j\sin(\beta l)}{\cos(\beta l)+j\bar{Z}(l)\sin(\beta l)}\right\}}\\ &= \frac{1}{\bar{Z}(l)}\end{aligned}$$

We can draw an observation that for every $\frac{\lambda}{4}$ distance moved, the normalized impedance inverts itself[8]. If the absolute impedance is inverted, it is **admittance**. The normalized impedance does not have a unit, it is dimensionless. So if I have an impedance greater than Z_0 along a transmission line, after $\frac{\lambda}{4}$ distance, it becomes less than Z_0 because the normalized impedance is the inverse of the one at the previous location. So at $\frac{\lambda}{4}$ distance, impedance inverts after another $\frac{\lambda}{4}$ again it re-inverts becoming the original impedance, which is equivalent to repeating itself at $\frac{\lambda}{2}$ distance.

Hence, when we talk about the periodicity of the impedance on the transmission line, at $\frac{\lambda}{2}$ the absolute or normalized impedance repeats itself, whereas every distance of $\frac{\lambda}{4}$, the normalized impedance inverts itself. Next on is the study of impedance matching characteristics; this property is used extensively for finding out the impedance transformation which can match impedance on transmission lines.

4.6.3 The matching condition characteristics

If the transmission line is terminated by its characteristic impedance, then the impedance seen at every point on the transmission line is equal to the characteristic impedance. So if $Z_L = Z_0$. This implies that $\bar{Z}_L = \frac{Z_L}{Z_0} = 1$. Then equation (4.17) becomes:

$$\begin{aligned}\bar{Z}(l) &= \left\{\frac{\bar{Z}_L\cos\beta l + j\sin\beta l}{\cos\beta l + j\bar{Z}_L\sin\beta l}\right\}\\ &= \left\{\frac{\cos\beta l + j\sin\beta l}{\cos\beta l + jsin\beta l}\right\}\\ &= 1\end{aligned}$$

Therefore, irrespective of the length of the transmission line, if the line is terminated in its characteristic impedance, then the impedance seen at every point on the transmission line is equal to the characteristic impedance. This means that once a line is terminated in its characteristics impedance, there is no need to compute the impedance transformation on the line, we can use any length of transmission and the impedance will always be the same along the length of the line.

$Z_L = Z_0$ that is, when the reflection coefficient is zero, there is no reflected wave on the transmission line and we have only a forward travelling wave on the transmission line. The forward-travelling waves always see an impedance that is equal to the characteristic impedance. This result is not new, it is what we had discussed earlier when we talked about transmission lines and that was, if a line is terminated in its characteristic impedance, the impedance seen at every point on the transmission line is equal to the characteristic impedance.

So these are the three very important characteristics of a lossless transmission line.

(i) Impedance transformation repeats at every $\frac{\lambda}{2}$ distances.

(ii) Normalized impedance inverts at every $\frac{\lambda}{4}$ distance.

(iii) Finally if the line is terminated in its characteristics impedance, the impedance seen at every point of the line is equal to the characteristics impedance.

With this understanding of impedance transformation, now we can go to the power transfer calculation of the transmission line.

4.7 Power Transfer on Transmission Line

Initially, our idea was to transfer power from the generator to load effectively. In the lossless case, it should be that all generator power is completely transferred to the load. However, we have seen that if the impedance is not equal to the characteristic impedance, then there will always be a reflection on the transmission line and whatever energy the generator supplies, part of this energy will get reflected back to the generator. Now when we talk about matching condition or maximum power transfer condition, there are two cases to consider;

(i) When the power is generated by the generator, it should be maximally transferred to the load.

(ii) When the reflected power comes back to the generator, the generator is not capable of absorbing power, so when the reflected power comes back with a different amplitude and phase, it negatively affects the generator's performance (destructive interference). So it is desirable that the generator should not see any power coming back at it.

For these two reasons that the generator power should be completely delivered to the load and that no reflected power should come back to the generator, we must make sure always that the impedance that the generator sees is always equal to the characteristic impedance. We will study these two issues later but

[8] Note the word **normalized**, it is not absolute impedance

let's study a general case and suppose we have a transmission line that is connected to a generator at one end and a load at the other end. Then, *how much power will be delivered to the load?* Again using the voltage and current equation, we can write down power at the location of the load that is, the power delivered to the load will be given as follows.

At load end, $L = 0$ therefore, $e^{-2\beta(l)} = e^{-2\beta(0)}$

$$\begin{aligned} V(0) &= V^+ \left\{1 + \Gamma_L e^{-2\beta(0)}\right\} \\ &= V^+ \{1 + \Gamma_L\} \\ I(0) &= \frac{V^+}{Z_0}\left\{1 - \Gamma_L e^{-2\beta(0)}\right\} \\ &= \frac{V^+}{Z_0}\{1 - \Gamma_L\} \end{aligned}$$

So from the general voltage and current relationship on the transmission line, we have found out the voltage and current values at the load end.

Z_0 is real for a lossless line, thus the conjugate of the current equation at the load end is:

$$I^*(0) = \frac{V^{+(*)}}{Z_0}\{1 - \Gamma_L\}$$

Power delivered to load.

$$\begin{aligned} P &= \frac{1}{2}\mathfrak{Re}\{V(0)I^*(0)\} \\ &= \frac{1}{2}\mathfrak{Re}\left\{V^+(1+\Gamma_L) \times \frac{V^{+(*)}}{Z_0}(1-\Gamma_L)\right\} \end{aligned}$$

$$P = \frac{1}{2}\frac{|V^+|^2}{Z_0}\left\{1 - |\Gamma_L|^2\right\}^9 \tag{4.19}$$

Γ_L is a real value since it is the absolute value taken here.

Recall that,

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0}.$$

Once load impedance is known, the reflection coefficient at the load end is known. Then we can calculate the modulus of the reflection coefficient so that the power delivered to the load can be calculated if V^+ is known. *How do we find V^+?* From the relationship here, if we know the amplitude of the forward travelling wave as well as the load impedance, we can calculate the power from the circuit perspective point of view. We can use a different argument in arriving at the same answer. That is on the transmission line the power supplied from the generator, in the form of a travelling wave goes towards the load and we already said that travelling waves always see impedance equal to characteristics impedance. So if the travelling wave has amplitude V^+, it is as if this wave is supplying the power to Z_0 in the lossless case. So one can say now that a wave having amplitude V^+ going forward direction sees impedance Z_0.

Power carried by forward wave:

$$P_{\text{for}} = \frac{1}{2}\frac{|V|^{+2}}{Z_0}$$

[9] Given $V = a + jb \qquad V^* = a - jb$
Then, $V \times V^* = (a+jb)(a-jb) = a^2 - jab + jab + (jb)(-jb) = a^2 + b^2$
$|V| = \sqrt{a^2+b^2}$ thus, $|V|^2 = a^2 + b^2$
This implies that $|V|^2 = V \times V^*$

When this gets to the load, part of the energy will get reflected back, and this backward travelling wave has amplitude V^- which also sees characteristics impedance.

The power carried by this wave:

$$P_{\text{ref}} = \frac{1}{2}\frac{|V^-|^2}{Z_0}$$

which is the power reflected by the load in the backward wave.

Therefore,

$$\begin{aligned} \text{Power delivered to load, } P \\ &= \text{Power transferred to load} \\ &\quad - \text{Power reflected} \\ &= \frac{1}{2}\frac{|V^+|^2}{Z_0} - \frac{1}{2}\frac{|V^-|^2}{Z_0} \\ &= \frac{1}{2}\left\{\frac{|V^+|^2}{Z_0} - \frac{|V^-|^2}{Z_0}\right\} \\ &= \frac{1}{2}\frac{|V^+|^2}{Z_0}\left\{1 - \left|\tfrac{V^-}{V^+}\right|^2\right\} \end{aligned}$$

Recall that $\Gamma_L = \frac{V^-}{V^+}$.

$$P_L = \frac{1}{2}\frac{|V^+|^2}{Z_0}\left\{1 - {\Gamma_L}^2\right\}.$$

So when we compute the power transfer on a transmission line, it is either done using the circuit theory concept or the wave theory concept. Thus we have found the real part that is, the power actually supplied to the load. One then asks *how do we calculate for complex part or put differently, how do we compute the complex power at any location on the transmission line?.*

If we calculate the power flow along any point on the transmission line not necessarily at the load end, what will that indicate? To test that theory, let's get the voltage and current at any arbitrary location on the transmission line that is,

$$\begin{aligned} V(l) &= V^+ e^{j\beta l}\left\{1 + \Gamma_L e^{-j2\beta l}\right\}, \\ I(l) &= \frac{V^+}{Z_0}e^{j\beta l}\left\{1 - \Gamma_L e^{-j2\beta l}\right\} \end{aligned}$$

Complex Power at location l is thus given as:

$$\begin{aligned} P(l) &= \frac{1}{2}(VI^*) \\ &= \frac{1}{2}\left(V^+ e^{j\beta l}\left\{1 + \Gamma_L e^{-j2\beta l}\right\}\right) \\ &\quad \times \left(\frac{V^+}{Z_0}e^{-j\beta l}\left\{1 - \Gamma_L e^{j2\beta l}\right\}\right) \\ &= \frac{1}{2}\frac{|V^+|^2}{Z_0}\left(1 + \Gamma_L e^{-j2\beta l}\right)\left(1 - \Gamma e^{j2\beta l}\right) \\ &= \frac{1}{2}\frac{|V^+|^2}{Z_0}\left(1 - \Gamma_L e^{j2\beta l} + \Gamma_L e^{-j2\beta l} - |\Gamma_L|^2\right) \\ &= \frac{1}{2}\frac{|V^+|^2}{Z_0}\left\{1 - |\Gamma_L|^2 - \Gamma_L\left(e^{j2\beta l} - e^{-j2\beta l}\right)\right\} \end{aligned}$$

Recall, $\jmath 2\sin\theta = e^{\jmath\theta} - e^{\jmath\theta}$, so,

$$\begin{aligned} P(l) &= \frac{1}{2}\frac{|V^+|^2}{Z_0}\left\{1 - |\Gamma_L|^2 - j2\Gamma_L\sin(2\beta l)\right\} \\ &= \frac{1}{2}\frac{|V^+|^2}{Z_0}\left\{1 - |\Gamma_L|^2 - \mathfrak{Im}\{\Gamma_L e^{-j2\beta l}\}\right\} \end{aligned} \tag{4.20}$$

Thus, given the complex power, the real part is $1 - |\Gamma_L|^2$ which is the resistive power, and the imaginary part is $j2\Gamma_L \sin(2\beta l) = \mathfrak{Im}\{\Gamma_L e^{-j2\beta l}\}$ which is the reactive power.

Therefore, at any location of the transmission line, the power is complex, the interesting thing to note is that the resistive power at any point is equal to the power which we earlier calculated that is delivered to the load end. So for a lossless line, the resistive power at any point along the line is the same as the power which is delivered to the load[10]. Hence at any point on the transmission line, power transfer of the resistive part is completely delivered to the load and as such the resistive power which is the actual power flow should be independent of any location on the line.

However, the reactive power is a function of l, which shows the energy stored at different locations along the transmission line. Now we have two parts when we calculate the power on a transmission line. *There is a resistive power which is a measure of power flow that ultimately gets delivered to the load and the reactive power which measures the amount of energy storage at different locations along the line and depends on the value of voltage and current at that location.* Recall that the voltage and current equations are standing waves and thus the energy stored also varies at different locations along the transmission line. Therefore, in conclusion, the reactive power will vary at different locations across the line while the resistive power which is the power delivered to the load will be independent of the location of the transmission line.

Example 4.7.1 Power Flow on a Lossless Line

A generator with voltage $V_g = 300\angle 0^o V$ and internal resistance, $Z_g = 50\Omega$ is connected to a load, $Z_L = 75\Omega$ through a 50-Ω lossless transmission line of length $l = 0.15\lambda$.

(a) Compute the input impedance of the line at the generator, Z_{in}.

(b) Compute the time-average power delivered to the line, P_{in}.

(c) Compute the time-average power delivered to the load, P_L. How does P_{in} compare to P_L? Explain.

Solution

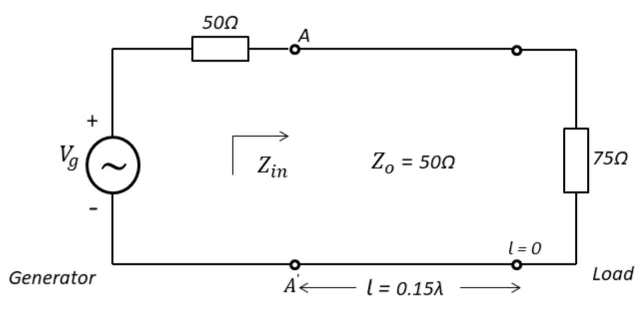


Figure 4.7: Circuit Diagram of Worked example

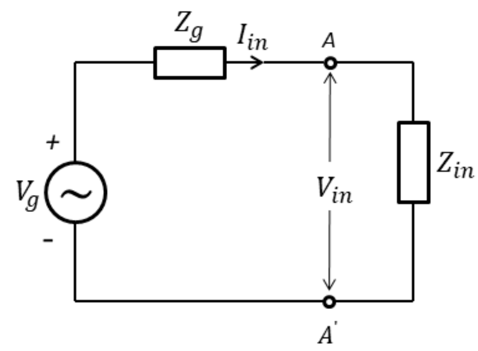


Figure 4.8: Transformed Circuit Diagram of Worked example

From the transmission line shown in figure 4.7, to find Z_{in} we will use the impedance transformation relation in equation (4.17) *such that*

$$\begin{aligned} Z_{in} &= Z_0 \left[\frac{Z_L \cos\beta l + jZ_0 \sin\beta l}{Z_0 \cos\beta l + jZ_L \sin\beta l}\right] \\ &= 50 \left[\frac{75\cos(0.3\pi) + j50\sin(0.3\pi)}{50\cos(0.3\pi) + j75\cos(0.3\pi)}\right]^{11} \\ &= (41.25 - j16.35)\Omega \end{aligned}$$

The time-average power delivered to the line is the power at point AA' in figure 4.7 which is given as

$$P_{in} = \frac{1}{2}\mathfrak{Re}\left\{V_{in} I_{in}^*\right\}$$

The voltage and current at point AA' can be found using the voltage and current variations in a lossless transmission line from equations (4.7) *and* (4.8) *but $|V^+|$ is not known. Instead, we turn to some circuit laws and analyze the transformed circuit shown in figure 4.8*[12].

Thus we can determine V_{in} and I_{in} from the circuit diagram is figure 4.8 as

$$\begin{aligned} I_{in} &= \frac{V_g}{Z_g + Z_{in}} = \frac{300\angle 0^o}{50 + (41.25 - j16.35)} \\ &= 3.24\angle 10.16^o A \\ V_{in} &= I_{in} Z_{in} = 3.24\angle 10.16^o (41.25 - j16.35) \\ &= 143.6\angle -11.46^o V \end{aligned}$$

Thus, the power delivered to the line is

$$\begin{aligned} P_{in} &= \frac{1}{2}\mathfrak{Re}\{V_{in} I_{in}^*\} \\ &= \frac{1}{2}\mathfrak{Re}\{143.6\angle -11.46^o \times 3.24\angle -10.16^o\} \\ &= \frac{143.6 \times 3.24}{2}\cos(21.62^o) \\ &= 216W \end{aligned}$$

Now, let us determine the power delivered to the load end, P_L, given as

$$P_L = \frac{1}{2}\mathfrak{Re}\{V_L I_L^*\}$$

[10] This can be explained with the line of thought that the resistive power finally gets to the load if the line is lossless since there is no absorption of power at any point along the line and so all the resistive power gets to the load because the load is the part where you have resistive component and power can be absorbed in that location.

[11] $\beta l = \frac{2\pi}{\lambda} \times 0.15\lambda = 0.3\pi$

[12] This approach will be dealt with appropriately in section 4.8

How do we determine V_L and I_L? *We can determine V_L and I_L using equations* (4.7) *and* (4.8) *where $l = 0$ such that we get*

$$V(0) = V_L = V^+\{1+\Gamma_L\}$$
$$I(0) = I_L = \frac{V^+}{Z_0}\{1+\Gamma_L\}$$

So we need to evaluate V^+ using the same voltage equation but with $l = L = 0.15\lambda$, that is

$$\begin{aligned}
V_{in} &= V^+ e^{j\beta l}\{1+\Gamma_L e^{-2j\beta l}\}\\
V^+ &= \frac{V_{in}}{e^{j\beta l}\left(1+\left(\frac{Z_L-Z_0}{Z_L+Z_0}\right)e^{-2j\beta l}\right)}\\
&= \frac{143.6\angle-11.46^o}{e^{j0.3\pi}\left(1+\left(\frac{75-50}{75+50}\right)e^{-j0.6\pi}\right)}\\
&= \frac{143.6\angle-11.46^o}{e^{j0.3\pi}\left(1+0.2e^{-j0.6\pi}\right)}\\
&= 150\angle-54^o V
\end{aligned}$$

Thus the values for V_L and I_L are

$$\begin{aligned}
V_L &= V^+\{1+\Gamma_L\} = 150\angle-54^o(1+0.2)\\
&= 180\angle-54^o V\\
I_L &= \frac{V^+}{Z_0}\{1+\Gamma_L\} = \frac{150\angle-54^o}{50}(1-0.2)\\
&= 2.4\angle-54^o A
\end{aligned}$$

Now we determine the power delivered to the load which is

$$\begin{aligned}
P_L &= \frac{1}{2}\mathfrak{Re}\{V_L I_L^*\}\\
&= \frac{1}{2}\mathfrak{Re}\{180\angle-54^o \times 2.4\angle 54^o\}\\
&= 216W
\end{aligned}$$

We can see that $P_{in} = P_L$ is expected because the line is lossless so the input power should equal the delivered power at the load point.

4.8 Evaluation of V$^+$

So far, our transmission line analysis (voltage and current equations, impedance transformation, power transfer analysis and so on) has been derived in terms of V^+. V^+ is a final arbitrary constant in the solution of the differential equation of the transmission line that we are yet to evaluate. It is the amplitude of the incident voltage, which we have assumed is known a priori. Now the question is *how do we evaluate V^+?*.

Considering the circuit in figure 4.9, Z_L is transformed to $Z_L^{'}$, from position BB' to AA'. $Z_L^{'}$ is the impedance of the load as seen by the generator end. With Z_L transformed to $Z_L^{'}$, the whole circuit is reduced to a lumped circuit as shown in figure 4.9.

Considering the lumped circuit, the voltage and current equations are[13]

$$V_A = \frac{Z_L^{'}}{Z_L^{'}+Z_s}.V_s \tag{4.21}$$
$$I_A = \frac{V_A}{Z_L^{'}} \tag{4.22}$$

[13]The voltage and current relations take the limit as the size of the circuit tends to zero so it is valid for any arbitrary frequency.

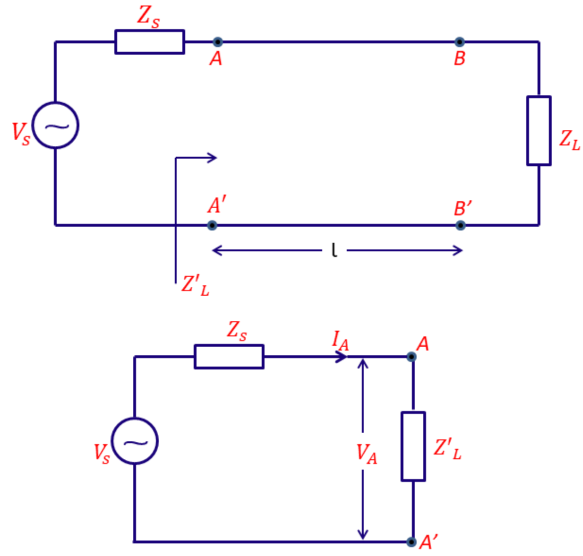


Figure 4.9: Transformation of load impedance from Z_L to $Z_L^{'}$

We can then determine the voltage and current at the transformed location using the transmission line equations, such that,

$$V_A = V(l) = V^+ e^{j\beta l}\left\{1+\Gamma_L e^{-j2\beta l}\right\}$$
$$I_A = I(l) = \frac{V^+}{Z_0}e^{j\beta l}\left\{1-\Gamma_L e^{-j2\beta l}\right\}$$

So we know now the value of V_A and I_A from the lumped element side of view and the transmission line element side of view. We can equate the two to get,

$$V_A = \frac{Z_L^{'}}{Z_L^{'}+Z_s}.V_s = V^+ e^{j\beta l}\left\{1+\Gamma_L e^{-j2\beta l}\right\} \tag{4.23}$$

$$I_A = \frac{V_A}{Z_L^{'}} = I(l) = \frac{V^+}{Z_0}e^{j\beta l}\left\{1-\Gamma_L e^{-j2\beta l}\right\} \tag{4.24}$$

So solving equation 4.24 making V^+ the subject of the formula, we get the expression of V^+ as follows.

$$V^+ = \frac{Z_L^{'} V_s e^{-j\beta l}}{(Z_S+Z_L^{'})(1+\Gamma_L e^{-j2\beta l})}, \tag{4.25}$$

$\alpha, \beta, l, \Gamma_L$ and Z_0 are all known values, so now we have a known value for V^+.

From here V^+ can be determined and then substituted into the power equation and so power delivered to the load can be calculated or power at any point along the line. This is a complete solution to voltage and current on the transmission line.

Summary

Table 4.1 shows the transmission line parameters for lossless and low-loss conditions that have been discussed and derived.

Other parameters that are derived are as follows and they apply equally to lossless and low-loss conditions.

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L+Z_0} = \frac{\bar{Z}_L - 1}{\bar{Z}_L+1} = \frac{V^-}{V^+}$$

Table 4.1: Transmission line parameters for lossless and low-loss conditions

Lossless	$\leftrightarrow$	**Low-loss**
$\alpha = 0$		$\alpha = \frac{1}{2}R\sqrt{\frac{C}{L}} + \frac{1}{2}G\sqrt{\frac{C}{L}}$
$\beta = \omega\sqrt{LC}$		$\beta = \omega\sqrt{LC}$
$\gamma = j\beta$		$\gamma = \alpha + j\beta$
$Z_0 = \sqrt{\frac{L}{C}}$ = real		$Z_0 = \sqrt{\frac{L}{C}}\left(1 - j\frac{R}{2\omega L} - j\frac{G}{2\omega C}\right)$

$$\text{VSWR}, \rho = \frac{1+|\Gamma_L|}{1-|\Gamma_L|}$$

$$R_{\max} = Z_0\left(\frac{1+|\Gamma_L|}{1-|\Gamma_L|}\right) = Z_0\rho$$

$$Z_{\max} = \frac{|V_{\max}|}{|I_{\min}|}$$

$$R_{\min} = Z_0\left(\frac{1-|\Gamma_L|}{1+|\Gamma_L|}\right) = \frac{Z_0}{\rho}$$

$$Z_{\min} = \frac{|V_{\min}|}{|I_{\max}|}$$

For Impedance transformation, we have

$$Z(l) = Z_0\left\{\frac{Z_L\cos\beta l + jZ_0\sin\beta l}{jZ_0\sin\beta l + Z_L\cos\beta l}\right\}$$

$$= Z_0\left\{\frac{Z_L\cosh\gamma l + Z_0\sinh\gamma l}{Z_0\cosh\gamma l + Z_L\sinh\gamma l}\right\}$$

Lastly, the power flow equations

$$P_L = \frac{1}{2}\frac{|V^+|}{Z_0}\{1-|\Gamma_L|^2\} \quad \text{Real power}$$

$$P(l) = \frac{1}{2}\frac{|V^+|}{Z_0}\{1-|\Gamma_L|^2\} - \mathfrak{Im}\{\Gamma_L e^{-j2\beta l}\} \quad \text{Complex power}$$

$$V^+ = \frac{Z_L' V_s e^{-j\beta l}}{(Z_s + Z_L')(1+\Gamma_L e^{-j2\beta l})}$$

Exercises

Ex. 63 — A 50Ω lossless transmission line is connected to a load of $50 + j50\Omega$. The maximum voltage measured on the line is 50V. Find the power delivered to the load and the peak voltage at the load end of the line.

Ex. 64 — If the two-antenna configuration shown in figure 4.10 is connected to a generator with $V_g = 250\angle 0^o$ and internal resistance $Z_g = 50\Omega$, how much average power is delivered to each antenna?

Ex. 65 — How does the presence of low losses in a transmission line impact the overall efficiency of power transfer?

Ex. 66 — Explain the concept of attenuation in a low-loss transmission line and its significance in power transmission.

Ex. 67 — In a low-loss transmission line, how do conductor materials influence power transfer efficiency?

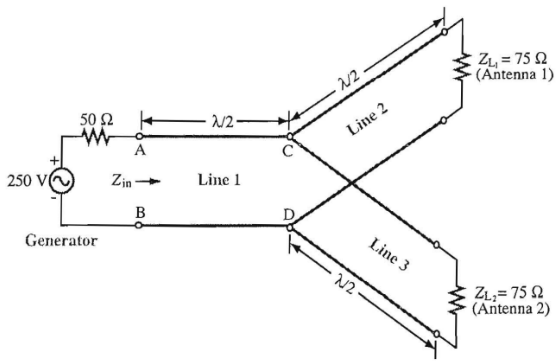


Figure 4.10: Antenna configuration for Ex. 64

Ex. 68 — Calculate the power loss in a low-loss transmission line with a length of 150 km, a current of 1000 A, and a resistance of 0.1 ohms/km.

Ex. 69 — Discuss the role of dielectric materials in minimizing losses and improving power transfer in low-loss transmission lines.

Ex. 70 — How does skin effect influence power transfer efficiency in low-loss transmission lines at high frequencies?

Ex. 71 — Explain the impact of frequency on power transfer capability in a low-loss transmission line.

Ex. 72 — Calculate the voltage regulation for a low-loss transmission line with a sending-end voltage of 220 kV and a receiving-end voltage of 210 kV.

4.9 Introduction to Smith chart

4.9.1 Objectives

At the end of this lecture, the student should have a full understanding of:

(i) Smith chart's characteristics using the constant reactance and resistance circles

(ii) Using the smith chart to solve transmission line analysis and impedance calculation

(iii) Using normalized impedance and admittance to analyze loads

(iv) identifying network parameters; maximum and minimum voltage and current, maximum and minimum resistance using the VSWR circle

(v) understand the concept of VSWR circle and movement along the transmission line using the VSWR cirlce

(vi) Identifying the types of load from the standing wave pattern: purely reactive loads, purely resistive loads, inductive loads and capacitive loads

Previously, we were able to derive equations for different parameters for a transmission line such as;

(i) Power delivered to the load,

(ii) Voltage Standing Wave Ratio (VSWR),

(iii) Maximum and minimum voltage and current,

(iv) Reflection coefficient,

(v) Impedance transformation ratio etc.

Up till now, we analyze transmission line characteristics using **analytical method**. But in this lecture, we utilize another approach for analyzing the problems of a transmission line which is the **graphical method**. This approach involves solving the problems of a transmission line from an image called a **Smith chart**. This approach is most preferred for solving transmission line problems than the analytical method owing to the following reasons;

(i) Images have a much longer-lasting impression than equations or text on the human mind.

(ii) The graphical approach is much simpler compared to the analytical approach where calculation can be reduced by significant amount

(iii) It is a very compact way of representing the impedance characteristics of transmission lines.

So in fact, whenever we are doing transmission line calculation even analytically, it will be appropriate to keep the graphical representation in mind and any analytical method used on transmission line should always be cross checked with the graphical method. This method if properly understood makes solving transmission line problems much easier and faster. It can also be used as a means of cross-checking the solution obtained using the analytical method so that one does not go conceptually wrong in solving transmission line problems.

It is very important to understand that the graphical method does not give the voltage and current solutions. It gives you only the representation of impedances on the transmission line or the standing wave characteristics of the transmission line. This means it can give you the **impedance transformation** relationship, and it can give you **reflection coefficient**, **VSWR**, location of **voltage minimum**, location of **voltage maximum** and so on.

4.9.2 Plane Transformation

In this section, the basic idea is to take the impedance which is normalized with respect to characteristics impedance and do transformation of these impedance into the complex reflection coefficient plane called **GAMMA** plane. By doing this then we can transform all impedance from impedance plane to reflection coefficient plane and then we see that this representation of impedances in the reflection coefficient plane makes the calculation much simpler than if we take impedances. The resistive part lies on the real and positive axis and the imaginary part can be positive or negative. Resistance is always real and positive, no negative resistances here. Positive reactance is for inductance and negative reactance for capacitance. So plotting impedance on the complex impedance plane, we get the the diagram below. We saw for the transmission line, what really matter is . Let us assume our transmission line is lossless. Then the characteristics impedance is a real quantity. All impedance are normalized to the characteristics impedance.

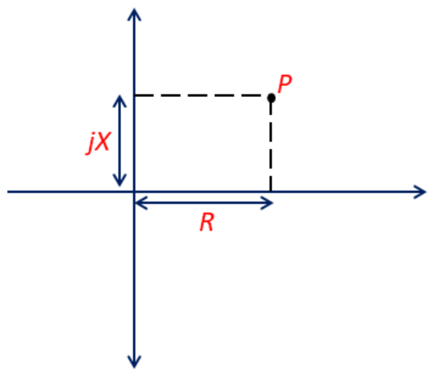


Figure 4.11: Absolute impedance on the complex Z-plane

R represents the real part called resistance and jX represents the imaginary part called reactance. Just like in the analytical approach, our analysis using the graphical method will be done with the normalized impedance rather than the **absolute impedance**. It, therefore, means that, given any impedance, it must be normalized with respect to the characteristic impedance (Z_0) such that;

$$\bar{Z}_L = \frac{Z_L}{Z_0} \quad \text{(normalised)}$$

Z_L is the absolute impedance. So our representation of the normalized impedance would be:

$$\bar{Z}_L = r + jx$$

Let us represent this expression graphically on the complex z plane as shown in figure 4.12.

From the figure 4.12, it is seen that the resistance **r** is always real and positive. Whereas the reactance could be positive which represents inductive reactance or negative which represents capacitive reactance. Considering all passive loads in the z plane, it can be deduced that all points on the imaginary axis

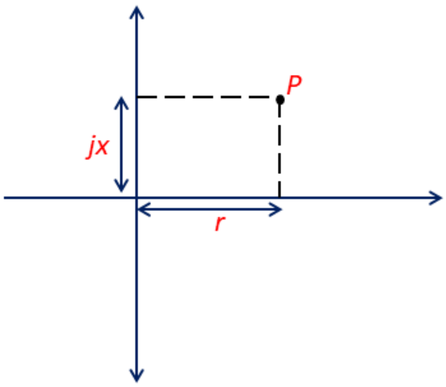


Figure 4.12: Normalized impedance on the complex Z-plane

plus all points on the right half plane covers all possible passive loads of a transmission line.

Graphically, the statement is represented in the figure 4.13

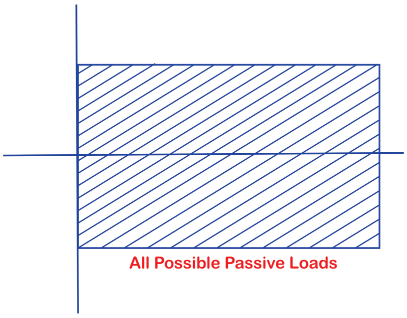


Figure 4.13: Representation of all passive loads on the complex Z- plane

Solving problems of transmission lines using the z-plane is quite difficult. For this reason, the impedance which was normalized with respect to the characteristic impedance is transformed into another plane which we shall call the **complex coefficient plane** or **Gamma plane** (Γ). The transformation from the Z plane to the Γ plane is made possible using the one-to-one relationship that exists between the impedance and the reflection co-efficient Γ rewritten in equation (4.26) in general the reflection coefficient will have a real and imaginary part. It can be written in either rectangular or polar form. Let the read part of the reflection coefficient be denoted by U and the imaginary part by v Γ = u + jv =

$$re^{j\theta}$$

in polar form, Γ =

$$\frac{\bar{Z}-1}{\bar{Z}+1}$$

shows that we have a one lo one relationship between the reflection coefficient and the normalized impedance. If we know normalized impedance we can find out the reflection coefficient, if we know reflection coefficient, we can find out the normalized impedance. $\bar{Z}$ =

$$\frac{1+\Gamma}{1-\Gamma}$$

So transformation between normalized impedance and reflection coefficient is a one to one relationship. If we plot reflection coefficient for all passive loads as we have seen earlier, for passion load, the magnitude of the reflection coefficient is always less than or equal to 1. At open or closed circuit, Γ = 1. For any other impedance n ¡ 1. So if we plot the reflection coefficient into the complex reflection coefficient plane, what we called the complex gamma plane we have the right half $\bar{Z}$ plane mapping into a unit Circle in the gamma plane. Now er have 2 planes, complex $\bar{Z}$ plane and complex gamma plane. Since there is a one to one transformation between these two planes, every point on the $\bar{Z}$ plane is mapped to a point on the Γ plane. So first exercise we do is to try and map all impedance into the gamma plane.

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0} \tag{4.26}$$

Normalizing we have;

$$\Gamma_L = \frac{\bar{Z}_L - 1}{\bar{Z}_L + 1} \quad \text{or} \quad \bar{Z}_L = \frac{1+\Gamma}{1-\Gamma}$$

With this transformation relationship, if the **normalized impedance** is known, we can find the **reflection coefficient** and vice versa. Hence in general, the reflection coefficient will have a real and imaginary part. It can be written in either rectangular or polar form. Let the real part of the reflection coefficient be denoted by u and the imaginary part by v. Γ u + jv = $re^{j\theta}$ in polar form = Γ = $\frac{1+\gamma}{1-\Gamma}$. So transformation between normalized impedance and reflection coefficient is a one to one relationship. If we plot reflection coefficient for all passive load as we have seen earlier, for passive load, the magnitude of the reflection coefficient is always less than or equal to 1 At open or closed circuit Γ = 1. For any other impedance, Γ ¡ 1. So if we plot the reflection coefficient into the complex reflection coefficient plane, what we called the complex gamma plane, we have the right half $\bar{Z}$ plane mapping into a unit circle in the gamma plane as shown in figure 4.14.

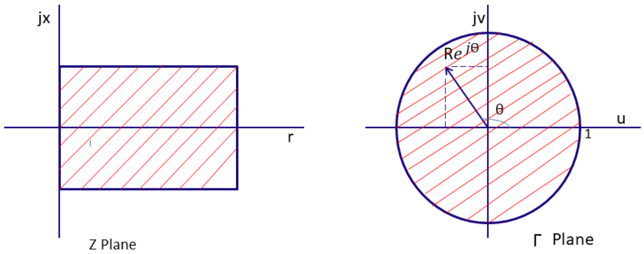


Figure 4.14: Z Plane and Complex Γ plane

In general, when solving graphically the problems of transmission lines, the basic idea is to take the impedance which is normalized with respect to the characteristics impedance and do a transformation of this impedance into the gamma plane.

4.9.3 Transformation analysis for complex plane

It was established that since there is a one-to-one transformation between the complex Z-plane and Γ plane every point on the Z-plane is mapped to a point on the Γ plane. In this section, we will be **mapping** the impedance from the Z-plane to the gamma plane.

Recall, $\bar{Z}_L = r + jx$ and also $\bar{Z}_L = \frac{1+\Gamma}{1-\Gamma}$.

$$r + jx = \frac{1+\Gamma}{1-\Gamma} \tag{4.27}$$

$$\Gamma = u + jv \tag{4.28}$$

Substituting equation 4.28 in 4.27 we have:

$$\begin{aligned} r + jx &= \frac{(1+u)+jv}{(1-u)-jv} \\ &= \frac{(1+u)+jv}{(1-u)-jv} \times \frac{(1-u)+jv}{(1-u)+jv} \\ &= \frac{1-u^2+jv(1+u)+jv(1-u)-v^2}{(1-u)^2+jv(1-u)-jv(1-u)+v^2} \\ &= \frac{1-u^2+2jv-v^2}{(1-u)^2+v^2} \\ &= \frac{1-(u^2+v^2)+2jv}{(1-u)^2+v^2} \end{aligned}$$

Thus the real and imaginary parts of the equation are:

$$r = \frac{1-(u^2+v^2)}{(1-u)^2+v^2} \quad \text{(real part)}$$

and

$$x = \frac{2v}{(1-u)^2+v^2} \quad \text{(imaginary part)}$$

Considering the real part, we can express it in the form of the equation of a circle.

$$\begin{aligned} r((1-u)^2+v^2) &= 1-(u^2+v^2) \\ r - 2ur + u^2r + v^2r &= 1 - u^2 - v^2 \\ u^2(r+1) - 2ur + (r-1) + v^2(r+1) &= 0 \\ u^2 - 2u\left(\frac{r}{r+1}\right) + v^2 + \left(\frac{r-1}{r+1}\right) &= 0 \end{aligned}$$

We further simplify using the method of completing the squares;

$$\left(u - \frac{r}{r+1}\right)^2 + v^2 + \left(\frac{r-1}{r+1}\right) - \left(\frac{r}{r+1}\right)^2 = 0$$

$$\left(u - \frac{r}{r+1}\right)^2 + v^2 = \left(\frac{r}{r+1}\right)^2 - \left(\frac{r-1}{r+1}\right)$$

$$\left(u - \frac{r}{r+1}\right)^2 + v^2 = \frac{r^2-(r^2-1)}{(r+1)^2}$$

$$\left(u - \frac{r}{r+1}\right)^2 + v^2 = \frac{1}{(r+1)^2} \tag{4.29}$$

Equation (4.29) represents the equation of a circle in the uv plane centered at $\left(\frac{r}{r+1}, 0\right)$ with radius $\left(\frac{1}{r+1}\right)$[14].

In a similar way, let's consider the imaginary part.

$$x = \frac{2v}{(1-u)^2+v^2}$$

$$x(1-2u+u^2+v^2) = 2v$$

$$v^2x - 2v + u^2x - 2ux + x = 0$$

We divide through by x and factorize

$$\begin{aligned} v^2 - \frac{2v}{x} + u^2 - 2u + 1 &= 0 \\ (u-1)^2 - 1 + \left(v - \frac{1}{x}\right)^2 - \frac{1}{x^2} + 1 &= 0 \end{aligned}$$

$$(u-1)^2 + \left(v - \frac{1}{x}\right)^2 = \frac{1}{x^2} \tag{4.30}$$

Equation (4.30) represents the equation of a circle in the uv plane centred at $\left(1, \frac{1}{x}\right)$ with radius $\left(\frac{1}{x}\right)$.

From equations (4.29) and (4.30), it can be deduced that for any value of r and x, we get a circle on the gamma plane. The circle that corresponds to **r** is called the *Circle of Constant Resistance* while the circle that corresponds to x is called the *Circle of Constant Reactance*.

So all point on the circle of constant resistance have same resistance value in the $\bar{Z}$ plane and all point on the constant resistance circle have some reactance in the $\bar{Z}$ plane. we draw these on a gamma plane taking note of centre and radius for each circle respectively, ($\frac{r}{r+1}$,0) lie on the real axis in the gamma plane. For the constant resistance circle, and the radius is $\frac{1}{r+1}$ and Γ varies from 0 to ∞. Similarly the constant reactance circle has centre (1,$\frac{}{x}$) i.e, in a vertical line passing the real axis at u = 1 and the radius is $\frac{1}{x}$ and x varies from -∞ to + ∞

4.9.4 Circle of Constant Resistance

In order to visualize the plot of the circles of constant resistance we need to transform the limits of the values of r, the real part of normalized impedance $\bar{Z}$. The limits of r are $[0, \infty]$ both inclusive and thus the corresponding values of the radius of these circles are $[1, 0]$. Plotting the circle for different values of r, we get the plot in figure 4.15.

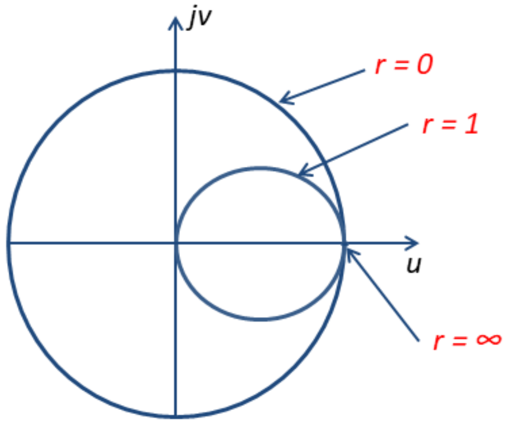


Figure 4.15: Plot of the Circles of Constant Resistance at the limits of the values of $r = 0$, $r = 1$ and at $r = \infty$

Let's plot for various values of r, by varying its value from 0 to ∞. At $r = 0$, from equation (4.29), the centre of the circle

[14] The general equation of a circle is $(x-a)^2+(y-b)^2 = r^2$ where (a, b) are the centre coordinates of the equation and r is the radius.

is $(0, 0)$ and the radius is 1. At $r = 1$ which is the condition $Z = Z_0$ we get the centre at $\left(\frac{1}{2}, 0\right)$ and radius of $\frac{1}{2}$ as shown in figure 4.15.

The complete plot of the circles of constant resistance for the values of increasing r is shown in figure 4.16. It can be said that the centre shifts towards the right and the radius keeps reducing towards zero. At $r = \infty$, the centre is (1,0) and the radius is zero, which is denoted as a dot on the real axis at 1.

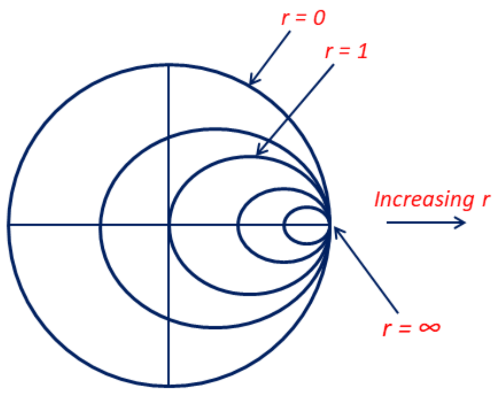


Figure 4.16: Plot of the circles of constant resistance for increasing values of r

We can observe that the circles of the constant resistance pass through $u = 1$, and $v = 0$.

4.9.5 Circle of Constant Reactance

Similarly, for the plots of the circles of constant reactance, we consider the limits of the values of x which are $[-\infty, +\infty]$. The corresponding values of the radius of the circle as given by equation (4.30) are both zero and the radius at $x = 0$ approaches $-\infty$ from the negative domain and $+\infty$ from the positive **domain** of x. Similarly, for the centre coordinates of these circles, considering the negative and positive domains separately, we deduce that the centre of these circles lies on a vertical plane passing through the u axis at $u = 1$ and varies from $-\infty$ to $+\infty$[15].

Figure 4.17 shows the plot of the circles of constant reactance.

As seen in figure 2.37, the centre always lies on a line $u = 1$ and as x increases in the **positive domain**, the centre shifts down along $u = 1$ line and the radius decreases. Similarly, as x increases in the **negative domain**, the centre shifts up along the $u = 1$ line and the radius decreases. The radius of the circles of constant reactance at $x = 0$ is shown with a line through the origin whose centre is at $\pm\infty$. Also, we can observe that the circles of constant reactance also pass through the coordinate $(u, v) = (1, 0)$. This signifies the uniqueness of the point.

The superimposed plots of the constant resistance and constant reactance circles are thus shown in figures 4.18 and 4.19.

It can be observed that the superimposed set of circles was masked by the outermost circle with a radius of 1. This circle is called the **limit circle** for the gamma plane which is the

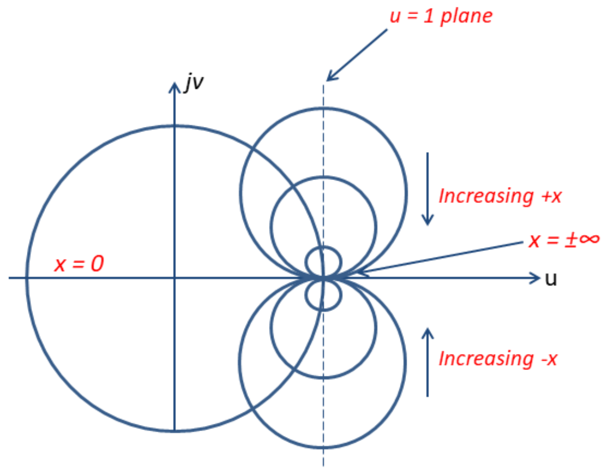


Figure 4.17: Plot of the circles of constant reactance at the limits $x = \pm\infty$ and $x = 0$ as well as inbetween.

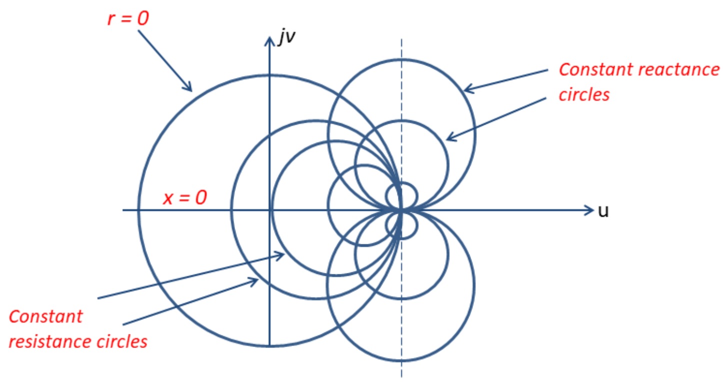


Figure 4.18: Plot of the two sets of circles superimposed

region where $\Gamma \leq 1$. This is so because in the analysis of transmission lines, we are not concerned with reflection coefficients that are greater than one, we are only interested in reflection coefficient that are less than or equal to 1 which is represented by the unit circle of radius of 1 which represent all possible passive loads. It therefore means that any other point outside this range of reflection coefficient, Γ, is not of practical relevance since it does not represent a passive load. Although the constant reactance circles fill the entire space, we only consider the portion that satisfies $\Gamma \leq 1$ and this lies within the unit circle of the complex gamma plane.

Some of the observations we made are summarized as follows:

(i) All circles of constant resistance have their centre lying on the positive axis and complex gamma plane.

(ii) All the circles passes through the point $u = 1$ and $v = 0$.

(iii) As the value of r increases, the centre of the circles of constant resistance shifts towards the right from the origin to point $u = 1$ and the radius decreases.

(iv) The centre of the circles of constant reactance circle line in a plane $u = 1$ and the upper half represents the inductive reactance and the lower half capactive reactance.

[15] At $x = \pm\infty$, the centre from equation (4.30) is $(1, 0)$ and at $x = 0$, the centre is $(1, \pm\infty)$.

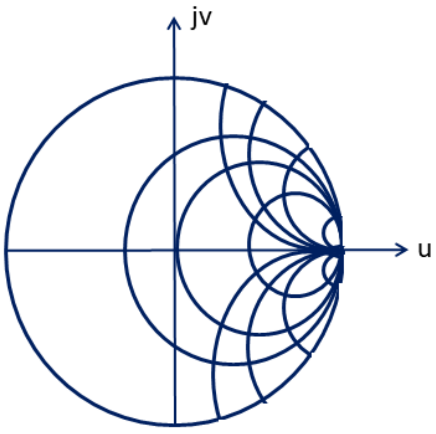


Figure 4.19: Masking of the superimposed set to relevant region

(v) As the reactance value increases in both the positive and negative domain, the centre of the circle shifts towards the right from the origin to $(1, 0)$ i.e it approaches the u-axis, and the radius of the circle decreases.Similarly, for the constant reactance cirlces, the constant reactance cirlce has a vertical line passing through u = 1 and the upper half represents the positive resistance and the lower half represents the negative reactance

(vi) When $x = \pm\infty$, the radius of the circle is zero which is a point at $(u, v) = (1, 0)$.

4.10

The superposition of the constant resistance circle and the constant reactance circle is now a coordinate system for impedance on the complex gamma plane. This is called the *Smith chart*[16]. It was developed by PHILIP HAGAR SMITH

The Smith chart (shown in figure 2.40 below), is a graphical aid designed for electrical and electronics engineers to assist in solving problems of the transmission lines and matching circuits. It is a product of superimposing the circle of constant resistance and the circle of constant reactance. It can be used to simultaneously display multiple parameters such as the impedances, admittances, reflection co-efficient etc. It is plotted on the complex reflection co-efficient plane in two dimensions and is scaled in normalized impedance (which is most common), normalized admittance or both. It has **circumstantial scaling** of wavelengths and degrees. The wavelength scale is used in distributed component problems and represents the distance measured along the transmission line connected between the generator and the load to the point under consideration.

Figure 4.20 shows the simplified version that will allow us to mark out some important points.

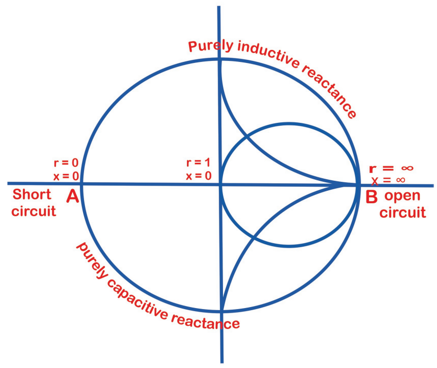


Figure 4.20: A Simplified Smith Chart

(i) The point A or any other point lying on the outermost circle correspond to $r = 0$ while any point lying on the horizontal axis represents $x = 0$

(ii) The intersection of this horizontal axis and the outermost circle corresponds to $r = 0, x = 0$ at A which is a load that is short-circuited from the impedance point of the line. It, therefore, means that point A is a short circuit.

(iii) At point B where the two sets of circles disintegrate into, corresponds to $r = \infty$ and $x = \pm\infty$ which is an open circuit.

(iv) Any point between A and B on the outermost circle represents pure reactances because for these points, r = 0. Above the u axis inside the limit circle, we have the inductive loads, and below the u axis in the limit circle, we have the capacitive loads (it, therefore means we can tell

PHILIP HAGAR SMITH (April 29, 1905 - August 29, 1987) was an electrical engineer who became famous for his invention of the SMITH CHART. He graduated from Tufts College in 1928 with a BS degree in electrical engineering while working for Bell Telephone Laboratories. He invented his eponymous Smith chart during the period of being faced with the challenge of how to show and evaluate multiple complex impedance parameters which can range from zero to infinity. When asked why he invented the chart, Smith explained *"From the time I could operate a slide rule, I've been interested in graphical representations of mathematical relationship"*. This interest in representing mathematical relationships in graphical form is what led to the invention of the Smith chart.

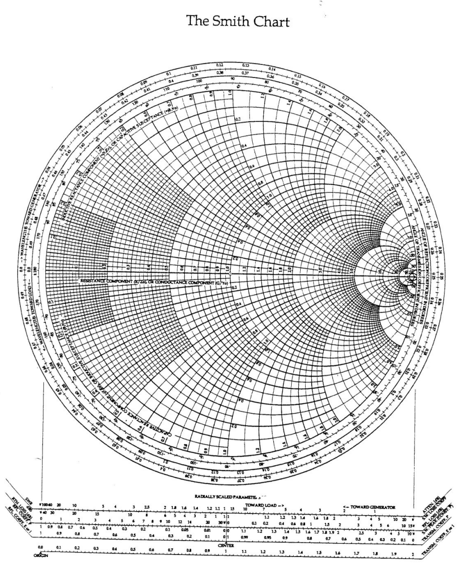
16 

Figure 2.40: The Complete Smith Chart

what kind of load we have by where it lies on the smith chart).

(v) At point C where $r = 0$ and $x = +1$, corresponds to purely reactance whose magnitude is equal to the of the line.

(vi) At point D where $r = 0$ and $x = -1$ represent a purely capacitive reactance whose magnitude is equal to the characteristic impedance of the line.

(vii) The centre of the has one more special point M which is the origin of the smith chart of the gamma plane is the intersection of $r = 1$ circle and $x = 0$ line.

At M the impedance is equal to the characteristic impedance of the line and that is the point of greatest interest to us because that point represents the **matched condition** of the transmission line. So when the impedance lies on the centre of the smith chart, which correspond to the magnitude of the reflection coefficent, $r_L = 0$. This point represent the matched conditon of the transmission line

These points if properly understood can be used to solve transmission line problems with ease.

To solve transmission line problems using the Smith chart, it should be held such that the most clustered portion of the circle should be towards the right hand of the user because the real axis at the complex gamma plane is in that direction and the complex imaginary plane is in the vertical direction. If done so then calculation on conversion from the complex reflection co-efficient plane to the impedance plane and vice versa can be done easily as well as marking of impedance points on the Smith chart to use to find co-ordinates of the point which gives the reflection co-efficient easily. In conclusion, the Smith chart is a handy tool for solving very complex problems.

Exercise

Ex. 73 — Why do we limit our analysis on the Smith chart to the limit circle?

Ex. 74 — As we move along a transmission line, how does a point move on the Smith chart?

Ex. 75 — At what point in solving transmission line problems do we consider the wavelength's scale?

Ex. 76 — Prove mathematically, the one-to-one relationship that is involved in impedance transformation from the Z-plane to the Γ-plane.

4.11 Concept of constant VSWR circle in lossless transmission line analysis

Previously, we developed a graphical tool by transforming the complex impedance of a transmission line into the gamma plane called the **Smith chart**. This we have seen as the superimposition of; circles of constant resistance on the circles of constant reactance, masked within the limit circle $|\Gamma| \leq 1$. Before we go into the use of the Smith chart for transmission line calculations, we develop one more set of circles called the constant and then superimpose it on the Smith chart.

We know that for a lossless transmission line for which attenuation constant is zero,

$$\Gamma(l) = \Gamma_L e^{-j2\beta l}$$

Where l is the distance from the load point towards the generator, Γ_L is the voltage reflection coefficient at the load end and β is the phase constant of the transmission line. The voltage reflection coefficient, Γ_L can be expressed in polar form as $|\Gamma_L|e^{j\theta_L}$ such that

$$\Gamma(l) = |\Gamma_L|e^{j\theta_L} \cdot e^{-j2\beta l}$$

Where θ_L is the phase angle of deflection of the load.

$$\Gamma(l) = |\Gamma_L|e^{j(\theta_L - 2\beta l)} \tag{4.31}$$

The total phase becomes $\theta_L - 2\beta l$. Hence the reflection coefficient $\Gamma(l)$ has magnitude $|\Gamma_L|$ on the complex gamma plane but a phase angle that varies with distance from the load end. Since l is positive when moving towards the generator, therefore, the phase increases in the negative direction. The distance of the point $|\Gamma_L|$ from the centre of the circle remains constant but its angle changes with respect to l. Hence $|\Gamma_L|e^{j(\theta_L - 2\beta l)}$ represents a *constant VSWR circle* having the same centre of origin with that of the complex gamma plane as shown in figure 4.21.

When moving towards the generator with increasing length, l, a corresponding anti-clockwise rotation in the gamma plane is experienced. That is on this circle, the magnitude of the reflection coefficient is the same no matter what point, but the angle differs and we know that,

$$VSWR, \rho = \frac{1 + |\Gamma_L|}{1 - |\Gamma_L|} \tag{4.32}$$

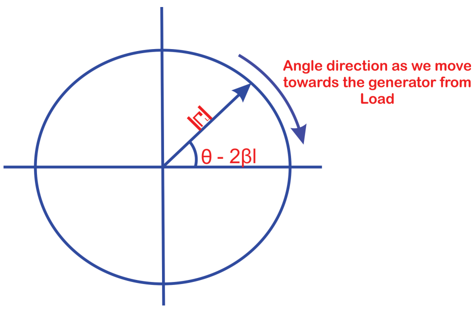


Figure 4.21: Constant VSWR Circle

Since $|\Gamma_L|$ is the same at any point on the circle, then the VSWR is the same for all points on the circle, hence we call the circle a **constant VSWR circle**. One should always note that the centre of all VSWR circles is the same as that of the origin of the complex gamma plane and the magnitude of the reflection coefficient is always less or equal to 1. A larger radius of $|\Gamma_L|$ gives more reflection coefficient with a lower impedance match and a smaller radius of $|\Gamma_L|$ gives a smaller magnitude of reflection coefficient and a better impedance match (see figure 4.22).

When we say that we have better impedance matching on the Smith chart, visually speaking, we mean that the point closest to the centre of the Smith chart is better for impedance matching because it represents a smaller magnitude of the reflection coefficient.

With the understanding of superimposing the constant VSWR circle on the Smith chart, we can therefore solve the transmission line problem.

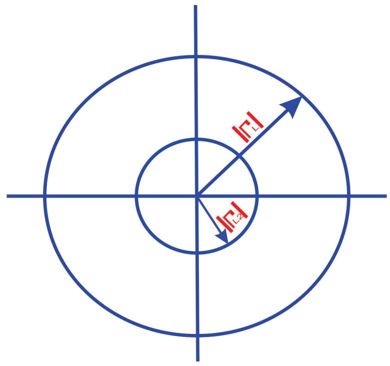


Figure 4.22: Impedance matching with different constant VSWR circles

4.12 Transmission Line Analysis using Admittance

However, as we have said, you may have connections in transmission lines that are in the form of parallel connections and we know that from circuit analysis, anywhere we have parallel connections, it is easier to deal with admittances rather than impedances. Before now, we have been discussing load impedances and load characteristics impedance of transmission lines. However, if we have to make parallel connections on the transmission line, we represent the load as admittances and before we carry out the analysis on the Smith chart.

Thus, in this section, we will determine how the Smith chart is used when we compute in terms of admittances. Since the Smith chart gives the normalized impedances of the transmission line, the same thing will be derived for admittances so we shall first define the normalized admittance of the transmission line and for that, we would require what is called the *characteristic admittance* of the line. is given in equation (4.33)

$$Y_0 = \frac{1}{Z_0} \tag{4.33}$$

Then, normalizing the admittance would yield

$$\bar{Y} = \frac{Y}{Y_0} \quad \text{thus, load admittance is} \quad \bar{Y}_L = \frac{Y_L}{Y_0}$$

The reflection co-efficient as derived in equation (4.26) is given as

$$\Gamma = \frac{Z_L - Z_0}{Z_L + Z_0}$$

To determine the reflection coefficient in terms of admittance we replace the impedance Z and characteristics impedance are then expressed as the inverse of admittance $\frac{1}{Y}$ and inverse of the $\frac{1}{Y_0}$ respectively.

$$\Gamma = \frac{\frac{1}{Y} - \frac{1}{Y_0}}{\frac{1}{Y} + \frac{1}{Y_0}}$$

We further simplify as follows;

$$\begin{aligned}\Gamma &= \frac{Y_0 - Y}{Y_0 + Y} \\ &= \frac{1 - \frac{Y}{Y_0}}{1 + \frac{Y}{Y_0}} \\ &= \frac{1 - \bar{Y}}{1 + \bar{Y}} \\ &= -1 \cdot \left(\frac{\bar{Y} - 1}{\bar{Y} + 1}\right)\end{aligned}$$

The negative sign indicates a phase change of 180°[17] thus, it can be expressed as

$$\Gamma = \frac{\bar{Y} - 1}{\bar{Y} + 1} e^{j\pi} \tag{4.34}$$

So the reflection coefficient written in terms of is the same as the reflection coefficient written in terms of normalized impedance except for the 180° phase change brought in by e^{jn}

This implies that the same value is derived for the reflection coefficient when calculated using the normalized impedance or admittance except there is a phase shift of 180° that would be experienced on the transmission line. In other words, on the complex gamma plane of the Smith chart, the 180° phase shift will correspond to a rotation of 180° (see figure 4.23). Essentially, the normalized impedance and admittance can be calculated in the same way except when carrying out calculations for normalized admittances, there is a rotation of 180° on the complex gamma plane otherwise, all other values and parameters remain unchanged.

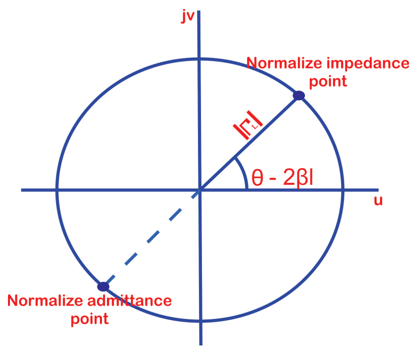


Figure 4.23: Impedance equivalent of Admittance on the gamma plane

Hence as far as the constant VSWR circles are concerned, any point on the Smith chart rotated by 180° corresponds to the normalized admittance with corresponding circles of constant resistance and circles of constant reactance.

[17] As -1 is a phase change of π, $e^{j\pi} = \cos\pi + j\sin\pi = -1$

What would be the complex representation of the admittance?

$$\begin{aligned}\bar{Y} &= \frac{G + jB}{Y_0} \\ &= g + jb \\ &= \text{conductance} + j(\text{susceptance})\end{aligned}$$

If we interchange **r** with **g** and **x**with **b**, then the circles through the point on the constant VSWR circle will be the **circle of constant conductance** and the **circle of constant susceptance** which are similar to the circle of constant resistance and circle of constant reactance except that they are rotated by 180° . This implies that our analysis with admittances will resemble one where every point on the Smith chart is flipped and so the clustered end goes to the left-hand side as shown in figure 4.24.

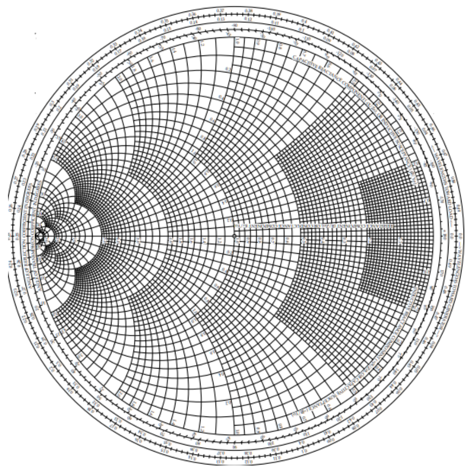


Figure 4.24: Inverted Smith chart

Alternatively, we can keep the Smith chart orientation the same and rotate the complex gamma axis by 180° before doing calculations for admittances. So, if we develop an understanding that we will not rotate the Smith chart, that means we want to use it in its original form, that is, the most clustered part of it to our right, then if we do the impedance calculations, the positive real axis is towards right and the positive imaginary axis goes up. However, during calculations using the Smith chart unchanged for admittances, the positive real axis is towards the left and the positive imaginary axis will be downwards. Normally, whenever we do the Smith chart calculations, we do not rotate the Smith chart. For the impedance, the gamma axis is the positive top right plane and for admittances, the bottom left plane is the positive plane. Depending on whether we are calculating for the impedance or admittances and if we require fixed measurement in the complex gamma plane, then the axis has to be rotated appropriately by 180° depending on whether we are using the impedance or admittance. This is important when finding the phase of the reflection coefficient, otherwise, the axis of the gamma plane is insignificant. So without worrying about the axis of the complex gamma plane, we can use the same Smith chart for the admittance as well as the impedance calculations. This is the reason why when you look at the Smith chart carefully, you will see that the upper half of the Smith chart is denoted by (x, b) and the lower part $(-x, -b)$. The circles are denoted by either r or g so any normalized value of r is equal to the same normalized value of g which represents the same circle.

Hence, as long we are dealing with a normalized quantity, the impedance and admittance can be treated exactly the same way on the Smith chart. However, the normalized values of g and r or b and x have different meanings physically. But do they represent the same physical conditions? The answer is NO! For example, $r = 0; x = 0$, corresponds to a short circuit condition— the impedance is zero— however, if we take a normalized value of admittance with $g = 0, b = 0$ which represents admittance equal to zero, it is not a short circuit but an open circuit condition on the line. Therefore, the normalized values of impedances and admittances can be treated exactly the same way in calculations but not the same interpretations for the physical conditions. In summary, mathematically, $g = 0 \quad , b = 0 \Longrightarrow$ open circuit and $g = \infty, b = \infty \Longrightarrow$ short circuit.

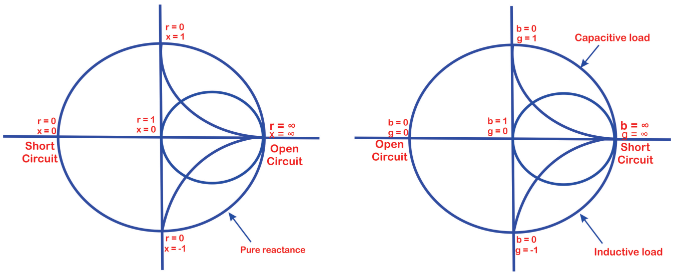


Figure 4.25: Physical variation of impedance and admittance

Let us, therefore, consider the points on the Smith chart for the admittance where the special points r and x are replaced with b and g respectively. Maintaining the original orientation of the Smith chart then for the admittance Smith chart, the upper half of the Smith chart represents capacitive loads and the lower half represents the inductive loads (see figure 4.25). With these in mind, the use of the Smith chart for impedance or admittance calculation is very straightforward.

Now let us make use of the Smith chart to solve transmission line problems.

4.13 Use of Smith chart for Transmission Line Calculation

Let us consider the simplest problem we can think of for the transmission line. Suppose for a given load, we want to find the reflection coefficient at the load point. Analytically, we can use the formula $\frac{Z_L - Z_0}{Z_L + Z_0}$ but in using the Smith chart, the problem is much simpler to solve. Remember the impedance and admittance you have on the Smith chart are all normalized quantities. So we take the following steps

Procedure:

(i) Normalize Z_L to get $\bar{Z}_L$, that is, $\frac{Z_L}{Z_0}$.

(ii) Determine this point on the Smith chart, that is, $r + jx$ by identifying the constant resistance and reactance circle with the value r and x respectively and the point of intersection (see figure 4.26).

(iii) Determine the distance of the point of intersection from the origin to find the magnitude of the reflection coefficient, $|\Gamma_L|$ and then the corresponding phase is given as θ_L.

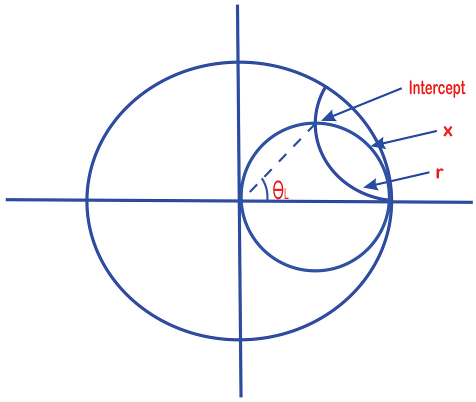


Figure 4.26: Point of intersection of circles of constant resistance and reactance

So without doing any calculation, just by measuring its distance, and the angle, we get the magnitude of reflection coefficient, $|\Gamma_L|$ and phase angle, θ_L.

Similarly, for the admittance, let $\bar{Y} = g + jb$. We determine the point of intersection on the Smith chart the same way as that of impedance, however, to find out the complex reflection coefficient, maintaining the orientation, then the positive real axis is to the left. So the phase angle will be measured from the positive real axis as shown in figure 4.27. We summarize the steps as follows.

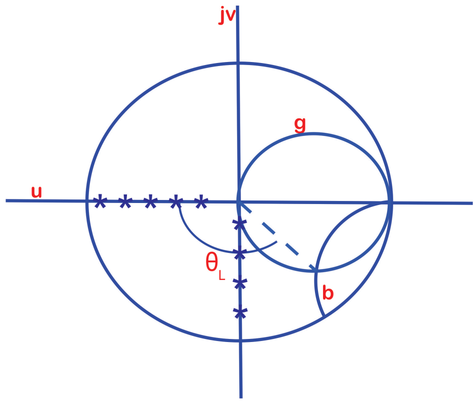


Figure 4.27: Determining the reflection coefficient for admittance on a Smith chart

Procedure:

(i) Normalize the admittance $\bar{Y}$ with the characteristic admittance, Y_0.

(ii) Determine the point on the Smith chart which is the point of intersection of the circle of constant conductance, g and the circle of constant susceptance, b.

(iii) Determine the distance of the point from the origin to find the magnitude of the reflection coefficient, $|\Gamma_L|$ but the phase angle, θ_L will be measured from the positive real axis (see figure 4.27).

Thus, when we are using normalized impedance or admittance, the appropriate rotation has to be made on the coordinate axes of the Smith chart. Once this is done, the calculation of the complex reflection coefficient is very straightforward. Mark the normalized points on the Smith chart. Find the distance from the origin to the marked point and then measure the angle from this marked point from the positive real axis which differs for impedance and admittance while maintaining the orientation of the Smith chart.

In another simple transmission line problem, let us suppose the magnitude of the reflection coefficient and phase angle is given, but we need to find the load. The steps are summarized below.

Procedure:

(i) Draw a circle of radius equal to the magnitude of the reflection coefficient from the origin.

(ii) Determine phase angle from the positive real axis depending on whether the impedance or admittance value at the load is required.

(iii) Mark the point on the circle and find the corresponding r and x circles or g and b circles.

(iv) Multiply the normalized impedance or admittance by the characteristic impedance or admittance to get the load impedance or admittance.

4.13.1

A common problem in transmission lines is impedance transformation. We have shown how it is done analytically but *how can we perform graphically?*

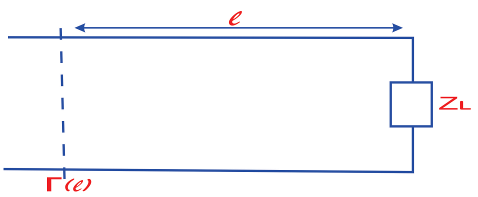


Figure 4.28: Tranformed impedance at distance l from the load

If the load impedance is given as shown in figure 4.28 and we are asked to find the reflection coefficient and load impedance at another point l on the transmission line, that is, given Z_L or Y_L, we are to find Z_L or Y_L at distance l from the load. We will use a hypothetical sketch of the Smith chart shown in figure 4.29.

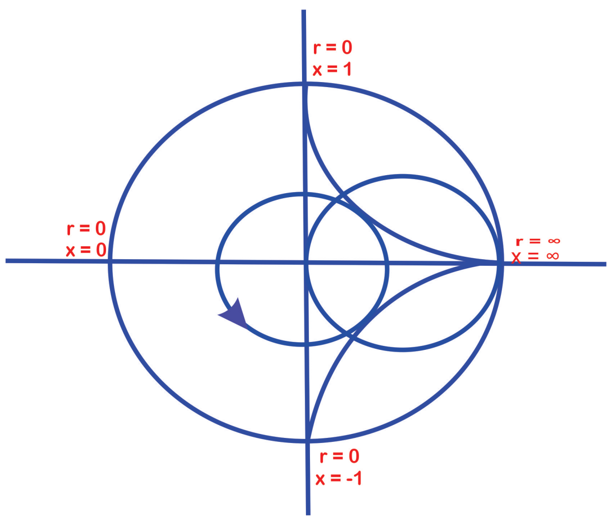


Figure 4.29: Points along the constant r and x Circles

Procedure:

(i) If the impedance is not normalized, we normalize all impedances.

(ii) Determine the point, $\bar{Z}_L$, on the Smith chart and draw a circle passing through the point from the origin. As we have seen earlier, the magnitude of the reflection coefficient remains the same as the point moves on the circle which is a constant VSWR circle.

(iii) Determine the point that translates to a distance l. In a Smith chart, one rotation, 2π translate to a phase change of $2\beta l$[18], that is $2\beta l = 2\pi \implies l = \frac{\pi}{\beta} = \frac{\lambda}{2}$[19] (see figure 4.30). Therefore, a full rotation around the Smith chart corresponds to a distance of $\frac{\lambda}{2}$ on a constant VSWR circle[20]. So we determine the corresponding $x\lambda$ for the value of l moving clockwise along the VSWR from the point $\bar{Z}_L$.

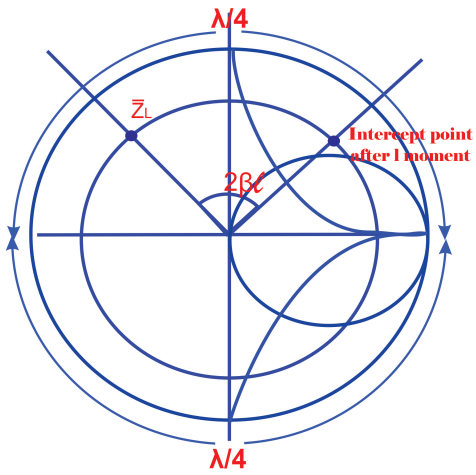


Figure 4.30: Tranformed impedance at distance l from the load

(iv) Mark the new point on the Smith chart and determine the magnitude of the reflection coefficient and phase angle as shown in figure 4.31 or the circle of constant resistance and circle of constant reactance to determine the normalized impedance at l, that is, $\bar{Z}_l = r + jx$.

(v) Multiply the normalized impedance at l by the characteristic impedance Z_0 to get $\bar{Z}_l$ at the new location.

A more general overview of a transmission problem is one where the impedance is given at a particular location and we need to transform the impedance to another location. If the new location is to the left of the previous impedance (towards the generator) then we follow the aforementioned steps, otherwise, we reverse the direction of the angular rotation instead of $-2\beta l$, we have $+2\beta l$ which is counter-clockwise rotation. Hence, moving towards the generator will be a clockwise rotation and moving away from the generator will be an anti-clockwise rotation. The sense of rotation in impedance calculations is extremely important because that tells us whether to move toward the generator or away from the generator. Therefore, in all transmission line calculations, the direction of the

[18] As we move along the transmission line, the phase angle changes by $2\beta l$ in the clockwise direction (from load towards generator).

[19] $\beta = \frac{2\pi}{\lambda}$

[20] The above analysis makes sense since one characteristic of the transmission line is that its impedance characteristics repeat every $\frac{\lambda}{2}$ distance on the line.

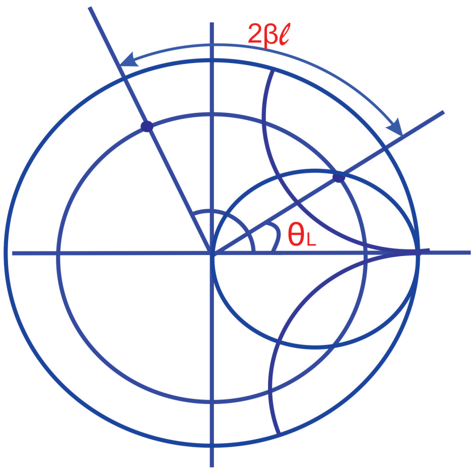


Figure 4.31: New phase angle after rotation

generator should be given much consideration because that will decide the movement along the transmission line.

For admittance calculations, the same steps apply except in the sense of the position of the positive real axis (see figure 4.32).

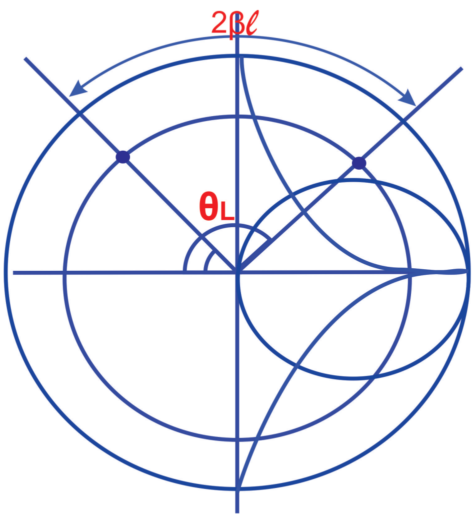


Figure 4.32: Phase angle measurement with admittance Smith chart

Analytically, the impedance transformation requires the calculation of sine and cosine functions that are unusually very complicated. With the help of the Smith chart, impedance transformation is easily achieved.

4.13.2 Voltage Standing Wave Ratio (VSWR)

Another quantity that is a measure of reflection is the VSWR and we established previously that $R_{\max} = Z_0\rho$ and $R_{\min} = \frac{Z_0}{\rho}$. If we normalize $R_{\max}$ and $R_{\min}$ then we get:

$$\bar{R}_{\max} = r_{\max} = \rho \tag{4.35}$$

and

$$\bar{R}_{\min} = r_{\min} = \frac{1}{\rho} \tag{4.36}$$

Therefore, the , $r_{\max}$ gives the VSWR, ρ. Similarly, on the left side of the VSWR circle on the Smith chart, we have $r_{\min}$. Hence $r_{\min} = \frac{1}{\rho}$. Thus, *how do we determine the VSWR graphically?* On the real axis, we have only resistance and no reactance so points of intersection of the constant VSWR circle with the real axis give the maximum and minimum normalized resistance (see figure 4.33). The reason why we call it the constant VSWR circle is that as determined in equation (4.14), it is given as $\frac{1+|\Gamma|}{1-|\Gamma|}$ which is constant for a lossless transmission line and these values of $r_{\max}$ and $r_{\min}$ can easily be deduced from the VSWR circles without a need to calculate. The VSWR is taken as the ratio of maximum impedance on the line to characteristics impedance.

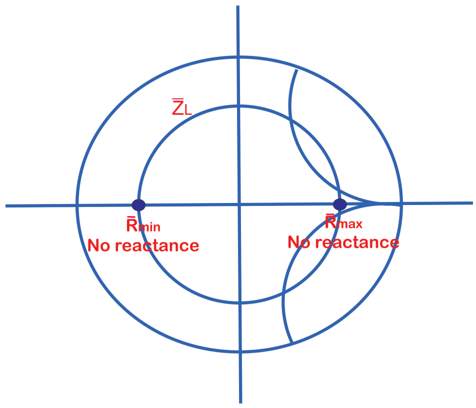


Figure 4.33: Maximum and Minimum Points of Resistance

Hence, the maximum value of the normalized resistance gives VSWR. The in terms of normalized quantity is seen when the VSWR circle intersects the real axis on the right side. The rightmost side corresponds to the impedance which is $r_{\max}$ and the reactance for that is zero— $r_{\max}$ is nothing but ρ. Similarly, the minimum value which we see on the transmission line is $r_{\min} = \frac{1}{\rho}$. The minimum resistance which corresponds to the leftmost intercept on the real axis gives $\frac{1}{\rho}$. Thus once the load impedance or any other impedance is marked on the Smith chart and the constant VSWR circle is drawn, the calculation of the VSWR is straightforward same as the reflection coefficient and transformed/normalized impedance. It is just a matter of plotting and reading different values on a Smith chart.

Next is to find out the locations of the on the transmission line.

4.13.3 Location of

Suppose we want to find out the distance of current or at the load end of the transmission line. From our knowledge of transmission line, at the point of maximum voltage, we experience minimum current which gives $R_{\max}$ then at , we experience maximum current which gives $R_{\min}$ hence $R_{\max}$ corresponds to $V_{\max}$, $I_{\min}$ and $R_{\min}$ corresponds to $V_{\min}$, $I_{\max}$. To find out these locations from the load, we move from point, Z_L to $r_{\max}$ and $r_{\min}$ as shown in figure 4.34. The corresponding clockwise angles, $\theta_{\max}$ and $\theta_{\min}$, covered indicate the distances $l_{\max}$ and $l_{\min}$ towards the generator.

We compute the corresponding distances $l_{\max}$ and $l_{\min}$, recall that $2\beta l = \theta \Longrightarrow l = \frac{\theta}{2\beta}$. Hence,

$$l_{\max} = \frac{\theta_{\max}}{2\beta} \quad \text{and} \quad l_{\min} = \frac{\theta_{\min}}{2\beta}$$

These give the distances from the load point. From figure 4.34, $\theta_{\max} < 180^o(\pi)$ and $270^o\left(\frac{3\pi}{2}\right) < \theta_{\min} > 360^o(2\pi)$ and they,

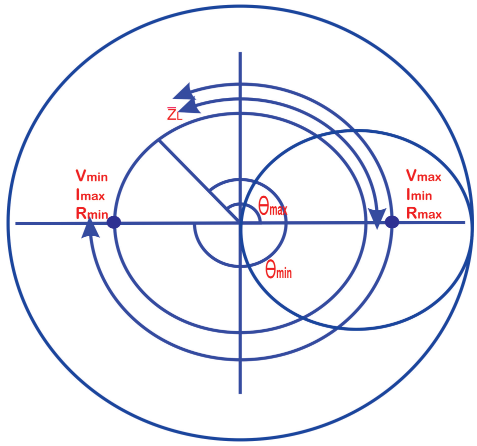


Figure 4.34: Location of the maximum and minimum values of voltage and current

therefore, are the angles that are moved to get from load to $r_{\max}$ and $r_{\min}$ respectively.

More generally,

$$\begin{aligned} l &= \frac{\theta_{\min}}{2\beta} \\ &= \frac{\theta_{\max} + \pi}{2\beta} \\ &= \frac{\theta_{\max}}{2\beta} + \frac{\lambda}{4} \end{aligned} \tag{4.37}$$

Example 4.13.1 Locate the points on the Smith chart

Locate the following points on the Smith chart (Take characteristic impedance, Z_0 = 50Ω).

(a) 50+j75Ω

(b) 10+j0Ω

(c) 0-j80Ω

(d) $\Gamma = 0.3\angle 60^o$

(e) VSWR, $\rho = 2.5$

(f) $R_{\min}$ on ρ = 1.5 circle

Solution

All impedances seen on the Smith chart are normalized impedances. Therefore, for impedance (a), (b), and (c), we will first normalize these points with the characteristic impedance such that

$$\begin{aligned} \bar{A} &= \frac{A}{Z_0} = \frac{50 + j75}{50} \\ &= 1 + j1.5 \\ \bar{B} &= \frac{B}{Z_0} = \frac{10 + j0}{50} \\ &= 0.2 + j0 \\ \bar{C} &= \frac{C}{Z_0} = \frac{0 - j80}{50} \\ &= 0 - j1.6 \end{aligned}$$

To locate point A on the Smith chart, we first identify the ***constant resistance circle*** *of $r = 1.0$. It is the circle that passes through the centre of the Smith chart from the right-hand side (see figure 4.35). Next, we identify the* ***constant reactance circle*** *of $x = 1.5$. Recall that the circles above the real axis passing through the centre are the* ***inductive reactance****, while the circles below the line are the* ***capacitive reactance****. Since the value $x = 1.5$ is positive, it should be the constant inductive reactance circle. The point of intersection of these circles represents the normalized impedance, $\bar{A}$ = $1 + j1.5$. The same steps apply when identifying points B and C on the Smith chart.*

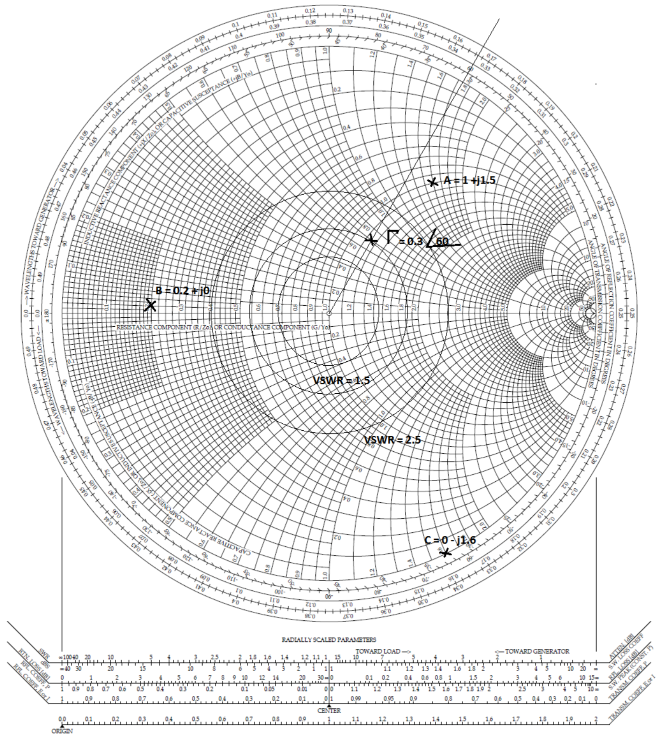


Figure 4.35: Worked Example

To locate the point in (d) which is the reflection coefficient, the approach is quite different. First, we measure the length of the diameter of the constant resistance circle $r = 1.0$ and record its value. In our case, we measured $7.6cm$. Thus we can determine the radius of a reflection coefficient of magnitude, 0.3, that is, by multiplying the length, $7.6cm$ with the value of the reflection coefficient which is 0.3, we will get a value, $2.28cm$.

Next, we draw a circle of radius $2.28cm$ from the origin on the Smith chart (see figure 4.35). Note that the degree of rotation (circumference of the gamma plane—limit circle) is calibrated in both degrees and distance. To locate the corresponding phase, we look along the circumference showing "ANGLE OF REFLECTION COEFFICIENT IN DEGREES" we will locate the angle 60° anticlockwise. Then draw a straight line through the centre of the plane to the located angle 60° on the Smith chart. The point of intersection of the circle and the line is the point, $\Gamma = 0.3\angle 60^o$ on the Smith chart (see figure 4.35).

To locate the point in (e), which is the VSWR of value 2.5, we will locate the point that reads $r = 2.5$ along the real axis on the Smith chart and draw a circle passing through that point from the origin. The VSWR circle of 2.5 is drawn through that point.

To locate the point in (f), first, we draw the VSWR circle of value 1.5 using the steps discussed above. The rightmost side along the circle where the real axis intersects the VSWR circle is the normalized $R_{\max}$ or $r_{\max}$ and the leftmost side is the *normalized $R_{\min}$ or $r_{\min}$. Therefore we record the value at the leftmost side and multiply the value by the characteristic impedance*

$$\begin{aligned} r_{\min} &= \bar{R}_{\min} = \frac{R_{\min}}{Z_0} \\ R_{\min} &= \frac{r_{\min}}{Z_0} = 0.75 \times 50 \\ &= 37.5\Omega \end{aligned}$$

Exercises

Ex. 77 — Locate the following points on the Smith chart for a 70-Ω high frequency lossless line.

(a)$140 + \mathrm{j}91\Omega$

(b)$\Gamma = 0.5\angle 29^o$

(c)$R_{\max}$ on $\rho = 3$ circle

Example 4.13.2 Transmission line analysis using a Smith chart

A 50Ω line is terminated in a load impedance $25 + j35\Omega$. Find the following using a Smith chart.

(a) Reflection coefficient in cartesian and polar form,

(b) Reflection coefficient and impedance at a distance of 0.2λ from the load end of the line.

(c) VSWR on the line.

Solution

We are given $Z_L = 25 + j35\Omega$ and $Z_0 = 50\Omega$, so first, we normalize the load impedance

$$\bar{Z_L} = \frac{25 + j35}{50} = 0.5 + j0.7$$

We mark this point on the Smith chart as shown in figure 4.36.

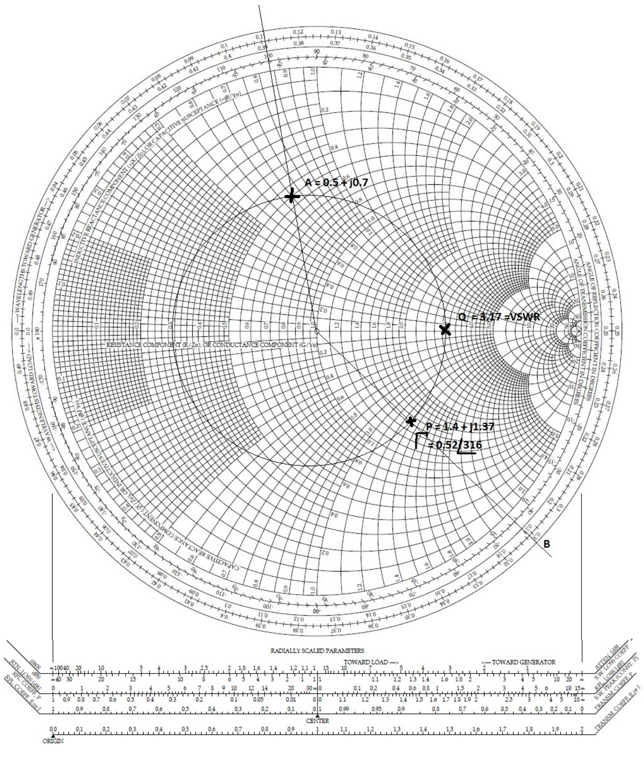


Figure 4.36: Worked Example

To get the reflection coefficient, we draw a circle from the origin of the Smith chart through the marked point (normalized load impedance), A. The distance between point A and the centre gives the magnitude of the reflection coefficient. [21]. *We read the value as;*

$$|\Gamma_L| = 0.52$$

To determine the angle ϕ_L, we draw a straight line from the origin through point A to the outermost circle as shown in figure 4.36. From the scale that reads "ANGLE OF REFLECTION COEFFICIENT" , read the angle the line makes from the positive real axis. The value gives the angle of the reflection coefficient in polar form so that the reflection coefficient at the load is $0.52\angle 100^o$ in polar form.

The cartesian form is obtained by converting the polar form to the cartesian form as shown below

$$\begin{aligned}\Gamma_L &= 0.52\angle 100^o \\ &= 0.52e^{j100} \\ &= 0.52\cos 100 + j\sin 100 \\ &= -0.09 + j0.512\end{aligned}$$

The next step is to get the reflection coefficient at a distance of 0.2λ from the load point, that is, $\Gamma(0.2\lambda)$.

First, we identify the scale on the outermost circle that reads "WAVELENGTHS TOWARDS GENERATOR" and measure the value measured at point A which we obtained as 0.11λ. To move a distance 0.2λ towards the generator means to move clockwise along the scale[22]. *We mark this point as "B", which has the value $0.11\lambda + 0.2\lambda = 0.31\lambda$. We draw a straight line through the origin of the Smith chart and point B and determine the point of intersection with the VSWR circle. The point is marked "P" (see figure 4.36) which is the point of the reflection coefficient moved 0.2λ towards the generator.*

Recall the magnitude of the reflection coefficient is given by the distance between the centre of the smith chart and the point "P". And the angle is measured from the positive real axis. We read the value as

$$\Gamma(0.2\lambda) = 0.52\angle 316^o$$

To find the normalized impedance at point P, we locate the constant resistance and reactance circles through the point. In our case we obtained

$$\bar{Z_L} = 1.4 - j1.37$$

$$\begin{aligned}\text{Thus, } Z_L &= Z_0 \times \bar{Z_L} \\ &= 50(1.4 - j1.37) \\ &= 70 - j68.5\Omega\end{aligned}$$

Next, we obtain the value ρ by reading the value at the rightmost side of the VSWR circle (see figure 4.36).

$$VSWR, \rho = 3.8$$

Example 4.13.3 Transmission line analysis with Admittances

A line is terminated in a normalized admittance of $\bar{Y_L} = 0.2 - j0.5\Omega$. Find;

(a) The location of the maximum voltage from the load end.

(b) The reflection coefficient, normalized admittance and normalized impedance at a distance 0.12λ from the load.

Solution

Now, in this problem, we are dealing with admittances instead of impedances. The admittance we are given here is already normalized, so we mark the point, $\bar{Y_L} = 0.2 - j0.5$, on the Smith chart as shown in figure 4.37.

Next, we draw a circle from the origin of the Smith chart through that point $\bar{Y_L}$ which should help us determine the position of maximum voltage, reflection coefficient, normalized admittance and normalized impedance at 0.12λ from the load.

[21] The distance is read from the scale for reflection coefficient E or I at the bottom of the Smith chart page

Note: This method is different from that used in example 4.13.1 and both methods are correct

[22] Similarly, moving towards the load on the Smith chart means to move in the anticlockwise direction along the scale that reads "WAVELENGTHS TOWARDS LOAD" .

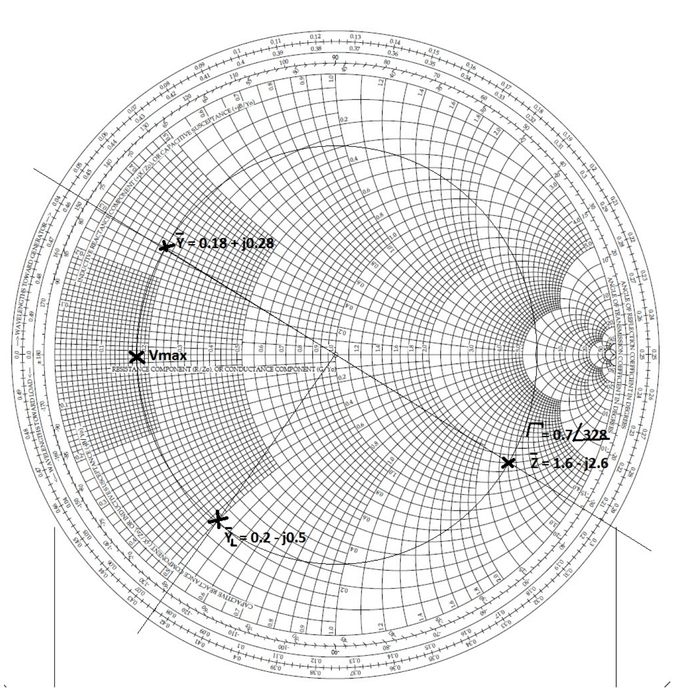


Figure 4.37: Worked Example

(a) Location of the voltage maximum from the load end

For the admittance Smith chart, the maximum voltage will now lie on the leftmost side of the VSWR circle of the Smith chart. This point is marked as $V_{\max}$ on the Smith chart (see figure 4.37). The distance between $\bar{Y}_L$ and $V_{\max}$ on the Smith chart is the location of the maximum voltage, moving towards the generator. [23].

(b) Reflection coefficient, normalized admittance and impedance at a distance 0.12λ from the load

First, we note the distance corresponding to $\bar{Y}_L$ which we measured as 0.424λ on the Smith chart. Moving a distance of 0.12λ towards the generator implies that the new location is $0.424\lambda + 0.12\lambda = 0.544\lambda$. But the highest value on the wavelength scale is 0.5λ so the value on the scale which corresponds to 0.544λ is 0.044λ[24]*. Next, we draw a line from the origin of the Smith chart through that point and mark the point of intersection of the line with the VSWR circle as $\bar{Y}$ (see figure 4.37). We read the corresponding value of normalized admittance as:*

$$\bar{Y} = 0.18 + j0.28\Omega$$

To get the normalized impedance at the distance of 0.12λ from the load point, we transform the admittance by $\frac{\lambda}{4}$[25]*. Put differently, the normalized admittance is the reciprocal(inverse) of the normalized impedance. A transformation by $\frac{\lambda}{4}$ can easily be calculated on the Smith chart by drawing a straight line from the normalized admittance, $\bar{Y}$ through the origin such that it divides the Smith chart into halves. We mark this point as $\bar{Z}$ (see figure 4.37) which is 180°from the point $\bar{Y}$. This point reads the values of normalized impedance as:*

$$\bar{Z} = 1.6 - j2.6\Omega$$

[23]Recall we move clockwise when moving towards the generator.
[24]$0.544\lambda - 0.5\lambda$
[25]Recall that the normalized impedance inverts itself every $\frac{\lambda}{4}$

Finally, we need to measure the reflection coefficient. To get the magnitude, we measure the distance between the origin of the Smith chart and $\bar{Z}$[26]*. While we get the phase angle by measuring the angle subtended for the positive real axis. That is,*

$$\Gamma = 0.7\angle 328^o$$

All these, demonstrate the effectiveness of the Smith chart in moving from impedance to admittance and vice versa. We switch from impedance to admittance by diagonally switching on the Smith chart. That is, at the load point for instance, $\bar{Y_L}$, if we go diagonally opposite, we get $\bar{Z}_L$. So basically, it helps us invert complex numbers.

4.13.4 Identifying from Standing Wave Pattern

We have so far discussed how to solve some simple common transmission line problems using both impedances and admittances and then we discussed the constant VSWR circle. In this section, we will discuss how to identify the from standing wave patterns. **Standing wave pattern** has two important characteristics which are:

(i) The location of maximum and minimum (current/voltage)

(ii) The VSWR circle.

So *how can we quickly identify the type of load*[27] *without calculating the VSWR?*

Recall that:

$$VSWR = \frac{|V|_{\max}}{|V|_{\min}} \tag{4.38}$$

From equation (4.38), the smaller the value of $|V|_{\min}$, the larger the value of $VSWR$, that is, as $|V|_{\min}$ tends to zero, $VSWR$ tends to infinity. When $|V|_{\max} \approx |V|_{\min}$, then $VSWR = 1$.

First, let us observe the variation of the impedance on the transmission line with the Smith chart then we will revisit our early question. Figures 4.38 and 4.39 are simplified Smith charts showing the VSWR circle for inductive and capacitive loads at the top and bottom areas respectively.

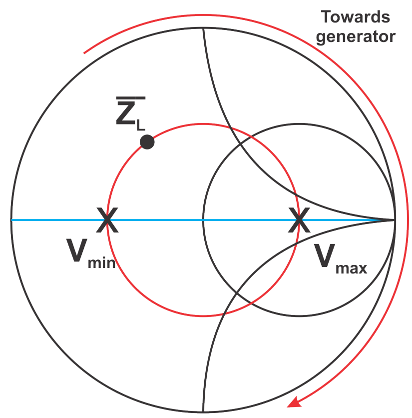


Figure 4.38: A Smith chart representation of inductive load

As shown in figure 4.38, if we move from the load towards the generator, that is, a clockwise rotation, for the inductive load we encounter $V_{\max}$ first before $V_{\min}$ in another $\frac{\lambda}{4}$ or πrad

[26]This is done either with a pair of dividers preferably to take the length and read its distance on the scale at the bottom of the Smith chart page. Alternatively, divide the distance measured with a pair of dividers by the distance measured for the $r = 1.0$ constant resistance circle.
[27]The type of load is not the exact value of the load

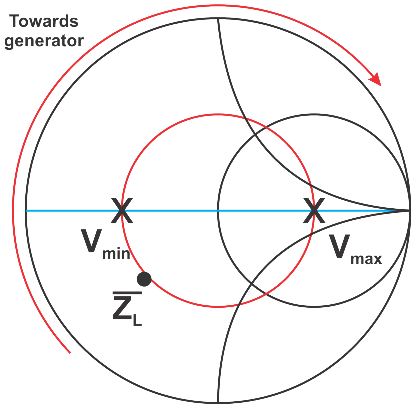


Figure 4.39: A Smith chart representation of capacitive load

movement. Similarly, if we move from $\bar{Z}_L$ in clockwise direction for capacitive load, we encounter $V_{\min}$ first before $V_{\max}$, in another $\frac{\lambda}{4}$ or πrad movement. Also, $V_{\min} \neq 0$ which means that for a complex inductive load that is not purely reactive, $VSWR \neq \infty$.

Complex Reactive Loads

Now, let us analyze some standing wave graphs and determine the type of load which they represent. The transmission line circuit is shown in figure 4.40 and we can analyse the standing wave pattern based on the new observation we have discussed. As the standing wave pattern moves from load towards the generator from right to left, it gets to a maximum first and its minimum is not zero so it $|V_{\min}|$ does not touch the real axis. This indicate a complex inductive load which is similar to being on the upper half of the smith chart.

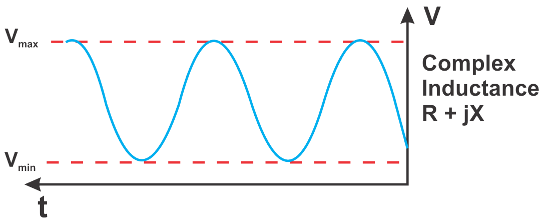


Figure 4.40: Standing wave pattern variation of a complex inductive load

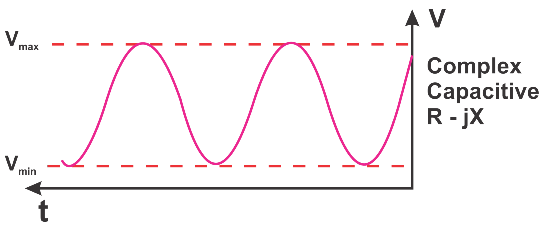


Figure 4.41: Standing wave pattern variation of a complex capacitive load

Similarly, for a complex capacitive load, $V_{\min} \neq 0$ and we encounter $V_{\min}$ first before $V_{\max}$ as shown in figure 4.41.

Purely Reactive Loads

Let us observe the with pure reactive components at the load end. As shown in figure 4.42, the voltage minimum is zero or it touches the horizontal axis that is $VSWR = \frac{V_{\max}}{V_{\min}} = \infty$ as $V_{\min} = 0$.

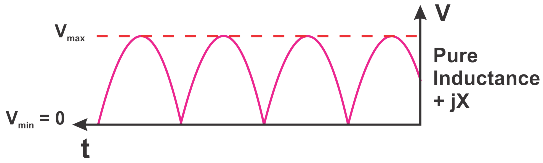


Figure 4.42: Standing wave pattern variation of a purely inductive load

As we move from the load end we meet $V_{\max}$ first indicating it is purely inductive and lies on the upper half of the Smith Chart.

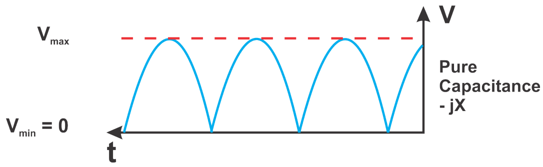


Figure 4.43: Standing wave pattern variation of a purely capacitive load

For purely capacitive loads, the voltage minimum is also zero, but we encounter a minimum first as we go from the load end towards the generator indicating a purely capacitive load. So it lies in the lower half of the Smith chart and $VSWR = \infty$.

Resistive Loads

Lastly, there is the case when the load is neither capacitive nor inductive. If so, then the load must lie on the real horizontal axis of the Smith Chart, so the location of the load itself will be at either the voltage minimum point or the voltage maximum point. If the load is purely resistive as in figure 4.44. Recall that on the Smith Chart, the points of intersection of the VSWR circle on the real axis will be the locations of the voltage maximum and minimum. So there will either be maximum or minimum voltage at the load end as shown in figure 4.44. The solid curve has $V_{\max}$ at load end meaning $R > Z_0$ while the dashed curve has $V_{\min}$ at load end which means $R < Z_0$. Z_0 is the characteristics impedance.

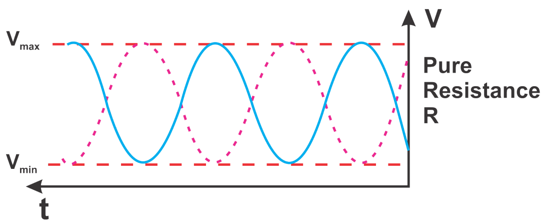


Figure 4.44: Standing wave pattern variation of a pure resistive

So looking at the standing wave pattern, one can quickly identify the types of loads because the information about the load is completely available from the standing wave pattern. Hence, $V_{\max}$ and $V_{\min}$ location and lowest value of $V_{\min}$ which is related to VSWR can help us identify loads very quickly.

So in this section, we have shown how to identify the load by looking at . Next, we go to applications of transmission lines since we make use of sections of transmission lines in realizing various circuit elements in high-frequency circuits.

Exercises

Ex. 78 — Why is the phase measurement crucial in impedance measurement

Ex. 79 — What factors makes the measurement of phase difficult at high frequencies

Ex. 80 — At high frequencies, what type of transmission line is used for measuring unknown impedance

Ex. 81 — Highlight the step-wise approach used to determine the impedance from the standing wave pattern

Ex. 82 — From what axis must a load which is neither inductive nor capacitive lie on a smith chart

Ex. 83 — During impedance calculations, what is the relevance of the ”sense of rotation”

Ex. 84 — Starting from the VSWR relationship, show the mathematical expression for impedance calculation, highlighting its real and imaginary parts

Ex. 85 — In the graphical method of analyzing and solving transmission line problems, what is the reason behind using normalized values

4.14 Applications of transmission lines

This section deals with the applications of transmission lines at high frequencies. Previously, we dealt with the analysis of transmission lines that is, the derivation of transmission line equations, power flow in a transmission line, voltage and wave characteristics of transmission lines, use of a graphical tool called a Smith chart to analyze the transmission line.

There are many applications of transmission lines but at high frequencies, many of the reactive elements are replaced by transmission line sections. This is the part we shall discuss in this section.

(i) Measurement of Unknown Impedance

(ii) As a Circuit Element

(iii) As a Resonant Circuit

(iv) Step Up Transformer

(v) Matching Impedance

4.14.1 Measurement of

The measurement of **phase** is a very difficult task at high frequency. Phase measurement is crucial in impedance measurement because the impedance of a circuit consists of both the resistance and the reactance, and the reactance component depends on the phase relationship between the current and the voltage in the circuit. At high frequencies, it becomes difficult to measure the phase of signals accurately due to several factors, such as signal distortion, noise, and the limited frequency response of the measurement instrument. For instance, the phase shift caused by the circuit under test becomes very small at high frequencies. Additionally, the **propagation delay** in cables, connectors and other components used in the measurement setup becomes significant at high frequencies and can introduce errors in the phase measurement. Therefore, in order to measure complex impedances, we have to measure the complex voltage and complex current. Since the measurement of phase is not that simple, the measurement of **complex impedance** becomes difficult at high frequencies. In transmission lines, the temporal phase between the voltage and current gets translated into the spatial phase in the form of a standing wave pattern which means we can assume that the temporal phase between voltage and current is gotten by measuring the standing wave pattern on the transmission line. This is the method used for measuring unknown impedance at high frequency, using a special transmission line called **slotted transmission line**.

The slotted transmission line as shown in figure 4.45 consists of a movable probe inserted in a slot in a transmission line. It determines the impedance of the high-frequency circuit by using the interaction between the electromagnetic wave and the slots in the transmission line. One advantage of slotted transmission lines is that they can be used to measure the impedance of a wide range of circuits, including active and passive circuits, without the need for direct electrical contact. Additionally, slotted transmission lines can provide accurate measurements of both the magnitude and phase of the impedance, making them useful for a variety of high-frequency circuit applications. In the measurement of unknown impedance, it gives access to the transmission line in real time, to measure the voltage amplitude of the standing wave found on it. Thus by observing the standing wave (signal amplitude), one can more easily obtain the phase of the complex unknown load impedance.

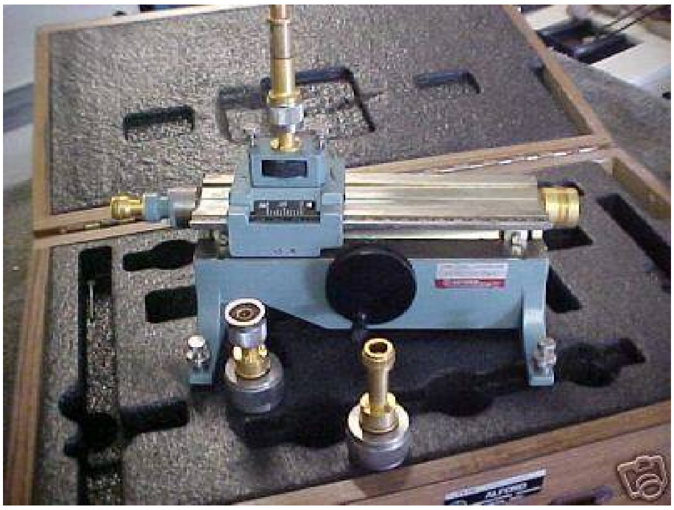


Figure 4.45: Slotted transmission line

The groove in the slotted transmission line is used for measuring the amplitude of the voltage along the transmission line. Hence, a voltage probe slides along the transmission line, measuring the magnitude variation of the voltage from one end of the transmission line. In order to measure the unknown impedance, the voltage source has to be at one end (that is, generator end) while the load (unknown impedance) is at the load end of the transmission line as shown in figure 4.46.

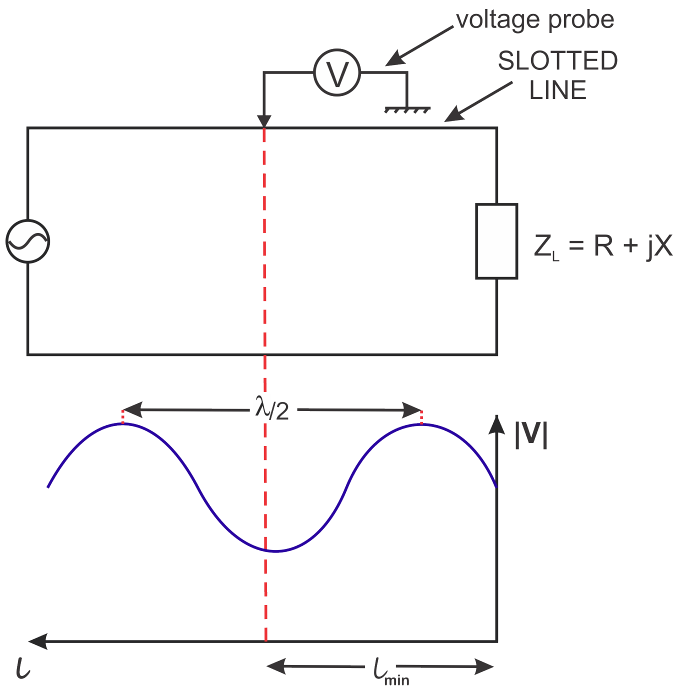


Figure 4.46: Measurement using the Slotted transmission line

The setup for the measurement is depicted in figure 4.46. The unknown impedance which is to be measured is connected to the end of the transmission line and then the transmission line is excited with the source of desired frequency. From the standing wave patterns, the steps taken to determine the impedance are outlined as follows:

(i) Measure the separation between the two maxima or minima on the transmission line to estimate the value of the wavelength. Since the distance between two maxima or minima is $\frac{\lambda}{2}$.

(ii) Find the phase constant "β" using $\beta = 2\frac{\pi}{\lambda}$.

(iii) Measure $l_{\min}$, which is the location of the voltage minimum from the load end of the transmission line.

(iv) Measure $V_{\max}$ and $V_{\min}$ to determine the VSWR ρ

(v) Calculate the impedance using impedance transformation relation

Impedance Calculation

Recall, VSWR is given as

$$\rho = \frac{|V|_{\max}}{|V|_{\min}} \tag{4.39}$$

Once we take measurement of $|V|_{\max}$ and $|V|_{\min}$, then we calculate for ρ. We then find the maximum and minimum impedance which one can see in the transmission line. So at $l_{\min}$ we have $R_{\min}$ at that point and $R_{\min} = \frac{Z_0}{\rho}$. Once $R_{\min}$ is known at $l_{\min}$, the calculation of unknown impedance is simply an impedance transformation problem. We recall that once the impedance at any point on the transmission line is known, the impedance at any other point can be calculated using the impedance transformation relationship.

Since we know $R_{\min} = \frac{Z_0}{\rho}$ at $l_{\min}$ from the load, we can transform this impedance by a distance $l_{\min}$ away from the generator, so that its impedance is equal to terminating impedance Z_L. So the unknown impedance is nothing but the transformation of $R_{\min} = \frac{Z_0}{\rho}$ by $l_{\min}$ to the load side of the generator. Equation (4.40) is the impedance transformation relation with l substituted with $l_{\min}$.

$$Z_L = Z_0 \frac{R_{\min}\cos(-\beta l_{\min}) + jZ_0 \sin(-\beta l_{\min})}{Z_0 \cos(-\beta l_{\min}) + jR_{\min} \sin(-\beta l_{\min})} \tag{4.40}$$

$l_{\min}$ is negative since we are moving away from the generator. Finding $Z_{\min}$ is a matter of separating the real and imaginary parts, we get the value of resistance and reactance[28] as shown below.

$$Z_L = Z_0 \frac{R_{\min}\cos\beta l_{\min} - jZ_0 \sin\beta l_{\min}}{Z_0 \cos\beta l_{\min} - jR_{\min} \sin\beta l_{\min}} \tag{4.41}$$

Rationalize into proper form by multiplying the top and bottom by the conjugate of the denominator.

$$\begin{aligned} Z_L &= Z_0 \left(\frac{R_{\min}\cos\beta l_{\min} - jZ_0 \sin\beta l_{\min}}{Z_0 \cos\beta l_{\min} - jR_{\min} \sin\beta l_{\min}}\right) \\ &\quad\times \left(\frac{Z_0\cos\beta l_{\min} + jR_{\min} \sin\beta l_{\min}}{Z_0 \cos\beta l_{\min} + jR_{\min} \sin\beta l_{\min}}\right) \\ &= Z_0 \frac{Z_0 R_{\min}\cos^2\beta l_{\min} + Z_0 R_{\min}\sin^2\beta l_{\min}}{Z_0^2\cos^2\beta l_{\min} + R_{\min}^2 \sin^2\beta l_{\min}} \\ &\quad + j\frac{(R_{\min}^2 - Z_0^2)\cos\beta l_{\min}\sin\beta l_{\min}}{Z_0^2\cos^2\beta l_{\min} + R_{\min}^2\sin^2\beta l_{\min}} \end{aligned} \tag{4.42}$$

Recall $\cos^2(A) + \sin^2(A) = 1$ and $\frac{\sin(A)}{\cos(A)} = \tan(A)$

$$Z_L = Z_0 \times \frac{Z_0 R_{\min} + j(R_{\min}^2 - Z_0^2)\cos\beta l_{\min}\sin\beta l_{\min}}{(Z_0^2\cos^2\beta l_{\min} + R_{\min}^2\sin^2\beta l_{\min})} \tag{4.43}$$

Divide numerator and denominator by $\cos^2\beta l_{\min}$:

$$Z_L = Z_0 \frac{\frac{Z_0 R_{\min}}{\cos^2\beta l_{\min}} + \frac{j(R_{\min}^2 - Z_0^2)\cos\beta l_{\min}\sin\beta l_{\min}}{\cos^2\beta l_{\min}}}{\frac{(Z_0^2\cos^2\beta l_{\min} + R_{\min}^2\sin^2\beta l_{\min})}{\cos^2\beta l_{\min}}}$$

[28]We are dealing with a lossless transmission line unless otherwise stated.

Recall $\frac{1}{\cos(A)} = \sec(A)$

$$Z_L = Z_0 \left\{\frac{Z_0 R_{\min}\sec^2\beta l_{\min} + j(R_{\min}^2 - Z_0^2)\tan\beta l_{\min}}{Z_0^2 + R_{\min}^2\tan^2\beta l_{\min}}\right\}$$

But $\frac{Z_0}{R_{\min}} = \rho$; dividing the numerator and denominator by $R_{\min}^2$

$$\begin{aligned} Z_L &= Z_0 \left\{\frac{\frac{Z_0 R_{\min}}{R_{\min}^2}\sec^2\beta l_{\min} + \frac{j(R_{\min}^2 - Z_0^2)}{R_{\min}^2}\tan\beta l_{\min}}{\frac{1}{R_{\min}^2}(Z_0^2 + R_{\min}^2\tan^2\beta l_{\min})}\right\} \\ &= Z_0\left\{\frac{\rho\sec^2\beta l_{\min} + j(1-\rho^2)\tan\beta l_{\min}}{\rho^2 + \tan^2\beta l_{\min}}\right\} \\ &= Z_0\left\{\frac{\rho(1+\tan^2\beta l_{\min}) + j(1-\rho^2)\tan\beta l_{\min}}{\rho^2 + \tan^2\beta l_{\min}}\right\} \end{aligned} \tag{4.44}$$

Thus, the normalized impedance to be measured is

$$\begin{aligned} \bar{Z_L} &= \frac{Z_L}{Z_0} \\ &= \frac{\rho(1+\tan^2\beta l_{\min})}{\rho^2 + \tan^2\beta l_{\min}} + \frac{j(1-\rho^2)\tan\beta l_{\min}}{\rho^2 + \tan^2\beta l_{\min}} \end{aligned}$$

Where the real part is $R = Z_0\left[\frac{\rho(1+\tan^2\beta l_{\min})}{\rho^2+\tan^2\beta l_{\min}}\right]$ and the imaginary part is $X = Z_0\left[\frac{(1-\rho^2)\tan\beta l_{\min}}{\rho^2+\tan^2\beta l_{\min}}\right]$.

Hence $Z_L = R + jX$ can easily be calculated from all parameters that we have measured viz; $|V|_{\max}$, $|V|_{\min}$, VSWR, $\rho = \frac{|V|_{\max}}{|V|_{\min}}$, $l_{\min}$, and β from $\beta = 2\frac{\pi}{\lambda}$. It can be seen that we have transformed $R_{\min}$ to Z_L without necessarily knowing the value of $R_{\min}$, as seen in equation (4.44), the expression for X that is, reactance and R that is, resistance, $R_{\min}$ is not calculated. However, in practice, when connecting the unknown impedance to the slotted line section, we use conductors as well as other forms of connectors. This means the location of the impedance is not accurately known, so the measurement of $l_{\min}$ sometimes becomes inaccurate. If we know $l_{\min}$ precisely, then the solution is straightforward and very accurate in finding out what the unknown impedance is. If $l_{\min}$ is not known, there would be an error in the unknown impedance.

To avoid error, we define the location of the load by replacing it with a short circuit. We then carry out the measurement of the transmission line, and find out the location of $l_{\min}$ with the line terminated with a short circuit. Then replace the short circuit with a load and find out the new standing wave pattern with the load. So we have two standing wave patterns on our transmission line,

(i) With the line terminated with a short circuit and

(ii) With the unknown load terminating the line.

For short circuits, we identify the exact location of minimum voltage by the standing wave, which becomes our origin for the unknown load impedance measurement of $l_{\min}$.

Hence the superposition of the load standing wave pattern on the short circuit standing wave is shown in figure 4.47.

The minimum voltage for the short circuit standing wave pattern repeats itself at $\frac{\lambda}{2}$, so we can start our measurement at position 1, 2, or 3 as shown in figure 4.47 and measure the corresponding distance $l_{\min}$ from either of these points towards the left. So instead of starting measurement at position 1 which was not accurate because of conductors used in the connection or the presence of connectors, we can start from 2 or 3 at the origin and then measure $l_{\min}$ of the load that occurs after the

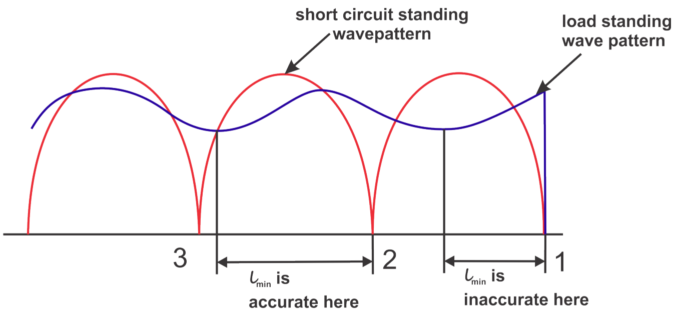


Figure 4.47: Superposition of the load standing wave pattern and the short circuit standing wave

first point. Thus, in practice, to find $l_{\min}$, we measure both the short circuit and the unknown load standing wave patterns, to remove the inaccuracies introduced by the connecting conductors or connectors.

This technique of impedance measurement is very useful. In fact at high frequencies (e.g., microwave frequencies) without measurement like this on a slotted line, one will not be able to measure the unknown impedance. So at high frequencies, the transmission line is suitable for the measurement of unknown impedance.

Example 4.14.1 Unknown Impedance Measurement

We have a slotted transmission line to which an unknown impedance Z_L is connected as shown in figure 4.48. The measurement of the standing wave pattern on the transmission line gave $\frac{\lambda}{2} = 30cm$ and the measurement of the variation of voltage as a function of length l on the transmission line gave $V_{\max} = 15V, V_{\min} = 5V$. Find the unknown impedance Z_L. (Take $Z_0 = 50\Omega$)

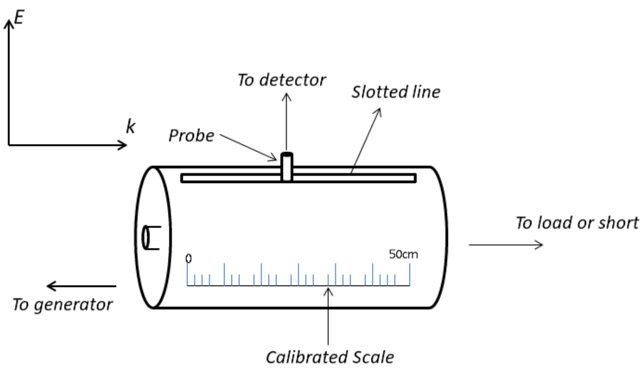


Figure 4.48: Slotted Transmission Line.

(a) Phase Estimation

Recall,

$$\begin{aligned} \phi &= 2\beta l \quad \textit{and } \beta = \frac{2\pi}{\lambda} \\ &= 2 \cdot \frac{2\pi}{\lambda} l \\ &= 2 \cdot \frac{360}{60} \cdot 27 \\ &= 324^o \end{aligned}$$

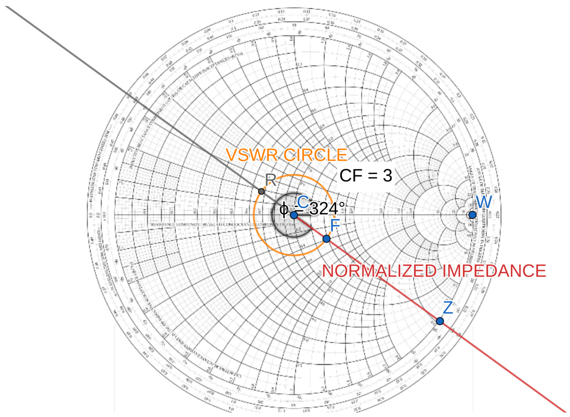


Figure 4.49: Worked Problem

(b) VSWR

$$\rho = \frac{V_{\max}}{V_{\min}} = \frac{15}{5} = 3$$

With all this information, we proceed to the Smith chart to calculate the unknown impedance. It is clear from here that moving a distance of $27cm$ from the load towards the generator, that is, clockwise movement on the Smith chart will correspond to 324°.

We should arrive at the maxima point on the impedance Smith chart, which is the intersection of the constant VSWR circle on the positive real axis. From the origin, we draw a circle with a radius, $\rho = 3$. It touches the real axis at $V_{\min}$ and $V_{\max}$ as shown in figure 4.49. The point where the angle cuts the constant VSWR circle is the unknown impedance $\bar{Z}_L$ (see figure 4.49).

Thus the unknown impedance according to the Smith chart is $\bar{Z}_L = 1.15 - j1.23$. With the characteristics impedance of the line given as 50Ω, then $Z_L = \bar{Z}_L \times Z_0 = 50(1.15 - j1.23) = (57.5 - j61.5)\Omega$. So unknown impedance can be measured very easily with a Smith chart without doing much complex calculation.

4.14.2 As a Circuit Element

Assuming we wound the inductor as shown in figure 4.50 at high frequencies such that its inductance is L. At high frequency, there is a capacitance between the turns of the inductor that almost shorts out current as $X_C = \frac{1}{2\pi f_c}$. This was not the case at low frequencies when the turns appeared separated from one another. Hence, at high frequency, we then have distributed capacitors between the turns of the inductor.

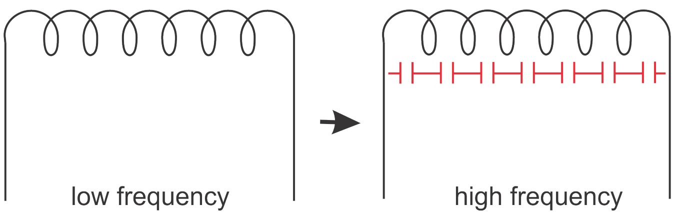


Figure 4.50: An inductor at Low frequency and then at High frequency

As the frequency increases, the inductor starts having capacitive reactance. It shows for a typical wire-wound inductor, the resonance frequency lies for distributed capacitances and inductances in the range of 100— 200MHz. It means if we wound an inductor, beyond about a few 100MHz, the inductor will not exhibit its ideal characteristics, but rather that of a capacitor because we have already crossed the resonance frequency. So the effect of capacitance is more dominant compared to the inductance.

Similarly, suppose we take a capacitor at a low frequency, it will exhibit its ideal characteristics. As the frequency increases, the connecting leads or wires which have inductance starts to predominate beyond certain frequencies, and the capacitor no longer exhibit its ideal characteristics (see figure 4.51).

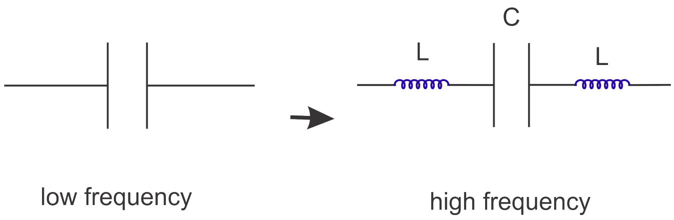


Figure 4.51: An capacitor at Low frequency and then at High frequency

Hence we can reliably design an inductor at low frequencies. But at high frequencies, its reliability is very poor because the inductor is no longer ideal. Similarly, for a capacitor at high frequencies, the inductance at the leads starts to predominate and the capacitor no longer exhibits its ideal characteristics. So at high frequencies, realizing a reactive element is not that easy because we do not have a reliable circuit element which will guarantee its use as a capacitor or inductor. When the frequency is increasing, the wavelength is becoming smaller as we have seen earlier. However, if you consider a short or open circuit transmission line, the input impedance of these lines would behave like a reactance.

So as the frequency increases and the wavelength gets smaller, the size of a section (a small lineal element) of a transmission line which can give you impedance will be reactive; this becomes more physically realizable. Two things have become clear when working with transmission lines and they are;

(i) The lumped circuits are becoming more difficult to realize at high frequencies.

(ii) Realization of reactance by using transmission lines section is easier because as the wavelength becomes smaller, the sections of the transmission line become more feasible to realize an ideal reactive element.

At high frequencies, most reactive sections are replaced by transmission lines as shown in figure 4.52. In the impedance Smith Chart, the open circuit is at the rightmost part and the short circuit is at the leftmost part of the outermost circle[29]. Moving along this outermost circle shows how much length (in terms of wavelength) is needed between load and generator to realize any pure reactance. We can assume the presence of that reactance at that point, from the open or short circuit end towards the signal source or generator.

Hence for an open or short circuit load, by using different lengths of transmission line, we can get different reactance at the input. Therefore, *what should be the length of the transmission line, so that one gets a particular value of reactance at the input?*

[29]The outermost circle has $r = 0$ and it implies that as we move along that circle we trace out pure reactance.

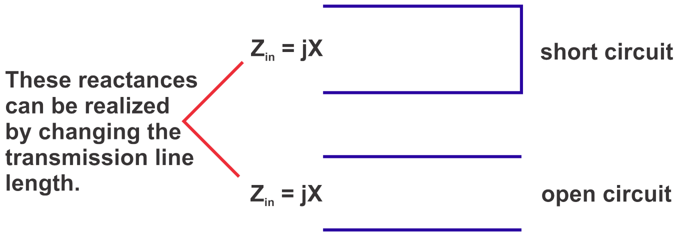


Figure 4.52: Transmission Line as a Circuit Element

Analytical approach of using transmission line as a circuit element

Analytically, l is the distance towards the generator and recall that for a lossless transmission line which we assume unless stated otherwise.

Considering a short circuit load at the end of a transmission line, then using the impedance transformation relation,

$$\begin{aligned} Z_{in} &= Z_0 \left\{ \frac{Z_L \cos\beta l + jZ_0 \sin\beta l}{Z_0 \cos\beta l + jZ_L \sin\beta l} \right\} \qquad Z_L = 0 \text{ for short circuit} \\ &= Z_0 \left\{ \frac{0 + jZ_0 \sin\beta l}{Z_0 \cos\beta l + 0} \right\} \\ &= jZ_0 \tan\beta l \end{aligned} \tag{4.45}$$

Similarly, considering an open circuit load at the end of a transmission line, $Z_L = \infty$ so we take limits as Z_L tends to infinity.

$$\begin{aligned} Z_{in} & \\ &= Z_0 \left\{ \frac{Z_L \cos\beta l + jZ_0 \sin\beta l}{Z_0 \cos\beta l + jZ_L \sin\beta l} \right\} \qquad Z_L = \infty \text{ for open circuit.} \\ &= Z_0 \left\{ \frac{\cos\beta l + j\frac{Z_0}{Z_L} \sin\beta l}{\frac{Z_0}{Z_L} \cos\beta l + j\sin\beta l} \right\} \\ &= Z_0 \left\{ \frac{\cos\beta l + j \times 0}{0 + j\sin\beta l} \right\} \\ &= -jZ_0 \cot\beta l \end{aligned} \tag{4.46}$$

Thus,

$$Z_{in} = jZ_0 \tan\beta l \Rightarrow \text{ short circuit} \tag{4.47}$$
$$Z_{in} = -jZ_0 \cot\beta l \Rightarrow \text{ open circuit} \tag{4.48}$$

Hence the problem of realizing a given reactance from an open or short circuit load involves finding the l that gets the appropriate Z_{in}. Let l_{SC} and l_{OC} represent short circuit and open circuit length respectively, then we have $X = jZ_0 \tan\beta l_{sc}$ and $X = -jZ_0 \cot\beta l_{oc}$ for short circuit and open circuit respectively. Since the phase constant β is known for the transmission line, we make l_{SC} and l_{OC} the subject of the equations, we already know the reactance, X we require, so that

$$l_{SC} = \frac{1}{\beta} \tan^{-1} \frac{X}{Z_0} \tag{4.49}$$
$$l_{OC} = \frac{1}{\beta} \cot^{-1} \frac{-X}{Z_0} \tag{4.50}$$

If the frequency of operation, as well as the velocity of the wave propagation, is known, then we can find λ and use $\beta = \frac{2\pi}{\lambda}$. Therefore, in conclusion, we can realize the unknown reactance by using sections of the transmission line.

As we know, for such calculations as this, the use of the Smith chart comes in handy. The top half of the impedance Smith chart at the outermost circle is pure inductance, and the bottom half is for capacitance. However, before we get into details of Smith chart usage, we have to be sure that $X = Z_0 \tan \beta l_{SC}$ and $X = -Z_0 \cot \beta l_{OC}$ for all variations of βl_{SC} from 0 to 2π, either $\tan \beta l_{SC}$ or $\cot \beta l_{OC}$ varies from $-\infty$ to $+\infty$. Essentially, we want to measure the values of X as we change the values of l_{SC} or l_{OC} in the transmission line, and then we can realize any arbitrary value of the reactance. There is no limit to this. So any value of reactance between $-\infty$ to $+\infty$ can be realized either by short-circuited line or open-circuited line.

The choice of whether to use a short circuit or open circuit line will depend upon the system you want to employ the circuit element. For example suppose we have a parallel wire transmission line, then connecting the two ends of the transmission line is rather easy, as we can short the end of the transmission line. However, a transmission line is like a Printed Circuit Board (PCB), with the ground plane on one side and the line on top separated by a dielectric. To short-circuit the line, we will have to drill a plated-through hole (or a signal-carrying conductive via) into the PCB which requires a lot of precision, so realizing a short circuit line on that configuration is rather difficult. For this reason, an open circuit line is preferred. There may be situations where the short circuit line will be preferred over the open circuit line and vice versa. Irrespective of the one used, we will be able to realize all the reactances from $-\infty$ to $+\infty$, that is, we can realize any capacitive or inductive reactance by using an open-circuited or short circuit section of a transmission line.

Graphical approach of using transmission line as a circuit element

Graphically, we want to find the length, the problem turns out to be the opposite of the impedance calculation. Since we know the input impedance Z_{in} at the generator end and we know terminating impedance, Z_L at the end of the transmission line to be 0 or ∞ (that is, it is a short circuit or open circuit point). So the challenge is that we want to find l_{SC} or l_{OC} where the impedances are known at both ends are known.

The steps taken on the Smith chart as outlined as follows:

(i) Mark the normalized impedance $\bar{Z}_{in}$ position, $\bar{Z}_{in} = jx$

(ii) Mark the point for a short circuit or open circuit, $jx = 0$ or $jx = \infty$ (see figure 4.53).

(iii) Measure the distance from the normalized impedance $\bar{Z}_{in}$ to the short or open circuit moving counter clockwise.

Figure 4.54 shows the distances between the normalized impedance $\bar{Z}_{in}$ and the short circuit end and open circuit end respectively for an inductive reactance. Essentially we move away from the generator to reach the short circuit point, that we move from the point, **A**, till we reach, **B**. On the Smith Chart, clockwise movement is towards the generator and anti-clockwise movement is away from the generator.

Similarly, to realize a capacitive reactance which will be located at the bottom half of the Smith chart for a start. Moving clockwise from short circuit point by no more than $\frac{\lambda}{4}$ we have inductive reactance, but beyond that point from $\frac{\lambda}{4}$ to $\frac{\lambda}{2}$ we have capacitive reactance (see figure 4.55).

Hence if the length lies within 0 and $\frac{\lambda}{4}$ for a short-circuited line, we have inductive reactance while if the length lies from $\frac{\lambda}{4}$

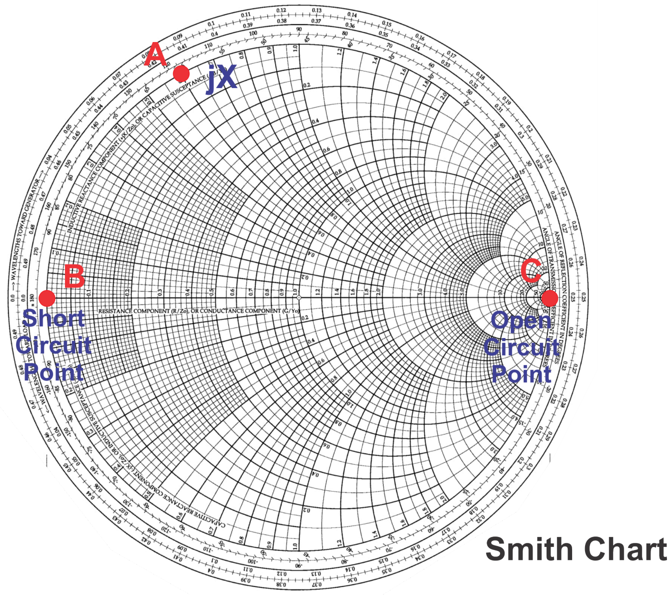


Figure 4.53: Open and Short Circuit points on the Smith Chart

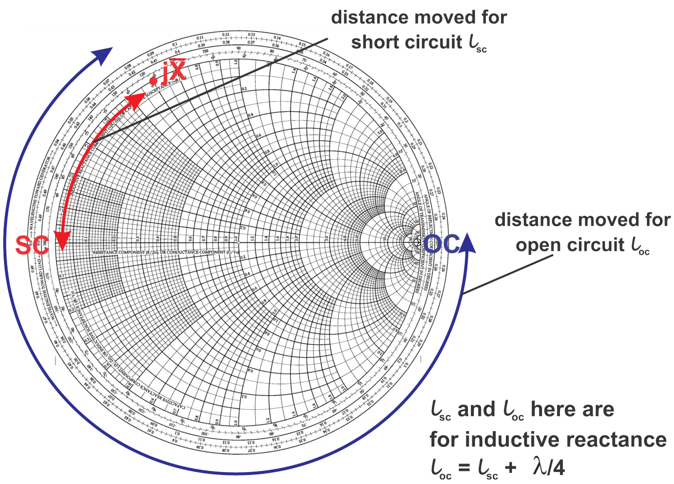


Figure 4.54: l_{SC} and l_{OC} for inductive reactance on the Smith Chart.

to $\frac{\lambda}{2}$ we have capacitive reactance. Similarly for an open circuit line, if the length lies within 0 to $\frac{\lambda}{4}$, we have capacitive reactance and if it lies within $\frac{\lambda}{4}$ to $\frac{\lambda}{2}$ we have inductive reactance. It can be written explicitly in the figure 4.56:

So any section of a transmission line of size 0 to $\frac{\lambda}{2}$ can realize any arbitrary value of capacitance and inductance at a particular frequency. We are not realizing a capacitor or inductor in the actual sense, but we are realizing capacitive and inductive reactance.

Example 4.14.2 Trasmission line as a circuit element

A lossless transmission line is terminated in a short circuit. How long (in wavelengths) should the line be in order for it to appear as an open circuit at its input terminal?

Solution

From equation (4.45)*, we have:*

$$Z_{in} = jZ_0 \tan \beta l$$

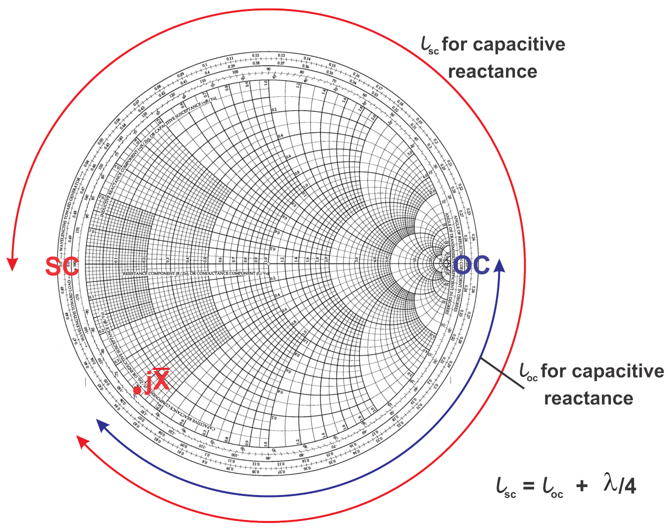


Figure 4.55: l_{SC} and l_{OC} for capacitive reactance on the Smith Chart

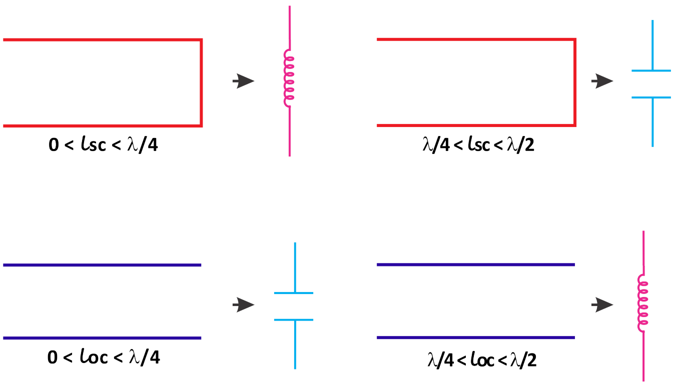


Figure 4.56: Nature of impedance for l_{SC} and l_{OC} after moving the different lengths

If $\beta l = (\frac{\pi}{2} + n\pi)$ *when* n *is an positive integer, then* $Z_{in} = j\infty\Omega$*. So we have:*

$$\beta l = \frac{\pi}{2} + n\pi$$
$$l = \frac{\pi}{2\beta} + \frac{n\pi}{\beta}$$
$$l = \frac{\lambda}{4} + \frac{n\lambda}{2}$$

Therefore the length of the line can be any odd multiple of $\frac{\lambda}{4}$*. For example, if* $n = 1$*, then* $l = \frac{3\lambda}{4}$*, if* $n = 3$*, then* $l = \frac{7\lambda}{4}$ *and so on.*

4.14.3 As a Resonant circuit

At $\frac{\lambda}{4}$ the behaviour of the line changes from X_L to X_C or X_C to X_L. Such a behaviour is similar to the characteristics of resonance, that is, at $\frac{\lambda}{4}$ location, we have characteristics that are more like resonance characteristics. Let us suppose we have a circuit, at resonant frequency[30], a part of the circuit has inductive behaviour while the other part has capacitive behaviour.

So a section of a transmission line may not only be used as a reactive element but it can also be used as a resonant LC circuit[31]. Now we see that suppose you take a line at $\frac{\lambda}{4}$ to $\frac{\lambda}{2}$, it will have a behaviour similar to the LC resonant circuit. It means that if we change the frequency for a given length of a transmission line, the impedance seen between the terminals of the line is going to change.

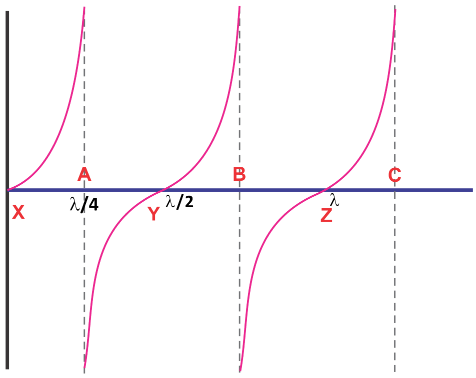


Figure 4.57: Resonance Characteristics for a short circuit end transmission line

For a short circuit whose length is $\frac{\lambda}{2}$, the impedance is equal to zero since the impedance repeats every $\frac{\lambda}{2}$. At $\frac{\lambda}{4}$ the line will be an open circuit and the impedance will be ∞. When the frequency is such that the length is zero (0) or $\frac{\lambda}{2}$ or λ we get zero impedance (see figure 4.57). However, if we go to the frequency for which the length is $\frac{\lambda}{4}$, the impedance seen between the terminals of the lines is ∞ since it will appear like an open circuit.

So for a given frequency as we change the line length, when the line length is $0, \frac{\lambda}{2}, \lambda$ we see impedances at X, Y and Z. At frequency for which the length is $\frac{\lambda}{4}, \frac{3\lambda}{4}, \frac{5\lambda}{4}$ we see impedance at A, B and C[32]. Starting from the short circuit (SC) point on the Smith chart $\frac{\lambda}{4}, \frac{3\lambda}{4}, \frac{5\lambda}{4}$ movement will correspond to an open circuit or ∞ impedance, hence A, B and C on the graph. Also, $\frac{\lambda}{2}, \lambda, \frac{3\lambda}{2}$ will correspond to an SC point, which is the point we started from[33]. Hence in the frequency range close to X, Y and Z, we have characteristics similar to that of a series LC circuit at resonance (see figure 4.58). However, in the vicinity of A, B and C the behaviour is similar to that of a parallel LC circuit at resonance (see figure 4.59).

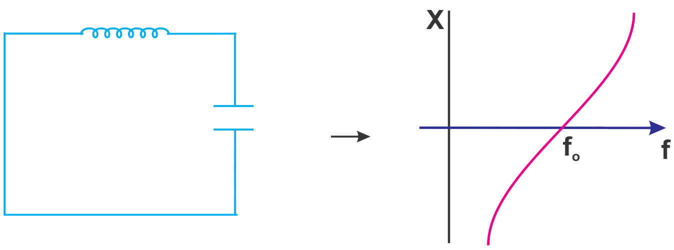


Figure 4.58: Series LC Circuit

Resonance for Series LC Circuit

From the characteristics of a series LC circuit, we know that at the resonance frequency, $X_L = X_C$. As frequency increases, then $X_L > X_C$ then we have a positive X, and conversely, as

[30] The resonance frequency is a frequency point where the inductive reactance of the inductor becomes equal in value to the capacitive reactance of the capacitor $X_L = X_C$

[31] It consists of an inductor and capacitor connected together. It is used for generating signals at a particular frequency or picking out a signal at a particular frequency from a more complex signal

[32] We are dealing with a short circuit for this analysis

[33] Recall that a complete 2π movement correspond to $\frac{\lambda}{2}$ on the Smith Chart

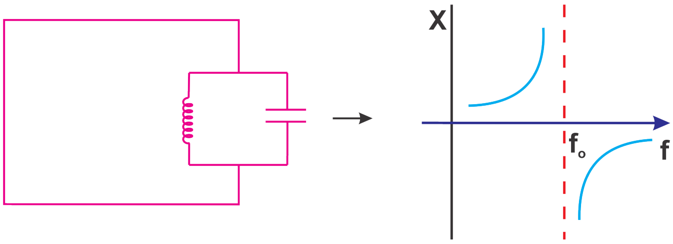


Figure 4.59: Parallel LC Circuit

frequency decreases, then $X_C > X_L$ therefore we have negative X (see figure 4.58).

Resonance for Parallel LC Circuit

At the resonance frequency, the equivalent impedance, X is ∞ that is

$$\frac{1}{X} = \frac{1}{X_C} + \frac{1}{X_L} = 0$$

At lower frequency, $\frac{1}{X_L}$ is larger than $\frac{1}{X_C}$. So X_L component dominates in the equivalent X. At higher frequencies, the $\frac{1}{X_C}$ is larger than $\frac{1}{X_L}$ (see figure 4.59).

So a transmission line can behave like a series resonance circuit of a series LC circuit and can behave like a parallel resonance circuit of a parallel LC circuit. For a given length of transmission line, those frequencies for which the length of the line tends to $\frac{\lambda}{2}$ the line will behave like a series resonance circuit. On the other hand for those frequencies for which the length of the line tends to $\frac{\lambda}{4}$, the line behaves like a parallel resonance circuit. The analysis of the transmission line as a resonant circuit was done for a short circuit end and as such the opposite observations will be made for an open circuit end transmission line.

For an open circuit whose length is $\frac{\lambda}{2}$, the impedance is equal to infinity. At $\frac{\lambda}{4}$ the line will be a short circuit and the impedance will be equal to zero. One side of that frequency will be capacitive that is, negative. So when the frequency is such that the length is zero (0) or $\frac{\lambda}{2}$ or λ we get an impedance equal to infinity. However, if we go to the frequency for which the length is $\frac{\lambda}{4}$, the impedance seen between the terminals of the lines is zero. So for a given frequency as we change the line length, when the line length is 0, $\frac{\lambda}{2}$, λ we see impedances at A, B and C as shown in figure 4.60. At frequency for which the length is $\frac{\lambda}{4}$, $\frac{3\lambda}{4}$, $\frac{5\lambda}{4}$ we are at X, Y and Z.

For an open circuit line for length around $\frac{\lambda}{2}$, the line will behave like a parallel resonant circuit because the input impedance will be at infinity. At about $\frac{\lambda}{4}$, the line is like a short circuit and so behaves like a series resonant circuit. If the length of the line is $\frac{\lambda}{2}$ the line will behave like a parallel resonant circuit. So depending on whether a short circuit line or open circuit line is used, the series or parallel resonance circuit can be realized at different frequencies. So invariably, at high frequencies, the transmission line can be used for realizing circuit elements and series or parallel resonance circuits.

Quality Factor of a resonant transmission line circuit

So *what is the quality factor of the resonant circuit?* By definition, the quality factor is related to the losses of the circuit (the higher the loss, the smaller the quality factor). So whenever we have a reactive element like an inductor or capacitor, the losses in these elements characterize the quality factor. Quality Factor is a parameter that describes how under-damped an oscillator or resonator is and characterizes a resonator's bandwidth relative to its centre frequency. It is a dimensionless quantity. It is a measure of the quality of a resonant circuit.

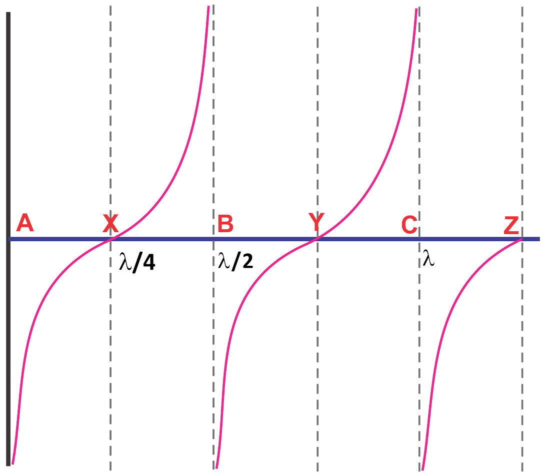


Figure 4.60: Resonance Characteristics for an open circuit end transmission line

If we are dealing with a lossless transmission line, then the quality factor for the resonance circuits should be infinite. There is no loss, however, in practice, there is always a small loss in the transmission line. Hence we will shift our analysis from a lossless transmission line to a low-loss transmission line. This analysis will be treated next.

Exercises

Ex. 86 — A 75-Ω resistive load is preceded by a $\frac{\lambda}{4}$ section of a 50-Ω lossless transmission line, which is itself preceded by another $\frac{\lambda}{4}$ section of a 100-Ω line. What is the input impedance?

Let us consider a transmission line with losses at the length of $l = \frac{\lambda}{4}$ so that we can then determine the actual value of the quality factor for this case. The transmission line can be either an open circuit or a short circuit as shown in figure 4.61.

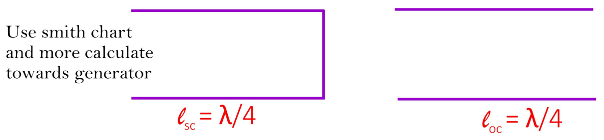


Figure 4.61: Short and Open Circuit end transmission line of length $\frac{\lambda}{4}$

At $\frac{\lambda}{4}$, the open circuit will appear as a short circuit to the other end of the line and similarly, the short circuit will appear as an open circuit. Therefore, a short circuit end transmission line will appear as a parallel resonance circuit at $l_{SC} = \frac{\lambda}{4}$ and an open circuit end transmission line will appear as a series resonance circuit at $l_{OC} = \frac{\lambda}{4}$. So, in order to find the quality factor, we first calculate the actual input impedance of these lines with losses at the given length.

Impedance of resonant circuit with losses

To calculate the input impedances of the lines Z_{OC} and Z_{SC}, for a length $l = \frac{\lambda}{4}$ with losses, recall the impedance transformation relationship is

$$
\begin{aligned}
&Z(l)\\
&= Z_0 \left(\frac{Z_L \cosh \gamma l + Z_0 \sinh \gamma l}{Z_0 \cosh \gamma l + Z_L \sinh \gamma l} \right) \quad \text{for short circuit } Z_L = 0\\
&\quad = Z_0 \left(\frac{0 \times \cosh \gamma l + Z_0 \sinh \gamma l}{Z_0 \cosh \gamma l + 0 \times \sinh \gamma l} \right)\\
&\quad = Z_0 \left(\frac{Z_0 \sinh \gamma l}{Z_0 \cosh \gamma l} \right)\\
&\quad = Z_0 \left(\frac{\sinh \gamma l}{\cosh \gamma l} \right)
\end{aligned}
$$

From trigonometry functions $\tanh \theta = \frac{\sinh \theta}{\cosh \theta}$

$$Z(l)_{SC} = Z_0 \tanh \gamma L \tag{4.51}$$

We adjust the expression for open circuit length at $Z_L = \infty$ and take limits at it tends to ∞, thus

$$
\begin{aligned}
Z(l)_{OC} &= Z_0 \left(\frac{\cosh \gamma l + \frac{Z_0}{Z_L} \sinh \gamma l}{\frac{Z_0}{Z_L} \cosh \gamma l + \sinh \gamma l} \right)\\
&= Z_0 \left(\frac{\cosh \gamma l + 0 \times \sinh \gamma l}{0 \times \cosh \gamma l + \sinh \gamma l} \right) \qquad (4.52)\\
&= Z_0 \left(\frac{\cosh \gamma l}{\sinh \gamma l} \right)
\end{aligned}
$$

From trigonometry functions $\frac{1}{\tanh \theta} = \frac{\cosh \theta}{\sinh \theta} = \coth \theta$

$$Z(l)_{OC} = Z_0 \coth \gamma l \tag{4.53}$$

$Z(l)_{SC} = Z_0 \tanh \gamma l$ and $Z(l)_{OC} = Z_0 \coth \gamma l$, but with $\gamma = \alpha + j\beta$, we have $\alpha \ll \beta$ for low loss transmission line[34], i.e $\gamma \neq j\beta$, the α part must be taken into consideration as well. We know that at $l = \frac{\lambda}{4}$, the input impedance of a short circuit line is an open circuit and that of an open circuit line is a short circuit. However, in the presence of a loss, that will not be true. The impedance will neither be infinity for Z_{SC} nor zero for Z_{OC}. If the loss is present, the input impedance will give

$$Z_{SC} = Z_0 \tanh \gamma l = Z_0 \tanh(\alpha + j\beta)l \tag{4.54}$$

Where $\gamma = \alpha + j\beta$ for a low-loss transmission line.

From trigonometry $\tanh(A + B) = \frac{\tanh A + \tanh B}{1 + \tanh A \tanh B}$

$$\tanh(\alpha + j\beta)l = \frac{\tanh(\alpha l) + \tanh(j\beta l)}{1 + \tanh(\alpha l)\tanh(j\beta l)} \tag{4.55}$$

Recall that $\tanh jA = j \tan A$.
Therefore,

$$Z_0 \tanh(\alpha + j\beta)l = Z_0 \left(\frac{\tanh \alpha l + j \tan \beta l}{1 + j \tanh \alpha l \tan \beta l} \right) \tag{4.56}$$

If α is small compare to β for a length of $\frac{\lambda}{4}$ of transmission line, αl is much smaller than 1. For a low loss line, $\alpha \ll \beta$ and $\beta = \frac{2\pi}{\lambda}$. Therefore $\tanh \alpha l \approx \alpha l$, then we have

$$Z_{SC} \approx Z_0 \left(\frac{\alpha l + j \tan \beta l}{1 + j\alpha l \tan \beta l} \right) \tag{4.57}$$

For $l = \frac{\lambda}{4}, \beta = \frac{2\pi}{\lambda} \implies \beta l = \frac{2\pi}{\lambda} \times \frac{\lambda}{4} = \frac{\pi}{2}$
Dividing both the numerator and denominator of equation 4.57 by $\tan \beta l$

$$
\begin{aligned}
Z_{SC} &= Z_0 \left(\frac{\alpha l + j \tan \beta l}{1 + j\alpha l \tan \beta l} \right) \times \left(\frac{\frac{1}{\tan \beta l}}{\frac{1}{\tan \beta l}} \right)\\
&= Z_0 \left(\frac{\frac{\alpha l}{\tan \beta l} + \frac{j \tan \beta l}{\tan \beta l}}{\frac{1}{\tan \beta l} + \frac{j\alpha l \tan \beta l}{\tan \beta l}} \right)\\
&= Z_0 \left(\frac{\frac{\alpha l}{\tan \beta l} + j}{\frac{1}{\tan \beta l} + j\alpha l} \right) \qquad \tan \beta l = \tan \frac{\pi}{2} = \infty\\
&= Z_0 \left(\frac{\frac{\alpha l}{\infty} + j}{\frac{1}{\infty} + j\alpha l} \right) \qquad \frac{\alpha l}{\infty} = 0,\ \frac{1}{\infty} = 0\\
&= Z_0 \left(\frac{j}{j\alpha l} \right)\\
&= Z_0 \left(\frac{1}{\alpha l} \right)
\end{aligned}
$$

$$Z_{SC} = \frac{Z_0}{\alpha l} \tag{4.58}$$

Therefore, this means that the input impedance of a short circuit line, if the length is $\frac{\lambda}{4}$ is $\frac{Z_0}{\alpha l}$, which is not infinity. Recall $Z_{SC} = \infty$ for a lossless transmission line, but $\alpha l \ll 1$ means Z_{SC} is large but not infinity.

Similarly, the open circuit impedance is

$$
\begin{aligned}
Z_{OC} &= Z_0 \coth \gamma l\\
&= \frac{Z_0}{\tanh \gamma l}\\
&= Z_0 \coth \gamma l\\
&= Z_0 \coth(\alpha + j\beta)l
\end{aligned}
$$

For low loss transmission line $\alpha \ll \beta$.

Recall,

$$\tanh(\alpha + j\beta)l = \frac{\tanh(\alpha l) + \tanh(j\beta l)}{1 + \tanh(\alpha l)\tanh(j\beta l)}$$

[34] A low-loss transmission line is not the same as a lossless transmission line

But,

$$\frac{1}{\tanh(\alpha + j\beta)l} = \frac{1 + j\tanh(\alpha l)\tanh(j\beta l)}{\tanh(\alpha l) + \tanh(j\beta l)}$$

And lastly, $\tanh jA = j\tan A$

$$\begin{aligned} Z_{OC} &= \frac{Z_0}{\tanh(\alpha + j\beta)} \\ &= Z_0\left(\frac{1 + j\tanh\alpha l\tan\beta l}{\tanh\alpha l + j\tan\beta l}\right) \\ &\approx Z_0\left(\frac{1 + j\alpha l\tan\beta l}{\alpha l + j\tan\beta l}\right) \end{aligned}$$

Divide both the numerator and denominator by $\tan\beta l$

$$\begin{aligned} Z_{OC} &= Z_0\left(\frac{\frac{1}{\tan\beta l} + \frac{j\alpha l\tan\beta l}{\tan\beta l}}{\frac{\alpha l}{\tan\beta l} + \frac{j\tan\beta l}{\tan\beta l}}\right) \\ &= Z_0\left(\frac{\frac{1}{\tan\beta l} + j\alpha l}{\frac{\alpha l}{\tan\beta l} + j}\right) \qquad \tan\beta l = \tan\frac{\pi}{2} = \infty \\ &= Z_0\left(\frac{\frac{1}{\infty} + j\alpha l}{\frac{\alpha l}{\infty} + j}\right) \qquad \frac{\alpha l}{\infty} = 0,\ \frac{1}{\infty} = 0 \\ &= Z_0\left(\frac{j\alpha l}{j}\right) \end{aligned}$$

$$Z_{OC} = Z_0(\alpha l) \tag{4.59}$$

Therefore, for low loss transmission line where $\alpha \ll \beta$, $Z_{SC} \approx \frac{Z_0}{\alpha l}$ and $Z_{OC} \approx Z_0\alpha l$.

Ideally $Z_{OC} = 0$, but this is not the case with the low-loss line where we have $Z_{OC} = Z_0\alpha l$. What we note here is that the input impedance of the resonance section of a transmission line is ideally zero or infinity for series or parallel resonance circuits respectively. In practice, we see that we have a small input impedance for a series resonance circuit and a large input impedance for a parallel resonance circuit.

With $Z_{OC} = Z_0\alpha l$ and $Z_{SC} = \frac{Z_0}{\alpha l}$, we can vary the frequency and measure the variation of Z_{SC} or Z_{OC} with frequency. Recall that α depends on frequency, length (l) also depends on frequency since $l = \frac{\lambda}{4}$. The variation of Z_{SC} or Z_{OC} with frequency gives us what is called the frequency response of the circuit.

Quality factor

Quality factor can be calculated in two ways:

- From the frequency response of the circuit
- From the basic definition of quality factor.

The quality factor (Q factor) can be computed from the frequency response of a system. The Q factor is a measure of the damping in a system and is related to the sharpness of the resonance peak in the frequency response. If we plot the current or voltage response when a current or voltage source is applied to the input of the section of the transmission line and we measure the 3dB bandwidth of the response, the centre frequency divided by the 3dB bandwidth of the frequency response gives the quality factor.

The voltage response is shown in figure 4.62. It doesn't matter if we use a series or parallel resonance circuit. At resonance frequency f_0, we get the maximum response, below and beyond that frequency, the amplitude drops. The 3dB frequency

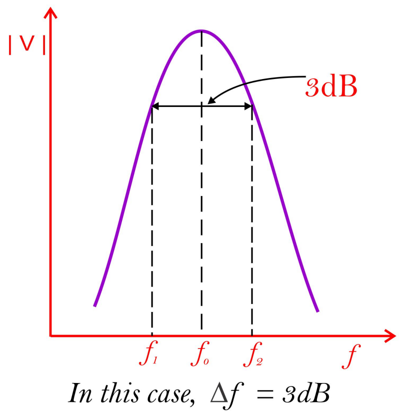


Figure 4.62: Frequency response of a resonant circuit

is where the amplitude reduces to $\frac{1}{\sqrt{2}}$ of its maximum value. In this case, Δf = 3dB bandwidth $= f_2 - f_1$, so the quality factor is given as equation (4.60)

$$Q = \frac{f_0}{\Delta f} = \frac{f_0}{f_2 - f_1} \tag{4.60}$$

Alternatively, we can take a section of the transmission line, find the voltage distribution and current distribution in the transmission line, and then calculate the power loss. Then use this information to calculate the quality factor.

$$\text{Quality factor, } Q = 2\pi\left(\frac{\text{Energy stored in the circuit}}{\text{Energy lost per cycle}}\right)$$

If the resonant frequency is f_0, the energy lost per cycle is f_0 cycles per second. Since in 1 second, we have f_0 number of cycles;

$$\text{Quality factor, } Q = 2\pi f_0\left(\frac{\text{Energy stored in the circuit}}{\text{Energy lost per second}}\right)$$

The energy lost per second = Power loss in the circuit

$$\text{Quality factor, } Q = 2\pi f_0\left(\frac{\text{Energy stored in the circuit}}{\text{Power loss in the circuit}}\right)$$

So to calculate the quality factor, we calculate the energy stored in a section of the transmission line and the power loss in the transmission line. Then we can easily find out the value of the quality factor for that section of the transmission line.

let us consider a section of a short-circuited transmission line of length, $l = \frac{\lambda}{4}$. Since the length depends on frequency, let us say frequency = f_0 when $l = \frac{\lambda_0}{4}$ and carry out our analysis. Hence, the resonant frequency of a circuit is $f_0 = \frac{v}{\lambda_0}$ (where v = velocity of the wave). We now find voltage and current expression on the transmission line with the load end short-circuited. With the load end short-circuited, $Z_L = 0$, the reflection coefficient at the load point is -1 due to total reflection.

So,

$$\Gamma_L = \frac{Z_L - Z_0}{Z_L + Z_0} = \frac{0 - Z_0}{0 + Z_0} = -1$$

A reflection coefficient of -1 means that the amplitude of the reflection coefficient is 1 with a phase change of 180°.

At this point, we can write down the voltage and current equation with the known reflection coefficient. We can also find out what the variation of voltage and current is on the transmission line from equation (2.36)

$$\begin{aligned} V(l) &= V^+e^{j\beta l} + V^-e^{-j\beta l}, \quad \frac{V^-}{V^+} = \Gamma, V^- = -V^+ \\ &= V^+e^{j\beta l} - V^+e^{-j\beta l} \text{Recall, } \frac{e^{j\beta l}-e^{-j\beta l}}{2j} = \sin\beta l \\ &= j2V^+\left(\frac{e^{j\beta l}-e^{-j\beta l}}{2j}\right) \\ &= j2V^+\sin\beta l \quad \text{Let, } j2V^+ = V_0 \\ &= V_0 \sin\beta l \end{aligned}$$

$$V(l) = V_0 \sin\beta l \tag{4.61}$$

In a similar manner, we can find the current distribution using equation (2.37),

$$\begin{aligned} I(l) &= \frac{V^+}{Z_0}e^{j\beta l} - \frac{V^-}{Z_0}e^{-j\beta l}, \quad \frac{V^-}{V^+} = \Gamma, V^- = -V^+ \\ &= \frac{V^+}{Z_0}e^{j\beta l} + \frac{V^+}{Z_0}e^{-j\beta l} \quad \text{Recall, } \frac{e^{j\beta l}+e^{-j\beta l}}{2} = \cos\beta l \\ &= \frac{2V^+}{Z_0}\left(\frac{e^{j\beta l}+e^{-j\beta l}}{2}\right) \\ &= \frac{2V^+}{Z_0}\cos\beta l \quad \text{Let, } 2V^+ = V_0 \end{aligned}$$

$$I(l) = \frac{V_0}{Z_0}\cos\beta l \tag{4.62}$$

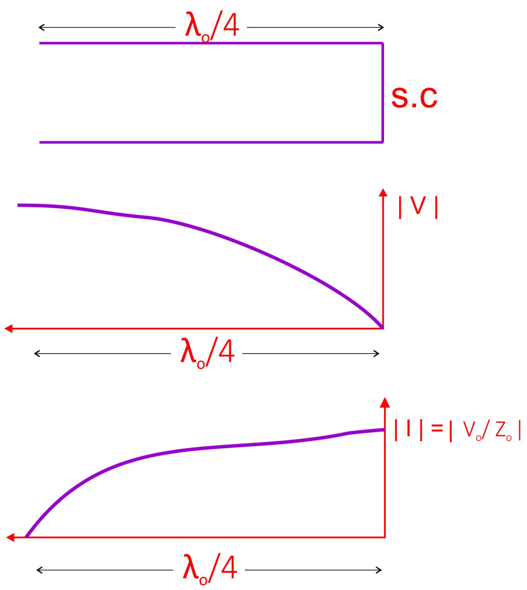


Figure 4.63: Voltage and Current distribution of the short-circuited transmission line

The plot of the voltage and current distributions along the short-circuited transmission line is shown in figure 4.63. It is shown that at $l = \frac{\lambda}{2}$ the voltage distribution is maximum while that of the current is zero. At $l = \frac{\lambda}{4}$, $\beta l = \frac{\pi}{2}$, so that $\sin\beta l = \sin\frac{\pi}{2} = 1$ and $\cos\beta l = \cos\frac{\pi}{2} = 0$. For the voltage distribution, we have a peak at $l = \frac{\lambda}{4}$ and a minimum of $l = 0$ and for the current distribution we have a peak value at $l = 0$ and a minimum value at $l = \frac{\lambda}{4}$.

Now we know the voltage and current distributions on the section of the transmission line, we can find the energy stored at different points of the transmission line. To find the energy stored, the capacitance of an infinitesimal section of the transmission line is $C\Delta l$ for Δl section since R, G, C and L are always stated per unit length. Hence, the capacitive and inductive energy stored in that section are

$$\begin{aligned} U_{\text{capacitor}} &= \frac{1}{2}C\Delta l V^2 \\ U_{\text{inductor}} &= \frac{1}{2}L\Delta l I^2 \end{aligned}$$

Hence, the total energy stored in a transmission line can be written as:

$$\begin{aligned} U &= \frac{1}{2}C\int_0^{\frac{\lambda_0}{4}} |V(l)|^2 dl \\ &+ \frac{1}{2}L\int_0^{\frac{\lambda_0}{4}} |I(l)|^2 dl \quad \text{from equations (4.61) and (4.62)} \\ &= \frac{1}{2}C\int_0^{\frac{\lambda_0}{4}} |V_0\sin\beta l|^2 dl + \frac{1}{2}L\int_0^{\frac{\lambda_0}{4}} \left|\frac{V_0}{Z_0}\cos\beta l\right|^2 dl \\ &= \frac{1}{2}CV_0^2\int_0^{\frac{\lambda_0}{4}} (\sin^2\beta l)dl + \frac{1}{2}L\frac{V_0^2}{Z_0^2}\int_0^{\frac{\lambda_0}{4}} (\cos^2\beta l)dl \end{aligned}$$

Recall from trigonometry

$$\sin^2\beta l = \frac{1-\cos 2\beta l}{2}, \quad \cos^2\beta l = \frac{1+\cos 2\beta l}{2}$$

$$\begin{aligned} U &= \frac{1}{2}CV_0^2\int_0^{\frac{\lambda_0}{4}} \left(\frac{1}{2} - \frac{\cos 2\beta l}{2}\right) dl \\ &+ \frac{1}{2}L\frac{V_0^2}{Z_0^2}\int_0^{\frac{\lambda_0}{4}} \left(\frac{1}{2} + \frac{\cos 2\beta l}{2}\right) dl \\ &= \frac{1}{2}CV_0^2\left[\frac{l}{2} - \frac{\sin 2\beta l}{4\beta}\right]_0^{\frac{\lambda_0}{4}} + \frac{1}{2}L\frac{V_0^2}{Z_0^2}\left[\frac{l}{2} + \frac{\sin 2\beta l}{4\beta}\right]_0^{\frac{\lambda_0}{4}} \\ &= \frac{1}{2}CV_0^2\left[\frac{\lambda_0}{8} - \sin\frac{2(\frac{2\pi}{\lambda_0})\frac{\lambda_0}{4}}{4\beta}\right] + \frac{1}{2}L\frac{V_0^2}{Z_0^2}\left[\frac{\lambda_0}{8} + \frac{\sin 2(\frac{2\pi}{\lambda_0})\frac{\lambda_0}{4}}{4\beta}\right] \\ &= \frac{1}{2}CV_0^2\left[\frac{\lambda_0}{8} - \frac{\sin\pi}{4\beta}\right] + \frac{1}{2}L\frac{V_0^2}{Z_0^2}\left[\frac{\lambda_0}{8} + \frac{\sin\pi}{4\beta}\right] \\ &= \frac{1}{2}CV_0^2\left[\frac{\lambda_0}{8}\right] + \frac{1}{2}L\frac{V_0^2}{Z_0^2}\left[\frac{\lambda_0}{8}\right] \end{aligned} \tag{4.63}$$

Where $\frac{1}{2}CV_0^2\left[\frac{\lambda_0}{8}\right]$ is the energy stored in line capacitance and $\frac{1}{2}L\frac{V_0^2}{Z_0^2}\left[\frac{\lambda_0}{8}\right]$ is the energy stored in line inductance.

We see that the two quantities are equal since

$$Z_0 = \sqrt{\frac{L}{C}} \Longrightarrow CZ_0^2 = L, \quad C = \frac{L}{Z_0^2}$$

$$\begin{aligned} U &= \frac{1}{2}CV_0^2\left[\frac{\lambda_0}{8}\right] + \frac{1}{2}CV_0^2\left[\frac{\lambda_0}{8}\right] \\ &= \frac{1}{2}CV_0^2\left[\frac{\lambda_0}{4}\right] \end{aligned} \tag{4.64}$$

Equation (4.64) is the total energy in the short-circuited transmission line of length $l = \frac{\lambda_o}{4}$. To calculate the quality factor, we also need to calculate the power loss in the transmission line. If we take the section of a transmission line $\frac{\lambda_0}{4}$ shown in figure 4.64.

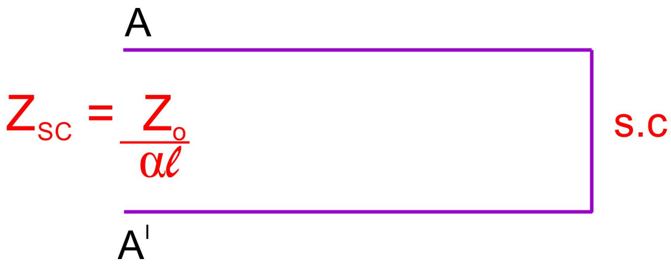


Figure 4.64: Input impedance of a short-circuited transmission line

The energy stored in the line is connected at point AA'. Finding out what the power loss at AA' is, the power loss is equivalent to the power loss in the line because the line is short-circuited. Therefore, the load is not consuming any power and the power supplied will be equal to the loss in the transmission line. So without revisiting the primary constant of the transmission line, just by calculating the input impedance of the line, we can find out what the equivalent resistance would be at the input point of the line and find out the losses through this resistance which will be the power loss in the line. Hence, the input impedance seen at AA' will be

$$Z_{SC} = \frac{Z_0}{\alpha l} \quad \text{from equation (4.57)}$$

And the power loss is

$$\begin{aligned} P_{loss} &= \frac{V_0^2}{Z_{SC}} \quad V_0 \text{ is the supplied voltage} \\ &= \frac{V_0^2}{\frac{Z_0}{\alpha l}} \\ &= \frac{\alpha l \times V_0^2}{Z_0} \\ &= \frac{V_0^2}{Z_0} \times \alpha \times \frac{\lambda_0}{4} \end{aligned} \tag{4.65}$$

Then, the quality factor, Q from equation (4.60) is given as follows by substituting equations (4.64) and (4.65).

$$\begin{aligned} Q &= 2\pi f_0 \left(\frac{\frac{1}{2} \times CV_0^2 \times \frac{\lambda_0}{4}}{\frac{V_0^2}{Z_0} \times \alpha \times \frac{\lambda_0}{4}} \right) \\ &= 2\pi f_0 \left(\frac{\frac{1}{2}C}{\frac{\alpha}{Z_0}} \right) \\ &= \frac{2\pi f_0 Z_0 C}{2\alpha} \quad \text{But, } Z_0 = \sqrt{\frac{L}{C}} \\ &\Rightarrow Z_0 C = \sqrt{LC}, 2\pi f_0 = \omega_0 \\ &= \frac{\omega_0 \sqrt{LC}}{2\alpha} \quad \omega_0\sqrt{LC} = \beta, \text{ phase constant} \\ &= \frac{\beta}{2\alpha} \end{aligned}$$

$$Q = \frac{\beta}{2\alpha} \tag{4.66}$$

Equation (4.66) is the same expression we would have obtained by finding out the frequency response of the transmission line and measuring the input and its variation as a function of frequency. It implies that the quality factor is related to the phase β and attenuation constant α of the transmission line. For a low loss line $\alpha \ll \beta$, thus Q is very large. This means that $Q \gg 1$ is for a low loss line. This is very important because it implies that at high frequencies, a section of a transmission line can give a high-quality factor circuit. Since the quality factor is related to 3dB bandwidth, the higher the quality factor, the smaller the 3dB bandwidth or the more tuned the circuit is. This is the frequency selectivity of the circuit and a high-quality factor signifies a very frequency-selective circuit. We can get a few hundred to a few thousand quality factor values for some low-loss transmission lines and a very good frequency sensitivity section of a transmission line is used as a resonant circuit.

4.14.4 As a Step Up Transformer

Let us consider a small section of transmission line of length, $l = \frac{\lambda}{4}$, shorted at the end shown in figure 4.65. Let's say by some means, voltage is induced along the section of the transmission line not from open circuit or short circuit end but between these two ends which we will take as the middle location as shown by the dotted lines.

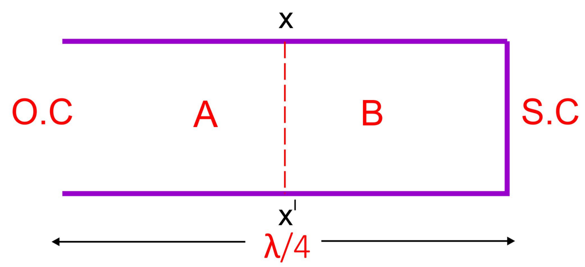


Figure 4.65: Voltage induced at the centre XX' of a transmission line of length $\frac{\lambda}{4}$ with one end open and the other end shorted

At any time, it would appear that the two sections, A and B, of the transmission line are connected in parallel. With the induced voltage, a standing wave signal will get induced in the transmission line. This standing wave (voltage and current distribution) does not see any impedance but the characteristic impedance of the transmission line. So the voltage at the middle point XX' sends two travelling waves along the transmission line as if the energy is supplied to the characteristic impedance of the two sections of the transmission line. The waves as shown in figure 4.66 are going in directions M and N respectively.

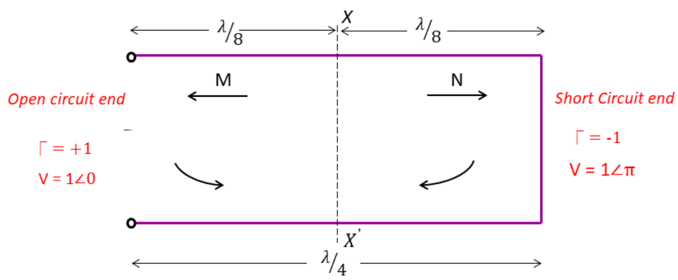


Figure 4.66: Standing wave propagation in the transmission line with one end open and the other end shorted

N travelling wave sees a short circuit and M travelling wave sees an open circuit. N wave sees a reflection coefficient of

-1, hence, at the short circuit point, the wave is reflected completely as shown by the arrow. With a reflection coefficient of -1, the amplitude is unity but the phase is 180°(or π). Hence, the wave travelling toward N after reflection undergoes a phase change of π and travels all the way to the open circuit side. At the open circuit point, the reflection coefficient = 1[35]. The reflected N wave reverses with π phase change gets to an open circuit with a reflection coefficient of 1, turns around and moves to point XX'. The distance traveled by the wave is $\frac{\lambda}{2} = \frac{\lambda}{8} + \frac{\lambda}{4} + \frac{\lambda}{8}$ and has gone through a phase change of 2π. This is because a travel distance of $\frac{\lambda}{4}$ corresponds to a phase change of π plus another π phase change due to a negative reflection coefficient. It means N wave has undergone a phase change of π due to $\frac{\lambda}{4}$ distance travelled + π due to the negative reflection coefficient at the short circuit end, resulting in a total phase change of 2π. Hence the wave that is induced at XX' is the same as the wave N which has been taken through a phase change of 2π. This is some kind of positive feedback because the induced voltage will add to the reflected voltage N, increase in amplitude and then the resulting constructive interference re-travels and repeats. That means the voltage essentially starts growing in the transmission line.

Exactly the same thing happens to the travelling wave M. It will be reflected at the open circuit end, travel $\frac{\lambda}{4}$ to the short circuit end, reverse with a phase of π and travels towards XX'. At the end has travelled $\frac{\lambda}{2}$ distance with a total phase change of 2π. Thus at point XX' the returned wave adds to the induced voltage at XX'. Then the resulting voltage at XX' grows as the travelling waves travel a trip one more time. Hence the standing wave of the setup will grow.

Now we ask the question, *how far will an induced voltage at XX' grow?* With no losses in the transmission line, it will grow up to infinity. The voltage will keep growing to infinite amplitude because there is nothing controlling this amplitude on the transmission line. Hence, if even a small voltage is induced on a section of the transmission line with resonant length, the resulting standing wave in the transmission line is much larger compared to the induced voltage on the transmission line. This resonant section of the transmission line can then be used as a voltage step-up transformer. Hence, we would measure a much larger voltage at the open section of the line i.e. at an open circuit compared to what we induced at point XX'. In theory, with no losses in the transmission line, this voltage amplitude should grow to infinity. However, that does not happen because as voltage and current amplitude is increasing in the transmission line, the ohmic losses also are increasing. When the losses in the line due to resistance and conductance (R and G) become equal to the energy source, at that point there is energy balance and the growth of the standing wave in the section of the transmission line stops.

Its application as a step-up transformer involves the connection of a small coupling source to generate a very large voltage and current when the losses are small. This is a useful application whenever we want to step up a voltage or current at high frequencies. It can be shown that this voltage growth is related to the quality factor of a transmission line. The higher the quality factor, the lower the losses in the transmission line and that will give higher voltages on the terminal of the transmission line.

The phenomena used here for voltage step-up could be harmful in many cases. Consider a situation where a small energy source gets unknowingly coupled to a small section of the transmission line. If the section of the transmission line is of resonance length, then the voltage developed on this line will be much larger than the circuit can handle and this can damage the circuit. Thus while designing high-frequency circuits, we should be careful not to couple some sections to the transmission line especially if they are of resonant length, otherwise, we develop a large voltage in that section that the circuit can barely handle.

With the applications of transmission lines we have discussed so far when sections of transmission lines are used for high-frequency circuits, we rarely see capacitors and inductors because they have been replaced by sections of transmission lines. What we see instead are active devices like transistors and FETs with reactive components completely realized by sections of the transmission line. So most microwave circuits will appear as transmission lines across active devices.

Exercises

Ex. 87 — A transmission line of length $\frac{\lambda}{4}$ is shorted at one end and open at the other end. If a voltage of 1V is induced at the center of the transmission line, what is the voltage at the open circuit end?

Ex. 88 — Calculate the input Impedance of the lines Z_OC and Z_SC, for a length $l = \frac{\lambda}{4}$ with losses

Ex. 89 — Show that the Total energy in a short circuited transmission line of length $\frac{\lambda}{4}$ is $\frac{1}{2}CV_0^2[\frac{\lambda}{4}]$

Ex. 90 — Show the behavior of a capacitor in a high frequency circuit and what is your observation.

Ex. 91 — What factors influence the transmission line impedance and what is the significance of characteristics Impedance in Transmission lines

Ex. 92 — How does the choice of dielectric material impact the performance of a transmission line?

Ex. 93 — Draw the frequency Response of a resonant circuit and show from it the expression for quality factor

Ex. 94 — With the aid of a graph. show the voltage and current distribution of a short circuited transmission line

35

$$\begin{aligned}\Gamma_L &= \frac{Z_L - Z_0}{Z_L + Z_0} \quad Z_L = \infty \text{ and taking limits} \\ &= \frac{1 - \frac{Z_0}{Z_L}}{1 + \frac{Z_0}{Z_L}} \\ &= \frac{1 - \frac{Z_0}{\infty}}{1 + \frac{Z_0}{\infty}} \\ &= \frac{1 - 0}{1 + 0} \\ &= 1\end{aligned} \tag{4.67}$$

4.14.5 Matching Of Impedances

In previous sections, we have dealt with other applications of transmission lines. In this section, we will discuss the most important application of transmission lines, which is *"Impedance Matching"*.

As we know when the load impedance is equal to characteristic impedance Z_0, there is no reflection on the line, hence there is maximum power transfer. But this *matching condition* is not always feasible. Thus, we need a module or device which will transform the load impedance to the characteristic impedance. This module is known as a **matching transformer unit** as shown in figure 4.67.

Figure 4.67: Matching Transformer unit

Let us suppose we want to convert the output impedance Z with the matching unit (transformer) to the characteristic impedance Z_0, which is shown in figure 4.67. At the input (leftmost side), we should see a characteristic impedance Z_0 and the module should be completely lossless (ideal) to allow maximum power transfer to load Z. First, let us discuss various methods by which different impedances can be matched to the characteristic impedance.

Figure 4.68: Application of transformer

In figure 4.68a we have two different resistances R_1 and R_2 which are to be matched, that is, $R_1 \neq R_2$. If they are connected directly there will be a mismatch and no maximum power transfer from R_1 to R_2. We then introduce some device in between that will make R_1 appear like R_2 when seen from the left side and R_2 appear like R_1, when seen from the right side, so from both sides it will appear that they are matched to the load and there is maximum power transfer. In figure 4.68b, we are matching R to Z_0, and like the previous case study we place in a module to match the impedances between R and Z_0. Hence on the left side, we should see Z_0 and vice versa. lastly, in figure 4.68c, which is the most important case. We have two long transmission lines which are to be connected together. If the length of each transmission is very large, the input impedance of transmission line one (1) will appear as its characteristics impedance Z_{01} and transmission line two (2) will appear as its characteristics impedance Z_{02}. For maximum power transfer, we bring a transforming device in between that will make Z_{01} see Z_{02} as Z_{01} and vice-versa. From both sides, it seems as if the line is terminated in its characteristic impedance. Hence there will be no reflection on either side of the transmission line.

Quarter Wavelength Transformer Technique

On the Smith Chart, if we have an impedance which is resistive, *how much can we move on the transmission line so that the impedance always appears like a resistance?* Let us introduce some section of the transmission line in between both impedances (input and output) which has a characteristic impedance Z_{0x} in figure 4.68a and 4.68b and is different from the characteristic impedance Z_0 to which the resistive impedance R is to be matched.

Hence, we find out Z_{0x} and the length of this introduced transmission line for which R can be matched with Z_0. Resistive impedance lie on the horizontal axis of the Smith Chart, $Z > Z_0$ lies on the right and $Z < Z_0$ lies on the left of the Smith chart. So the question is, *what should be the value of Z_{0x} and length l, that will match the resistive impedance to the characteristic impedance Z_0?* From the Smith Chart, to transform from R to another resistive value, we need to move by either $\frac{\lambda}{4}$ or $\frac{\lambda}{2}$. But a $\frac{\lambda}{2}$ movement brings us back to the same impedance and thus we don't use this value because it will transform the impedance to itself. A $\frac{\lambda}{4}$ movement on the other hand, will have a resistive impedance which is different from the original resistive impedance(that is, inverts itself), thus $l = \frac{\lambda}{4}$ will be most appropriate because at this length we will have a different resistive value.

Figure 4.69: The quarter wavelength transformer technique

From figure 4.69, at the transformer (output), the normalised impedance seen will be $\frac{R}{Z_{0x}}$. At the transformer (input), the normalised impedance seen will be $\frac{Z_{0x}}{R}$[36]. So the absolute characteristic impedance Z_0 seen at the input will be equal to:

$$\begin{aligned} Z_0 &= \left(\frac{Z_{0x}}{R}\right) \times Z_{0x} \\ &= \frac{Z_{0x}^2}{R} \end{aligned}$$

Therefore,

$$Z_{0x} = \sqrt{RZ_0} \tag{4.68}$$

Equation (4.68) implies that if we have a mismatch condition, we can input a matching transformer (an example is the BALUN[37]). Introducing this section of the transmission line, we will have a matched condition where R will appear as Z_0 when seen from the input and Z_0 will appear as R when seen from the output. Since the transmission line is lossless, there is a match on the transmission line from Z_0 to R, therefore, we will see a maximum power transfer. This technique of matching impedance is known as the **Quarter Wavelength Transformer Technique** which derives its name from the fact that we have to travel a length which is equal to a quarter, that is, 1/4 of the wavelength λ. This matching transformer is also used for matching lumped circuits.

Complex Load Impedance Matching

Now, we would like to find out *can the technique (quarter wavelength transformer technique) only be used to match resistive impedance?* The answer is no, we can use this method to solve complex impedances but it requires us to do some thinking because, at first look, it appears impossible. This is because on the Smith Chart, to derive Z_{0x}, we use only the horizontal axis of the chart. Now, *how can we use both the real and imaginary parts of the chart?*

We solve it by first transforming the complex normalized impedances to a real impedance. This is done by moving along the length of a transmission line. That means if you introduce

[36] We know from the property of the transmission line, that for every distance of $l = \frac{\lambda}{4}$, the impedance inverts itself

[37] A BALUN is a type of electrical transformer used to connect a balanced transmission line, such as a twisted pair cable, to an unbalanced transmission line, such as a coaxial cable. The name BALUN is short for " balanced to unbalanced.". It has a length $l = \frac{\lambda}{4}$, and a characteristic impedance Z_{0x}. BALUNs are commonly used in radio and television systems, as well as in data communication networks. They are available in a variety of designs, including ferrite-core BALUNs, air-core BALUNs, and transmission line BALUNs.

some length of transmission line with characteristic impedance Z_0, then the impedance seen at the other end is real if the length is such that you get a voltage (maximum or minimum) for that length.

Figure 4.70: The quarter wavelength transformer technique for complex load Z

From figure 4.70, we have to move Z which is complex by a distance l on the transmission line to make its transformed impedance real at point B.

Procedure:

(i) Mark the normalized point of the impedance, **P**, $r + jx$[38] on the Smith Chart as shown in figure 4.71.

Figure 4.71: Simplified Smith chart for complex load Quarter wavelength transformer

(ii) Draw a constant VSWR circle at point **P** and find the first point of intersection with the real axis moving clockwise. In figure 4.71 it is marked as **T**. The point **T** is a real impedance $r_{\max}$. The next point of intersection with the real axis is **S** and it is $r_{\min}$.

(iii) The respective distances of these points from the load impedance are $l_{\max}$ and $l_{\min}$ respectively. $l_{\max}$ is the distance from point **P** to **T**, while $l_{\max}$ is the distance from point **P** to **S**.

(iv) Determine the VSWR, ρ, and determine the corresponding resistive values $R_{\max}$ and $R_{\min}$ at points **T** and **S** respectively.

(v) Considering the additional transmission line section of length $l_{\max}$, the absolute impedance Z' seen at point **T** is

$$Z' = R_{\max} = Z_{01} \times \rho \tag{4.69}$$

Similarly for $l_{\min}$, the absolute impedance Z' seen at point **S** is

$$Z' = R_{\min} = \frac{Z_{01}}{\rho} \tag{4.70}$$

Therefore the absolute impedance Z' at point B in figure 4.70 can be found using either $R_{\max}$ or $R_{\min}$ and the characteristics impedance of the quarter wavelength transformer Z_{0x} would be

$$Z_{0x} = \sqrt{Z_{01}\rho Z_0} \quad \text{or} \quad Z_{0x} = \sqrt{Z_{01}\frac{Z_{01}}{\rho}}$$

Any solution is acceptable depending on whether we want the length l to be small or large. If we are taking cost into consideration, for instance, we will choose a shorter length of the transmission line.

In conclusion, we can say that *Quarter Wavelength Transformer technique* can be used for matching complex impedance Z to the characteristics impedance Z_0. However, it has a big drawback in that you require a unique characteristic impedance Z_{0x} for the matching device for different loads. This is not desirable and realizable in nature because, in practice, we do not get a transmission line whose characteristic impedance can be varied easily. The characteristic impedance Z_{0x} depends on the physical dimensions of the transmission line as will be seen in chapter 15. Typically, we have cables of transmission lines with standard characteristic impedance. For example for co-axial cable, the characteristic impedances are 50Ω and 75Ω. For parallel wire transmission lines, the characteristic impedances are 300Ω and 600Ω.

Example 4.14.3 Quarter Wavelength Matching

A load impedance $Z_L = (75 - j35)\Omega$ is to be matched to 50Ω using a quarter wave transformer. Design the matching set-up.

Solution

We recall that the quarter wave transformer can match a resistive load impedance to the characteristic impedance (that is, two resistive impedances). However, the load impedance given is complex, so we use a section of a transmission line of length l as shown in figure 4.72, to first convert the impedance to a real value, from that point onward a quarter wave transformer can then be used for the matching.

Figure 4.72: Load Impedance Matching Design with Quarter Wavelength Trasfromer

The length of the quarter wave transformer is $\frac{\lambda}{4}$ at XX' (see figure 4.73), $Z_L = (75 - j35)\Omega$ should be real.

Figure 4.73: Quarter Wavelength Transformer Design

On the Smith chart, we draw the VSWR circle for normalized Z_L[39] and move clockwise to the point of intersection with the real axis which gives either $R_{\min}$ or $R_{\max}$.

*Figure 4.74 shows the Smith chart analysis. $\bar{Z}_L = 1.5 - j0.7$ is marked at **X** and the constant VSWR circle is drawn using **OX** as radius and it intersects the real axis at **P** and **Q**. If we move from **X** to **P** (**XP**) we get a resistive impedance corresponding to $r_{\min}$ and if we move from **X** through **P** to **Q** (**XPQ**), we get another resistive impedance corresponding to $r_{\max}$.*

Figure 4.74: Quarter Wavelength Design— graphical analysis

*The distance and resistance for **XP** are*

$$\begin{aligned} XP &= 0.5\lambda - 0.306\lambda \\ &= 0.194\lambda \\ &= l_1 \end{aligned}$$

$$\begin{aligned} r_{\min} &= \bar{R}_{\min} \quad \textit{From figure 4.74} \\ &= 0.5\Omega \end{aligned}$$

Thus, $R_{\min}$ is

$$\begin{aligned} R_{\min} &= Z_0 \times \bar{R}_{\min} \\ &= 50 \times 0.5 \\ &= 25\Omega \end{aligned}$$

[38] The normalized impedance $\bar{Z}$ is normalized with respect to characteristic impedance of the line Z_{01}, that is $\bar{Z} = Z/Z_{01}$

[39] Normalization is done using the characteristic Impedance of the section of the transmission line of length l.

*Similarly, the distance and resistance for **XPQ** are*

$$\begin{aligned} XPQ &= 0.194\lambda + 0.25\lambda \\ &= 0.444\lambda \\ &= l_2 \end{aligned}$$

$$\begin{aligned} r_{\max} &= \bar{R}_{\max} \quad \textit{From figure 4.74} \\ &= 2\Omega \end{aligned}$$

Thus, $R_{\max}$ is

$$\begin{aligned} R_{\max} &= Z_0 \times \bar{R}_{\max} \\ &= 50 \times 2 \\ &= 100\Omega \end{aligned}$$

Hence $l = l_1 = 0.194\lambda$ or $l_2 = 0.444\lambda$ which is the length of the section of the transmission line for transforming the complex impedance to real impedance just before the Quarter Wavelength Transformer.

The outward circle of the Smith chart is calibrated in λ, so one can take the angle subtended by the movement and do a direct conversion to λ, λ at X is 0.306 moving to 0.5 is a difference of 0.194λ[40]*. The top half is 0.25λ, added to the 0.194λ calculated initially brings us to 0.444λ*[41] *for $R_{\max}$. Going back to the problem we have that for $l_1 = 0.194\lambda$, $R_{\min} = 25\Omega$.*

$$Z_{0x}{}^{42} = \sqrt{Z_0 \times R} = \sqrt{50 \times 25} = 35.35\Omega$$

For $l_2 = 0.444\lambda, R_{\max} = 100\Omega$

$$\begin{aligned} Z_{0x} &= \sqrt{Z_0 \times R} \\ &= \sqrt{50 \times 100} \\ &= 70.7\Omega \end{aligned}$$

Here we see how to match a complex impedance to the characteristic impedance (real impedance) by using a section of the transmission line and the Quarter Wavelength Transformer. The Quarter Wavelength Transformer can match only impedance which is real values, that is why we introduce a section of transmission line before the Quarter Wavelength Transformer to help convert our complex load impedance to real impedance.

Stub Matching Technique

The question then is, *can we use the standard transmission line which is available with their standardised characteristic impedances?* Yes we can and the approach used is known as **Stub Matching Technique**. Let's define a Stub.

A stub is an auxiliary section of a transmission line which is either short-circuited (S.C) or open-circuited (O.C) that is attached to the main transmission line either in series or parallel. There are four(4) types of stub matching techniques, which are

(i) Series O.C stub matching technique

(ii) Series S.C stub matching technique

(iii) parallel O.C stub matching technique

(iv) Parallel S.C stub matching technique

[40] $l_1 = 0.5\lambda - 0.306\lambda = 0.194\lambda$
[41] $l_2 = 0.25\lambda + 0.194\lambda = 0.444\lambda$

In this chapter, we are going to be concentrating on the **Parallel Short Circuit Stub Matching Technique.**

By using the stub at the proper location on the main transmission line, it is possible to match a complex impedance Z to the characteristic impedance Z_0. Recall the advantage of the stub technique is that we do not require a unique characteristic impedance Z_0 for each impedance matching. The characteristic impedance of the main and auxiliary transmission lines remains the same. However, we change the length l_l from load to the location of the stub and the length of the stub (or auxiliary transmission line) attached to the main transmission line l_s (see figure 4.75).

There are three (3) techniques for attaching a stub to the transmission line to match the load impedance to the characteristic impedance and they are:

(i) Single Stub matching technique - One attached auxiliary line.

(ii) Double Stub matching technique - Two attached auxiliary lines.

(iii) Triple Stub matching technique - Three attached auxiliary lines.

Single Stub Matching Technique

In this section, we will discuss the single stub matching technique in the *parallel short circuit stub matching technique configuration*. Figure 4.75 shows the structure of a single stub matching technique. Again, our aim is to find a real impedance at point A that is equal to the characteristic impedance, Z_0. The main transmission line with characteristics impedance Z_0 is connected to load impedance Z.

Figure 4.75: Single stub matching technique with one auxiliary transmission line attached

The stub length l_s is connected in parallel to the main transmission line at a distance l_l from the impedance to be matched. We know that as voltage or energy moves from the generator to the load, it sees two (2) parts: the stub section and load Z section. At the load, there is a reflection and at the stub, there is a total reflection which is in the opposite direction ($\Gamma_{SC} = -1$) since the stub is short-circuited. If we make sure that both reflected energy from the stub and from the load match each other in amplitude and are opposite in phase, there will be no reflection beyond point A. Since there is no reflection beyond A that is, the input impedance is always going to be equal to the characteristic impedance.

The theory, therefore, is as follows; *we need to find the length l_l such that when the load impedance Z is transformed by this distance, the resistive part becomes equal to Z_0 and the parallel combination of the transformed impedance and the reactive part from the stub cancels out.* We choose l_s, such that the reactive part of the stub cancels the reactive part of the load after its transformation. The purpose of the length l_l (location of stub), is to make impedance Z to be equal to characteristic impedance Z_0 and the purpose of the length of stub l_s is to neutralize the reactive part. We will proceed to solve the single stub matching technique graphically using the Smith Chart.

For the single stub matching, we will work in terms of admittance since we have parallel connections. Therefore, let the load admittance and characteristic admittance be

$$Y = 1/Z \quad \text{(load admittance)}$$
$$Y_0 = 1/Z_0 \quad \text{(characteristic admittance)}$$

Hence, the normalized load admittance is

$$\bar{Y} = \frac{Y}{Y_0}$$

Figure 4.76 shows the simplified Smith Chart for the single stub matching problem. Let P be the point of the normalized admittance to be matched and the dashed circle passing through P is the constant VSWR circle. Where this circle intersects $g = 1$ circle denoted by points B and D, we have that the real part of the transformed admittance, Y equals the characteristic admittance Y_0. That is, at points B and D, the distance from P to B or P to D will give us a conductive part equal to unity (same as Y_0) and then the unmatched reactive part should be at those locations. These locations give the value of l_l which corresponds to point A in figure 4.75, such that from the load section, $\bar{Y} = 1 + jb'$. Therefore, the stub must have $-jb'$ as its admittance so that when they are connected in parallel, the susceptance cancels out leaving only the normalised admittance of $1 + j0$. The process involved is summarized as follows:

Figure 4.76: Simplified Smith Chart for Single Stub Matching

(i) First locate the given admittance value to be matched on the Smith Chart.

(ii) Draw your constant VSWR circle through the located point on Smith Chart.

(iii) Move along the constant VSWR circle in a clockwise direction and mark the two points it intersects with the constant $g = 1$ circle.

(iv) The distance from the first location (point P) to either of these new points (point B or D) is the location of stub l_l.

(v) Choose a point from the new points say point B (1+jb'), and take a mirror image which is point D (1 - jb'). Draw a constant susceptance circle passing through this point D. Where this circle touches the outermost circle of the Smith Chart is $-jb'$.

(vi) The distance moving anticlockwise from this point S to the SC point on the Smith Chart gives us the length of the stub (see figure 4.76).

Similarly, when the stub is connected in series, we use the same pattern but we will use impedance values instead of admittances. The single stub matching technique is extremely useful in matching impedances to the characteristic impedance of the line. In this technique, we are matching impedance without any unique transmission line unit as in the case of a Quarter Wavelength Transformer Technique. However, this technique has a small drawback which is that the location of the stub depends on the impedance to be matched. Though the single stub technique can match all possible loads, for every load to be matched, the location of the stub has to be changed which may not be that easy once the stub is connected. To solve these limitations as we did for quarter wavelength transformer technology, we will use the Double Stub Matching Technique.

Example 4.14.4 Single Stub Matching

Design a single stub matching unit to match the load impedance shown in figure 4.77 to the characteristic impedance, $Z_0 = 50\Omega$. Calculate for l_s and l_1.

Figure 4.77: Single Stub Matching design

Solution

We normalize $Z_L = 90 - j25$ to get $\bar{Z}_L = 1.8 - j0.5$. Since stub and Z_L are in parallel, we convert $\bar{Z}_L$ to its admittance value. $\bar{Y}_L$ should be transformed from $\bar{Y}_L$=$g + jb$ to $1 + jb_1$ hence it lies on the circle of constant conductance $g = 1$. Draw a circle of constant VSWR with OP as the radius. We get the point where this VSWR circle intercept the g=1 circle moving clockwise(towards the generator)from Q to M. The distance QM corresponds to L_1. $L_1 = 0.152\lambda - 0.032\lambda = 0.12\lambda$.

Mirror M on the VSWR circle to get $1 - jb1$. On the constant susceptance curve of jb, find where it intercepts the circle at $g = 0$. This is the outermost circle. R represent the pure susceptance value of jb_1.

Figure 4.78: Single Stub Matching Design— graphical analysis

Hence we have to move from the short circuit (SC) end of the stub towards the generator to see this Susceptance. The distance l_s is the movement from SC to -jb1 towards the generator and that distance represents the length of the stub. For admittance on the Smith chart, the rightmost side is the SC part and the leftmost part is the open circuit part. This is opposite to the convention in the impedance Smith chart.

At SC $l_{sc} = 0.25\lambda$ moving towards the generator.At R, $l = 0.402\lambda$,the distance $l_s = 0.402 - 0.25 = 0.152\lambda$. Now we know the length of the stub and the location of the stub as well. These are some of the problems which can be solved easily with the help of the Smith chart. Using an analytic approach for the same process is an extremely tedious task. This example clearly demonstrates the use of a Smith chart for solving complex transmission line problems.

Double Stub Matching Technique

This technique involves the use of two (2) stubs. In this method, we change only the length of the stub while the location of the stub is fixed. The separation between the stubs is $3\frac{\lambda}{8}$. The distance between stub 1 and load is l_l as shown in figure 4.79

Figure 4.79: Double Stub Matching Technique with two auxiliary lines attached

From figure 4.79, starting from the load end, the initial admittance is $\bar{Y} = g + jb$. After a distance l_l moving clockwise to point B (towards generator), at this point B^{-}, the admittance becomes $g_1 + jb'$ and when this is combined with the susceptance(or admittance) from stub 1, $-jb_1$, we get $g_1 + jb_1^{'}$ at B$^+$. When the admittance at B$^+$ is further transformed along the length $3\frac{\lambda}{8}$(clockwise, towards the generator), we get $1 + jb_2$ at point C$^-$. When this is combined with the susceptance of stub 2 $-jb_2$, we get $1 + 0j$ beyond point C, C$^+$. As for the single-stub matching technique, we will solve the double-stub matching technique using the Smith Chart.

Figure 4.80 shows the analysis for the structure in figure 4.79. We have already seen in the single stub matching

Figure 4.80: Simplified Smith Chart for Double Stub Matching Technique

technique that at the location where a stub is connected if the admittance seen towards the load is of the form $1 + jb$ then the stub can cancel the reactive part of the admittance and we get a matching condition. In the case of the double stub matching technique, if the admittance at point C is of the form $1+jb$, the matching condition can be achieved. This means at location B the admittance must be a transformed version of $1 + jb$ by a distance $3\frac{\lambda}{8}$ away from the generator. The translation of $1 + jb$ by a distance of $3\frac{\lambda}{8}$ counter-clockwise is the same as the rotation of the $g = 1$ circle by 270°counter-clockwise around the centre of the Smith Chart.

Procedure:

(i) Mark the normalized admittance $g + jb$ denoted by A in figure 4.80.

(ii) Move on the constant VSWR circle by l_l to get to point B$^-$ $(g_1 + jb')$.

(iii) At point B$^+$, since we are adding the admittance of stub 1, $-jb_1$, we are changing only the reactive part. So we move on the circle of constant conductance, g_1, from B$^-$ till we get to the rotated $g = 1$ circle where it intersects at B$^+$(Notice that we will intersect the $g = 1$ circle at two points, this gives two possible solutions for B$^+$ but we usually we pick one). So essentially, stub 1 helps in bringing the admittance to lie on the rotated circle.

(iv) Now to find the length of the first stub, we find the difference between the two (2) susceptance values B$^-$ and B$^+$ since the conductance is the same. Mark the difference on the outermost circle of the Smith Chart S_1. The distance from this point to the SC point moving anticlockwise gives the length of the first stub.

(v) At point B$^+$, we now move on the constant VSWR circle of $g_1 + jb_1^{'}$ in a clockwise direction by $3\frac{\lambda}{8}$ to transform the admittance $g_1 + jb_1^{'}$ to $1 + jb_2$ at C$^-$.

(vi) To cancel the reactive part, jb_2 we take a mirror image, $1 - jb_2$, then find susceptance circle of $-jb_2$ at the outermost circle of the Smith Chart, S_2. We measure the distance from this point to the SC point in an anticlockwise direction to give us the length of the second stub, l_{s2}.

We have successfully tackled the drawback of the flexible location of the stub in the single stub matching technique. However, the double stub matching technique has a drawback which is, the whole essence of matching is possible provided that by moving in a constant conductance circle we can intersect the $g = 1$ circle. If by the movement through the constant conductance circle, we can not reach the rotated $g = 1$ circle, then matching is not possible. Now we know the nature of the constant g circle (that they are within one another), so it is possible that if the admittance at point B$^-$ lies in a circle smaller than the $g = 1$ circle, then by our movement inside the constant conductance circle, we will never be able to intersect the rotated $g = 1$ circle, which means matching is not possible. What this means is that if $g_1 + jb'$ lies within the shaded forbidden region shown in figure 4.80, then impedance matching is not possible. The impedance we want to match can lie in the forbidden region, but the transformed impedance after distance l_l should be out of the forbidden region.

So by choosing a proper value for l_l, we can avoid the forbidden region. Therefore, the double stub matching technique has a small limitation that it cannot match those impedances whose transformed value after l_l lies within the forbidden region. However, this has been solved by introducing the concept of *three stub matching*.

Example 4.14.5 Double stub matching technique

For a load impedance, $Z_L = (100+j100)\Omega$, and characteristic impedance, $Z_0 = 50\Omega$, design a double stub matching unit to match the load for the transmission line. Note $l_1 = 0.4\lambda$.

Solution

Procedure:

(i) Determine the normalized value for the load resistance Z_L.

$$\begin{aligned}\bar{Z_L} &= \frac{Z_L}{Z_0}\\ &= 2 + j2\end{aligned}$$

Figure 4.81: Worked Example

(ii) Mark the point on the Smith Chart $\bar{Z_L} = 2 + j2$

(iii) Draw the constant VSWR circle from the normalized impedance point

(iv) Move 180° from the normalized impedance point on the constant VSWR circle to the normalized admittance Y_L

(v) Then move from point Y_L by distance l_1(in this case 0.4λ) clockwise(towards generator). Where the line meets the VSWR circle, mark the point and name it Y_{d1}.

(vi) Draw a rotated 270° circle of radius 1. ***Note:*** *$\frac{3\lambda}{8}$ correspond to 270° rotation on the Smith Chart because 0.5λ correspond to 360°.*

(vii) At point Y_{d1}, move along the constant g circle to intersect the rotated $g = 1$ circle. Mark the points of intersection E and D. Point E and D are Y_{11}

(viii) Solve for Y_{s1} (Using either point E and D as Y_{11}). Y_{s1} = Y_{11} - Y_{d1}

(ix) Mark point Y_{s1} on the Smith chart, extend the point to the outermost circle and mark the point l^1

(x) Moving from short circuit point to point Y_{s1}. Calculate the distance $l_{s1} = 0.25 + l^1$

(xi) To get l_{s2}. With the radius from the centre to point E, draw a circle and with the radius from the centre to point D draw a circle

(xii) Mark the 270°point of radius point E circle, where it meets with the unit resistance circle, mark the point and call it Y_{d2}. With radius from centre to point, D mark the 270 deg point, where it meets with the unit resistance circle and calls it $Y_{d2}^{'}$

(xiii) With point Y_{d2} or $Y_{d2}^{'}$ we can then get Y_{s2}. Note, $Y_{22} = 1$ thus, $Y_{s2} = Y_{22} - Y_{d2}^{'}$

(xiv) Mark the point Y_{d2} on the Smith Chart and extend to the outermost circle. Mark the point l^{11}

(xv) Measure the length from the short circuit point to the l^{11}
$Y_{d2} = 0.25 + l^{11}$

Three Stub Matching Technique

Figure 4.82: Three stub matching technique with three auxiliary lines attached

From figure 4.82, if the matching is done with stub 2 and 1 and stub 3 is made an open circuit (so that the l_{s3} equals quarter wavelength, $\frac{\lambda}{4}$), we find that stub 2 and 3 makes the transformed impedance to lie in the forbidden region. So we disconnect stub 3 and make use of stubs 1 and 2. Depending on whether the transformed impedance lie within the forbidden region or not, we can make use of stub 1 and 2 or stub 2 and 3. The final solution for impedance matching is the three-stub technique. By using this method, all possible impedances can be matched without moving the location of the stub on the transmission line.

In conclusion, we have dealt with the major application of transmission lines which is impedance matching in which we covered the different methods with which impedance can be matched. These are:

1. The quarter wavelength matching technique.
2. Single stub matching technique.
3. Double stub matching technique and
4. The triple stub matching technique.

In the end, we resolved that the triple stub matching technique is the best technique among all the techniques.

Exercises

Ex. 95 — A load impedance of $Z_L = (100 + j100)\Omega$ is to be matched to a transmission line of characteristic impedance $Z_0 = 50\Omega$ using a double stub matching technique. The length of the line is $l_1 = 0.4\lambda$. Determine the length of the stubs and the distance of the stubs from the load.

Ex. 96 — A 50Ω transmission line with an air core operates at 100MHz and is connected to a load impedance of $Z_L = (27.5 + j35)\Omega$. Design a single stub tuner.

Ex. 97 — Find the matching network to match $Z_L = (100 + j100)\Omega$ to a line of characteristic impedance 50Ω using:

(a)a single series O.C stub

(b)a single parallel(shunt) O.C stub

(c)a single series S.C stub

(d)a single parallel S.C stub

Ex. 98 — The terminating impedance of a transmission line, Z_L is $(100 + j100)\Omega$ and the characteristic impedance, Z_0 of the line is 50Ω. The first stub is placed at 0.4λ away from the load. The distance between the two stubs is $\frac{3}{8}\lambda$. Determine the length of the short-circuited shunt stubs for matching to be achieved. What terminations are forbidden for matching by the double stub?

Ex. 99 — Determine the values of d and l in the single stub matching setup shown in figure 4.83.

Figure 4.83: Ex. 99

4.15 Lossy Transmission Line

In previous sections, we have discussed the lossless and low-loss transmission lines. We have studied the characteristics of lossless transmission lines and have seen various applications of the lossless transmission line. In practice, however, as frequency increases, the loss increases and the line becomes very lossy. In this section, we will see briefly the characteristics of a lossy line.

Obviously, if a line is very lossy then it is not a very efficient medium for transfer of power, so when we say a line is lossy in practice it is not very lossy but it is moderately lossy. Let us briefly examine the characteristic impedance and the propagation constant of a very lossy line, after which we will examine that of a moderately lossy line.

4.15.1 Very Lossy Transmission Line

For a very lossy line, we will define the relationship between the primary constants as:

$$R \gg \omega L$$
$$G \gg \omega C$$

For this line, the characteristics impedance, Z_0 is given as

$$\begin{aligned} Z_0 &= \sqrt{\frac{R+j\omega L}{G+j\omega C}} \\ &\approx \sqrt{\frac{R}{G}} \end{aligned} \tag{4.71}$$

Equation (4.71) is a real quantity. Similarly, the propagation constant given in equation (4.72) is also a real quantity.

$$\gamma = \sqrt{(R+j\omega L)(G+j\omega C)} \approx \sqrt{RG} \tag{4.72}$$

We observe that the characteristic impedance is real and the propagation constant is also real. Recall that when we studied the lossless transmission line the characteristic impedance was real, so looking at Z_0 does not tell you whether the line is very lossy or lossless. However, when we look at the propagation constant γ for a lossless line, γ was purely imaginary (α = 0 and γ = jβ) but for a very lossy line γ is a real quantity(β = 0 and γ = α) which means there is no phase variation in space for whatever voltage or current variation we have on the transmission line. A lossy line does not represent the wave phenomena, so essentially the structure is not representing a medium which is carrying voltage or current waves. There is a voltage and current variation on the structure but it does not represent the wave phenomena instead you have a voltage that varies exponentially with the attenuation constant α = $\sqrt{RG}$ and there is no phase variation which implies no travelling wave.

Obviously, we are not interested in this case we are investigating. Let us however consider the case where R and G are comparable to ωL and ωC and we refer to this transmission line as the moderately lossy line.

4.15.2 Moderately Lossy Transmission Line

Mathematically, the relationship between the primary constants can be expressed as follows:

$$R \approx \omega L$$
$$G \approx \omega C$$

The propagation constant for the case of the moderately lossy line γ is given in equation (4.73) is a complex quantity.

$$\gamma = \alpha + j\beta \tag{4.73}$$

Where α is comparable to β since R and G are comparable to ωL and ωC respectively. From the voltage equation, we derived in equation (2.36) written as follows:

$$\begin{aligned} V &= V^+e^{\gamma l} + V^-e^{-\gamma l} \quad \text{substituting } \gamma = \alpha + j\beta \\ &= V^+e^{\alpha l}e^{j\beta l} + V^-e^{-\alpha l}e^{-j\beta l} \end{aligned}$$

Such a travelling wave exponentially decays with the attenuation constant α and it has a phase constant β. Let us consider the forward travelling wave, $V^+e^{\alpha l}e^{j\beta l}$ we see that the wave exponentially grows towards the generator or in other words decays as we move towards the load.

Similarly, $V^-e^{-\alpha l}e^{-j\beta l}$ exponentially decays as we move towards the generator or exponentially grows as we move towards the load. Depending upon the value of the ratio of V^- and V^+ which is the reflection coefficient at the load point, Γ_L the two travelling waves propagating in opposite directions decay or grow exponentially as we move towards the generator. We recall from the lossless case that the amplitude of Γ_L was the same at every very point on the transmission line, however, that is not true for the moderately lossy line because the amplitude of the reflected wave to the incident wave varies with αl, as Γ is now a function of location on the transmission line.

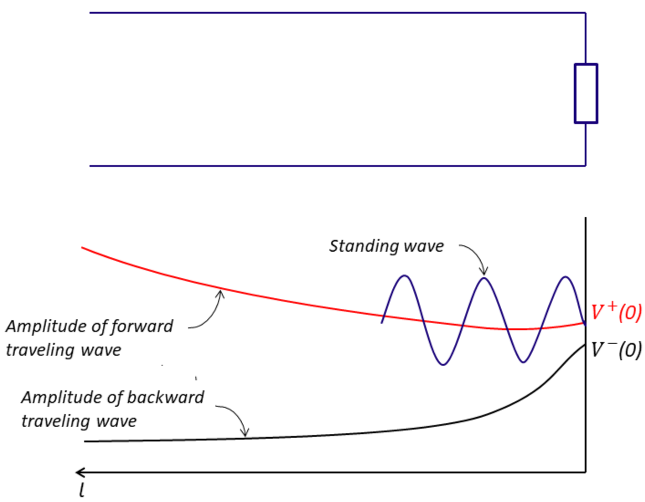


Figure 4.84: A plot of the attenuation of the forward and backward travelling waves of the moderately lossy line from the load point.

The incident wave exponentially decays as we move towards the load and depending on the load value there is a certain reflection coefficient at the load so the value of the reflected wave exponentially decays as we move towards the generator. So as we move to the generator the amplitude of the standing wave changes because, at the load end, the reflected wave dominates while at the generator end the incident wave dominates or the wave behaves more and more like a travelling wave instead of a standing wave.

From a different view, it implies that if there is a large mismatch at the load side, as we move towards the generator the mismatch becomes weaker and the matching improves and we

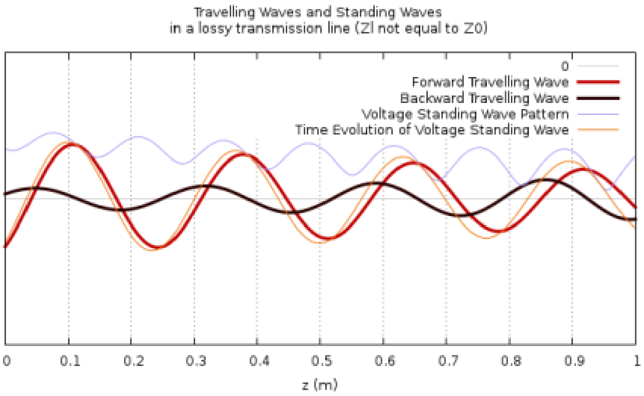


Figure 4.85: Standing wave of a moderately lossy line

see an impedance at the generator end which is very close to the characteristic impedance Z_0 of the transmission line because only the forward travelling wave is seen at the generator end. We conclude that for a lossy line, irrespective of what the load impedance is, there will always be a matching seen at the generator. But this is not a very good situation, because the power which is supplied by the generator is not delivered to the load. A substantial amount of power has been lost in the transmission line. So many times in transmission line design, a deliberately lossy line is introduced so that even in any experiment we connect some arbitrary load and get some really strong reflections, at least these reflections will not go and damage the generator. So in this case the purpose certainly is not maximum power transfer but it is to protect the generator from any unwanted reflections.

Using the Smith chart for lossy lines

Let us examine the effect of a lossy line on the reflection coefficient of the transmission line. For the lossy line, the expression for the reflection coefficient is given as;

$$\begin{aligned}\Gamma(l) &= \frac{V^-}{V^+}e^{-2\alpha l}e^{-j2\beta l} \quad \text{But, } \Gamma_L = \frac{V^-}{V^+}\\ &= \Gamma_L e^{-2\alpha l}e^{-j2\beta l}\end{aligned}$$

We would notice that if we know Γ_L, as we move towards the generator there will be a change *but in what manner?* It is immediately clear when Γ_L is expressed in this form $|\Gamma_L|e^{j\theta_L}$ where θ_L is the phase of the reflection coefficient at the load end and $|\Gamma_L|$ is the magnitude of the reflection coefficient at the load end.

Then,

$$\Gamma(l) = |\Gamma_L|e^{-2\alpha l}e^{j(\theta_L - 2\beta l)} \tag{4.74}$$

Equation (4.74) shows that the total phase at a distance l is $(\theta_L - 2\beta l)$ and the amplitude of the reflection coefficient at location l is $|\Gamma_L|e^{-2\alpha l}$. As we move towards the generator l is positive and the amplitude of the reflection coefficient goes on reducing exponentially.

The expression $|\Gamma_L|e^{j(\theta_L - 2\beta l)}$ we have seen earlier in the complex gamma plane, traces a curve which is a circle. However, now there is the term $e^{-2\alpha l}$, the radius of the circle is reducing continuously and the expression $|\Gamma_L|e^{-2\alpha l}e^{j(\theta_L - 2\beta l)}$ essentially draws a spiral on the reflection coefficient plane as shown in figure 4.86(a).

The VSWR which we have used to measure the contribution of the reflected wave, is not meaningful in the case of a

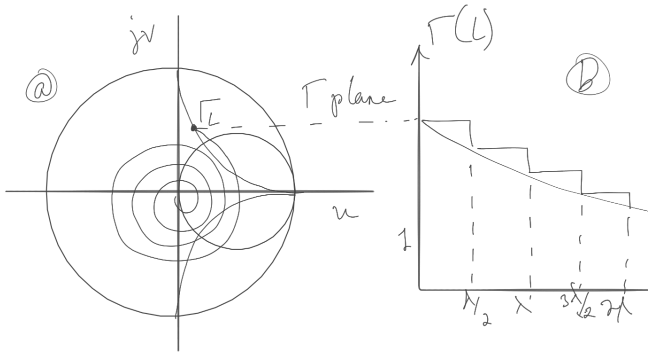


Figure 4.86: (a) A plot of the reflection coefficient on the Γ-plane. (b)Variation of the reflection coefficient along the transmission line towards the generator and the step value of $l = \frac{\lambda}{2}$ for $\Gamma(l)$ to approximate the spiral to a series of concentric circles.

moderately lossy transmission line because it is no more characteristic of the load. It also has become a function of the transmission line characteristics, that is, the VSWR is a function of the length along the transmission line. Also, the standing wave expression is no longer a sinusoidal function and hence the separation between two adjacent minima is not exactly equal to $\frac{\lambda}{2}$. However, the separation is approximated to $\frac{\lambda}{2}$ because in practice most of the transmission lines have a loss which is reasonably small.

The process of analyzing moderately lossy lines using the Smith chart is hard without software because the correct variations of the reflection coefficient or impedance variation need to be drawn. If the transmission line we are using has a small loss then we can approximate the spiral to a series of concentric circles. Each circle that follows is drawn after we move a distance of $\frac{\lambda}{2}$. From figure 4.86(b) we apply the approximation and a constant value of $\Gamma(l)$ is used for every $\frac{\lambda}{2}$ before correcting the magnitude of the reflection coefficient. The radii of the circles will differ by δ given by

$$\begin{aligned}\delta &= \Delta\Gamma(l)\\ &= \Gamma_L(1 - e^{-2\alpha\cdot\frac{\lambda}{2}})\end{aligned}$$

A plot of these circles is shown in figure 4.87, however, if we want to have a very accurate analysis then ideally we have to really draw this spiral on the Smith chart.

The analytic calculation of the impedance on a lossy line is as general as we have discussed earlier in the previous chapters. So those impedance transformation relationships using the hyperbolic cosines and sines are applicable for the calculation of the impedances on the transmission line. Therefore, with this minor modification to the analysis, the Smith chart can be used for moderately lossy transmission lines.

4.15.3 Characteristic Impedance and Propagation Constant measurement

Here we would like to measure the characteristic impedance and the propagation constant which are the secondary constants. Normally for calculations of transmission lines, we generally do not estimate the primary constants rather we estimate the secondary constants. Assuming there is a setup which can measure an unknown impedance at an unknown frequency.

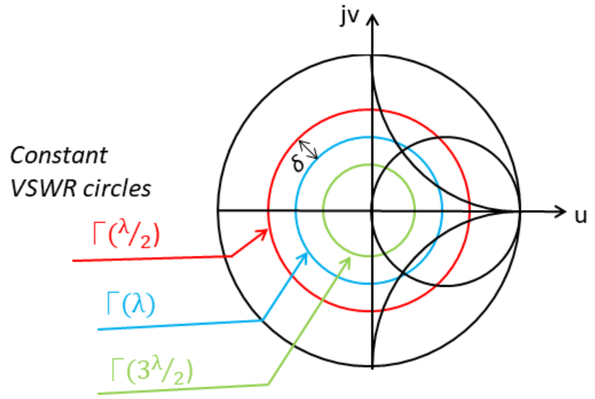


Figure 4.87: Series of concentric circles of reflection coefficient $\Gamma(l)$ for $l = \frac{\lambda}{2}, \lambda, \frac{3\lambda}{2}$ after we apply the approximation.

Then we can measure the characteristic impedance of a moderately lossy line or a low loss line. This measurement can be done by conducting a short circuit and open circuit test of a section of a transmission line. Consider the length of the transmission line, l, one end of the transmission line is connected to the impedance measurement setup as shown in figure 4.88.

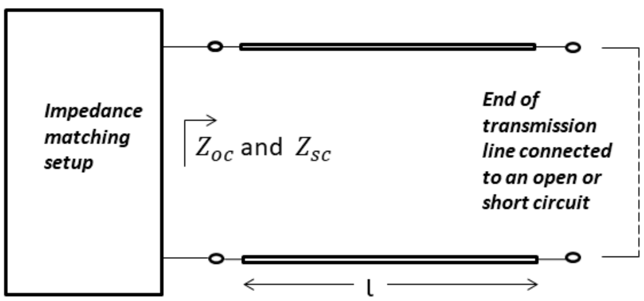


Figure 4.88: Impedance measurement setup

We conduct the two tests and measure the input impedances of this length of the transmission line by making the other end a short circuit and an open circuit. From the length, we get Z_{oc} and Z_{sc}. Recall,

$$Z_{sc} = Z_0 \coth \gamma l \tag{4.75}$$
$$Z_{oc} = Z_0 \tanh \gamma l \tag{4.76}$$

So with l known, we can measure these impedances. We can get the expression for the characteristic impedance and the propagation constant as follows;

Multiplying equation (4.75) and (4.76), gives

$$\begin{aligned} Z_{sc} Z_{oc} &= Z_0 \tanh \gamma l Z_0 \coth \gamma l \\ &= {Z_0}^2 \quad \text{Recall, } \coth \theta = \frac{1}{\tanh \theta} \end{aligned}$$

Therefore,

$$Z_0 = \sqrt{Z_{oc} \cdot Z_{sc}} \tag{4.77}$$

Also dividing equation (4.75) by (4.76) gives

$$\begin{aligned} \frac{Z_{sc}}{Z_{oc}} &= \frac{\tanh \gamma l}{\coth \gamma l} \\ &= \tanh^2 \gamma l \end{aligned}$$

Thus,

$$\begin{aligned} \tanh \gamma l &= \sqrt{\frac{Z_{sc}}{Z_{oc}}} \quad \text{Let } \sqrt{\frac{Z_{sc}}{Z_{oc}}} = A \\ &= A \quad \text{But } \tanh \theta = \frac{e^{\theta} - e^{-\theta}}{e^{\theta} + e^{-\theta}} \\ &= \frac{e^{\gamma l} - e^{-\gamma l}}{e^{\gamma l} + e^{-\gamma l}} \end{aligned}$$

Dividing the numerator and denominator by $e^{-\gamma l}$

$$\begin{aligned} \tanh \gamma l &= \frac{(e^{\gamma l} - e^{-\gamma l})/e^{-\gamma l}}{(e^{\gamma l} + e^{-\gamma l})/e{-\gamma l}} \\ &= \frac{e^{2\gamma l} - 1}{e^{2\gamma l} + 1} \\ &= A \end{aligned}$$

Making $e^{2\gamma l}$ the subject of the formula

$$\begin{aligned} e^{2\gamma l} &= \frac{1+A}{1-A} \equiv \Re e^{j\theta} \\ e^{2(\alpha + j\beta)l} &\equiv \Re e^{j\theta} \quad \text{Since } \gamma = \alpha + j\beta \end{aligned}$$

$$e^{2\alpha l} e^{j2\beta l} \equiv \Re e^{j\theta} \tag{4.78}$$

$$\begin{aligned} \Re &= \left| \frac{1+A}{1-A} \right| \\ \Rightarrow e^{2\alpha l} &= \left| \frac{1+A}{1-A} \right| \end{aligned}$$

$$\alpha = \frac{1}{2l} \ln \left| \frac{1+A}{1-A} \right| \tag{4.79}$$

Hence attenuation constant can be calculated once A is known, where $A = \sqrt{\frac{Z_{sc}}{Z_{oc}}}$. Next we find β from equation (4.78).

$$e^{j2\beta l} = e^{j\theta}$$

However, now there is an ambiguity of multiples of 2π since the same magnitude of a complex variable can be gotten after we transverse 2π in either direction. So, $e^{2\beta l} = e^{(\theta \pm 2m\pi)}$ and

$$\beta = \frac{1}{2l}(\theta \pm 2m\pi) \tag{4.80}$$

Where m is an integer quantity and θ is the angle of $\frac{1+A}{1-A}$. Hence β is calculated with some level of uncertainty in terms of what should be m for our solution. The calculation of the attenuation constant as we have noticed is straightforward from the short circuit and open circuit test.

If we take the length of the line to be $\frac{\lambda}{2}$, it is clear that $m = 0$ for sure, and then we have a unique value of β or a correct value for β. However, if we take $l < \frac{\lambda}{2}$, especially at high frequencies, the length of the line is very small and since attenuation is very small on the transmission line the α generally is very small. So for a small length of transmission line, the losses are not very significant, as a result when you try to calculate the equivalent value of α, you do not get a very accurate value for α, because of the small length of transmission line which is less than $\frac{\lambda}{2}$. The line behaves more or less like a lossless line. Hence the equation of attenuation constant becomes rather unreliable if we take a small length of line which is less than $\frac{\lambda}{2}$.

On the other hand to improve the reliability of α, if we take a long length of cable, suddenly we have many periods of the wavelength on this transmission line. While $\alpha = \frac{1}{2l}\ln\left|\frac{1+A}{1-A}\right|$ becomes more accurate at this point, we have to resolve the $\beta = \frac{1}{2l}(\theta \pm 2m\pi)$ problem of many solutions within the line length in the measurement of the phase constant.

To resolve this issue, using the setup we measure two frequencies. The first measurement of Z_{oc} and Z_{sc} is taken at one frequency to get β which is ambiguous with $2m\pi$. Then the frequency is changed slowly so that we get the same value of phase variation. At that point, the number of cycles which we have on the transmission line has just changed by one. So by slowly changing the frequency and making sure only one cycle change takes place in θ, then the m is changed from m to $m+1$ and then we get a value of β. From here the two values of β is used to find the correct value of β as follows:

Let f_1 and f_2 be the frequencies in which we carried out the measurement of Z_{oc} and Z_{sc}, let us assume at f_1 we get Z_{oc} and Z_{sc} and get

$$\beta_1 = \frac{1}{2l}(\theta + 2m\pi)^{43} \tag{4.81}$$

Next, we slowly change frequency to get to where Z_{oc} and Z_{sc} again become the same. Assuming that the length of the transmission line is very large, by changing the frequency by a small amount, the attenuation constant does not change significantly. What it essentially means is that if we increase the frequency by a small amount, the number of wavelengths set up on the line is changed by a cycle($\frac{\lambda}{2}$), then the loss does not change significantly and that is the reason we get the same value of Z_{oc} and Z_{sc}.

At f_2,

$$\beta_2 = \frac{1}{2l}(\theta + 2(m+1)\pi) \tag{4.82}$$

Subtracting equation (4.81) from (4.82)

$$\begin{aligned}\beta_2 - \beta_1 &= \frac{(m+1)\pi}{l} - \frac{m\pi}{l}\\ &= \frac{\pi}{l}\end{aligned}$$

β_2 and β_1 are related to the velocity or wavelength of the transmission line as ($\beta = \frac{2\pi}{\lambda} = \frac{2\pi}{\frac{v}{f}} = \frac{2\pi f}{v}$) so that

$$\begin{aligned}\frac{2\pi f_2}{v} - \frac{2\pi f_1}{v} &= \frac{\pi}{l}\\ \frac{2\pi(f_2 - f_1)}{v} &= \frac{\pi}{l}\\ \beta = \frac{2\pi f}{v} &= \frac{2\pi f}{2l(f_2 - f_1)}\end{aligned}$$

$$v = 2l(f_2 - f_1) \tag{4.83}$$

But $\beta = \frac{2\pi f}{v}$, so substituting equation (4.83) gives:

$$\beta = \frac{2\pi f}{2l(f_2 - f_1)}$$

$$\beta = \frac{\pi f}{l(f_2 - f_1)} \tag{4.84}$$

So what we have done is carry out this measurement at two frequencies which are closely spaced in such a way that changing the frequency from f_1 to f_2, the number of cycles on the section of the line is only increased by one. Then from there, we calculate the value of the velocity of the wave in the transmission line and then we get the phase constant.

At this point, one may ask *why are we estimating the velocity on the line? Does the wave not travel with the velocity of light on the line or if we know β can we not find what λ is and then find the velocity?* In fact, the problem is exactly the opposite. The problem depends on the structure of the line, the velocity of the wave actually changes, FREQUENCY is the one that is SACRED but depending upon propagation characteristics, velocity changes or the wavelength changes and therefore the phase constant changes. So it is not that we know the wavelength from the velocity and we are trying to find out β. In fact, β is the quantity which is the most unknown quantity. So for a given structure we first estimate the value of β, then wavelength ($\frac{2\pi}{\beta} = \lambda$) and velocity ($\lambda f = v$). So in practice, the measurement of the attenuation constant and the phase constant have to go through all these steps that have been stated.

This is the method which we mentioned in the open or short circuit test. This is the most widely used test in practice for measuring the characteristic impedance and propagation constant of the line. Of course, there are certain practical difficulties when we try to apply the open circuit or short circuit to the end of the transmission line, even if you short the two conductors of a transmission line, there will always be some conductance at the end. If the two ends are left open for the open circuit test, there will be some fringing capacitance at the end of the line. So realizing a perfect open or short circuit at a very high frequency is not that straightforward. People make extra effort to develop modules called short circuit modules and open circuit modules which can be connected to the end of the line to realize a good open or short circuit.

Figure 4.89: Open circuit and short circuit load

In this section, we have discussed the characteristics of a lossy line and that of a moderately lossy line. We have also studied how to make use of the Smith chart for this moderately lossy transmission line and the use of a practical method that brought in δ change for $\frac{\lambda}{2}$ movement in length when drawing the approximation for the reflection coefficient on the gamma plane. Lastly, we saw a practical method for estimating the characteristic impedance and complex propagation constant of a transmission line.

Exercises

Ex. 100 — A lossy transmission line has a characteristic impedance of 50Ω and a propagation constant of $0.02 + j0.1$. Find the wavelength, velocity of propagation and the loss in dB/m.

4.16 Various Types of Transmission Lines

A final aspect of our discussion on transmission line is the characteristic impedance of common transmission lines in practice. At the beginning of our study, we saw various kinds of transmission line such as coaxial line, parallel wire, micro-strip structure etc., one would need to calculate the characteristic impedance of the line. Let us see the formulas used for calculating the characteristic impedance of the various transmission line. The most commonly used transmission line is the coaxial line commonly used for connection between electric equipments.

4.16.1 Coaxial line

The coaxial cable has an inner diameter d outer diameter D and dielectric constant ϵ_r for the medium separating the outer and inner conductor (see figure 4.90). Teflon is mostly used as the dielectric material to separate d from D and it has range of values from 2 to 4.

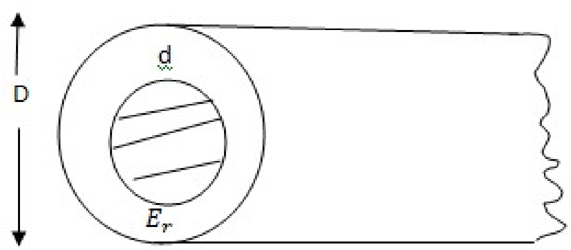


Figure 4.90: coaxial line

Figure 4.91: Typical image of a coaxial line

For this structure the characteristic impedance Z_0 is given by equation (4.85).

$$Z_0 = \frac{138}{\sqrt{\epsilon_r}} \log\left(\frac{D}{d}\right) \tag{4.85}$$

It shows that the characteristic impedance of a coaxial line depends on the ratio of outer to inner diameters, and the dielectric constant separating the inner and the outer diameters.

Let us take some typical values just to get a feel of what kind of parameters are required to realize certain characteristic impedance, say ϵ_r=4,

$$\begin{aligned} Z_0 &= \frac{138}{\sqrt{4}} \log\left(\frac{D}{d}\right) \\ &= 69 \log\left(\frac{D}{d}\right) \end{aligned}$$

This implies that

$$\frac{D}{d} = 10^{\left(\frac{Z_0}{69}\right)}$$

What we see from this relationship is that the ratio of $\frac{D}{d}$ increases very rapidly as the characteristic of impedance Z_0 increases. Suppose that $Z_0 = 69\Omega$ then, $\dfrac{D}{d} = 10$ or $D = 10d$. Similarly, when $Z_0 = 138$

$$\begin{aligned} \frac{D}{d} &= 10^{\left(\frac{138}{69}\right)} \\ &= 10^2 \\ &= 100 \end{aligned}$$

Therefore $D = 100d$.

Thus, the size of the outer conductor compared to the inner conductor increases very rapidly as the characteristic impedance of the line increases. This implies that the coaxial structure is more suited for realizing low characteristic impedances and that is the reason why typically the lines which are used as coaxial lines have characteristic impedance that lie around 50Ω to 75Ω. If we say let us realize a characteristic impedance of 200 or 300 Ohms with the same structure, $\frac{D}{d} = 10^{\left(\frac{300}{69}\right)} = 22275.4$ which is obviously an outrageous value of $\frac{D}{d}$ and this size will be physically unreliable. Hence, the coaxial structure is intrinsically more suited for realizing low characteristic impedances of order 50Ω or 75Ω. Typically the coaxial cables are standardized for 50Ω, however when we go for antenna application we get a cable that has 75Ω characteristic impedance and the reason is that the antenna that is mostly used in practice that is, the *half wave dipole* has input impedance very close to 75Ω as we shall see in later chapters. So just from the compatibility point of view of the impedances, whenever we use the coaxial cables for the antenna, we use the cable which is having a characteristic impedance of 75Ω. However for most of other application of high frequency equipment the impedance has been almost standardized to 50Ω.

4.16.2 Parallel wire transmission line (balanced)

As the name suggest it has two conductors which are parallel with conductor diameter d and separation between the two conductors as D (see figure 4.92). Normally this arrangement is used in the air, so that the dielectric material separating these two conductors is air with $\epsilon_r = 1$. However there are some applications where they are insulated from each with an insulating material and then encapsulated in a material casing; in this case, $\epsilon_r \neq 1$. Such a structure is common in a overhead power line or a *flat ribbon cable* used to connect a *yagi antenna* to a television.

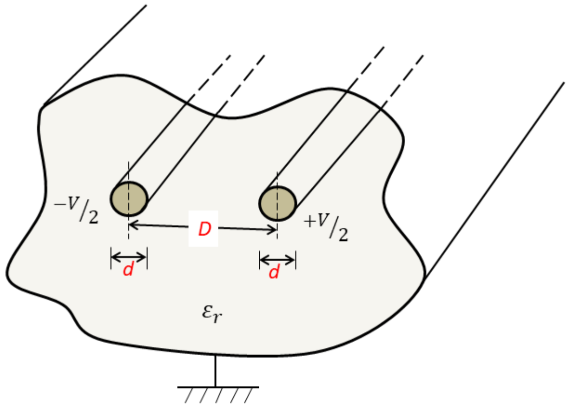


Figure 4.92: Parallel Wire Transmission Line

In the case where there is no air medium, we have a plastic encapsulation with $\epsilon_r \neq 1$. The characteristic impedance of this structure is given by equation 4.86.

$$Z_0 = \frac{276}{\sqrt{\epsilon_r}} \log\left(\frac{2D}{d}\right) \tag{4.86}$$

When air is the dielectric, ϵ_r=1, so that

$$Z_0 = 276 \log\left(\frac{2D}{d}\right) \tag{4.87}$$

We can only make D as small as possible to vary Z_0. if $D \approx d$ that is, almost close to each other then

$$\begin{aligned} Z_0 &= 276 \log\left(\frac{2d}{d}\right) \\ &= 276 \log(2) \\ &= 82.8\Omega \end{aligned}$$

So minimum realizable value of D is when $D = d$, and at that point $Z_{\min} = 82.2\Omega$. Suppose $\frac{D}{d} = 5$, then

$$\begin{aligned} Z_0 &= 276 \log(2 \times 5) \\ &= 276 \log(10) \\ &= 276\Omega \end{aligned}$$

As we can see for the parallel wire (balanced) structure, realizing low impedance is more difficult compared to coaxial cables. This is because we cannot realize impedance less than 82.8Ω. However we can realize high impedance easily. Since we are talking about two conductor which are parallel, separating them to vary D is very easy.

This kind of transmission line intrinsically is used for realizing high characteristic impedance. They are also used in telephone lines— all the telephone lines which are seen in towers are in the form of parallel wire lines. So generally, the parallel wire transmission line are used in that application where one would like to realize high impedances or conversely, when one would have to give such a structure a high impedance. So the coaxial structure and parallel wire structure are complimenting in terms of characteristic impedance. The coaxial cable structure can give low impedances easily and are difficult to get high impedances from. Whereas the parallel lines will easily give high impedance but are difficult to get low impedance from. Typically for a parallel wire transmission line the characteristic impedance is $Z_0 = 300\Omega$ or $Z_0 = 600\Omega$ so the parallel wire transmission line are also standardized to a fixed characteristic impedance.

4.16.3 Microstrip line structure

The third structure in practice is the microstrip line structure and this structure is realized in practice at microwave frequencies for making circuits. Also, whenever we have a printed circuit board configuration the microstrip line is the kind of transmission line used (see figure 4.93). The characteristic impedance of the structure depends on the dielectric ϵ_r and the ratio of $\frac{W}{h}$. The characteristic impedance for this structure is given as equation 4.88

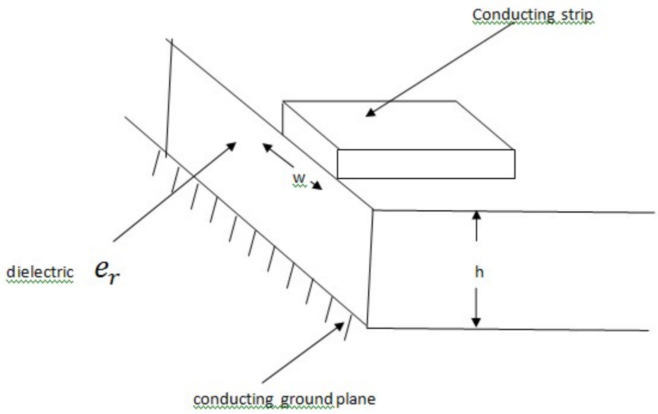


Figure 4.93: Microstrip line

$$Z_0 = \frac{377}{\left(\sqrt{\epsilon_r}\frac{W}{h} + 2\right)} \tag{4.88}$$

At microwave frequencies, one can use a substrate like alumina as the dielectric which has a high value of dielectric constant (as high as 9.8). The dielectric constant value may have very wide range and $\frac{W}{h}$ can vary by a very wide range, so that we can realize a wide range of characteristic impedance. Using high ϵ_r we can realize low Z_0 and high Z_0 with small ϵ_r.

We can vary the impedance by a large amount by changing $\frac{W}{h}$ ratio. Such that at high frequencies, we standardize the frequency to 50Ω, if we are to connect it to a coaxial cable or make it have high impedance (maybe if we want to connect it to the parallel wire transmission line). There is a small difference however between these two connections and that is if we take a structure which is the coaxial cable structure with inner and outer conductor, we can connect voltage to the inner conductor and ground the outer conductor as shown in figure 4.94.

It shows that a coaxial cable has a centre conductor which is connected to a voltage source and the outer conductor is grounded. This configuration is the same for the micro-strip line structure. We have the ground plane and the strip which carries voltage relative to the ground plane. So in both cases, above the ground is defined and the voltage is applied with respect to ground. However comparing this with the parallel wire structure, this is a completely floating structure and the ground is not defined for this structure. We can either say one of these

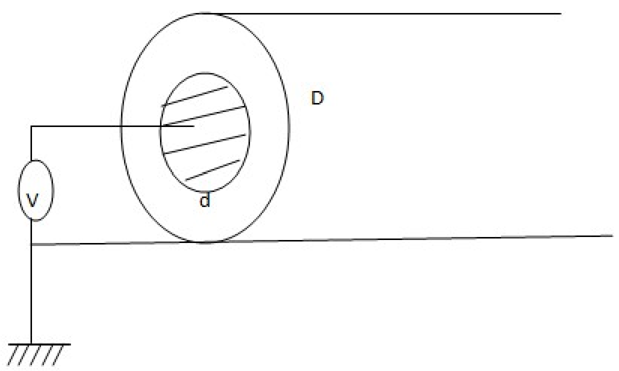


Figure 4.94: Variable Microstrip Line Structure

conductors is at ground or zero potential while the other is at V or we say that there is a virtual ground in between them somewhere at the middle with one cable at $\frac{+V}{2}$ and the other at $\frac{-V}{2}$. Hence this structure for which the ground is not defined is a floating structure while the other two have a well defined ground and the voltage is applied between the ground and the conductor.

The other two structures are called *unbalanced structures* and the parallel wire transmission line is called a *balanced structure*. Whenever we make connection between balanced and unbalanced lines,two things we see is that first impedance has to be matched at the junction, secondly since you are bringing now a connection between a balanced floating structure and an unbalanced structure,you require some transformer in between which can connect voltage from one bias structure to another bias structure. This device is called a *balance to unbalanced transformer* or *BALUN*.So a structure which matches the impedance as well as the nature of the cables is called BALUN[44] as shown in figure 4.95

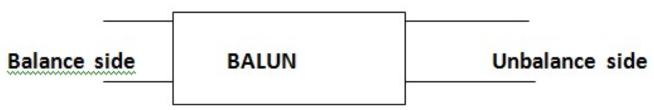


Figure 4.95: Boxview model of a BALUN

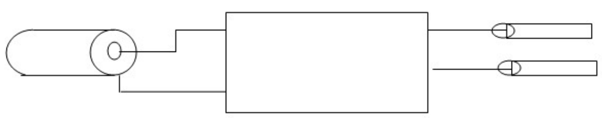


Figure 4.96: Typical Application of a BALUN

A typical application as shown in figure 4.96 is when we make a connection between a *yagi antenna* and a television. This is the most visible application you can see around, of course there are many other high frequency application of the *BALUN*. Practically when using a *yagi antenna* connected to a television, the signal from the antenna is sent using a *flat ribbon cable* connected to the *BALUN* structure which is a black-box seen behind a television set. Typically, the connector at the back of a television set is a coaxial connector so it is unbalanced. The impedance of the coaxial connector is 50Ω, but for the *flat ribbon cable* coming from the antenna, it is much different compared to 50Ω. So normally we introduce a small box which is the *BALUN* and it transforms the impedance from the *flat antenna ribbon cable* balanced structure to the unbalanced structure. If this is not done, there is the possibility that the reflection from the junctions will appear in form of ghosts on the television. So appropriate transformation of impedance on the line and also connecting balanced to unbalanced side of the network properly with the help of a *BALUN* improves the quality of the reception with any high frequency signal. This essentially conclude our discussion on the transmission line.

[44] An acronym from the words BALanced and UNbalanced

4.17 Conclusion

Let us recap what we have done in this very important aspect of high frequency circuit called transmission line. We saw the limitation of the analysis of the lumped circuits that is as we go to high frequency,the size of the component becomes comparable to wavelength and then the voltage and current cannot be assumed constant along any electronic component, so we introduced the concept of the *distributed element*. We studied the *transit time* effect and from there we derived the relationship between voltage and current for high frequency circuits that is, for circuits where the wavelength becomes comparable to the dimensions of the components— transit time effect cannot be neglected. We saw that the relationship were in the form of differential equations whose solution turned out to be that of a wave equation. So in general at high frequencies the voltage and current exist in the form of waves in the circuits.

Then we studied the superposition of these waves and we saw that in general we got a standing wave. We also investigated what conditions lead to two waves which are traveling in opposite directions or which lead to standing waves. So we introduced the concept of voltage *reflection coefficient*, which we discussed as the measure of energy reflected away from the load.

Then we introduced the concept of *matching condition* which means that if a load impedance is equal to the characteristic impedance, the energy is transferred from the generator to the load efficiently and there is no reflection on the line. We then studied the impedance transformation characteristic of the transmission lines. Then we looked at many applications of transmission line which are *as a circuit element*, *as a resonant circuit*, *as a voltage/current step up transformer* and *as a matching transformer*.

Lastly, we looked at the characteristic impedance for various structures which are used as transmission line at high frequencies. For the coaxial cable, we concluded that its characteristic impedance is rather low and that it is difficult to realize high characteristic impedance using the coaxial structure. On the contrary the parallel wire structure is more capable of giving high characteristic impedances. We also looked at some applications where the parallel and coaxial transmission line could be used and lastly, was the microstrip line structure which we studied to be normally seen at microwave frequencies of few GHz, its connection on a printed circuit board, and lastly using a *BALUN* to connected balanced and unbalanced structures in other to prevent reflections.

Today computer speed reach GHz and transmission line effects is going to play a very prominent role in the circuit design. Few decades back when the frequencies where not very high, electronics circuit design was quite simple, all the transmission line effect not playing a role in the circuit. However when circuit chips are operated at frequencies of few GHz (computers are soon operating at frequencies of a few GHz), the reflections, the mismatching and so on which we discussed become very vital in designing the electronic circuits. So in today's electronic circuits, the concept of transmission line plays a very important role. The subject of transmission line has become very important for recent years because of these high speed electronic circuits which are now part of our everyday life.

4.18 Exercises

Ex. 101 — Find the characteristic impedance each of the following transmission lines:

(a)a coaxial cable using a solid polyethylene dielectric having $\epsilon_r = 2.3$ with inner radius 0.5mm and outer radius 1.5mm;

(b)a parallel wire line with wire radius 0.5mm and spacing 1mm;

(c)a microstrip line with substrate thickness 1mm, dielectric constant 4.5, and strip width 1mm.

Ex. 102 — A 600-Ω transmission line is 150m long, operates at 400kHz with $\alpha = 2.4 \times 10^{-3}$ Np/m and $\beta = 0.0212$ rad/m, and supplies a load impedance $Z_L = 424.3\angle 45°\,\Omega$. Find the length of line in wavelengths, $\Gamma_L, \Gamma(L)$, and $Z(L)$. For a received voltage $V(0) = 50\angle 0°$V, find $V(L)$, the position on the line where the voltage is maximum, and the value of $|V|_{\max}$.

Ex. 103 — A 70-Ω high frequency lossless line is used at a frequency where $\lambda = 80cm$ with a load at $x = 0$ of $(140 + j91)\Omega$. Use the Smith chart to find the following:

(a)The reflection coefficient Γ_L

(b)VSWR

(c)The distance to the first voltage maximum from the load

(d)The distance to the first voltage minimum from the load

(e)The impedance at the first voltage maximum

(f)The impedance at the first voltage minimum

(g)The input impedance for a section of line of length that is $54cm$ long, and

(h)The input admittance for a section of line of length that is $54cm$ long.

1.An experiment conducted by Bells lab by Engr. Hagar smooth showed that when a transmission line having the characteristic impedance of 50 ohms was used to transmit and electromagnetic wave at a frequency of 30MHz across a load or $100 + j100$. The power delivered was insufficient due to an uneven matching conditions. By placing two stubs, where the first is at a distance of 0.1λ from the load, and the distance between the two stubs is $3\lambda/8$. Find the:

i) Open circuit distances

ii) Short circuit distances The fact that he can make use of any for his design for maximum power transfer.

2.As a new engineer working for Intel incorporation, you know you have been assigned an immediate tax to work on an antenna design where your major role involves matching the impedance between the field line and the antenna. In order to avoid wave reflections from the transmission line having characteristic impedance of 50ohms across the antenna of impedance $100 + j100$ohms. If two shunt stubs are placed on the line where the distance between the load and the first stub is negligible and the spacing circle is $\frac{3\lambda}{8}$. At what lengths are the stubs short circuited to have an equivalent matching condition across the line (Note: Negligible means l1 = 0).

3.For the transmission line diagram shown below, using double stubs matching, find the open-circuit lengths of the stubs and draw the equivalent diagrams.

4.At what length does a shunt stub have to be placed from the transmission line of normalized admittance of $0.25 + 0.25j$ ohms to obtain maximum power transfer when the characteristic impedance is 50 ohms?

5.Also, from the above question, what is the load impedance and what is the limitation to this type of matching technique?

6.Find the short circuit length and the distance at which a series stub has got to be placed in order to match a transmission line of load admittance, $0.25+0.25j$, and characteristic impedance of 50ohms. What type of matching is best applied for the above question? And draw the equivalent circuit diagram?

7.As an IT student working for JLC PCB, you have been consulted by a client who wants to design a PCB at a frequency of 50 MHz across a load impedance of $100+j100$ ohms. The client is unable to find a perfect matching of impedance after using the single stub matching technique. With your profound knowledge of double stub matching and by placing the stub at a negligible distance from the load, while the spacing circle is $\frac{\lambda}{8}$. Find the open-circuit lengths at which the clients stub must be situated to have perfect matching across the transmission line having a characteristic impedance of 50ohms.

8.Given: $Z_L = 100 + j100\Omega$

$Z_O = 50\Omega$

$l_1 = 0.1\lambda$

$l_2 = 0.375\lambda$ Find the

aOpen circuit length

bShort circuit length

9.An 80MHz electromagnetic signal was transmitted to an RF antenna having load impedance of $100 + j100$ ohms and characteristic impedance of 50 ohms. If two shunt stubs are used in order to minimize reflection across the transmission line, where the distance from the first stub is 0.1λ and the spacing circle is $\frac{3\lambda}{8}$, find the open circuit lengths, L_{S1} and L_{S2}, of the stub that is necessary to achieve maximum power transfer and draw the equivalent diagrams of the circuit.

Chapter 5

Coordinate System and Vectors

5.1 Objectives

By the end of this chapter, the you should be able to:

(i) To introduce the concept of coordinate systems and their importance in electromagnetism and antenna theory.

(ii) To explain the different types of coordinate systems, such as Cartesian, cylindrical, spherical, and others, and their properties, such as origin, axes, unit .

(iii) To demonstrate how to convert between different using transformation equations and matrices.

(iv) To show how to perform basic operations, such as dot product, cross product, , , and indexcurl, on and vector fields in different coordinate systems.

(v) To apply the basic operations to solve problems in electromagnetism and antenna theory.

(vi) To enable understanding of the advantages and disadvantages of different coordinate systems for different applications in electromagnetism and antenna theory.

(vii) To develop the skills and intuition to choose the most appropriate coordinate system for a given problem or situation in electromagnetism and antenna theory.

(viii) To review and revise the mathematical background and tools needed to work with coordinate systems, such as trigonometry, calculus and linear algebra.

(ix) To provide examples and exercises that reinforce the learning outcomes and objectives of the coordinate system topic.

Up to this point, we have discussed one important case of transmission lines. We saw that when we brought in the concept of space, the voltage and current came to exist in the form of waves in the circuit. However, the concepts of voltage and current apply to **bound structures** like electrical circuits where you have conductors separated by dielectrics, coaxial cables, parallel wire transmission lines and so many others. In media that are **unbounded** or semi-infinite in extent or a media which is only di-electric, we find that the use of voltage and current is not very attractive. In fact, in many applications, it is very difficult to find quantities like voltage and current. In this situation, we have to go to the more fundamental quantities like the electric field and magnetic field. So having now got some feel for the wave phenomenon, especially for special cases like current and voltage, we depart to the more generalized phenomenon of electromagnetic waves and that is the waves in the form of electric and magnetic fields. Henceforth, we would discuss the phenomenon of electromagnetic waves in terms of electric field and magnetic field. Electric field and magnetic field are vector quantities so we have to deal with $\boldsymbol{E}$ and $\boldsymbol{H}$ analysis on three-dimensional space. To achieve this, we would revise our concept of vector calculus and vector algebra.

Let us first see how to represent the vector quantities in 3D space before going into vector algebra and calculus. Mathematically, we have to represent the quantities $\boldsymbol{E}$ and $\boldsymbol{H}$ as in the three major coordinate systems.

5.2 Coordinate systems

The three major coordinate systems used for vector representation in three-dimensional space are :

(i) Cartesian coordinate system (x, y, z)

(ii) Cylindrical coordinate system (r, ϕ, z)

(iii) Spherical coordinate system (r, θ, ϕ)

5.2.1 Cartesian coordinate system

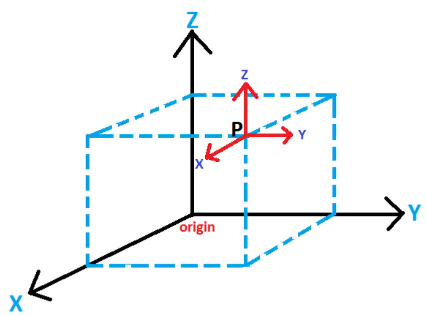


Figure 5.1: The Cartesian coordinate system

Imagine if we have a three-dimensional unbound space, like a box, the three axes will be the three edges of the box x, y, and z in the sequence x to y to z. The coordinate of a point P in the Cartesian[1] coordinate system can be written as $\boldsymbol{P}(x, y, z)$.The is the convention we follow when defining the axis.

[1]The adjective ˇCartesian˝refers to the French mathematician and philosopher Rene Descartes (1596 - 1650), born in France, he published this idea in 1637, seventeenth century. He is considered the father of modern philosophy. Many other coordinate systems have been developed since Descartes, such as the polar, cylindrical and spherical coordinates.

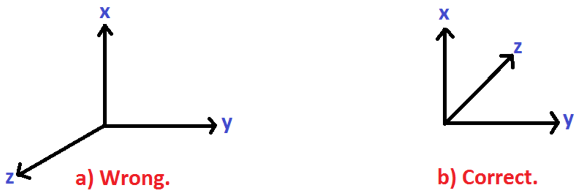


Figure 5.2: The right-hand coordinate system

By following the **right-hand convention**, if we point the fingers of the right hand going from X to Y, then the thumb points in the **positive** Z direction. Note that the right-hand fingers wrapped around an object shows the rotation from X to Y by the curled fingers wrapped around the object. This is X rotating towards Y to get the position Z (thumb pointing up), that is, $\bar{X} \times \bar{Y} = \bar{Z}$; assume $\bar{X}$ is a vector in the x direction, $\bar{Y}$ is a vector in the y direction and $\bar{Z}$ is a vector in the z-direction.

If the curl of the right-hand goes from Y to Z, the thumb will point in the direction of x or $\bar{Y} \times \bar{Z} = \bar{X}$ in vector notation form.

If the fingers curl from $\bar{Z}$ to $\bar{X}$, the thumb should point in the direction of $\bar{Y}$. The right-hand convention will enable us to resolve some ambiguity in finding the direction of the , so in this case, we visualize the three-dimensional space as a box and apply the right-hand convention. As we can see here if we take a horizontal plane which is XY, you get your right-hand fingers curled from X to Y on this plane, the thumb will point upward to the Z direction. At location P for instance, we can define a vector which is represented by the arrows shown at P.

The three $\bar{\mathbf{X}}, \bar{\mathbf{Y}}, \bar{\mathbf{Z}}$ are called the components of vector P at the location usually pointing in the three coordinate axes. Hence for any vector in space, we have the component along the x direction, along the y direction and the z direction so that any vector in space can be represented by a set of three elements. We require a lot of imagination to visualise a vector that can be the electric or magnetic field in three-dimensional space. We can write these expressions mathematically, but ultimately, it is required for one to visualize these in three-dimensional space in order to make the study of electromagnetic waves more interesting to us. The idea of visualizing the vectors in three-dimensional space is to get a physical feel for the vectors and then the problem is solved. The Cartesian coordinate system is the simplest coordinate system and the important feature of this system is that no matter where you go in space, the direction of vector P along the x, y, and z directions remains the same. This means that if we go along the Y axis, the X axis remains the same and perpendicular to the Y axis at any point in space. This may not be true for other coordinate systems, so we will write down the vector relationship for the Cartesian coordinate system since it is the coordinate system that is very easy to visualize.

5.2.2 Cylindrical coordinate system

Imagine a three-dimensional space like a cylinder, with z as the axis of the cylinder, then if we write the same cartesian coordinate x, y, z, the plane passing through XY will be perpendicular to the axis of the cylinder. The point P has a radius r, if we draw a perpendicular line from P to meet the XY plane, to make distance r, the angle that this radius vector makes with the x-axis is ϕ. The distance we travel from the origin to meet P along

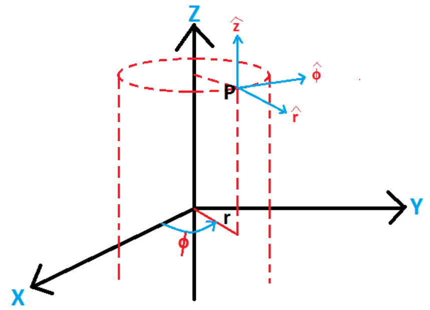


Figure 5.3: The cylindrical coordinate system

the axis of the cylinder is the z point, so we have got a coordinate system with a sequence (r, ϕ, z). Again we follow the right-hand coordinate system, so if we have a vector, then with fingers pointed from r to ϕ (curling from r to ϕ), we must get z direction and if we point from ϕ to z, we get r direction and so on.

How do we define ϕ direction?

An arrow pointing in the direction of z gives the Z vector. Similarly, a vector pointing in the radius vector r outward from the origin is the r vector. The ϕ vector is that which is tangential to the surface of the cylinder which is passing through point P. Here we can see the relationship between the Cartesian coordinate system and the cylindrical coordinate system. The geometry of the problem determines the choice of the coordinate system to be used in the analysis of the problem. This means that we use the rectangular coordinate system for rectangular geometry, and the cylindrical coordinate system will be applied to cylindrical geometry (found in coaxial cable, optical fibre, circular waveguide, or many other structures where the geometry looks more like a cylinder). We can analyse the problem in any coordinate system we choose, but it is easier to work in a coordinate system that is closer to the geometry of the problem, that is why when we do analysis, we first choose the appropriate coordinate system and then solve the problem of electromagnetics in that coordinate system.

The one thing we note in the cylindrical coordinate system compared to the Cartesian coordinate system is that in the Cartesian coordinate system, the direction of the x,y, and z coordinate remains the same everywhere in space no matter where the point moves, the X always orient in the same direction physically. However, in the cylindrical coordinate system, the $\hat{z}$ vector is always in the same direction no matter where you are in space but the $\hat{r}$ and $\hat{\phi}$ vector will keep changing direction as we go to different locations or points in space.

5.2.3 Spherical coordinate system

In this coordinate system, the three-dimensional space is imagined as a big sphere. We shall consider an octant of the sphere, with a point P marked on the surface of the sphere. P can be represented in terms of three quantities viz the radial distance of point P from origin r, the angle ϕ made by a line projected on the XY plane from P with the X axis, the angle θ made by a tangent to the surface in the vertical plane passing through P and this is the angle the radius vector makes with the Z axis. In

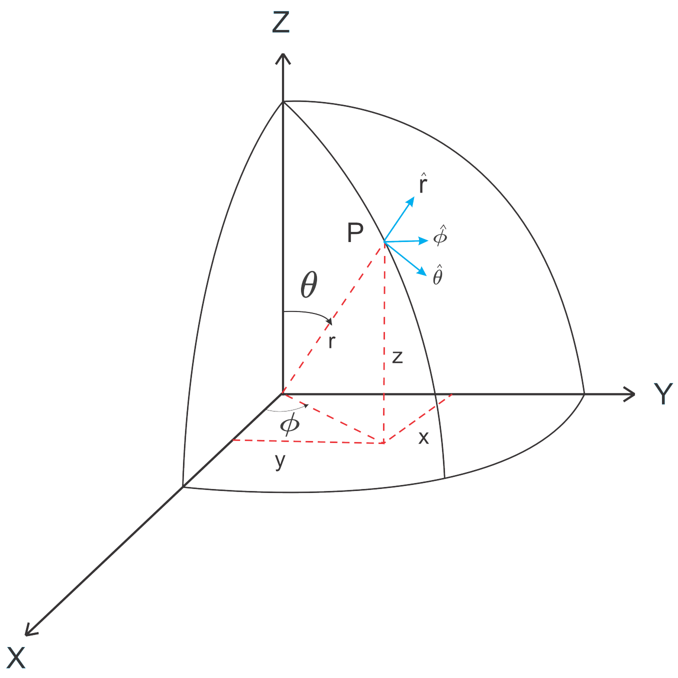


Figure 5.4: The spherical coordinate system

this coordinate system, the point P is represented as (r, θ, ϕ) in that sequence. The r represents the radius vector which is the distance from the origin; θ is the angle which is measured from the z-axis to the radius vector; ϕ is the angle which is measured from the x-axis to the projection of the radius vector on the XY plane.

In a spherical coordinate system, the point is defined by radial distance and two angles compared to the Cartesian or cylindrical coordinate systems. In the cartesian coordinate system, the location was defined by three distances (x,y,z) but in the cylindrical coordinate system, it was defined by two distances (r,z) and one angle ϕ.

Using the right-hand convention; if the fingers curl from r to θ, then the thumb must point in the positive ϕ direction, if the fingers curl from θ to ϕ, then the thumb must point in the direction of $\hat{r}$. For the spherical coordinate system, point P has $\hat{r}$ radially pointing outward and normal to the surface of the sphere; if we take an angle θ and draw a tangent passing through P at that angle, going away from θ, that is, the positive direction of vector θ. The right-hand rule must be followed when developing a coordinate system because when we do vector analysis, we assume that certain conventions like the right-hand convention are being applied. The figures above show how the coordinate axes are drawn when the right-hand convention is followed.

The spherical coordinate system has a special property, that is when the point called origin is defined, all measurements can be taken from there. In the Cartesian coordinate system, you can shift the origin anywhere in space whereas in the cylindrical coordinate system, the cylindrical axis line is defined and all measurements are taken from there. So whenever we have a problem like an antenna kind of problem, where we have a source of energy sending out electromagnetic waves and this source is more like a localized point or region in space, then we use the spherical coordinate system for this problem. For a problem where the energy is going to flow along the length of the structure (structures like coaxial cable, waveguide or transmission line), the cylindrical coordinate system is more appropriate or then in some general cases, the Cartesian coordinate system will be the best if we are considering a closed structure like a closed box like a resonator or a cavity. Let's go to the basic definition of vectors and their operators since we have learnt about coordinate systems.

5.3 Vector analysis

A vector is made up of a set of three components, so any vector can be represented by three components say (a,b,c) irrespective of the coordinate system we are using. Mathematically, this description of a vector is enough. However, when we go to the solution of physical problems of electromagnetic waves, we would like to visualize these vectors in three-dimensional space even though this is abstract. Electric field and magnetic field are very abstract concepts because we do not know how to visualize these quantities, so we have to give some physical picture of these vectors by representing three components (a,b,c).**The most commonly used convention for this is to represent a vector like an arrow**, the arrow is represented with a head and a tail. The arrow direction shows the direction of the vector and the length of the arrow is the magnitude of the vector in three-dimensional space. Suppose the arrow was going away from you or coming towards you, as it goes away from us, we see a circle and an "X "covering the span of the circle and if it comes towards us we see a dot at the centre of the circle. In the arrow representation of a vector, the circle's size denotes the vector's magnitude, and a small circle represents the small magnitude of the vector.

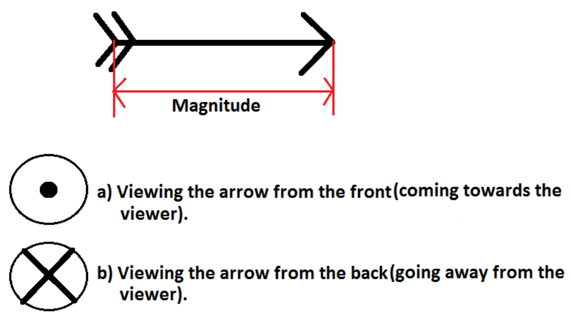


Figure 5.5: The arrowhead representation of vector

This convention shall be followed when we are visualizing three-dimensional vector fields in 3D, we see it as a distribution of these vectors or these arrows in 3D space. Abstract quantities like electric and magnetic fields will have some physical means of visualizing them and we have explained here the framework for visualizing these abstract quantities in 3D space.

5.3.1 Basic operation of vectors

Let's say a vector A denoted as $\bar{A}$ in the Cartesian coordinate system is given as :

$$\bar{A} = A_x\hat{x} + A_y\hat{y} + A_z\hat{z} \tag{5.1}$$

where $\hat{x}, \hat{y}, \hat{z}$ are the unit vectors in the coordinate axis x,y, and z if we imagine a vector like an arrow in the direction of the X axis with unity length, then that vector is denoted as $\hat{x}$. Similarly, a vector of unit length oriented in the y direction is denoted by $\hat{y}$ and a vector of unit length oriented in the Z direction is denoted as $\hat{z}$. A_x, A_y, A_z are called components of vector A in the x,y, and z directions respectively, so a general vector $\bar{A}$ in the three-dimensional cartesian coordinate system can be represented by the x component multiplied by the unit vector in the direction in x plus the Y component multiplied by the unit vector in the y direction, plus the z component multiplied by the unit vector in the z-direction.

Now we can define certain operations on the vector $\bar{B}$ with :

$$\bar{B} = B_x\hat{x} + B_y\hat{y} + B_z\hat{z} \tag{5.2}$$

The addition or subtraction of $\bar{A}$ and $\bar{B}$ is obtained by adding or subtracting individual components of $\bar{A}$ and $\bar{B}$.

$$\bar{A} + \bar{B} = (A_x + B_x)\hat{x} + (A_y + B_y)\hat{y} + (A_z + Bz)\hat{z} \tag{5.3}$$

$$\bar{A} - \bar{B} = (A_x - B_x)\hat{x} + (A_y - B_y)\hat{y} + (A_z - Bz)\hat{z} \tag{5.4}$$

or in general form,

$$\bar{A} \pm \bar{B} = (A_x \pm B_x)\hat{x} + (A_y \pm B_y)\hat{y} + (A_z \pm Bz)\hat{z} \tag{5.5}$$

Another important operation these two vectors can perform is the multiplication operation. Two product operations are defined for vector quantities: a and a .

5.3.2 Scalar product or dot product

$$\bar{A} \cdot \bar{B} = A_x \cdot B_x + A_y \cdot B_y + A_z \cdot Bz \tag{5.6}$$

The dot product of two vectors $\bar{A}$ and $\bar{B}$ is a scalar quantity which is the sum of the product of the components of the two vectors. This is the reason it is called the of two vectors.

5.3.3 Vector or cross product

$$\bar{A} \times \bar{B} = \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ A_x & A_y & A_z \\ B_x & B_y & B_z \end{vmatrix}$$

We solve the determinant to get

$$\begin{aligned} \bar{A} \times \bar{B} = (A_yB_z - A_zB_y)\hat{x} - (A_xB_z - A_zB_x)\hat{y} \\ + (A_xB_y - A_yB_x)\hat{z} \end{aligned} \tag{5.7}$$

This is called a because it gives another vector. We can see that $\bar{A}.\bar{B} = \bar{B}.\bar{A}$ for the , the order does not matter but $\bar{A} \times \bar{B} \neq \bar{B} \times \bar{A}$, the magnitude of the vector remains the same but the vector $\bar{A} \times \bar{B}$ is in the opposite direction of $\bar{B} \times \bar{A}$. $\bar{A} \times \bar{B}$ represents a vector which is perpendicular to the plane containing the vector $\bar{A}$ and $\bar{B}$ if $\bar{A}$ and $\bar{B}$ were represented by two arrows, consider a plane passing through these two arrows, then the cross product vector will be a vector perpendicular to these two arrows (as shown in figure 16.6 below).

The question is how do we know the direction of the arrow which is the cross product? With the fingers curling from $\bar{A}$ to $\bar{B}$, the direction of the thumb shows the direction of the arrow of the vector $\bar{A} \times \bar{B}$, interchanging to $\bar{B} \times \bar{A}$, the fingers will go from $\bar{B}$ to $\bar{A}$ and the direction of the thumb will be opposite so that the magnitude of the vector will remain the same, but the direction changes. These are the two important operations

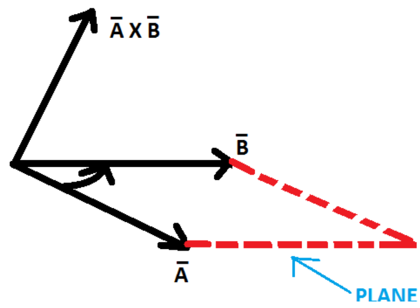


Figure 5.6: The representation of the cross product; $\bar{A} \times \bar{B}$ represents a vector which is perpendicular to the plane containing the vector $\bar{A}$ and $\bar{B}$. $\bar{A}$ and $\bar{B}$ are represented by two arrows, the cross-product vector will be a vector perpendicular to these two arrows.

of vectors which we will encounter when we go to the analysis of electromagnetic waves. We require two operators on the vectors which are the differential operators. Consider a field which is a vector quantity, and at every location in space, we define this quantity which is a vector quantity. So I go to space at any point and measure it. This quantity will have magnitude at that location and this quantity will have orientation if you imagine the vector like an arrow. Let us say we have a quantity like a velocity distribution, let's say air velocity in a medium, if we go around this medium, at every location if we measure the velocity of the wind, and find out in which direction the wind is flowing, you will get the direction of the wind. So we know the extent of the wind movement and the direction in which the wind is moving. These two quantities together can be put in the form of an arrow. The extent of the wind movement is denoted as the length of the vector, and the direction in which the wind was flowing, which can be shown by the arrow. At every location, we have this quantity which is the velocity of air around that medium which can be denoted by a vector. Similarly, let's say we have a flow of some liquid, if we go to various locations, we again have a quantity which is a measure of the rate of flow of liquid at a point and if we know the direction the fluid is flowing, we again represent it with an arrow in that location. A vector field (electric field, magnetic field) is a quantity which can be represented by its strength and also its direction. In 3D space, the vector fields will be of different magnitudes and also they will have different orientations. Assuming we have a scalar field like temperature variation, we measure the variation of temperature in different directions. Suppose we take temperature variation right above the surface of the earth, we must note that temperature drop is more at a higher altitude. Along the horizontal direction, the temperature variation is almost negligible, if we travel a distance of 100 km on the earth's surface, the temperature will not vary significantly but 100 km above the surface of the earth, the temperature variation will be significant by about 40 or 50 degrees. Although temperature quantity is a scalar, its variation depends on direction, it does not have much variation in the horizontal direction but it has significant variation in the vertical direction. A scalar quantity with three dimensions can have a significant variation in different directions, the variation is a vector quantity, so we define a differential operator called the **gradient operator**. The gradient operator is used for the scalar field and the outcome of this is a vector quantity.

Let f be a function of x, y and z i.e f(x,y,z) the gradient of the scalar function f(x,y,z) represents the maximum rate of change of this function with respect to 3D space if we find the rate of change of the function in three-dimensional space x,y,z, then we find out the direction in which the function is changing maximally. The vector of the rate of change of the function f is called the gradient vector which is defined by a ∇

$$\nabla = \frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z} \tag{5.8}$$

Equation 16.8 is a differential operator acting on a scalar function to give a vector. The gradient of a scalar field is

$$\nabla f = \frac{\partial f}{\partial x}\hat{x} + \frac{\partial f}{\partial y}\hat{y} + \frac{\partial f}{\partial z}\hat{z} \tag{5.9}$$

From equation 16.9,

$$\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z}$$

It tells how far the scalar field f travels in the x, y and z directions respectively. To find out the on the scalar function, we find out the rate of change of these quantities with respect to the three-dimensional space and this is the expression for the Cartesian coordinate system.

$$\bar{F} = F_x\hat{x} + F_y\hat{y} + F_z\hat{z} = \bar{F}(x, y, z) \tag{5.10}$$

F_x, F_y, F_z is a scalar quantity in terms of x, y and z. Now we can define the differential operators which are

(i) **Divergence operator** which is like dot product of the and the function $\bar{F}$ i.e $\nabla.\overrightarrow{F}$ gives

$$\nabla.\bar{F} = \frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z} \tag{5.11}$$

(ii) **The cross product** of ∇ and $\bar{F}$ i.e $\nabla \times \bar{F}$ called the curl of vector

$$\begin{aligned} \bar{F} &= \nabla \times \bar{F} \\ &= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} \end{aligned}$$

The curl of a vector is a vector quantity whereas the divergence of a vector is a scalar quantity. Next, we will get the physical feel of these quantities of divergence and curl. It will become obvious if we want to capture those physical effects, then the appropriate concepts will be divergence and curls. Divergence and curl are the basic concepts applied to solve electromagnetic problems.

5.4 Exercises

Ex. 104 — In the description of a vector using arrow notation, the length of the arrow may represent what property of the vector?

Ex. 105 — What is the relationship between the direction of the cross product of vectors and the plane of vectors?

Ex. 106 — The outcome of the gradient of a scalar field is what quantity?

Ex. 107 — Contrast the outcome of the divergence of a vector field and the curl of a vector field.

Ex. 108 — What are the basic importance of scalar and in the basic concept applied to solve electromagnetic problems?

Ex. 109 — If you switch on a torch and point it in any direction, what happens to the rays of light?

Ex. 110 — If $\bar{A} = 2\hat{x} + 3\hat{z}$ and $\bar{B} = 4\hat{x} - \hat{y} + 3\hat{z}$, find: (i) $\bar{A} \times \bar{B}$ (ii) $\bar{B} \times \bar{A}$

Ex. 111 — What is the gradient of the given scalar field: $F = \frac{\cos(5x)}{y}$

Ex. 112 — A field is said to contain only magnitude, and we need to perform an operation on this field to produce a vector field. Carry out this analysis on this given field: $T = x^2yz^3 + xy^2z^2$.

Ex. 113 — Check if the fields given below have a tendency to curl or diverge and write a brief discussion on what actually happens. $\bar{P}$ = xsin(2y)$\hat{x}$, $\bar{Q}$ = e^ycos(z)$\hat{x}$ + $\sinh(x)\hat{y}$.

5.5 Divergence and Stokes Theorem

Here we try and get the physical feel for the divergence and curl of a vector $\bar{F}$ and then find out the basic theorems which are used in vector operations. To get a feel for the dot product, essentially if we consider a vector field like a fluid flow, we draw a small box or small volume and ask how much is the net flow of the fluid per unit volume. That quantity is nothing but the divergence of the vector.

$$\nabla \cdot \bar{F} = \frac{\partial Fx}{\partial x} + \frac{\partial Fy}{\partial y} + \frac{\partial Fz}{\partial z} \tag{5.12}$$

Similarly for the curl of the vector $\bar{F}$, as the name suggests, it is either something curling or rotation is involved here. If we consider a vector field and keep an object in this field such as fluid flow on the surface of a river, then that keeps the object turning on the surface of the river because of the differential flow of the layers of the water, they sometimes have a rotational effect on the object. That rotational effect is what is captured by the curl operator. So if we define the net rotation created on an object per unit area of the object, then that quantity is essentially the curl of that vector field.

$$\begin{aligned} \text{Curl of vector} \quad \bar{F} &= \nabla \times \bar{F} \\ &= \begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ F_x & F_y & F_z \end{vmatrix} \end{aligned} \tag{5.13}$$

5.5.1 Divergence or Curl Producing Vector Fields

To add a bit of a feel, let us ask what kind of field will give you divergence or curl. We can draw a vector field shown below.

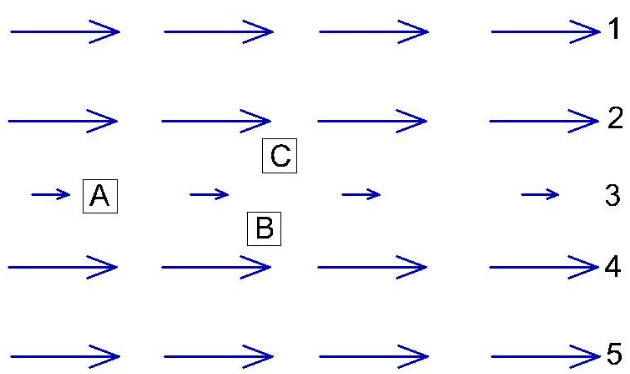


Figure 5.7: Horizontal Vector Field

Along the horizontal direction, the vectors are the same in magnitude and direction. Going down vertically across, the magnitude changes for 2 to 3 and 3 to 4. 1 to 2 and 4 to 5 remain the same. A small area **A** placed at the location shown, can create some kind of shear on the object depending on the strength of the vector through the top and bottom of **A**. **B** on the other hand will turn counterclockwise while **C** will turn clockwise as the force created by the arrow magnitude is more at the top for **C** compared to the bottom of **C** and more at the bottom of **B** compared to the top. Hence rotation is experienced in that region. If we however consider a field shown below. That is the field whose magnitude is changing in the **x** direction.

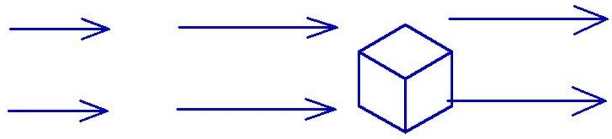


Figure 5.8: Changing Field Magnitude

If we put a small volume in this region and treat the field like fluid flow velocity with increasing velocity of the fluid as we move towards the right as shown by the magnitude of the arrow. The fluid coming out of the box has more value than that going into it. Hence there is a net outward flow of fluid from this volume. This kind of field will have a divergence.

Divergence is the best way to capture the effect of something either flowing out of a volume (oozing out) and the direction can be positive (net flow out of the volume) or negative (net flow into the volume).

Looking at this phenomenon where some rotation is involved, the phenomenon which can capture the effect is the *curl*. Divergence and curl are the mathematical operations which capture these phenomena.

We see later when we go into electromagnetic fields, that the concept of physics is mathematically captured by the curl and divergence operators.

5.5.2 Laplacian operator

Another operator defined in terms of the del operator ∇ is called the **Laplacian Operator**. It is a second-order operator denoted as $\nabla^2 = \nabla \cdot \nabla$.

If we take the ∇ as a vector operator, then the dot product of the two ∇ operators is given by ∇^2 and it becomes a scalar operator.

$$\nabla \cdot \nabla = \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z} \right) \cdot \left(\frac{\partial}{\partial x}\hat{x} + \frac{\partial}{\partial y}\hat{y} + \frac{\partial}{\partial z}\hat{z} \right) \tag{5.14}$$

$$\begin{aligned} \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2} &= \nabla^2 \\ &= \nabla \cdot \nabla \end{aligned} \tag{5.15}$$

So the Laplacian operator is a second-order differential operator and the operator is a scalar operator. So to operate on a scalar quantity, it becomes $\nabla \cdot \nabla$(a scalar quantity).

For a Scalar function F, $\nabla^2 f = \nabla \cdot (\nabla f)$ but $\nabla f =$ gradient of scalar function of f. And the dot product in $\nabla \cdot (\nabla f)$ stands for the divergence of the gradient of f.

$\nabla^2 f = \nabla. \cdot (\nabla f) =$ Divergence of gradient of the scalar function. The ∇^2 operator is not restricted to the scalar function only, it can operate also on vectors as $\nabla^2 \bar{F} = \nabla \cdot \nabla \bar{F}$ but this will not have meaning as we have not defined what $\nabla \bar{F}$ is.

The DEL operator ∇ can operate as either divergence or a curl. So in this case, we take ∇^2 and operate it directly on the vector quantity so that:

$$\nabla^2 \bar{F} = \frac{\partial^2}{\partial x^2}(\bar{F}) + \frac{\partial^2}{\partial y^2}(\bar{F}) + \frac{\partial^2}{\partial z^2}(\bar{F}) \tag{5.16}$$

That is we take the second order differential of all the components that make up $\bar{F}$ i.e

$$\bar{F} = F_x\hat{x} + F_y\hat{y} + F_z\hat{z}$$

So that

$$\nabla^2 \bar{F} = \frac{\partial^2}{\partial x^2}(F_x\hat{x} + F_y\hat{y} + F_z\hat{z}) + \frac{\partial^2}{\partial y^2}(F_x\hat{x} + F_y\hat{y} + F_z\hat{z}) + \frac{\partial^2}{\partial z^2}(F_x\hat{x} + F_y\hat{y} + F_z\hat{z}) \tag{5.17}$$

Then we combine individual components to have

$$\nabla^2 F = (\frac{\partial^2 Fx}{\partial x^2} + \frac{\partial^2 Fx}{\partial y^2} + \frac{\partial^2 Fx}{\partial z^2})\hat{x} + (\frac{\partial^2 Fy}{\partial x^2} + \frac{\partial^2 Fy}{\partial y^2} + \frac{\partial^2 Fy}{\partial z^2})\hat{y} + (\frac{\partial^2 Fz}{\partial x^2} + \frac{\partial^2 Fz}{\partial y^2} + \frac{\partial^2 Fz}{\partial z^2})\hat{z}$$

$$\nabla^2 F_x\hat{x} + \nabla^2 F_y\hat{y} + \nabla^2 F_z\hat{z} = \nabla^2 \bar{F} \tag{5.18}$$

Later we shall see that this operation will be required when solving the problem of electro-magnetics.

5.5.3 Integral Operators

We have been dealing with differential operators up to this point now we have to deal with integral operators which have to do with the integration of the vector fields. If we have a vector field, there is a possibility that we can take the integral of this vector field in a plane or along the contour or along a path. It is possible that we can take the integration of this vector on a surface which can be an open or closed surface. Or we can take the integral over a volume.

When we do the integration along a path, it is called **Contour or Line Integration**, while integration over a surface we call **Surface Integration**. And since a surface is two-dimensional in nature, we have a double integral.

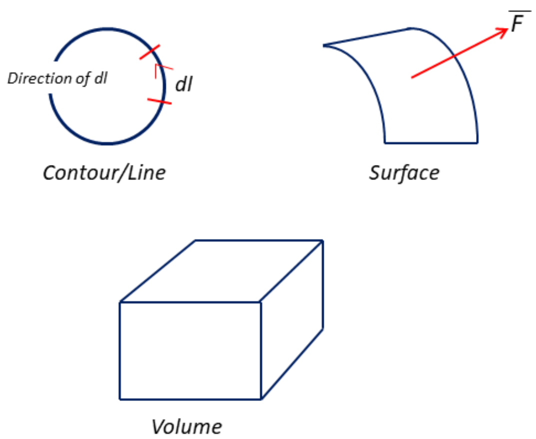


Figure 5.9: line surface and volume integration

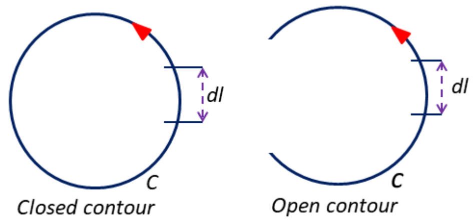


Figure 5.10: Open and closed contour

So the contour integration is a single integration, for a volume integration, it will be a triple integration i.e. in the three-dimensional space.

We have important theorems which connect line integrals to surface integrals and surface integrals to volume integrals. We may need to change from one type of integral to another integral type e.g. from line integral to surface integral or from surface integral to volume integral.

It should be kept in mind however that integral is a Scalar quantity. Thus the final answer after integration will always be a Scalar quantity. We can define line, surface and volume integration if we take some vector $\bar{A}$ i.e. $\bar{A}$ is a vector field.

5.5.4 Line Integral

The line integration around some path C is $\int \bar{A}\cdot d\bar{l}$. If we integrate along the contour in the direction shown, dl has both magnitude and direction. If the path is a closed path, or mathematically we say if the contour is a closed contour, we have a closed line integral = $\oint \bar{A}\cdot d\bar{l}$.

$$\int \bar{A}\cdot d\bar{l} = open\ contour \quad while, \oint \bar{A}\cdot d\bar{l} = closed\ contour \tag{5.19}$$

So at every location, we find out the product $\bar{A}\cdot d\bar{l}$ and sum it all up. This is the basis for the line integral.

5.5.5 Surface Integral for Infinitesimally small Area

We can carry out a similar operation for the surface integral as having an area A with an infinitesimally small area $d\bar{A}$. $d\bar{A}$ has a direction perpendicular to that small area, if the vector field in this area is $\bar{A}$, we can define the surface integration as: (the surface may be closed or open, here it is an open surface).

Surface Integral =

$$\iint \bar{A}\cdot d\bar{A} = \text{open surface}$$

$$\oiint \bar{A}\cdot d\bar{A} = \text{closed surface} \tag{5.20}$$

$\bar{A}\cdot d\bar{A}$ is a scalar quantity so that the integration is also a scalar quantity. Now the direction for $d\bar{A}$ is thus; if it is a closed

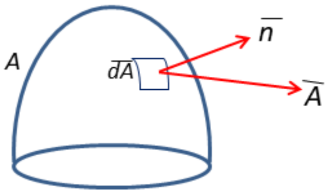


Figure 5.11: Surface Integral

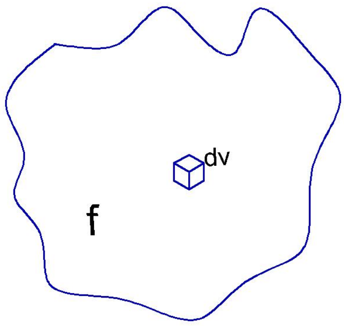


Figure 5.12: volume integral of scalar f

surface, the direction is taken as the outward normal for the volume created by the closed surface and it is in the positive normal direction.

If it is an Open Surface there is no preference for defining the $unit\ normal$. We can define the $unit\ normal$ in either direction.

n_1 and n_2 are valid unit normal for Surface **S**.

However, later on when we try to connect the contours to the surface for an open surface, then at that time we follow the convention of the right-hand rule. That is for any open surface to know the positive unit normal, just let the curl of the finger go counter-clockwise around the surface, and the direction to the thumb points in the positive direction of the unit normal to that surface.

5.5.6 Volume Integral

The third is the volume integral, if we have some function which is a Scalar f:

$$\text{Volume integral} = \iiint f dV$$

Now when we try to relate the different integrals, ie, the line integral to the surface integral and surface integral to a volume integral, essentially the vector fields are operated on by the DEL ∇ operators, then there is a relationship between these del operated fields in the integration domain. So two important theorems essentially relate this i.e. they are called the **Divergence theorem** and the **Stokes theorem**.

5.5.7 Divergence Theorem

This converts a surface integral for a vector field to a volume integral. If we have a vector field $\bar{A}$, then the divergence theorem

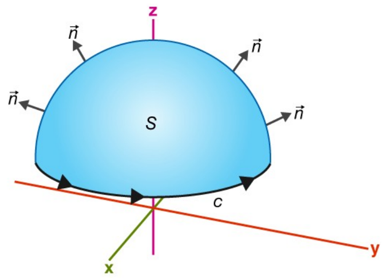


Figure 5.13: Stokes theorem

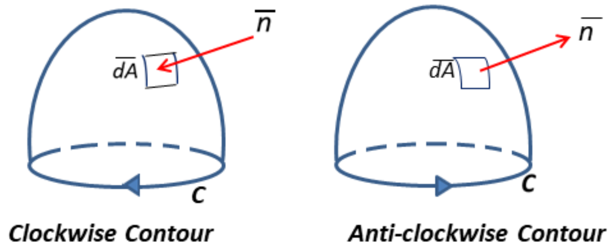


Figure 5.14: Contour Integration

states that for a **closed surface**, we have that:

$$\oiint \bar{A} \cdot \bar{da} = \iiint (\nabla \cdot \bar{A}) dV \tag{5.21}$$

So if we have a volume bounded by surface S, if we have a vector $\bar{A}$ and we know the surface integral of this vector i.e $\oiint \bar{A} \cdot \bar{da}$, it can be converted into a volume integral.

So $\oiint \bar{A} \cdot \bar{da} = \iiint (\nabla \cdot \bar{A}) dV$ is called the divergence theorem. So when we have to convert between closed surface integral and volume integral, this theorem comes in very handy.

5.5.8 Stokes Theorem

It converts a closed line integral to an open surface integral or vice-versa.

The surface shown is open and we have the boundary defined by contour C. Now we can have a small area with unit vector $\hat{n}$.

Then we can have the contour C for which line integral was defined for vector $\bar{A}$. So *how do we define the path C with respect to $\hat{n}$?*. The convention is to follow the right-hand rule. If our contour as shown is going anticlockwise, then $\hat{n}$ goes outward, if it goes clockwise $\hat{n}$ goes inward.

Note that the bottom of the surface is open as compared to when this was not the case in the divergence theorem example. For these open surface S bounded by closed contour C the Stokes theorem states that:

$$\oint_c \bar{A} \cdot \bar{dl} = \iint_s (\nabla \times \bar{A}) \cdot \bar{da} \tag{5.22}$$

So this theorem can be used to convert a line integral to a surface integral or vice-versa.

5.5.9 Summary

To summarize all these again, if we have a closed surface, the closed surface integral and the integral of the volume enclosed by that closed surface can be related by the divergence theorem.

$$\oiint \bar{A} \cdot d\bar{A} = \iiint_v (\nabla \cdot \bar{A}) dV \tag{5.23}$$

For an open surface whose boundary at the open end is defined by a closed line path or contour, the line integral of that closed contour and the surface integral in the open surface are related by *Stokes Theorem*.

$$\oint \bar{A} \cdot d\bar{l} = \iint_s (\nabla \times \bar{A}) \cdot d\bar{a} \tag{5.24}$$

Later on for solving the problem of electro-magnetics, when we write *Maxwell's Equation* in integral and differential forms, the two theorems become very important and become handy in converting from integral to differential forms. Having understood the basics of vectors, now we can go to the basic quantities of electro-magnetics which we are going to make use of in further analysis, that is the quantity like electric and magnetic fields.

The origin of the electromagnetic phenomenon is based on the basic *CHARGE*. So if we consider a charge, we know in the form of *Coulombs* law that the charge has an effect all around it. If you put another charge in the vicinity of that charge, it experiences a force. This quantity is measured by an effect called the *electric field*.

So if we take a charge and go in the vicinity of that charge, we experience a force which is characterized by a quantity called the electric field. However, if we put this charge in motion, then it constitutes a current as the current is nothing but the movement of charge or rate of change of charge.

So if the charge starts moving, that means you have a sustained flow of charges that gives you current and then you have magnetic fields. So some charge when it is steady or stationary, it gives you an electric field, when it starts moving, it gives you current and that gives you a magnetic field.

The charge can accelerate also, then if we accelerate the charges, then it gives you an electric and magnetic field. So essentially we are dealing with these quantities here; the charges, current, electric and magnetic fields and try to establish the relationship between these quantities. The relationship we have between these quantities is given by what is called *Maxwell's Equations*. So essentially now starting with the basics of these quantities, we would try to establish a relationship between these laws of physics in mathematical form using vector notation and vector theorems which we establish, we get Maxwell's equation. So the quantity which we have now as the first quantity is electric field $\bar{E}$ which is nothing but the force per unit charge. The electric field is a vector quantity. It has both magnitude and direction. It has V/m as its unit. Then we have a media property. Let's say we measure the electric field in a vacuum, we experience a certain force on a unit charge. If we measure the same thing and change the medium parameter to some other dielectric, then the force measured will change. So the quantity which does not depend on any parameter is the electric displacement vector $\bar{D}$. So electric field is a quantity which is related to the charge producing the field and is also affected by the medium parameter which is called the permittivity of the medium. So we have a medium parameter called

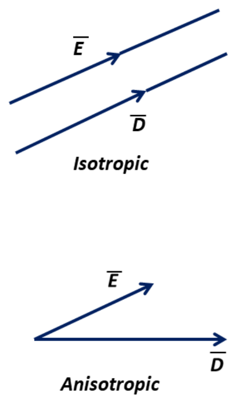


Figure 5.15: Isotropic and Anisotropic Medium

permittivity ε and has unit Farad/m. If the medium parameter changes, the permittivity changes and the electric field $\bar{E}$ measured at that point changes. Permittivity of vacuum or free space is denoted by $\varepsilon_0 = \frac{1}{36\pi} \times 10^{-9} F/m$. Sometimes written as $8.86 \times 10^{-12} F/m$. If we take another medium whose permittivity is ε, the ratio $\frac{\varepsilon}{\varepsilon_0}$ is what is called the dielectric constant for the material or relative permittivity. Dielectric constant $\varepsilon_r \equiv$ relative permittivity= $\frac{\varepsilon}{\varepsilon_0}$. The quantity which is independent of the medium parameter is the displacement vector $\bar{D} = \varepsilon \bar{E}$

For a given property change, the $\bar{E}$ changes, $\bar{D}$ remains the same, as it only depends upon the charge which is creating the field. The quantity ε in a general medium can be constant everywhere or it can vary as a function of space in three-dimensional space. It can also depend on the direction, which means if we measure permittivity in the x direction, its value may vary from that in the y and z direction. So in general, the quantity ε can be direction dependent, or it may be space dependent. When ε varies as a function of space, we call the medium **INHOMOGENOUS MEDIUM**. If ε is direction dependent, then the medium is called an ANISOTROPIC MEDIUM. If ε does not vary as a function of space, we call it **HOMOGENOUS MEDIUM**. If it does not vary with direction, we call that medium **ISOTROPIC MEDIUM**. However, we will be dealing with media that are homogenous and isotropic. This means that the dielectric constant of the medium, which we shall consider, has ε that is constant in space and does not change with direction. However, if ε was anisotropic, then ε becomes a 3×3 matrix of ε multiplied by vector $\bar{E}$. For isotropic medium ε is a scalar quantity. So for this course, we deal with the medium where ε is a scalar quantity. So in a homogeneous and isotropic medium, $\bar{D}$ is a scalar version of $\bar{E}$. Comparing $\bar{D}$ and $\bar{E}$, they have different magnitudes but with the same direction. However, if the medium were to be anisotropic, ε is a 3×3 matrix and it rotates the vector $\bar{E}$ and in general $\bar{D}$ and $\bar{E}$ are not in the same direction.

Now we have another quantity which is very useful that we have to define, the *Electric Potential* of the field. The electric field is related to the electric potential V (scalar) by $\bar{E}$= -∇ V. So if we know the potential at a point, we can take the gradient to get the electric field at that location. From here we get the unit of the electric field. So if we find the potential, the unit of V is Volts. The del operator is a differential spatial operator $\frac{d}{dx}$;

i.e $\frac{1}{l} or \frac{1}{m}$ that gives unit of electric field, which is V/m. So this relation of finding out the electric field from the potential comes in very handy whenever we want to find the electric field in a general complex distribution of charges. If we find the electric field $\bar{E}$ at a particular location and find the component of the electric field for each of the charges which are distributed in space, we carry out vector additions at that point for the electric fields. It is however easier to find the potential of the different charges and add them together. The negative gradient of that gives you the electric field. So these are the basic quantities for representing the electrostatic parameters; the electric field $\bar{E}$ and the electric potential V.

5.6 EXERCISES

Ex. 114 — How is the divergence of a vector field expressed mathematically?

Ex. 115 — How is the curl of a vector field expressed mathematically?

Ex. 116 — What is the Laplace operator? Also, what is its mathematical expression?

Ex. 117 — Integral operators have to do with the integration of the vector field. True/False

Ex. 118 — Explain the geometric interpretation of a surface integral and how it relates to the flux of a vector field through a surface.

Ex. 119 — State and explain the divergence theorem and its relationship to surface integrals.

Ex. 120 — Compare and contrast the physical interpretations of line integrals, surface integrals, and volume integrals. How do these integrals capture different aspects of a mathematical model, and in what scenarios would one choose volume integrals over the other types?

Ex. 121 — What is the Divergence Theorem, and how does it establish a relationship between a volume integral and a surface integral involving a vector field?

Ex. 122 — Explain how the Divergence Theorem is used in electromagnetics, specifically in the context of electric flux and Gauss's Law.

Ex. 123 — Discuss practical engineering scenarios where the Divergence Theorem is applied. How does it simplify problem-solving in fluid mechanics or heat transfer, for example?

Ex. 124 — Explain the statement and significance of Stokes' Theorem. How does it relate the circulation of a vector field around a closed curve to the behavior of the field on the surface enclosed by the curve?
Use the Divergence Theorem to find the outward flux of **F** across the boundary of the region.

Ex. 125 — What is the electric potential energy of two electrons separated by 5.0 mm? If the separation increases, does the potential energy decrease or increase?

Ex. 126 — The electric field inside a non-conducting sphere of radius R with charge spread uniformly throughout its volume is radially directed and has magnitude:

$$E(r) = \frac{q}{4\pi\varepsilon_0 R^3}$$

Taking $V = 0$ at the center of the sphere, find the electric potential V inside the sphere.

Ex. 127 — Calculate the average magnitude of the electric field between two points with a potential difference of 150 V between them and are separated by a distance of 2.5 cm.

Ex. 128 — An electric field is expressed as $\mathbf{E} = 2\mathbf{j} + 3\mathbf{j}$. Find the potential difference $(V_A - V_B)$ between two points A and B whose position vectors are given by $\mathbf{r}_A = \mathbf{i} + 2\mathbf{j}$ and $\mathbf{r}_B = 2\mathbf{i} + \mathbf{j} + 3\mathbf{k}$.

Ex. 129 — For an isotropic radiator, the electric field intensity at a distance R is measured as 3 V/m. What will be the electric field intensity at a distance $3R$?

Chapter 6

The Basic Laws To Maxwell's Equation (1)

6.1 OBJECTIVES

By the end of this chapter, the you should be able to:

(i) Derive Maxwell's equations.

(ii) Explain the fundamental laws behind Maxwell's equations.

(iii) Solve problems involving Maxwell's equations.

(iv) Should be able apply Maxwell's equation in real world scenario.

6.2 INTRODUCTION

It was established in the previous chapter that, the origin of the electric and magnetic field is the charge. The effect of this charge is measured by a quantity called the *electric field*. However, when the same charge is kept in motion, it constitutes a current and the presence of this current is felt by a quantity called *magnetic field* . So we can take current as the origin of the magnetic field.

In this chapter, we will be looking at some laws like ohms law and other laws of electromagnetism and see how we can formulate these laws in a mathematical form called [1] Maxwell's equations. And also see the relationship between current and magnetic field, magnetic flux density and magnetic field.

Now let's consider a current carrying element which is a small piece of wire carrying current(I) with length given by $d\bar{l}$ having direction shown by the arrow in fig 5.1. At some point in space, P is given by the distance r from the current element.

The magnetic field due to the current element $Id\bar{l}$ is given by the expression

$$\boxed{d\bar{H} = \frac{Id\bar{l} \times \hat{\mathbf{r}}}{4\pi r^2}} \quad (Amp/m) \tag{6.1}$$

This is the *Magneto-motive force (MMF) at that point P*magneto-motive force (mmf). Analysis From Equation 5.1

(i) The expression shows that the magnetic field is proportional to the current I and the length of the wire through which the current flows but inversely proportional to the square of the distance r from the current element (wire).

(ii) By applying the right-hand rule, the direction of the magnetic field ($d\bar{H}$) points in the clockwise direction looking in the direction of the current.

[1]James Clerk Maxwell, a Scottish mathematician and physicist. He was born on June 13, 1831, in Scotland and died on Nov 5, 1879, in Edinburgh. He studied electromagnetic radiation, electricity, light and gas. He is known for his formulation of electromagnetic theory and he's regarded by most modern physicists as the scientist of the 19th century who has the greatest influence on 20th-century physics, for the nature of his contribution.

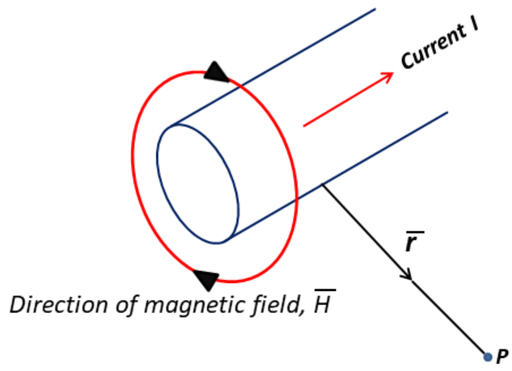


Figure 6.1: Current carrying element(wire)

The quantity related to the medium parameters is the magnetic flux density ($\bar{B}$) and not the magnetic field ($\bar{H}$). So no matter the medium, $d\bar{H}$ depends on I, $d\bar{l}$ and r which are not the medium parameters.

6.3 Magnetic Flux Density $\bar{B}$

This is the product of the permeability of the medium and the magnetic field strength.
Mathematically,

$$\boxed{\bar{B} = \mu\bar{H}} \quad (Wb/m^2)or(Tesla) \tag{6.2}$$

$$\mu = \mu_r \cdot \mu_0$$

where:

μ : Magnetic permeability of the material

μ_r : Relative permeability of the material

μ_0 : Permeability of free space (approximately $4\pi \times 10^{-7}$ H/m)

This relationship allows you to relate the magnetic properties of a material to those of free space. In a vacuum, μ_r is equal to 1, and μ is equal to μ_0. For materials with μ_r greater than 1, the material enhances the magnetic field, and for materials with μ_r less than 1, the material reduces the magnetic field.

Depending upon the magnetic properties of the material, the value of permeability changes. So the magnetic flux at point, P, is related to the permeability of the medium and the magnetic field strength. The unit of magnetic flux density is in Tesla or $Weber/m^2$. Analysis From Equation 5.2.

(i) The magnetic flux density $\bar{B}$ tells us the density of magnetic lines of forces at that particular location. Meaning it is essentially showing you some kind of vectors which are oriented and hence the reason it is a vector quantity.

(ii) As we define the electric field as the *force per unit charge due to presence of charges*, we can also define the magnetic field as the *force experienced by a unit magnetic pole placed in the vicinity of the current*, making $\bar{B}$ and $\bar{H}$ vector quantities.

(iii) In the case of the magnetic field, if the medium is [2]*Isotropic*, $\bar{B}$ and $\bar{H}$ will be in the same direction. Whereas if it is an [3]*Anisotropic* medium $\bar{B}$ and $\bar{H}$ will not be in the same direction. The next one we are looking at is Ohm's law.

6.4 Ohm's Law

We know that [4]Ohms law relating to electric field, is the relationship between voltage and current flowing in a medium or in a circuit. Ohm's law states that the current(I) through a conductor between two points is directly proportional to the voltage(v) across the two points. However, the general form of Ohm's law is the relationship between a quantity called *the conduction current density $\bar{J}$ and electric field $\bar{E}$* .

So if we take a medium where the conductivity of the medium is not constant, then it is not meaningful to define the total current for the medium. Normally what we do is, define a quantity called the conduction current density $\bar{J}$ (i.e. *current flowing per unit area in this conductor*). It has direction which is the direction the current is flowing.

This current is flowing due to an electric field because it has some electric potential. We have seen that the electric field is related to the gradient of potential difference.

So if we go from circuital Ohm's law, we'll have a relationship between voltage and current which have a proportional relationship and also a proportionality constant called *Resistance*(R).

Now let's establish a relationship between the conduction current density $\bar{J}$ and electric field $\bar{E}$.
Mathematically,

$$I \quad \alpha \quad V \quad \Rightarrow I = \frac{V}{R}$$

Making the 'R' subject of the formula, we have that;

$$R = \frac{V}{I} \tag{6.3}$$

Where R = The introduced constant of proportionality called the *resistance*.

For a conductor of length ℓ and the cross-sectional area 'A' The resistance is directly proportional to the length and inversely proportional to the cross-sectional area of the conductor.

[2]An isotropic medium is a medium that has uniform permeability in all directions of the medium

[3]Anisotropic medium is one such that the permeability and permittivity of the medium are not uniform

[4]Ohm's law was formulated by a German physicist and mathematician, George Simon Ohm (16 March 1789 - 6 July 1854) who began as a school teacher in Brandenburg-Bayreuth, Bavaria in Germany. He attended the University of Erlangen and won Copley Medal in 1841 for his research and is known for Ohm's law, Ohm's phase law, and Ohm's acoustic law. He began his scientific career in Physics(Electricity) at the University of Munich where Karl Christian Von Langsdorf was his Doctoral advisor.

Mathematically,

$$R\alpha\frac{\ell}{A} \quad R = \frac{\rho\ell}{A}$$

Substitute equation 5.3 into the above formula. We have that;

$$\frac{V}{I} = \frac{\rho\ell}{A}$$

Rearranging the equation we have that;

$$\frac{V}{\ell} = \frac{I\rho}{A} \tag{6.4}$$

Recall that $\vec{E} = \frac{V}{Ł}$ into 5.4, we have that;

$$\vec{E} = \frac{I\rho}{A}$$

But the ratio of current to the area is given by a quantity called conduction current density J (i.e $\frac{I}{A}$), Hence

$$\vec{E} = \vec{J}\rho \quad \Rightarrow \vec{J} = \frac{1}{\rho}\vec{E}$$

Where the conductivity σ is the the inverse of resistivity and verse versa ρ (i.e $\frac{1}{\rho}$),
Hence, $\sigma = \frac{1}{\rho}$

$\boxed{\vec{J} = \sigma\vec{E}}$ (A/m^2)

This is the *Conduction current density - The general form of Ohm's law*. Mathematically, it is the product of electric field intensity $\vec{E}$ and σ from a reference point where σ is the conductivity of the medium(proportionality constant).

The relative permeability of the medium μ_r and the conductivity of the medium σ can be represented as 3×3 matrices.

In the case of an isotropic medium, both μ_r and σ are scalar quantities. This implies that their values remain the same, regardless of the direction.

On the other hand, for an anisotropic medium, where properties vary with direction, μ_r becomes a 3×3 matrix, and σ also transforms into a 3×3 matrix.

This situation parallels the behavior observed in the electric field, where in an isotropic medium, the displacement vector and electric field align in the same direction. In contrast, for an anisotropic medium, the direction of the electric field may not coincide with that of the displacement vector.

Similarly, in the realm of the magnetic field, if the medium is isotropic, the magnetic flux density ($\bar{B}$) and conduction current density ($\bar{J}$) align in the same direction. Conversely, in an anisotropic medium, the direction of electric field ($\bar{E}$) and the conduction current density ($\bar{J}$) are not the same.

6.5 The Physical Laws To Maxwell's Equation

Now from this basic introduction of the parameters, we are set to establish the physical laws mathematically, which when compiled gives Maxwell's equation and these physical laws are the experimental verification of the postulates. There are two ways to get Maxwell's equation

(i) You treat it as a mathematical postulate.

(ii) Try to represent these mathematical laws in appropriate mathematical form to give us Maxwell's equation.

Both ways have merits, that is, if we say Maxwell's equations are mathematical postulates, then they are very exact. However one could ask a question, *How do you get these postulates?* One cannot simply come up with these mathematical formulae, without having a background. So the background is in the physical laws. First, the physical laws came from curiosity towards the relationship between electrical and magnetic fields. So if we go by the picture that Maxwell's equations were guided by the experimented laws, then the advantage is that we always try to see the phenomenon in physical terms. So if we have any electromagnetic phenomenon, and accept that the origin lies in the physical laws, it will always be useful to look at every phenomenon of electromagnetics physically.

However, if you treat these equations more like mathematical postulates then one may get lost in the mathematical manipulation of these equations. So from the exactness point of view, the mathematical postulates have merit, but seeing what is happening physically, first understanding the physical laws and then getting Maxwell's equations.

What we will be doing in this chapter is that would state first the physical laws and then use the vector identities and theorems such as divergence and Stokes theorems, we would try to get the mathematical form and then finally we get Maxwell's equations.

6.6 Gauss Law

This law states that the total displacement(electric flux) coming out of a closed surface is equal to the net charge enclosed by the closed surface.

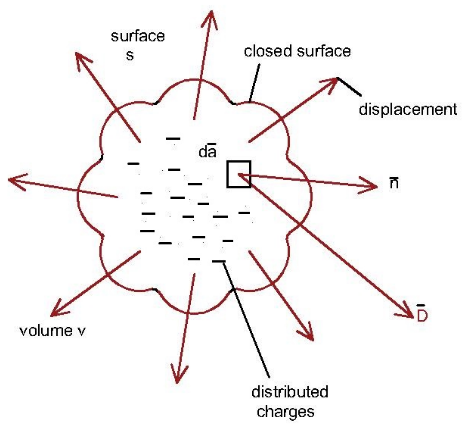


Figure 6.2: A closed surface showing the direction of displacement vector

Total displacement shown by each arrow = net charge enclosed by surface

Mathematically,

$$\Phi_E = Q_{\text{enclosed}} \tag{6.5}$$

$$\boxed{\oiint_s \bar{D}\cdot d\bar{a} = \iiint_v \rho dv} \tag{6.6}$$

This is *Gauss law in integral form*.Where $\oiint_s \bar{D}\cdot d\bar{a}$ is the total outward displacement from the total surface. This means using Gauss' law[5] to find out the total outward displacement from the total surface from the total charge integrated over the volume.

In general, the displacement vector will be varying as a function of the location on the surface, also it will be oriented in different directions. However, the net displacement which is coming out of the surface will be a component of each arrow normal to the surface at that location.

Hence considering a small incremental area ($d\bar{a}$) on the surface as shown in figure 5.3. ($d\bar{a}$) has a direction defined by a unit normal $\hat{n}$ with the direction of the displacement vector given by $\bar{D}$. The outward displacement coming out from the area is the dot product of $\bar{D}$ and $d\bar{a}$. If that is gotten, add all the contributions from others around the surface, and then we have the total displacement coming out of the close surface. Let us say in general, we have charges which are distributed inside this surface and are characterized by a charge density. This is denoted by volume charge density, ρ (which is representing the density of charges inside the volume). Its S.I. unit of measurement is in coulomb per cubic metres, (C/m^3).

In other words, if we add up all the charges inside the volume, we get the charge enclosed by this surface. So integrating the displacement vector over the surface, we get the total outward displacement from this surface. Applying[6] divergence theorem to $\iint_s \bar{D}\cdot d\bar{a}$ (Meaning converting surface integral to line volume integral), we will have

$$\oiint_s \bar{D}\cdot d\bar{a} = \iiint_v (\nabla\cdot\bar{D})dv$$

Subtitute into equ 5.6

$$\iiint_v (\nabla\cdot\bar{D})dv = \iiint_v \rho dv$$

$$\iiint_v (\nabla\cdot\bar{D})dv = \iiint_v \rho dv$$

Collect like terms and factorize out

$$\iiint_v (\nabla\cdot\bar{D})dv - \iiint_v \rho dv = 0$$

$$\iiint_v (\nabla\cdot\bar{D})dv - \iiint_v \rho dv = 0$$

$$\iiint_v (\nabla\cdot\bar{D} - \rho)dv = 0$$

This relationship in the integral can only be true for all arbitrary volume if and only if $\nabla\cdot\bar{D} - \rho = 0$

$$\nabla\cdot\bar{D} - \rho = 0$$

$$\boxed{\nabla\cdot\bar{D} = \rho} \tag{6.7}$$

Where $\bar{D}$ is the displacement vector and ρ is the volume charge density. Analysis From Equation 5.7

[5]Gauss law was formulated by a German mathematician, Carl Friedrich Gauss (30 April 1777 - 23 February 1855)who developed the divergence theorem. He was the first physicist to measure electric and magnetic quantities in absolute units. He began his scientific research at the University of Gottingen on mathematics and physics as his field where Johann Friedrich Ptaff and Johann Christian Martins were his doctoral advisor. He won the Lalande Prize in 1810 and Copley Medal in 1838. some of his own doctoral students include Carl Wolfgang Benjamin Goldschmidt (A co-author with J.C Eduard on Analytical Optics), and Johann Benedict Listing(He discovered properties of the half-twisted strip in 1858, and is known also by Listing's law in Ophthalmology).

[6]$\iint_A A\cdot d\bar{a} = \iiint(\nabla\cdot\bar{A})dv$. This is the divergence law that was formulated by Gauss.

(i) From the Gauss law in differential form, we saw that the divergence of the displacement vector at any point is related to the charge density at that point. This relationship is called a point relation.

(ii) The differential form of Gauss law has a limitation as it assumes that, we have a [7]continuous media, meaning it will not hold for a [8]discontinuous media (i.e. the derivative at that point does not exist).

(iii) While the integral form of Gauss law is applicable in all situations.

In conclusion, if the medium is continuous, we apply the differential form. But if not, we apply integral form.

Considering the relationship between the displacement vector and the electric field, we have

$$\boxed{\bar{D} = \epsilon \bar{E}} \tag{6.8}$$

Substituting equation 5.8 into 5.7, we have

$$\nabla \cdot (\epsilon \bar{E}) = \rho$$

If we assume that the medium is homogeneous (i.e. ϵ is not a function of space). Then

$$\boxed{\nabla \cdot \bar{E} = \frac{\rho}{\epsilon}} \tag{6.9}$$

Gauss law for electric charges for homogeneous medium For a closed surface the net magnetic charge is always zero. If we apply Gauss law for the magnetic charges or poles, then there is no net charge enclosed by the closed surface. i.e

$$Q_{\text{enclosed}} = \iiint_v \rho dv = 0$$

Hence, the net charge coming out of a closed surface for a magnetic field is always equal to zero, because there are no net magnetic charges enclosed by the surface. i.e

$$\phi = \iint_s \bar{B} \cdot d\bar{a} = 0$$

Applying divergence theorem to $\iint_s \bar{B} \cdot d\bar{a}$ we have

$$\iiint_v (\nabla \cdot \bar{B}) dv = 0$$

For this to be true for any arbitrary volume $(\nabla \cdot \bar{B})$ must be equal to zero.

$$\boxed{\nabla \cdot \bar{B} = 0} \tag{6.10}$$

Gauss law for magnetic charges (magneto-statics)

6.7 Amperes Law

[9]Amperes law is also called *Amperes circuit law* and it states that the magneto-motive force(mmf) around a closed loop is equal to the total current enclosed by the loop.
Mathematically,
Magnetomotive Force (mmf) = $\oint_c \bar{H} \cdot d\bar{l}$

[7]Continuous medium refers to a material or substance that exhibits uniform properties throughout its volume.

[8]Discontinuous medium refers to a material or substance that exhibits variations in its properties.

[9]Ampere's law was discovered by Andre Marie Ampere in 1823

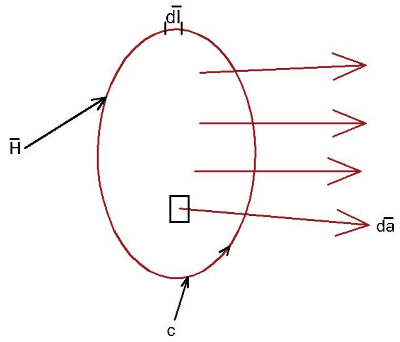


Figure 6.3: A loop of arbitrary size of non-uniform current

Considering a closed loop shown in Figure 6.3, we can calculate the mmf since it is the line integral of the magnetic field along the path created by the closed loop. i.e

$$\Sigma I_{enclosed} = \oint_c \bar{H} \cdot d\bar{l} = mmf$$

Let the total current enclosed by the loop be I.
Therefore $\Sigma I_{enclosed} = I$

$$\Rightarrow I = \oint_c \bar{H} \cdot d\bar{l}$$

Due to these changes, the varying current can be written as current density. The integration of this current density $\bar{J}$ over an enclosed area created by the closed path c? gives us the total current enclosed by contour c.
Mathematically,

$$I = \iint_A \bar{J} \cdot d\bar{a}$$

$$\boxed{\oint_c \bar{H} \cdot d\bar{l} = \iint \bar{J} \cdot d\bar{a}} \tag{6.11}$$

This is Amperes Law in *Integral Form*. Now to get the differential form of Amperes law, we will apply *stroke's theorem* (i.e. converting from line integral to surface integral).

$$\iint_A (\nabla \times \bar{H}) \cdot d\bar{a} = \iint_A \bar{J} \cdot d\bar{a}$$

$$\iint_A (\nabla \times \bar{H}) \cdot d\bar{a} - \iint_A \bar{J} \cdot d\bar{a} = 0$$

$$\iint_A (\nabla \times \bar{H} - \bar{J}) d\bar{a} = 0$$

For this to be true for all arbitrary area, $\nabla \times \bar{H} - \bar{J}$ must be equal to zero

$$\Rightarrow \nabla \times \bar{H} - \bar{J} = 0$$

$$\boxed{\nabla \times \bar{H} = \bar{J}} \tag{6.12}$$

This is Amperes Law in *Differential Form*. Analysis From Equation 5.12

(i) What this means is, the curl of $\bar{H}$ is equal to the total conduction current density at that point. Again the differential form of Amperes law is only valid for a continuous medium(meaning if we go to a point and the magnetic field at that point is known(*Point Relationship*), then we can apply the differential form of Amperes law). While

(ii) The integral form of Amperes law is applicable to the finite area enclosed by path c.

(iii) The above Amperes law in differential form does not show the complete equation for Maxwell's fourth equation. Hence, it will be looked further into in the next chapter.

6.8 Faraday's Law of Electromagnetic Induction

This law states that the total Emf around a closed loop is equal to the rate of change of magnetic flux enclosed by that loop, and that, the direction of emf is such that, the magnetic field produced by this current due to this emf opposes the original magnetic field. In other words, EMF is the electromotive force induced in the circuit (measured in volts, V). the rate of change of magnetic flux through the coil with respect to time (measured in Weber per second or volt). EMF $= -\frac{d\Phi_B}{dt}$

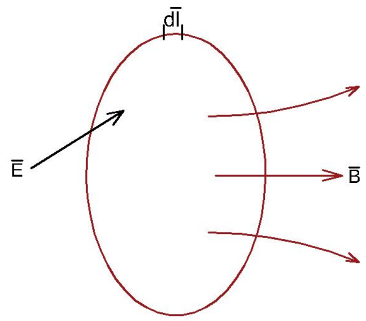


Figure 6.4: A closed loop used for calculating $\bar{B}$

Considering the figure 5.4, the total flux is found by integrating $\bar{B}$ over the area enclosed by that loop. Hence, the rate of change of this flux density is equal to the total emf produced around the loop.

Mathematically,

$$\begin{aligned}\text{EMF} &= \oint_c \bar{E}\cdot d\bar{l} \\ &= -\frac{\delta}{\delta t}(\phi)\end{aligned} \tag{6.13}$$

Where ϕ = Total flux enclosed by the loop.

$$\boxed{\oint_c \bar{E}\cdot d\bar{l} = -\frac{\delta}{\delta t}(\phi)} \tag{6.14}$$

This is the EMF produced by $d\bar{l}$. Analysis From Equation 5.13

(i) The negative sign shows that the emf produced is such that, the magnetic field produced by this emf will oppose the original magnetic field.

(ii) The magnetic flux can be gotten by integrating the flux density over the area created by the loop to give us the total flux(ϕ) mathematically,

$$\phi = \iint_A \bar{B}\cdot d\bar{a} \tag{6.15}$$

Where $d\bar{a}$ = Incremental Area and $\bar{B}$ = Magnetic flux density.

Substituting equation 5.15 into 5.14

$$\boxed{Emf = \oint_c \bar{E}\cdot d\bar{l} = -\frac{\delta}{\delta t}(\iint_A \bar{B}\cdot d\bar{a})}$$

This is a general relationship relating to the rate of change of flux

$$\boxed{\oint_c \bar{E}\cdot d\bar{l} = -\frac{\delta}{\delta t}(\iint_A \bar{B}\cdot d\bar{a})} \tag{6.16}$$

Now let's analyze the general relationship relating to the rate of change of flux.

The rate of change of flux can either happen in two ways.

(i) By varying the area of the loop with time while keeping the magnetic flux density $\bar{B}$ constant. This is called the *generator action* (this is how the generator works, where the magnetic flux density is kept constant and by rotating the loop in a magnetic field, the area of the loop changes and we get induced emf).

(ii) By changing the magnetic flux density $\bar{B}$ with time while the area of the loop remains constant. This is called *transformation action* (This is what happens in a transformer where the size of the coils are not varying as the function of time, but the magnetic flux density induced in this coil varies with time which makes us have induced emf)

Now from the integral form of Faraday's law, let's get its differential form. Hence, we are interested in the time-varying fields (i.e. $\bar{A}$ and $\bar{B}$ fields) and not space-varying fields (loop area) which makes the area not a function of time.

$$\oint_c \bar{E}\cdot d\bar{l} = -\frac{\delta}{\delta t}(\iint_A \bar{B}\cdot d\bar{a})$$

Applying [10]Stokes Theorem to $\oint_c \bar{E}\cdot d\bar{l}$, we have;

$$\oint_c \bar{E}\cdot d\bar{l} = \iint_A (\nabla\times\bar{E})\cdot d\bar{a}$$

$$\iint_A (\nabla\times\bar{E})\cdot d\bar{a} = -\iint_A \frac{\delta\bar{B}}{\delta t}\cdot d\bar{a}$$

$$\iint_A (\nabla\times\bar{E})\cdot d\bar{a} = -\iint_A \frac{\delta\bar{B}}{\delta t}\cdot d\bar{a}$$

$$\iint_A (\nabla\times\bar{E})\cdot d\bar{a} + \iint_A \frac{\delta\bar{B}}{\delta t}\cdot d\bar{a} = 0$$

$$\iint_A \{(\nabla\times\bar{E}) + \frac{\delta\bar{B}}{\delta t}\}d\bar{a} = 0$$

[10] $\oint_c \bar{A}\cdot d\bar{l} = \iint_s (\nabla\times\bar{A})\cdot d\bar{a}$. This is original form of stokes law.

For any arbitrary area $(\nabla \times \bar{E}) + \frac{\delta \bar{B}}{\delta t}$ must be zero for the relationship to be valid.

Hence;

$$\nabla \times \bar{E} + \frac{\delta \bar{B}}{\delta t} = 0$$

$$\nabla \times \bar{E} = -\frac{\delta \bar{B}}{\delta t} \tag{6.17}$$

This is *Faraday's Law in Differential Form*. For a time-varying medium, if the electric field at a particular point is known, we can find the rate of the magnetic flux density at the location by finding the curl of $\bar{E}$ at that location.

For a non-time varying medium (medium whose permeability μ does not vary with time), the curl of $\bar{E}$ can give the rate of change of magnetic field strength at that location.

What that means mathematically is that,

Recall that $\nabla \times \bar{E} = -\frac{\delta \bar{B}}{\delta t}$

We know that $\bar{B} = \mu \bar{H}$

Substituting $\bar{B}$ into the differential form of Faraday's law, we have

$$\nabla \times \bar{E} = -\frac{\delta}{\delta t}(\mu \bar{H})$$

$$\nabla \times \bar{E} = -\mu \frac{\delta \bar{H}}{\delta t} \tag{6.18}$$

This is the Differential form of Faraday's Law of Electromagnetic Induction for a non-time varying medium

6.9 Summary

(i) Note that the integral form of all the physical laws discussed in this chapter is applicable in all situations.

(ii) The origin of the electric and magnetic field is the charge. The effect of the charge is measured by a quantity called the ***electric field***.

(iii) Electric field can be defined as the *force per unit charge due to presence of charges*

(iv) When the same charge is kept in motion, it constitutes a current and the pressure of the current is felt by a quantity called ***magnetic field***.

(v) Magnetic field can also be defined as the *force experienced by a unit magnetic pole placed in the vicinity of the current*.

(vi) The magnetic field due to the current element $Id\bar{l}$ is given by: $\boxed{d\bar{H} = \frac{Id\bar{l} \times \hat{\mathbf{r}}}{4\pi r^2}}$ (Amp/m)

(vii) Magnetic field is proportional to the current I and the length of the wire through which the current flows, but inversely proportional to the square of the distance from the current carrying element (wire).

(viii) Magnetic Flux Density $\bar{B}$ is the product of the permeability of the medium and the magnetic field strength. And is given by the formula;

$$\boxed{\bar{B} = \mu \bar{H}} \quad (Wb/m^2)$$

(ix) The general form of Ohms law which is the relationship between conduction current density $\bar{J}$ and electric field Which is given by; $\boxed{\bar{J} = \sigma \bar{E}}$ (A/m^2)

(x) The two ways in getting Maxwell's equation

  (i) You treat it as a mathematical postulate

  (ii) Try to represent these mathematical laws in appropriate mathematical form to give us Maxwell's equation.

(xi) Gauss law states that *the total displacement coming out of a close surface is equal to the net charge enclosed by the closed surface*

(xii)

$$\iint_s \bar{D} \cdot d\bar{a} = \iiint_v \rho dv$$

Gauss law in integral form

(xiii)

$$\nabla \cdot \bar{E} = \frac{\rho}{\epsilon}$$

Gauss law for electric charges for homogeneous medium

(xiv)

$$\nabla \cdot \bar{B} = 0$$

Gauss law for magnetic charges

(xv) Amperes law also called *Amperes circuit law* states that the magneto-motive force(mmf) around a closed loop is equal to the total current enclosed by the loop.

(xvi)

$$\oint_c \bar{H} \cdot d\bar{l} = \iint \bar{J} \cdot d\bar{a}$$

Amperes Law in Integral Form

(xvii)

$$\boxed{\nabla \times \bar{H} = \bar{J}}$$

Amperes Law in Differential Form

(xviii) Faraday's Law of Electromagnetic Induction states that the total Emf around a closed loop is equal to the rate of change of magnetic flux enclosed by that loop, and that, the direction of emf is such that, the magnetic field produced by this current due to this emf opposes the original magnetic field.

(xix)

$$\oint_c \bar{E} \cdot d\bar{l} = -\frac{\delta}{\delta t}(\phi)$$

EMF produced by $d\bar{l}$

(xx) The negative sign shows that the emf produced is such that, the magnetic field produced by this emf will oppose the original magnetic field.

(xxi)

$$\nabla \times \bar{E} = -\frac{\delta \bar{B}}{\delta t}$$

Faraday's Law in Differential Form

In this chapter, we have seen the four basic laws (Gauss law of Electric charges, the Gauss Law of Magnetic charges, The Amperes Circuit Law and the Faraday Law of Electromagnetic induction) and how we were able to write each of them mathematically and in their integral and differential form using reactor algebra and theorem (Stokes's and Divergent Theorem).

6.10 EXERCISE

Ex. 130 — A non-uniform charge distribution within a sphere is given by $\rho(r) = \rho_0 \left(1 - \frac{r}{R}\right)$, where r is the radial distance from the sphere's center ($0 \leq r \leq R$) and ρ_0 is a constant charge density. Calculate the electric field at a distance r from the center of the sphere.

Ex. 131 — A long straight wire with a current of $5\,\mathrm{A}$ is surrounded by a cylindrical surface of radius $0.02\,\mathrm{m}$. Determine the magnetic field at a distance $0.05\,\mathrm{m}$ from the wire.

Ex. 132 — In free space, where there are no charges or currents ($\rho = 0, J = 0$), express Maxwell's equations in differential form.

Ex. 133 — A uniformly charged sphere with a total charge q and radius r generates an electric field (e) at a point outside the sphere. determine the expression for the electric field (e) as a function of the radial distance (r) from the center of the sphere.

Ex. 134 — Express Maxwell's first and second equation in integral and differential form.

Ex. 135 — Show the relationship that relates the current density with electric field

Ex. 136 — Show with diagram the magnetic field due to current element

Ex. 137 — State the laws used to derive Maxwell's equation.

Ex. 138 — Show that $\nabla \times \bar{E} = -\frac{\delta \bar{B}}{\delta t}$

Ex. 139 — Show that the general form of Ohm's law is $\bar{J} = \sigma \bar{E} \quad (A/m^2)$

Chapter 7

The Basic Laws To Maxwell's Equation (2)

7.1 OBJECTIVES

By the end of this chapter, the student should be able to do the following:

(i) Know the modification of Ampere's law.

(ii) Derive Maxwell's equation in it's differential and integral form.

(iii) Explain the concept of surface charge and surface charge density.

(iv) Explain the concept of surface current and surface current density.

7.2 INTRODUCTION

In the last chapter, the mathematical equations of the physical laws of electromagnetism were established. These laws include; ampere's circuit law, Gauss law, and Faraday's law of electromagnetic induction. Compiling these laws in mathematical form gives **Maxwell's Equations**, and this compilation was done by Maxwell. In compiling these laws, Maxwell discovered some inconsistencies in the ampere's circuit law. This inconsistency was examined and modified. This would be discussed shortly.

7.3 Ampere Circuit Law Modification and Maxwell's Equations

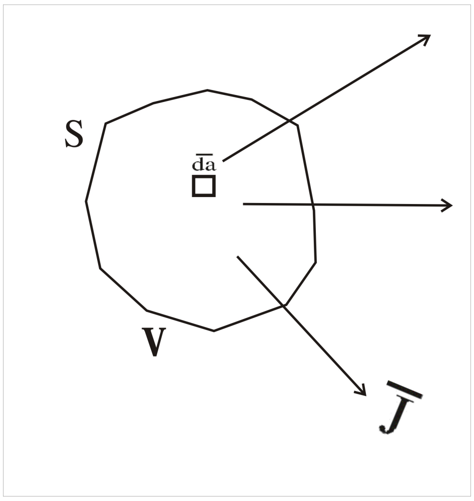


Figure 7.1: Conduction current density from a closed surface

Consider a closed surface S of a certain volume V shown in Figure 6.1. Assuming the charges are distributed inside the surface, there is the possibility that the charges would leave the surface and when this happens, there would be a rate of change of charge. This means there will be current flow from the surface and hence a current density is distributed on the surface. The charges would leave in the form of current leading to a reduction in the net charge inside the surface.

The concept of charge-current relation has been worked out and it also applies to the closed surface in Figure 6.1. The current produced on the surface would yield a surface current density. This current density is the vector field around the surface. Let's say we denote this current density of the surface with $\bar{J}$. If we take the divergence of $\bar{J}$, it would stand for the outflow of the electric vector field($\bar{J}$).

If we consider a small area on the closed surface S given by $d\bar{a}$, then the product of conduction current density $\bar{J}$ and the small area $d\bar{a}$ gives the outward current from the infinitesimally small area $d\bar{a}$. By summing over the entire surface(basically integrating), then we get the net outward current flow from the whole surface S.

$$\text{Net outward current} = \oiint_S \bar{J} \cdot d\bar{a}$$

, But as we know, the current is actually the rate of change of charge and since this current is coming out from the surface, it should be equal to the rate of charge from the total volume and since the current is coming outwards, the net charge must be decreasing inside the volume. So if the volume is said to have a charge density denoted by $\rho(C/m^3)$, then the net decrease or the rate of change of decrease of charge inside the volume is the same as the total current coming out from the surface. To get the total charge enclosed by the closed surface S, we get the integral of ρ over the volume created by the closed surface.

$$\text{Rate of charge decrease in volume} = -\frac{\partial}{\partial t}\iiint \rho dv$$

So we get that;

$$\oiint_S \bar{J} \cdot d\bar{a} = -\frac{\partial}{\partial t}\iiint_V \rho dv$$

If we say volume is not changing with time, so that it is only ρ that changes with time, we have that;

$$\oiint_S \bar{J} \cdot d\bar{a} = -\iiint_V \frac{\partial \rho}{\partial t} dv$$

We can convert a surface integral to a volume integral by using the divergence theorem. Establishing this, we get that;

$$\oiint_S \bar{J} \cdot d\bar{a} = \iiint_V (\nabla \cdot \bar{J}) dv$$

From this, we get that;

$$\iiint_V (\nabla \cdot \bar{J})dv = -\iiint_V \frac{\partial \rho}{\partial t} dv$$

this would give;

$$\iiint_V (\nabla \cdot \bar{J} + \frac{\partial \rho}{\partial t})dv = 0$$

and since this is the relation we have gotten in the equation above, it should be true for any arbitrary volume, so that the integral must be zero. Thus,

$$\nabla \cdot \bar{J} = -\frac{\partial \rho}{\partial t} \tag{7.1}$$

This is called **Continuity Equation** and should be duly noted.

So if we have time-varying charges, then $-\frac{\partial \rho}{\partial t}$ is a finite quantity and divergence of the conduction current density $\nabla \cdot \bar{J}$ is not zero. But if we take a case where the charges are not varying that is a static case, then the rate of change of charge will be zero ($\frac{\partial \rho}{\partial t} = 0$) and so $\nabla \cdot \bar{J} = 0$. This makes physical sense, especially when looking at it with the concept of divergence, which as we know measures the net quantity coming out from a unit volume, and since ($\frac{\partial \rho}{\partial t} = 0$), there would be no net flow and thus no divergence. So if the current is coming out from this volume it means that surface charge must be leaving the volume, then there must be a change in the total charge or change in charge density inside the volume.

If the charges are not changing inside the volume(static case), then whatever charge enters the volume is the same as the charge leaving the volume. So the net charge inside the volume is the same and in that case, the divergence of the conduction current is zero. So whenever we have 'time-varying quantities', the continuity equation must be satisfied by the conduction current density and the charges. This 'continuity equation' was what led to the difficulty in compiling other equations by Maxwell. Now let us examine this difficulty.

We had earlier seen from the previous chapter, that from Ampere's circuit law, we get the relation $\nabla \times \bar{H} = \bar{J}$, that is, the curl of the magnetic field is equal to the conduction current density. If we apply vector operation on this without taking note of the physical aspects(that is the physical meaning of $\nabla \times \quad and \quad \nabla \cdot$), we can expand both sides of the 'Ampere's circuit law' relation to get;

$$\nabla \cdot (\nabla \times \bar{H}) = \nabla \cdot \bar{J} \tag{7.2}$$

knowing that we can interchange $\nabla \times and \nabla \cdot$ then the relation becomes;

$$\nabla \times (\nabla \cdot \bar{H}) = \nabla \cdot \bar{J} \tag{7.3}$$

but $\nabla \times \nabla \cdot \bar{H}$ is actually zero by identification. This imply that $\nabla \cdot \bar{J} = 0$. So from Ampere's circuit law $\nabla \cdot \bar{J} = 0$ while from the continuity equation $\nabla \cdot \bar{J} = -\frac{\partial \rho}{\partial t}$.

The inconsistency arises from the fact that Ampere's circuit law does not satisfy the continuity equation. This was the difficulty that was encountered by Maxwell. He resolved this difficulty by introducing the concept of the **Displacement Current Density**. So what he said was that "in the continuity relation $\nabla \cdot \bar{J} = -\frac{\partial \rho}{\partial t}$ replace ρ from Gauss law with divergence of displacement vector $\bar{D}$", that is $\rho = \nabla \cdot \bar{D}$. So we have;

$$\nabla \cdot \bar{J} = -\frac{\partial}{\partial t}(\nabla \cdot \bar{D}) \tag{7.4}$$

Interchanging space and time operators, we have;

$$\nabla \cdot \bar{J} = -\nabla \cdot (\frac{\partial \bar{D}}{\partial t}) \tag{7.5}$$

$$\nabla \cdot \bar{J} + \nabla \cdot (\frac{\partial \bar{D}}{\partial t}) = 0 \tag{7.6}$$

we then integrate over the entire volume of the closed surface to get; where

$$\iiint_V \nabla \cdot \bar{J} = \oiint_S \bar{J} \cdot \tag{7.7}$$

and

$$\nabla \cdot (\frac{\partial \bar{D}}{\partial t})dv = \oiint_S (\frac{\partial \bar{D}}{\partial t}) \tag{7.8}$$

$$\iiint_V \nabla \cdot \bar{J} + \nabla \cdot (\frac{\partial \bar{D}}{\partial t})dv = \oiint_S (\bar{J} + \frac{\partial \bar{D}}{\partial t}) = 0 \tag{7.9}$$

Equation 6.4 is gotten by the divergence theorem

This surface integral implies that the current which is coming from the closed surface, S, is not only because of the 'conduction current density' $\bar{J}$ but also as a result of $\frac{\partial \bar{D}}{\partial t}$ which is the 'Displacement Current Density'. So for a homogenous equation, $\frac{\partial \bar{D}}{\partial t}$ has the same unit as current density. However, Displacement Current Density does not depend on the conductivity of the medium unlike conduction current density $\bar{J}$ that does. $\bar{J}$ is related to conductivity by ohms law $\bar{J} = \sigma \bar{E}$. However if $\sigma = 0$, then we have only $\frac{\partial \bar{D}}{\partial t}$ from $\bar{J} + \frac{\partial \bar{D}}{\partial t}$ to play with. The quantity $\frac{\partial \bar{D}}{\partial t}$ is equal to some current and hence it is called Displacement Current Density. So $\frac{\partial \bar{D}}{\partial t}$ is the quantity that was introduced by Maxwell to satisfy the continuity equation.

So the net current from any closed surface is not only due to conduction current density $\bar{J}$ alone, but the summation of $\bar{J}$ and the 'displacement current density' $\frac{\partial \bar{D}}{\partial t}$. So if we use this summation to define the total current, then Ampere's law can be modified to say that **the magnetomotive force around a closed loop is equal to the total current enclosed by that loop which includes the conduction current as well as displacement current**. It should be noted that current due to displacement current density is not due to charge flow, in fact, it can exist without the presence of charges(and hence conduction current density $\bar{J}$). $\frac{\partial \bar{D}}{\partial t}$ is related to electric field $\bar{E}$. So even without charges, if we have an electric field that is time-varying, $\frac{\partial \bar{D}}{\partial t}$ will equate to a current flow. So the quantity $\frac{\partial \bar{D}}{\partial t}$ represents the rate of change of the electric field and this is essentially current.

So if we take the total current which is a combination of conduction current and displacement current, this would give the $magnetomotive force$ around a closed loop. Ampere's circuit law is thus modified with current being the sum of conduction current and displacement current.

Amperes law is modified to;

$$\nabla \times \bar{H} = \bar{J} + \frac{\partial \bar{D}}{\partial t} \tag{7.10}$$

This essentially resolves the difficulty which was faced by Maxwell and this gives the complete description of the phenomenon of electro-magnetics. With these equations which we have gotten so far (that is Gauss's law of magnetic field, Faraday's law of electromagnetic induction and the modified amperes law), we have a complete set of equations which represent the static and time-varying electric and magnetic fields. These equations are called **Maxwell's Equations**. As we earlier mentioned, Maxwell's equations can be written in differential form or in integral form and depending on the suitability, the equations can be used in either form. Finally, we make a list of the four of Maxwell's Equations.

Table 7.1: **Maxwell's Equations**

Electromagnetic Laws	**Differential form**	**Integral form**
Gauss laws;	$\nabla \cdot \bar{D} = \rho$	$\oiint\limits_S \bar{D} \cdot d\bar{a} = \iiint\limits_V \rho dv$
	$\nabla \cdot \bar{B} = 0$	$\oiint\limits_S \bar{B} d\bar{a} = 0$
Faradays law;	$\nabla \times \bar{E} = -\frac{\partial \bar{B}}{\partial t}$	$\oint\limits_C \bar{E} \cdot d\bar{l} = -\iint\limits_S \frac{\partial \bar{D}}{\partial t} \cdot d\bar{a}$
Modified Amperes Law;	$\nabla \times \bar{H} = \bar{J} + \frac{\partial \bar{D}}{\partial t}$	$\oint\limits_C \bar{H} d\bar{l} = \iint\limits_S (\bar{J} + \frac{\partial \bar{D}}{\partial t}) \cdot d\bar{a}$

These are the set of equations which governs the total phenomenon of electro-magnetics for $static$ as well as $time\ varying\ fields$ so once we have these generalized equations, we can reduce the equation to get those for $Static\ Fields$ by putting all time derivatives to zero. So depending on whether the 'medium parameter' (permeability μ, permittivity ϵ) of the medium is varying as a function of space(inhomogeneous) or is constant as a function of space (homogeneous), we can get $\bar{D} = \epsilon\bar{E}$ and $\bar{B} = \mu\bar{H}$. So we can have various forms of these equations depending upon the condition applied to the medium; whether we are dealing with given electric fields, magnetic fields, or whatever parameters are associated with them. As we said, for static fields, the time-varying parameters become zero. Looking at Table 6.1, it essentially means that the equations with time derivatives become; $\nabla \times \bar{E} = 0$ and $\nabla \times \bar{H} = \bar{J}$.

However in this part of the course, which is on Electromagnetic Waves, we are dealing with time-varying quantities only, and so all the quantities ($\rho, \bar{B}, \bar{J}, \bar{H}$ and so on), will be considered as time-varying. Later on, we will look for solutions to these equations of time-varying fields.

As we mentioned earlier Maxwell's equations in differential form cannot be applied in a certain situation where a medium has a discontinuity. That means if we talk about media interfaces where medium properties suddenly change, like permittivity or permeability suddenly change, at those boundaries of abrupt change, the derivatives cannot be defined. So the differential form of Maxwell's equations is not useful in this situation. However, as we have mentioned the integral form is always useful and can be applied in any situation. However, if we apply the integral form to discrete media interfaces, we get relationships between the quantities $\bar{D}, \bar{B}, \bar{E}, \bar{H}$ in the two media just across the interface. That relationship is what we call the *Boundary Condition*. The same set of equations in differential form gives what is called *Point Relation* that is they are valid at every point in space. The equation in integral form when applied to discrete media interfaces gives what is called *Boundary Condition*. However before we go into boundary condition, we will introduce the concept of surface current and surface charges.

7.4 Concept of Surface Charges and Surface Current

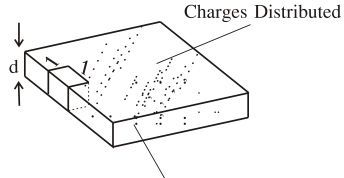


Figure 7.2: Model for studying surface charge

7.4.1 Surface Charge and Surface Charge Density

Let us look at a surface with thickness d, and a charge density $\rho C/m^3$ as shown in figure 7.2. Considering unit area on the surface, with length 1 and width 1, so if a volume of the unit area on the surface of the sheet is considered, the total charge will be in this volume which is having a height d and area 1, this is the charge density ρ multiplied by the elemental volume.

The total charge in the unit volume of the slab is;

$$= \rho \times (1 \times 1 \times d) = \rho d$$

the unit is Coulombs C.

If we now reduce the thickness of this slab and go to a limit when **d** tends to zero ($d \to 0$), then as a result, current density would tend to infinity ($\rho \to \infty$), so that the product 'ρd' would tend to finite quantity. in this situation ($d \to 0$, $\rho \to \infty$), we see a net charge which is just on the surface because the charge is now confined to a thickness of zero. This essentially implies that the charge will just be lying on the surface and that charge is in the unit area.

So if we take the quantity ρd and take the limit when $d \to 0$, we get a quantity which is charge distributed on the surface and thus there would be a charge density confined on the surface. The unit for $surface\ charge\ density$ ρ_s is actually C/m^2 unlike $volume\ charge\ density$ ρ (C/m^3).

So we say that;

$$\rho_s = \lim_{d \to 0} \{\rho d\} \tag{7.11}$$

This relation is called **Surface Charge Density**.

So two things we note here, if we go from volume charge density ρ to surface charge density ρ_s, when $d \to 0$, the surface charge density ρ_s is equal to infinite volume charge density($\rho \to \infty$). If that infinite volume charge density is confined to our thickness, that gives the distribution of charges on the surface. Later on, we would see that in situations like conducting boundaries, where the conductivity becomes infinite, you might get volume charge density which will be infinite and these charges here will truly be confined to the surface and at that time the concept of surface charge current density will be useful. So at the moment without getting into which media will have a surface charge density, we can say principally that when we have charges distributed truly on a surface with zero thickness, those charges are called *surface charges*. We can do a similar thing for the current also.

7.4.2 Surface Current and Surface Current Density

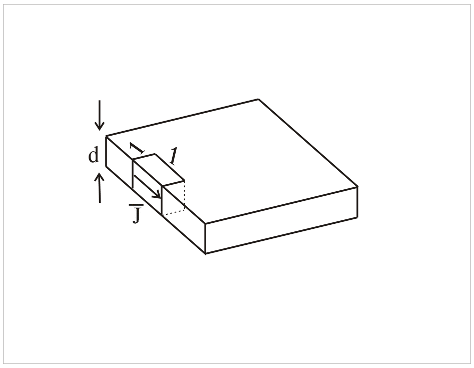


Figure 7.3: Model for studying surface current

Let's say we have a slab of thickness d carrying current $\bar{J}$ as shown in Figure 19.3. So the current flowing in the unit element is $\bar{J} \times$ {surface of the unit element through which $\bar{J}$ flows} (note that '$\times$' represents multiplication not cross product). This surface considered through which $\bar{J}$ flows has its direction parallel to that of $\bar{J}$ and is shown by the dotted line.
So we get that current flowing in the unit element is given by; $\bar{J} \times (d \times 1) = \bar{J}d$. Again if we make $d \to 0$ and $\bar{J}$ goes to infinity, you will have a current truly flowing on the surface and that current is the **surface current** $\bar{J}_s$.
So essentially surface current is given by;

$$\bar{J}_s = \lim_{d \to 0} \{\bar{J}d\} \tag{7.12}$$

Again this applies to boundaries which are conducting boundaries, so when the conductivity of the medium becomes infinite, then the current which is $\bar{J} = \sigma\bar{E}$ for a finite electric field becomes infinite and then you have what is called a 'surface current'. Since the unit of conduction current density $\bar{J}$ is A/m^2, the unit of surface current density is A/m($\bar{J} \times d$). That is the reason this quantity is also called the $linear\ surface\ current\ density$ because it is defined per unit length.

In all we are having quantities like charge density or volume charge density ρ, we have conduction current density $\bar{J}$, we have displacement current density $\frac{\partial \bar{D}}{\partial t}$, we have surface charge densityρ_s and we have surface current density $\bar{J}_s$. Generally one will say that these are the sources which are related to the electric and magnetic fields. So one can establish a relationship between these quantities which are the sources of the fields and these relationships are called **boundary conditions**. Now let us solve some problems to reinforce our understanding of these concepts.

7.5 Problems

Example 7.5.1 *The volume charge density inside a hollow sphere is*

$$\rho = 10e^{-20r} C/m^3.$$

Find the total charge enclosed within the sphere. Also, find the electric flux density on the surface of the sphere for a radius of 2m.

Solution

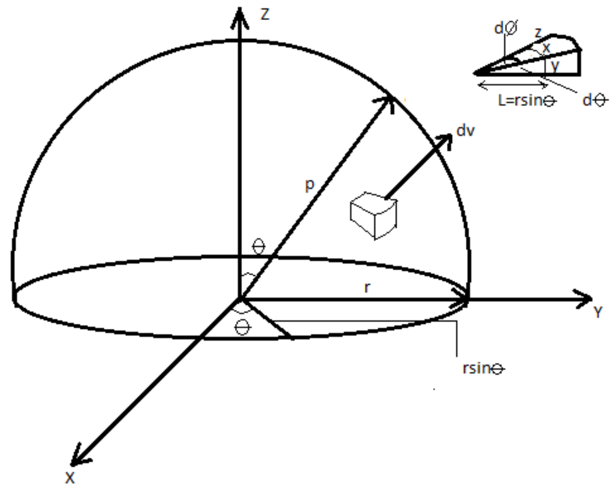


Figure 7.4: spherical coordinate system

The total charge enclosed in the sphere is;

$$Q = \iiint_V \rho dv$$

From Figure 19.4,

$$dv = dr \times (rsin\theta d\phi) \times rd\theta$$
$$dv = r^2 sin\theta dr d\theta d\phi$$
$$so, Q = \int_{\phi=0}^{2\pi}\int_{\theta=0}^{\pi}\int_{r=0}^{2} \rho r^2 sin\theta dr d\theta d\phi$$
$$Q = \int_{\phi=0}^{2\pi} d\phi \int_{\theta=0}^{\pi} sin\theta d\theta \int_{r=0}^{2} 10e^{-20r} r^2 dr$$

integrating this would give the value of Q and so we get that;

$$Q = \frac{\pi}{100} C$$

To find the electric flux density, we use Gauss law which states that the electric flux density over the entire area of the sphere is equal to the total charge enclosed. Since the surface is spherical, the charges are uniformly distributed in a spherically symmetric manner. So we can find it according to Gauss law;

$$4\pi r^2 D = Q = \frac{\pi}{100}$$
$$D = \frac{Q}{4\pi r^2} = 6.25 \times 10^{-4} C/m^2$$

Example 7.5.2 *The electric flux density is given as*

$$\bar{D} = x^3\hat{x} + x^2 y\hat{z}.$$

Find the charge density inside a cube of side 2m placed at the origin with its side along the coordinate axes

Solution

Here we use the differential form of Gauss law to find out first the charge density and then the charge enclosed inside the cube.

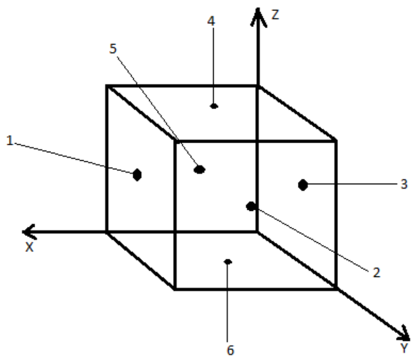


Figure 7.5: Model for electric field density inside a cube

We have the following area 1,2,3,4,5,6 that makes up the closed surface that gives the volume of the cube. Again, we make use of Gauss law;

$$\oiint_S \bar{D}\cdot d\bar{a} = \textit{Charge enclosed}$$

We denote $A_1, A_2, A_3, A_4, A_5 and A_6$ *as the area element on the faces labelled 1,2,3,4,5 and 6 respectively. Since the centre of the sphere is the origin, with side length 2m, we have a +1 and -1 offset from the origin which is the cube centre.*

$dA_1 = dzdy, x = +1, dA_2 = dxdz, y = +1, dA_3 = dydz, x = -1, dA_4 = dxdy, z = +1, dA_5 = dxdz, y = -1, dA_6 = dxdy, z = -1$

$\oiint_S \bar{D}\cdot d\bar{a}$ *will be the sum total of all* $D\cdot d\bar{a}$ *for individual areas.*

Taking the outward normal as position for the volumes

$A_1 = dzdyx = A_2 = dxdzy = A_3 = dydz - x = -1, A_4 = dxdyz = A_5 = dxdz - y = A_6 = dxdy,$

with unit vector added to show the direction of the ones with outward normal unit vector $\hat{n}$ *added*

$$\iint_S \bar{D}\cdot d\bar{a} = \iint_{A1} \bar{D}\cdot d\bar{A}_1 + \iint_{A2} \bar{D}\cdot d\bar{A}_2 + \iint_{A3} \bar{D}\cdot d\bar{A}_3 +$$

$$\iint_{A4} \bar{D}\cdot d\bar{A}_4 + \iint_{A5} \bar{D}\cdot d\bar{A}_5 + \iint_{A6} \bar{D}\cdot d\bar{A}_6$$

$$= \iint (x^3\hat{x} + x^2y\hat{z}).(dzdy\hat{x})|_{x=1} +$$

$$\iint (x^3\hat{x} + x^2y\hat{z}).(dxdz\hat{y})|_{y=1} +$$

$$\iint (x^3\hat{x} + x^2y\hat{z}.(-dydz\hat{x})|_{x=-1} +$$

$$\iint (x^3\hat{x} + x^2y\hat{z}).(dxdy\hat{z})|_{z=1} +$$

$$\iint (x^3\hat{x} + x^2y\hat{z}).(-dxdz\hat{y})|_{y=-1} +$$

$$\iint (x^3\hat{x} + x^2y\hat{z}).(-dxdy\hat{z})|_{z=-1}$$

$$= \int_{-1}^{1}\int_{-1}^{1} (\hat{x} + y\hat{z}).(dzdy\hat{x}) + \int_{-1}^{1}\int_{-1}^{1} (x^3\hat{x} + x^2\hat{z}).(dxdz\hat{y}) +$$

$$\int_{-1}^{1}\int_{-1}^{1} (-\hat{x} + y\hat{z}).(-dydz\hat{x}) + \int_{-1}^{1}\int_{-1}^{1} (x^3\hat{x} + x^2y\hat{z}).(dxdy\hat{z}) +$$

$$\int_{-1}^{1}\int_{-1}^{1} (x^3\hat{x} - x^2\hat{z}).(-dxdz\hat{y})| \int_{-1}^{1}\int_{-1}^{1} (x^3\hat{x} + x^2y\hat{z}).(-dxdy\hat{z})|$$

$$= \int_{-1}^{1}\int_{-1}^{1} dzdy + \int_{-1}^{1}\int_{-1}^{1} dydz + \int_{-1}^{1}\int_{-1}^{1} x^2ydxdy + \int_{-1}^{1}\int_{-1}^{1} -x^2ydxdy$$

$$= 4 + 4 = 8c$$

SIMPLER METHOD

using gauss law; $\nabla\cdot\bar{D} = \rho$

$$\frac{\partial D_x}{\partial x} + \frac{\partial D_y}{\partial y} + \frac{\partial D_z}{\partial z} = \rho$$

substituting $\bar{D}$ *components, we have;*

$$\rho = \frac{\partial}{\partial x}(x^3) + 0 + \frac{\partial}{\partial z}(x^2y)$$
$$\rho = 3x^2.$$

To get the total charge enclosed Q, we take the volume integral of the charge density ρ

$$\begin{aligned} Q &= \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} \rho dxdydz \\ &= \int_{-1}^{1}\int_{-1}^{1}\int_{-1}^{1} 3x^2dxdydz \\ &= 12\int_{-1}^{1} x^2dx \\ &= 12\frac{2}{3} \\ &= 8C \end{aligned}$$

This could also be solved using the integral form of Gauss law.

Example 7.5.3 *In a conducting medium, the magnetic field is given as*

$$\bar{H} = y^2z\hat{x} + 2(x+1)yz\hat{y} - (x+1)z^2\hat{z}.$$

Find the conduction current density at point (2,0,-1)m. Also find the current enclosed by a square loop y=1, $0 < x < 1$, $0 < z < 1$.

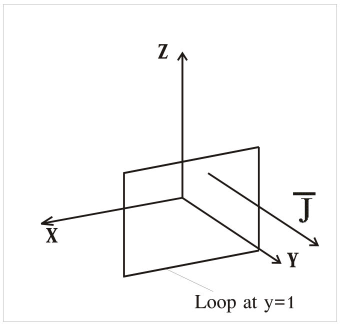


Figure 7.6: Loop at y=1

Solution

Here we will use amperes law(differential form) to solve for the conduction current density. Amperes law is given by;

$$\bar{J} = \nabla \times \bar{H}$$

This is essentially the curl of $\bar{H}$ *and is evaluated by solving for the determinant of the matrix formed by* ∇ *and* $\bar{J}$ *.*
so we have that $\bar{J} =$

$$\begin{vmatrix} \hat{x} & \hat{y} & \hat{z} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ H_x & H_y & H_z \end{vmatrix}$$

$$\bar{J} = (\frac{\partial H_z}{\partial y} - \frac{\partial H_y}{\partial z})\hat{x} + (\frac{\partial H_x}{\partial z} - \frac{\partial H_z}{\partial x})\hat{y} + (\frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y})\hat{z}$$

Solving, we get;

$$\bar{J} = -2(x+1)y\hat{x} + (y^2 + z^2)\hat{y}$$

At location (2,0,-1) the conduction current density is;

$$\bar{J} = \hat{y}$$

With conduction current density $\bar{J}$ *known, we can now find out the current enclosed by a loop at y=1 by integrating* $\bar{J}$ *over the area of the loop. The conduction current density is perpendicular to the plane created by that loop. So looking at the XZ*

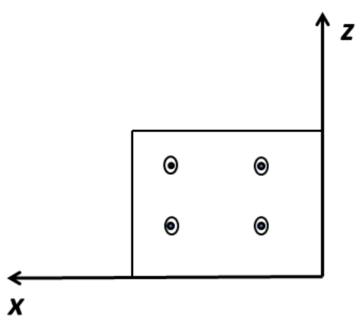


Figure 7.7: Direction of the conduction current in the XZ plane

plane as shown in Figure 19.6, if we go by the right-hand rule, we have to go in an anti-clockwise direction to get the current flowing in that direction(perpendicular to the plane created by the loop).

$$\begin{aligned} I &= \iint \bar{J} d\bar{a} \\ &= \iint \bar{J}\hat{y} dx dz \\ &= \int_0^1 \int_0^1 J_y dx dz \\ &= \int_0^1 \int_0^1 (y^2 + z^2) dx dz \end{aligned}$$

at y=1

$$\begin{aligned} I &= \int_0^1 \int_0^1 (1 + z^2) dx dz \\ &= \frac{4}{3} A \end{aligned}$$

So these are some simple problems which essentially give you some feel on how to apply Maxwell's equations in real life, either with the differential form or integral form.

7.6 EXERCISE

Ex. 140 — Explain the correction made by Maxwell to Ampere's Circuit Law.

Ex. 141 — What is Displacement Current Density.

Ex. 142 — Show that $\nabla \times \bar{H} = \bar{J} + \frac{\partial \bar{D}}{\partial t}$

Ex. 143 — Explain the concept of Surface charge and surface charge density.

Ex. 144 — Explain the concept of Surface current and surface current density.

Ex. 145 — Explain conduction current, convection current and displacement current

Ex. 146 — A square loop of side length 2 m lies in the yz-plane. The loop carries a current of 5 A in the counterclockwise direction when viewed from the positive x-axis. Determine the magnetic field at the center of the loop using Ampere's Circuital Law.
Solution:

Ex. 147 — A long straight wire carries a current of 8 A along the positive y-axis. The position vector of a point in space is given by $\mathbf{r} = \begin{bmatrix} -2 \\ 4 \\ 1 \end{bmatrix}$ meters. Find the magnetic field at this point using Ampere's Circuital Law.

Ex. 148 — Explain Boundary conditions.

Chapter 8

Medium of Finite Conductivity

8.1 Objectives

1. Derive the respective equations for zero conductivity and finite conductivity in a medium.
2. Understand the effects of frequency and attenuation constant on a medium.
3. Understand the concept of loss tangent.
4. Understand the concept of intrinsic impedance.

According to Maxwell's equation:

$$\nabla \cdot D = {\rho_v}^1 \tag{8.1}$$

$$\nabla \cdot B = 0^2 \tag{8.2}$$

$$\nabla X E = \frac{\delta B}{\delta t}^3 \tag{8.3}$$

$$\nabla X H = \frac{\delta \bar{D}}{\delta t} + J^4 \tag{8.4}$$

8.2 Maxwell's equation in relation to Ampere's Law

The conductivity of a medium appears in Maxwell's equation and it corresponds to the ampere's law.

We have two cases to deal with:

Case 1: When the conductivity of the medium = 0

The is 0 (J=0), therefore:

$$\nabla X \bar{H} = \frac{\delta \bar{D}}{\delta t} \tag{8.5}$$

Case 2: We take into consideration, the conduction current density which is related to the finite conductivity of the medium.

$$\nabla X \bar{H} = \tfrac{\delta \bar{D}}{\delta t} + J$$

We know from Ohm's law:

$$J = \sigma \bar{E}^5 \tag{8.6}$$

The simplest case is an (homogenous) which makes the conductivity uniform in the space and not dependent on the direction.

When an electric field is impressed in a medium of finite conductivity, then there are two types of current charges that will flow in the medium; One corresponds to the current density(σ), and the other corresponds to the displacement current density.

So in general, in any medium, there are two types of current that flow: **Conduction current** and **Displacement current.**

Let us investigate the equations in both Case 1 and Case 2 above:

Recall:

$$\bar{D} = \epsilon_o \epsilon_r \bar{E}^6 = \epsilon \bar{E} \tag{8.7}$$

$$\frac{\delta}{\delta t} = j\omega^7 \tag{8.8}$$

For a finite conductivity in a medium:

According to the equation above:

$$\nabla X \bar{H} = \tfrac{\delta \bar{D}}{\delta t} + J$$

Substituting for $\bar{D}$ in the equation above, we have:

$$\nabla X \bar{H} = \tfrac{\delta \epsilon \bar{E}}{\delta t} + J$$

Substituting $\epsilon = \epsilon_o \epsilon_r$

$$\nabla X \bar{H} = \frac{\delta}{\delta t}[\epsilon_o \epsilon_r \bar{E}] + J \tag{8.9}$$

Substituting $\frac{\delta}{\delta t} = j\omega$ in equation**??**

$$\nabla X \bar{H} = j\omega \epsilon_o \epsilon_r \bar{E} + J \tag{8.10}$$

According to Ohm's law, $J = \sigma \bar{E}$. We substitute this in equation8.10 to get:

[1] Gauss's Law for the electric field describes the static electric field generated by a distribution of electric charges

[2] Gauss's Law for magnetism states that the magnetic field B has divergence equal to zero

[3] Faraday's Law states that when the magnetic flux linking a circuit changes, an EMF is induced in the circuit proportional to the rate of change of the flux linkage

[4] Ampere - Maxwell law relates electric currents and magnetic flux

[5] In physics, the term Ohm's law is also used to refer to various generalizations of the law originally formulated by Ohm. Where J is the current density at a given location in a resistive material, E is the electric field at that location, and σ is a material-dependent parameter called the conductivity.

[6] Electric flux density is a measure of the strength of an electric field generated by a free electric charge, corresponding to the number of electric lines of force passing through a given area.

$$\nabla X\bar{H} = j\omega\epsilon_o\epsilon_r\bar{E} + \sigma\bar{E}$$
$$\nabla X\bar{H} = j\omega\epsilon_o\{\epsilon_r + \tfrac{\sigma}{j\omega\epsilon_o}\}\bar{E}$$

$$\nabla X\bar{H} = j\omega\epsilon_o\left\{\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}\right\}\bar{E} \tag{8.11}$$

For No conductivity:
According to the equation above:

$$\nabla X\bar{H} = \tfrac{\delta\bar{D}}{\delta t}$$

Substituting $\frac{\delta}{\delta t} = j\omega$ in the equation above

$$\nabla X\bar{H} = j\omega(\bar{D})$$

Substituting for $\bar{D} = \epsilon_o\epsilon_r\bar{E}$ in the equation above, we have:

$$\nabla X\bar{H} = j\omega(\epsilon_o\epsilon_r\bar{E}) \tag{8.12}$$

Thus,
For a medium with no conductivity:

$$\nabla X\bar{H} = j\omega(\epsilon_o\epsilon_r\bar{E})$$

For a medium of finite conductivity:

$$\nabla X\bar{H} = j\omega\epsilon_o\left\{\epsilon_r - j\tfrac{\sigma}{\omega\epsilon_o}\right\}\bar{E}$$

8.3 Complex Dielectric Constant

Now, when we compare the two equations for no conductivity and for a finite conductivity in a medium, and we treat the medium like a dielectric, we realize that:

$$j\omega(\epsilon_o\epsilon_r\bar{E}) \equiv j\omega\epsilon_o\left\{\epsilon_r - j\tfrac{\sigma}{\omega\epsilon_o}\right\}\bar{E}$$

$\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}$ becomes the Relative permittivity(dielectric constant) of the medium.

Because of the finite conductivity, the dielectric constant has become a complex quantity.

Hence, the conductivity of a medium can be accounted for by effectively introducing the concept of the complex dielectric constant.

$$\textbf{Complex dielectric constant}(\epsilon_{rc}) = \epsilon_r - j\frac{\sigma}{\omega\epsilon_o} \tag{8.13}$$

NOTE: When we have a finite conductivity in a medium, the dielectric constant becomes a complex quantity. The dielectric constant has become a function of frequency.

Earlier, the medium properties were not dependent on frequency but if we introduce the concept of complex dielectric constant, then the medium property which is the relative permittivity or dielectric constant(which has become a complex quantity) now depends on frequency.

Now, the question that comes to mind is, should we call this medium a conductor or a dielectric? The answer lies in the two terms in the equation for finite conductivity.

$$\nabla \times \bar{H} = j\omega\epsilon\bar{E} + \sigma\cdot\bar{E}$$

When an electric field is imposed firmly in the medium either the conduction current or displacement current dominates.

(i) If the displacement current dominates we say that this medium is a dielectric.

If the conduction current dominates we say this medium is a conductor.

Mathematically;

If $\sigma \gg \omega\epsilon$ then the medium is a Conductor.

If $\sigma \ll \omega\epsilon$ Then the medium is a Dielectric.

(ii) Since the term $\omega\epsilon$ is dependent on the frequency range and it is dependent on a given permittivity or conductivity, it might behave as a dielectric or conductor, depending on what frequency range we are operating on.

If the frequency is very low, σ would be much larger compared to $\omega\epsilon$.

If the frequency is very high, σ would be much lesser compared to $\omega\epsilon$.

Therefore, the lower we go on the frequency spectrum, the more conductive the medium becomes and the higher we go on the frequency spectrum, the more dielectric the medium becomes.

Hence, the behaviour of the medium depends on the frequency of the operation.

(iii) When $\sigma = \omega\epsilon = \omega\epsilon_o\epsilon_r$

$$\omega_T = \frac{\sigma}{\epsilon_o\epsilon_r} \tag{8.14}$$

ω_T is called the transitive frequency.

At this point, the conduction current density and the displacement current density become equal.

If $\omega > \omega_T$ then it is a dielectric medium.

If $\omega < \omega_T$ then it is a conductor.

Example 8.3.1 *What are the frequency ranges in which copper can act like a dielectric or conductor? Copper: $\sigma = 5.6 \times 10^7\ S/m$, $\epsilon_0 = 8.854 \times 10^{-12}\ F/m$, $\epsilon_r = 1$*

Recall,

$$\omega_T = 2\pi f_T \tag{8.15}$$

And $\omega_T = \frac{\sigma}{\epsilon_0\epsilon_r}$, so,

$$2\pi f_T = \frac{\sigma}{\epsilon_0\epsilon_r} \tag{8.16}$$

When we substitute the values into the formula, we have:

$$f_T = \frac{5.6 \times 10^7}{8.854 \times 10^{-12} \times 1} Hz \tag{8.17}$$

This implies that copper will behave as a conductor at this frequency and below. If we go higher than this in the frequency spectrum, then copper begins to behave like a dielectric medium.

Example 8.3.2 *The conductivity of sea water, $\sigma = 10^{-3}S/m$, $\epsilon_o = 8.854X10^{-12}Fm^{-1}$, dielectric constant $\epsilon_r = 80$. When we substitute these values into the formula we have, $f_T = 225KHz$*

Seawater is like a dielectric. However, at a frequency less than 225KHz, the sea water is more like a conductor.

8.4 Propagation of Electromagnetic Waves

We'll take two extreme cases for the propagation of electromagnetic waves for a medium of finite conductivity which are:

(i) Low Conductivity (good dielectric)

(ii) High Conductivity (good conductor)

Maxwell's equation for the two cases above is solved for a dielectric medium by replacing the dielectric constant of this medium with the complex dielectric constant for the conducting medium.

8.4.1 Wave equation

$$\nabla^2\left\{\bar{E},\bar{H}\right\} = -\omega^2\mu\epsilon_o\epsilon_{rc}\left\{\bar{E},\bar{H}\right\}^8 \tag{8.18}$$

This means that,

$$\nabla^2\bar{E} = -\omega^2\mu\epsilon_o\epsilon_{rc}\bar{E}$$

$$\nabla^2\bar{H} = -\omega^2\mu\epsilon_o\epsilon_{rc}\bar{H}$$

$$\gamma^2 = -\omega^2\mu\epsilon_o\epsilon_{rc}$$

γ = Propagation constant

Therefore,

$$\textbf{Propagation constant} = \sqrt{-\omega^2\mu\epsilon_o\epsilon_{rc}{}^9} \tag{8.19}$$

Let us assume the medium is not magnetic, ie $\mu_r = 1$, then we substitute $\epsilon_{rc} = \epsilon_r - j\frac{\sigma}{\omega\epsilon_o}$ in equation8.19 and then we have a new equation:

$$\gamma = j\omega\sqrt{\mu\epsilon_o}\left\{\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}\right\}^{\frac{1}{2}} \tag{8.20}$$

This is a complex quantity and no longer a purely magnetic quantity. First, we note that σ the conductivity of the material is introduced. If the medium was a pure dielectric with no conductivity, σ = 0. From $\gamma = \alpha + j\beta$, α tells you the change in amplitude of the wave as the wave travels. This is called **Attenuation constant.** So as soon as we have conductivity in the medium, we have an attenuation constant. So when the wave propagates in a medium with finite conductivity, its amplitude reduces. Physically when there was no conductivity in the medium, the wave was propagating, we had an electric and magnetic field. Now we have a finite conductivity in the medium, this leads to a finite conduction current. With finite conductivity, we have a finite resistivity as the inverse of conductivity is resistivity. Once you have conductivity, the conduction current flows, and since the medium has finite resistivity, then we have an ohmic loss in the medium or I^2R loss in the medium. As a result, when the wave propagates, the power or energy which the wave is carrying gets converted into heat. It heats up the medium as a result. This is the reason why as the wave propagates, its amplitude reduces, because the power carried is reduced as it travels along the medium. Physically, it makes sense that when we have a finite conductivity, we must have an attenuation of the wave.

Attenuation Constant:

$$\alpha = Re(\gamma) = Re\left\{j\omega\sqrt{\mu\epsilon_o}\left\{\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}\right\}^{\frac{1}{2}}\right\}$$

Phase constant

$$\beta = I_m(\gamma) = I_m\left\{j\omega\sqrt{\mu\epsilon_o}\left\{\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}\right\}^{\frac{1}{2}}\right\}$$

The result of the algebraic manipulation is:

$$\alpha = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\sqrt{1+\frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}} - 1\right\}^{\frac{1}{2}\,10} \tag{8.21}$$

$$\beta = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\sqrt{1+\frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}} + 1\right\}^{\frac{1}{2}\,11} \tag{8.22}$$

We will show how we got this result above and some procedure would be valid to get α and β

$$\gamma^2 = -\omega^2\mu\epsilon_o\left\{\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}\right\} \tag{8.23}$$

$$\gamma = \alpha + j\beta$$

$$(\alpha + j\beta)^2 = -\omega^2\mu\epsilon_o\left\{\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}\right\}$$

$$(\alpha^2 - \beta^2 + j2\alpha\beta) \equiv -\omega^2\mu\epsilon_o\left\{\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}\right\}$$

Open the bracket and compare the real terms and the imaginary terms and then we have:

$$\alpha^2 - \beta^2 = -\omega^2\mu\epsilon_o\epsilon_r$$

$$2\alpha\beta = \mu\omega\sigma \Rightarrow \alpha = \frac{\mu\omega\sigma}{2\beta}$$

Put α in the equation above:

$$\left\{\frac{\mu\omega\sigma}{2\beta}\right\}^2 - \beta^2 = -\omega^2\mu\epsilon_o\epsilon_r$$

$$\frac{\mu^2\omega^2\sigma^2}{4\beta^2} - \beta^2 = -\omega^2\mu\epsilon_o\epsilon_r$$

$$\mu^2\omega^2\sigma^2 - 4\beta^4 = -4\beta^2\omega^2\mu\epsilon_o\epsilon_r$$

Multiply through by the negative sign:

$$4\beta^4 - 4\beta^2\omega^2\mu\epsilon_o\epsilon_r - \mu^2\omega^2\sigma^2 = 0$$

Then, let β^2 = x

$$4x^2 - 4x\omega^2\mu\epsilon_o\epsilon_r - \mu^2\omega^2\sigma^2 = 0$$

[8] The wave equation is an important second-order linear partial differential equation for the description of waves - as they occur in classical physics.

[9] The propagation constant of a sinusoidal electromagnetic wave is a measure of the change undergone by the amplitude and phase of the wave as it propagates in a given direction.

[10] The real part of the propagation constant is the attenuation constant and is denoted by α(alpha). It causes signal amplitude to decrease along a transmission line. The natural units of the attenuation constant are Nepers/meter

[11] The imaginary part of the propagation constant is the phase constant and is denoted by β(beta).

The equation above follows the general expression of a quadratic equation of the form $ax^2 + bx + c = 0$
From the quadratic formula:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

Where a = 4, b = $-4\omega^2\mu\epsilon_o\epsilon_r$, c = - $\mu^2\omega^2\sigma^2$

$$x = \frac{4\omega^2\mu\epsilon_o\epsilon_r + \sqrt{16\omega^4\mu^2\epsilon_o^2\epsilon_r^2 + 16\mu^2\omega^2\sigma^2}}{8}$$

$$x = \frac{\omega^2\mu\epsilon_o\epsilon_r + \sqrt{\omega^4\mu^2\epsilon_o^2\epsilon_r^2 + \mu^2\omega^2\sigma^2}}{2}$$

Now,

$$\beta^2 = x = \frac{\omega^2\mu\epsilon_o\epsilon_r}{2}\left\{1 + \left\{1 + \frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}\right\}^{\frac{1}{2}}\right\}$$

Then,

$$\beta = \left\{\frac{\omega^2\mu\epsilon_o\epsilon_r}{2}\left\{1 + (1 + \frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2})^{\frac{1}{2}}\right\}\right\}^{\frac{1}{2}}$$

With relative permeability μ_r = 1, $\mu = \mu_o$

$$\beta = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\sqrt{1 + \frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}} + 1\right\}^{\frac{1}{2}}$$

α can be obtained in a similar way from $\alpha^2 - \beta^2 = -\omega^2\mu\epsilon_o\epsilon_r$

$$\alpha = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\sqrt{1 + \frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}} - 1\right\}^{\frac{1}{2}}$$

NOTE: The attenuation constant is a function of frequency ω and the conductivity of the medium. It is also true for the phase constant. With pure dielectric, β was only proportional to frequency ω. So in the presence of finite conductivity, the attenuation constant,α and phase constant β of the waves have become functions of frequency and conductivity.

Hence, we take the two extreme cases with this investigation made. If we take the conductivity which is very small compared to $\omega\epsilon$: A dielectric medium (low loss medium) has a conductivity and if the conductivity is more, the conduction current increases, therefore the loss will increase. In the normal electrical circuits we are used to, when the conductivity increases, there will be less loss in the circuit. More conductivity means less resistivity and the smaller the loss in the circuit. However, what we are seeing here is quite the opposite. We find that when the conductivity is higher, we have a higher attenuation, and thus more loss of energy in the medium.

8.5 Loss Tangent

When we are dealing with electrical circuits, we are dealing with conduction currents. We were dealing with components that were more like conducting load of components and in that situation when the conductivity is large or infinite (for a conductor) there is no loss, the resistivity is zero and even if the current flows, there is no loss of power. However, when we come to the case of a dielectric, then an ideal dielectric without any conductivity does not have any loss. So if we have a medium which is like a dielectric(low loss), it must have zero conductivity. In both situations, there is no loss. Then if we take a dielectric medium, the higher the conductivity, the higher the loss. If we take an ideal conductor, the lower the conductivity, the more the loss. So essentially, what we see here is that, since we are now dealing with a dielectric medium, the increase in the conductivity of the medium essentially gives us the Ohmic loss and because of that we have attenuation and a power loss in the medium. As a measure of how much power is lost in the medium, we define a parameter called the LOSS TANGENT which is the ratio of conduction current to displacement current.

$$\textbf{Loss tangent}(\tan\delta) = \frac{\sigma}{\omega\epsilon_o\epsilon_r} \tag{8.24}$$

So, whenever we talk about a dielectric, how good the dielectric is, is measured by the loss tangent. Since we are talking about a good dielectric, we assume that this medium is predominantly dielectric and the conduction current is much smaller than the displacement current. So generally, for a good dielectric material, the loss tangent is extremely small.

$$(\tan\delta) = \frac{\sigma}{\omega\epsilon_o\epsilon_r} \ll 1 \text{ for a good dielectric}$$

So, if we take a material which we want to use for dielectric, then $\tan\delta$ should be very small. Ideally we want $\tan\delta$ = 0, but in practice it ranges from 10^{-4} to 10^{-3}. The smaller the value, the better the dielectric, and the less the ohmic losses as the wave propagates in this medium.
Whenever we have a dielectric material, the loss tangent is mentioned instead of the frequency and $\tan\delta$ is a frequency-dependent quantity. So we must know the loss tangent at the frequency of operation. If we know these values at a certain frequency, we can always compare the loss tangent to the appropriate frequency. This is a useful parameter for characterizing dielectric materials and determining whether it is a good dielectric or not.
Let us now take the two extreme cases of wave propagation that is, when the medium is a very good dielectric and when it is a very good conductor.

For a good dielectric (low loss dielectric): $\omega\epsilon_o\epsilon_r \gg \sigma$

$$\alpha = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\sqrt{1 + \frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}} - 1\right\}^{\frac{1}{2}}$$

$$\alpha = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\left\{1 + \frac{1}{2}\frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2} + \ldots\right\} - 1\right\}^{\frac{1}{2}}$$

$$\alpha = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\frac{\sigma}{\sqrt{2}\omega\epsilon_o\epsilon_r}\right\} = \frac{\sigma}{2}\sqrt{\frac{\mu_o}{\epsilon_o\epsilon_r}}$$

Now for β, Recall:

$$\beta = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\sqrt{1 + \frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}} + 1\right\}^{\frac{1}{2}}$$

Applying binomial expansion to the term $\sqrt{1 + \frac{\sigma^2}{\omega^2\epsilon_0^2\epsilon_r^2}}$

$$\sqrt{1 + \frac{\sigma^2}{\omega^2\epsilon_0^2\epsilon_r^2}} = \left\{1 + \frac{\sigma^2}{\omega^2\epsilon_0^2\epsilon_r^2}\right\}^{\frac{1}{2}} = 1 + \frac{\sigma^2}{2\omega^2\epsilon_o^2\epsilon_r^2} + \ldots$$

$$\beta = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{1 + \frac{\sigma^2}{2\omega^2\epsilon_o^2\epsilon_r^2} + \cdots + 1\right\}^{\frac{1}{2}}$$

$$\beta = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{2 + \frac{\sigma^2}{2\omega^2\epsilon_o^2\epsilon_r^2}\right\}^{\frac{1}{2}}$$

$$\beta = \omega\sqrt{2}\cdot\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{1 + \frac{\sigma^2}{4\omega^2\epsilon_o^2\epsilon_r^2}\right\}^{\frac{1}{2}}$$

$$\beta = \omega\sqrt{\mu_o\epsilon_o\epsilon_r}\left\{1 + \frac{\sigma^2}{8\omega^2\epsilon_o^2\epsilon_r^2} + \ldots\right\}$$

$$\beta = \omega\sqrt{\mu_o\epsilon_o\epsilon_r}\left\{1 + \frac{1}{8}\frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}\right\} \cong \omega\sqrt{\mu_o\epsilon_o\epsilon_r}$$

So for $\omega\epsilon_o\epsilon_r \gg \sigma$, α is very close to zero because $\frac{\sigma}{2}\sqrt{\frac{\mu_o}{\epsilon_o\epsilon_r}}$ is very small. But however, since α is now accounting for the losses in the medium, no matter how small, we want to know its value. The phase constant here is the same as what we have gotten for the lossless medium or dielectric without finite conductivity. So this case is similar to the case of low-loss transmission lines when compared. In the low loss Transmission line case, we have said that the phase constant is the same as that of a lossless transmission line, and the attenuation constant is very small. So lines can be treated like lossless lines. Only when the losses are to be calculated, then we take into account the attenuation constant α. We can do the same thing for this medium. We can say that for a medium which is low loss ($\omega\epsilon_o\epsilon_r \gg \sigma$), it can be treated like a lossless medium without conductivity for all practical purposes. Only when we are to find out the amplitude of the wave we then make use of attenuation constant α and find the resultant amplitude over some distance. With $\frac{\sigma}{2}\sqrt{\frac{\mu_o}{\epsilon_o\epsilon_r}}$, for small values of σ, α almost increases linearly with σ. The wave amplitude decays exponentially with α (by $e^{-\alpha x}$). If α varies linearly with σ, the wave decays rapidly if the value of σ changes. So any small change in σ will affect the amplitude substantially over some distance. We can also define the intrinsic impedance η for this medium as: $\eta = \sqrt{\frac{j\omega\mu}{j\omega\epsilon}}$, which now becomes $\sqrt{\frac{j\omega\mu}{\sigma + j\omega\epsilon}}$, and for a good dielectric, $\omega\epsilon_o\epsilon_r \gg \sigma$. Now the intrinsic impedance is no more a real quantity when we have finite conductivity in the medium. For a free space, recall that $\eta_o = \eta$ = 377 or $120\pi\Omega$ and the medium was appearing like a resistance for a wave propagating in the medium. It felt as if it was going inside a resistance. This is not true when we have finite conductivity. In a medium with finite conductivity, it sees an impedance which is a complex quantity and this is precisely what we saw for the transmission line case also, when R and G were not zero, the characteristic impedance of the Transmission line became complex.

8.6 ExerciseList

Ex. 149 — Derive the expressions showing finite conductivity and zero conductivity in a medium using Maxwell's equation relating to Ampere's law?

Ex. 150 — The metal, Silver, is known to have the highest electrical conductivity known to man. Estimate the frequency range in which this metal can behave like a conductor or a dielectric medium given? [σ = 6.3 x 10^7 s/m, ϵ = 8.854 x 10^-12 Fm^-1, ϵ_r = 0.99]

Ex. 151 — How can the conductivity of a medium be accounted for using the concept of a complex dielectric constant? How does the medium's behaviour depend on frequency?

Ex. 152 — Using the necessary equations, show how electrical circuits analyzed in this chapter behave in an opposite manner to how normal circuits behave i.e high conductivity coming with lesser ohmic losses?

Ex. 153 — What is the effect of the attenuation constant (α) in a medium of finite conductivity?

Ex. 154 — Explain the significance of the complex dielectric constant in the context of a medium with finite conductivity. How does it relate to the concept of relative permittivity or dielectric constant?

Ex. 155 — Discuss the relationship between conductivity and power loss in dielectric and conductor materials. How does the loss tangent serve as a measure of power loss in a dielectric medium? Provide a qualitative explanation.

Ex. 156 — Why is the loss tangent crucial when evaluating the suitability of a material for use as a dielectric?

Ex. 157 — Explain the conditions under which a dielectric is considered "good" based on the value of the loss tangent?

Ex. 158 — Discuss the typical range of values for the loss tangent in practical applications.

Ex. 159 — Describe how the loss tangent is influenced by the frequency of operation and why it is considered a frequency-dependent quantity. Discuss its importance in characterizing dielectric materials for specific frequencies of operation?

Ex. 160 — Explain the concept of loss tangent ($\tan(\delta)$) and its significance in the context of dielectric materials.

Ex. 161 — Define loss tangent and provide the formula for calculating it in terms of conductivity (σ), angular frequency (ω), permittivity of free space (ϵ_0), and relative permittivity (ϵ_r).

Ex. 162 — For a material with a relative permittivity (ϵ_r) of 1 and a conductivity (σ) equal to 5.6×10^7 S/m, discuss the influence of frequency on the behavior of the material. Explain under what conditions this material can be considered as more conductive or more dielectric.

Ex. 163 — Investigate a material with $\epsilon_r = 1$, $\sigma = 5.6 \times 10^7$ S/m, and a permittivity of free space (ϵ_0) of 8.854×10^{-12} F/m. Determine the frequency range where this material exhibits characteristics of a dielectric and the range where it behaves as a conductor. Provide numerical values and explain the underlying principles.

Chapter 9

Wave Propagation in a Medium

9.1 Objective

1. To understand the fundamental principles of wave propagation.
2. To analyze the behavior of waves in different types of media.
3. To be able to predict wave phenomena.

In the previous chapter, we discussed Wave Propagation in a medium with finite conductivity. Now, we are going to be investigating the wave propagation in a medium that is a good conductor. (i.e. $\sigma \gg \omega\ \epsilon_o\ \epsilon_r$, and the conduction current is much or far larger compared to the displacement current).

Recall,

$$\beta = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\sqrt{1+\frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}}+1\right\}^{1/2}$$

$$\alpha = \omega\sqrt{\frac{\mu_o\epsilon_o\epsilon_r}{2}}\left\{\sqrt{1+\frac{\sigma^2}{\omega^2\epsilon_o^2\epsilon_r^2}}-1\right\}^{1/2}$$

This is what we saw earlier, but we want to take a different approach to the problem in a simpler way.

$$\gamma = j\omega\sqrt{\mu\epsilon_o}\left\{\sqrt{\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}}\right\} \tag{9.1}$$

Taking the square of both sides, $j^2 = -1$

and $\gamma^2 = -\omega^2\mu\epsilon_o\left\{\epsilon_r - j\dfrac{\sigma}{\omega\epsilon_o}\right\}$

multiplying the numerator and denominator of the RHS by $j\omega\epsilon_o$

$$\gamma^2 = \frac{-\omega^2\mu\epsilon_o}{j\omega\epsilon_o}\cdot j\omega\epsilon_o\left\{\epsilon_r - j\frac{\sigma}{\omega\epsilon_o}\right\}$$

$$\gamma^2 = -\{-j\}\omega\mu\cdot\{j\omega\epsilon_o\epsilon_r - j^2\sigma\}$$

$$\gamma^2 = j\omega\mu\epsilon_o\{\sigma + j\omega\epsilon_o\epsilon_r\}$$

$$\gamma = \sqrt{j\omega\mu\{\sigma + j\omega\epsilon_o\epsilon r\}}$$

But if,

$$\sigma \gg \omega\epsilon_o\epsilon r$$

Then

$\gamma \approx \sqrt{j\omega\mu\sigma}$

Recall from De Moivre's Theorem;

$$\sqrt{j} = \sqrt{e^{j\frac{\pi}{2}}} = e^{j\frac{\pi}{4}} = \frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$$

Therefore,

$$\gamma = \sqrt{\omega\mu\sigma}\{\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\}$$

Then we can deduce from the above equation that:

$$\alpha = \sqrt{\frac{\omega\mu\sigma}{2}} \text{ and } \beta = \sqrt{\frac{\omega\mu\sigma}{2}}$$

Therefore, $\alpha \approx \beta$

This means that for this medium, the attenuation constant α and phase constant β are almost equal.

If we recall earlier, we had established that the Transmission line is characterized as a Lossy line whenever we had the Phase constant comparable to the Attenuation constant. And we happen to have a similar situation in the wave propagation through a medium that is a good conductor as shown above. We can therefore say that for a good conductor, the medium is like a very lossy transmission line since the attenuation constant is equal to the phase constant.

This means that when the Wave propagates in this medium, the amplitude dies down very rapidly in the direction of propagation. So when the wave tries to enter the conducting medium, Its amplitude reduces very rapidly, since $\alpha = \beta$. So a good conductor is visualized as an extremely lossy transmission line.

However, there is a difference between a Lossy Transmission Line and a good conducting medium. In a lossy line, the power gets lost in the ohmic resistance of the line. So there is a loss of power when the wave is propagated along the Transmission line, but in the case of a good conducting medium, when the wave is attenuating, the power is not necessarily getting lost in the conductor. There is even no power loss when the thickness is small and consequently, we are going to look at the case of $\alpha = \infty$, over zero distance, the wave dies down as it tries to penetrate this medium.

So in the good conducting medium, the wave cannot penetrate the medium because α is very large due to large attenuation. This does not mean that the power is getting lost in the heating medium. Something else is happening, which we are not going to discuss now. However at this point, we can assume that a good conductor is like a lossy transmission line, and that is the reason we can have the quantity $\alpha = \beta$ and the new amplitude will vary as $e^{-\alpha x}$. So if we plot the wave amplitude as a function of x when we travel a distance $x = \frac{1}{\alpha}$, the wave amplitude will die down to $\frac{1}{e}$ of its initial value.

So we can say that when the wave tries to propagate in this conducting medium effectively, the propagation of the wave is over the distance $\frac{1}{\alpha}$, as beyond $\frac{1}{\alpha}$ distance the field is very

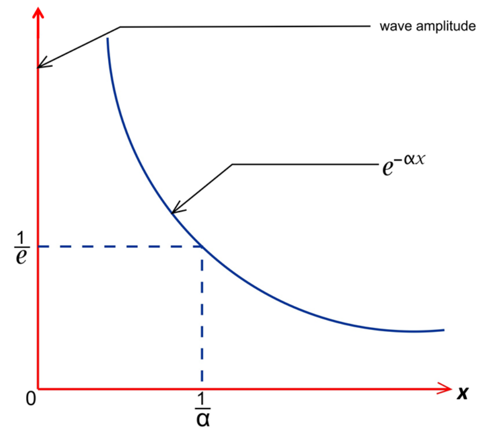


Figure 9.1: Wave amplitude as a function of x

small. As a rule of thumb, we can say that the wave will propagate over a distance of $\frac{1}{\alpha}$ into the medium. This is the effective length over which the propagation of the wave takes place from the beginning of the medium. We can therefore say that any electromagnetic wave cannot go deep into a conducting medium. It penetrates a little bit from the surface of the conductor. i.e. if the energy in an electromagnetic wave tries to go into a conducting medium, it goes only a short distance from the surface of the conducting medium, and deeper into the conducting medium, the fields are extremely small. This creates a surface phenomenon on the surface of the conductor. So if $\frac{1}{\alpha}$ is the effective length over which the wave propagates, we can say that beyond $\frac{1}{\alpha}$, the field does not exist, so only $\frac{1}{\alpha}$ thickness of the conductor will matter in the conduction of electromagnetic waves. Beyond $\frac{1}{\alpha}$ the field is assumed to be too small and can be neglected. Hence only $\frac{1}{\alpha}$ of the thickness of the conducting medium is what will decide the propagation characteristics of the electromagnetic wave. Deeper in the conducting medium is of no relevance as the field dies down rapidly as you go inside the conducting medium.

9.2 Effective Depth

Effective depth is always represented by δ and it is given by:

$$\delta = \frac{1}{\alpha}$$

Recall,

$$\alpha = \sqrt{\frac{\omega\mu\sigma}{2}}$$

Therefore,

$$\delta = \frac{1}{\sqrt{\frac{\omega\mu\sigma}{2}}} = \sqrt{\frac{2}{\omega\mu\sigma}}$$

And $\omega = 2\pi f$,

$$\delta = \sqrt{\frac{2}{2\pi f\mu\sigma}} = \sqrt{\frac{1}{\pi f\mu\sigma}}$$

Effective Depth,

$$\delta = \sqrt{\frac{1}{\pi f\mu\sigma}}$$

This can be seen to be inversely related to $\sqrt{f}$,$\sqrt{\mu}$, and $\sqrt{\sigma}$, therefore, when there is an increase in frequency or σ, the depth of penetration of the electromagnetic wave into the conducting medium decreases. With $\sigma = \infty$(i.e. Ideal Conductor), $\delta = 0$, that is the wave will not penetrate the medium at all. Similarly, if we go into very high arbitrary frequencies, the depth of penetration becomes extremely small. Let's take typical numbers just to get a feel for what depth we have when we operate at low frequencies.

Let us take an arbitrary medium of conductivity $\sigma = 10^7\Omega^{-1}m^{-1}$(this is the typical value for a good conductor), $f = 100Mhz$(typical value for radio transmission).

Recall,

$\delta = \sqrt{\frac{1}{\pi f\mu\sigma}} = \sqrt{\frac{1}{\pi \times 100 \times 10^6 \times 4\pi \times 10^{-7} \times 10^7}} = 1.6 \times 10^{-5}m = 16\mu m$

So, for a medium that is a good conductor with the conductivity of the order of $10^7\Omega^{-1}m^{-1}$ (copper has $5.6 \times 10^7\Omega^{-1}m^{-1}$), like copper or silver and a frequency of 100MHz to propagate our electromagnetic wave, It will not go deeper than $16\mu m$ from its surface. So if we go to frequencies of GHz used for satellites or microwave frequencies, the effective depth becomes even smaller. So effectively, for a good conductor, if we take a frequency which is a typical radio frequency, the depth of penetration is of few microns to tens of microns. So basically, this propagation is taking place on the skin of the material of the order of a few tens of microns. This is the reason we call this effective depth the **SKIN DEPTH OF THE MATERIAL**.

Skin Depth is a parameter that tells how electromagnetic energy is going to exist on a good conducting material or medium. Beyond this depth, the propagation does not matter as the field would have died out. So if you have a high-frequency component, since the energy is going to lie on the surface at a few microns or tens of microns, we have to make that layer very good, because any perturbation in the properties of that layer will affect the propagation of the electromagnetic wave. So if we have a conductor and want to send some high-frequency signal through this conductor, the field is essentially going to lie only on the surface of the conductor in the skin depth.

The current excited in this conductor will essentially be confined to the skin depth of the conductor. That is the region in which the field is going to exist and the current will flow in that skin depth.

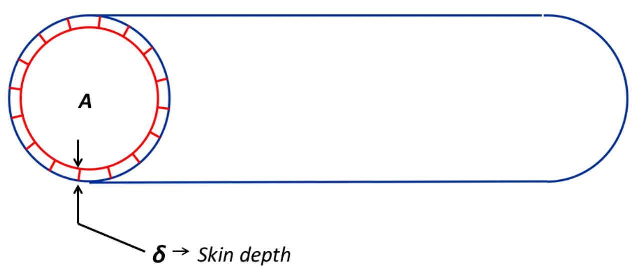


Figure 9.2: A conductor

The region marked A does not contribute to the flow of the current. So even if we make the conductor hollow, the propagation of the energy along the conductor will not be affected since the skin depth layer remains intact. To have good propagation, the property of the small thin layer should be very good, as any perturbation here will affect the propagation of the electromagnetic wave significantly.

For this reason, if we go to microwave frequency components, where the frequency is very high and the skin depth is very small in microns, the surface has to be perfect as any imperfection will affect the propagation of the electromagnetic wave significantly.

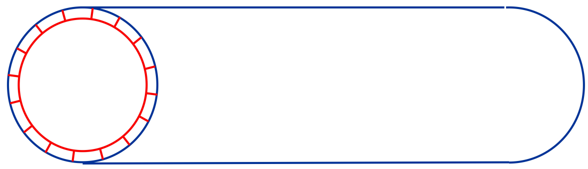


Figure 9.3: An hollow conductor

Let us say we have a conducting medium of finite conductivity as shown above to which we are sending signals along. Considering a certain area of the cross-section for this conductor, at extremely low frequencies, the current is distributed uniformly across the cross-sectional area of this conductor. So we can find out the area of the cross-section and then find out what is the resistance per unit length since the resistivity of the conductor is known.

However, if we increase the frequency, the concept of skin depth starts coming into the picture. The current now gets more and more confined to its surface which means, the area of cross-section over which the current flows reduces. As a result, the resistance of the conductor starts increasing as we go to higher frequencies. So because of the skin depth, the resistance now becomes a function of frequency. For a given conductivity as the frequency increases, the skin depth reduces, and as the skin depth reduces, the area of the cross-section over which the current flows reduces, and because of that, the resistance of the conductor increases.

So what we find out is that the resistance which we treated like the property of the wire is not so when you go to higher frequencies. As the frequency gets higher, the same conductor starts showing higher resistance, and the wave propagation is essentially confined to the surface of the conductor. For the intrinsic impedance of the medium, $\eta_c = \sqrt{\dfrac{j\omega\mu}{\sigma + j\omega\epsilon_o\epsilon r}}$

And as $\sigma \gg \omega\epsilon_o\epsilon_r$

$$\eta_c \approx \sqrt{\frac{j\omega\mu}{\sigma}} = \sqrt{j}\cdot\sqrt{\frac{\omega\mu}{\sigma}}$$

Recall,

$$\sqrt{j} = \frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}$$

η_c becomes;

$$\eta_c = \sqrt{\frac{\omega\mu}{\sigma}}\left\{\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}}\right\}$$

$\eta_c = \sqrt{\dfrac{\omega\mu}{\sigma}} < 45^o$ in terms of magnitude and phase.

Therefore, the Electric and Magnetic fields are related to the intrinsic impedance of the medium.

For a Dielectric Medium

$$\frac{|E|}{|H|} = \eta,$$

So in this case, $\dfrac{|E|}{|H|} = \eta_c = \sqrt{\dfrac{\omega\mu}{\sigma}} < 45^o$

So we see that the electric and magnetic fields are still perpendicular to each other but there is a time lag in the magnetic field compared to the electric field. The ratio $\frac{|E|}{|H|}$ is now dependent on frequency and not only on the medium parameter, as it used to be in the case of the dielectric medium. So when we talk about a good conducting medium, important things happen.

When the wave tries to penetrate, the wave is restricted to a very short distance called the conducting depth. Also, the electric and magnetic fields are not in the time phase anymore, there is a phase difference in time which is 45°.

So the wave propagation in a dielectric medium and a conducting medium are completely different. In a dielectric medium, the wave can go deep into the medium, and the electric and magnetic fields are in the time phase. The wave amplitude reduction is very small. Whereas in a conducting medium, the wave amplitude drops very rapidly, it can't go deep into the medium, and also, the electric and magnetic fields have a phase difference in time.

If we take a general medium which is a combination of a dielectric and a conductor, then all these relationships are going to be complex as we don't have any approximations. For those media, for which conduction current is comparable to displacement current, we have to do a rigorous analysis. For the two extreme cases i.e. good dielectric and good conductor, this approximation can be made easily and one can understand the wave propagation in a more comprehensible manner than in the general medium. Now we go into another aspect, *which is, what will be the velocity of the wave as it tries to move in this medium?* Let us take a medium that is a good dielectric and assume the losses are very small, so we can treat the dielectric almost like a lossless dielectric. Then we ask the question *"How is the wave going to propagate in this medium?"* If the medium is lossless, the propagation constant is purely imaginary, and then any field E can be written as:

$$\bar{E} = \bar{E}_o e^{-j\beta x} e^{j\omega t} = \bar{E}_o e^{j\{\omega t - \beta x\}}$$

This represents a wave travelling in a positive direction. Now if we stand on a particular location in space i.e. x is constant and look at the wave which is moving in space, the phase of the wave will vary with respect to time (with wt). That is, the phase is linearly increasing as a function of time. However, suppose we hold on to a particular phase point on the wave, say we are holding on to A since the wave is moving, we will move along with the wave with the velocity at which the wave is travelling if we don't want to lose that phase point we are holding on to.

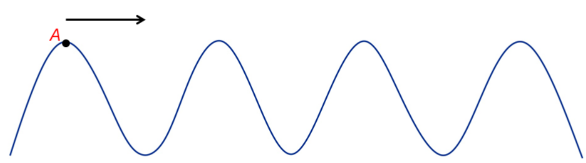


Figure 9.4: A wave form

So for an observer holding on to A, the phase does not change anymore. Hence, to the observer, the phase appears to be constant as a function of time. So by making the phase of the wave($wt - \beta x$) constant as a function of time, whatever speed the observer needs to move at to keep this phase constant is the velocity of the wave. So basically, we make the phase constant as a function of time to get the velocity of the wave.

$$\omega t - \beta x = constant$$

Differentiating, we have

$$\omega - \beta \frac{dx}{dt} = 0$$

Meanwhile,

$$\frac{dx}{dt} = velocity = \frac{\omega}{\beta}$$

9.3 Phase Velocity

The velocity we calculated from the phase being stationary, is called the **PHASE VELOCITY** of the wave. That means a constant phase point on the wave is moving with this velocity. This concept of naming the velocity is necessary because we are going to encounter another velocity of the wave with which the energy travels.

However, when we talk about waves moving in some arbitrary direction, this concept of saying that the phase is moving with this velocity comes into play because the energy might move in some other velocity. We call the velocity of this constant phase, the phase velocity v_p.

$$v_p = \frac{\omega}{\beta}$$

For a given frequency in any medium, we find out what its phase constant β is, then $\frac{\omega}{\beta}$ gives a quantity which is the velocity that is called Phase Velocity.

Essentially, the knowledge of propagation constant ($\gamma = \alpha + j\beta$) is important because that tells the phase of the wave when it travels through the medium. For a travelling wave, this is very straightforward, we get $\omega t - \beta x = constant$. However, the concept of phase velocity can be extended to any arbitrary wave propagation. For example, suppose we have two waves that are travelling in opposite directions, which forms a standing wave, we can still define the total phase which is a combination of space and time. By making that quantity stationary as a function of time, we get the expression for the phase velocity of the wave. So the concept is very general, Though for travelling wave it is very simple, and so we easily get its phase velocity as $v_p = \frac{\omega}{\beta}$. We will use this relation often when we go to the propagation of electromagnetic waves.

9.4 The Case of a Pure Dielectric

As we have seen earlier, if we take a dielectric medium, that is a PURE DIELECTRIC(ϵ_r)

The phase constant will be:

$$\beta = \omega\sqrt{\mu_o\epsilon_o\epsilon_r}$$

Let's consider free space (i.e. $\epsilon_r = 1$), then for free space, $\beta = \omega\sqrt{\mu_o\epsilon_o}$. Since phase velocity $v_p = \frac{\omega}{\beta}$, phase velocity in free space will become;

$$v_p = \frac{\omega}{\omega\sqrt{\mu_o\epsilon_o}} = \frac{1}{\sqrt{\mu_o\epsilon_o}} = \frac{1}{\sqrt{4\pi \times 10^{-7} \times \frac{1}{36\pi} \times 10^{-9}}} = 3 \times 10^8 m/s.$$

At this point, we noticed that the velocity of the wave when it is travelling in free space is the velocity of light in free space. Since light is a transverse electromagnetic wave, it is travelling with a velocity given by the phase velocity, $v_p = 3 \times 10^8 m/s$. So any wave, not only light which is a transverse electromagnetic wave travelling in free space, will have a velocity equal to $3 \times 10^8 m/s$. v_p is the velocity in vacuum denoted by c. We can therefore represent the velocity of the wave in a dielectric medium as;

$$v_p = \frac{\omega}{\omega\sqrt{\mu_o\epsilon_o\epsilon_r}} = \frac{1}{\sqrt{\mu_o\epsilon_o}} \times \frac{1}{\sqrt{\epsilon_r}} = \frac{c}{\sqrt{\epsilon_r}}$$

So we see that for $\epsilon_r \neq 1$ when an electromagnetic wave travels in a medium, its phase velocity is not exactly equal to $3 \times 10^8 m/s$ because of ϵ_r. If $\epsilon_r > 1$, then the phase velocity will always be less than c (that is, $3 \times 10^8 m/s$). Where c is the velocity of light in free space.

So the velocity of a wave in any dielectric medium is always less than the velocity in free space. We also know that the refractive index of the medium is defined as the ratio;

$$\textbf{Refractive Index}(\eta) = \frac{\text{Velocity of light in a vacuum}}{\text{Velocity of light in a medium}} \tag{9.2}$$

Therefore,

$$\frac{\text{Velocity of light in a vacuum}}{\text{Velocity of light in a medium}} = \frac{c}{v_p} = \sqrt{\epsilon_r}$$

So the Refractive index of a lossless medium is given by $\eta = \sqrt{\epsilon_r}$. So whenever we have a medium that is pure dielectric or good dielectric, the refractive index and dielectric constant are related by this relationship. The electromagnetic wave always slows down in a medium compared to a vacuum, when the medium has a dielectric constant greater than 1. Let us take a look at what happens when we apply the relationship to a conductor;

$$\beta = \sqrt{\frac{\omega\mu_o\sigma}{2}} \text{ and } v_p = \frac{\omega}{\beta} = \frac{\omega}{\sqrt{\frac{\omega\mu_o\sigma}{2}}} = \sqrt{\frac{2\omega}{\mu_o\sigma}}$$

So the velocity we had for a dielectric medium $\frac{C}{\sqrt{\epsilon_r}}$ was not a function of frequency, sine ϵ_r is not a function of frequency, so v_p is not frequency dependent for dielectric material. i.e. all frequency starts with the same speed in the medium. However, for a medium which is a conductor $v_p = \sqrt{\frac{2\omega}{\mu_o\sigma}}$, even if medium properties are not varying as a function of frequency, $\beta = \sqrt{\frac{\omega\mu_o}{2}}$ and v_p are functions of frequency.

This medium is thus called a DISPERSIVE MEDIUM because the velocity is varying as a function of frequency. Hence, all frequencies do not travel at the same speed and because of that, we have a dispersion in the medium, so we have an important contribution to draw. We got the phase velocity and realize that for a dielectric material, the phase velocity depends on the dielectric constant of the medium, and for $\epsilon_r > 1$, the v_p is always less than the phase velocity in vacuum. Though the wave slows down, it is not a dispersive medium since all frequencies travel at the same speed. We find that the refractive index, $\eta = \sqrt{\epsilon_r}$.

However, for the conductor, v_p is a function of frequency i.e. the medium becomes a dispersive medium. So a pure dielectric medium is a Non Dispersive medium, and a good conductor is a dispersive medium. This essentially summarizes the propagation of a transverse electromagnetic wave in an unbound medium. Before we proceed further, we have to investigate first the power flow associated with the electromagnetic waves and then treat the more complex analysis of the propagation of

Electromagnetic waves in a bound medium or in a medium with interfaces that are dielectric or conducting interfaces.

9.5 ExerciseList

Ex. 164 — How does real characteristic impedance and real propagation constant impact our ability to determine whether the line is lossy or lossless?

Ex. 165 — On a moderately lossy transmission line, how much does the reflection coefficient at the load point affect the standing wave's amplitude variation?

Ex. 166 — In what way does the purely imaginary propagation constant in a lossless transmission line differ from the real propagation constant in a very lossy line and what implication does this have for the representation of a wave phenomena in the transmission line?

Ex. 167 — How is wave propagation affected by the medium's properties, i.e. density and elasticity?

Ex. 168 — What is the relationship between properties of a medium and the velocity of a wave?

Ex. 169 — How does wave interaction with the medium affect its propagation?

Ex. 170 — What is the Skin depth of a material and how does it affect wave propagation?

Ex. 171 — Under what factors does a wave lose energy or attenuate when propagating through a medium?

Ex. 172 — What is the relationship between Refractive Index and dielectric constant?

Ex. 173 — What is a pure dielectric and how does it differ from a conductor?

Ex. 174 — What is a dispersive medium and why is it called a dispersive medium?

Ex. 175 — What is the requirement to have good wave propagation in a conductor?

Ex. 176 — What is the difference between wave propagation in a dielectric medium and a conducting medium?

Chapter 10

Power Flow in an Electromagnetic Wave

10.1 Objective

1. To understand the power flow associated with electromagnetic fields from Maxwell's equations
2. The importance of poynting factors in power flow

10.1.1 INTRODUCTION

Up till now, we have seen that a time-varying electric and magnetic field constitutes a wave phenomenon. Thus this wave requires some power or energy to flow. In this chapter, we look at the power flow associated with electromagnetic waves. We do some derivation starting with Maxwell's Equation and find out the power flow associated with a time-varying electromagnetic field.

We have also investigated a wave which is called uniform plane wave i.e. a wave propagation in an unbound medium. However, in developing the power flow equations associated with electromagnetic waves, we would do a general analysis and not restrict our focus to uniform plane waves, at the end we find out the power flow associated with uniform plane waves which we have discussed in chapter 8 and chapter 9. So whenever we are at the beginning of analysis in electromagnetics, we go back to Maxwell's Equations and find answers that are consistent with Maxwell's equation. The same thing we do have if we are asked the question, *"What power do we get for the power flow associated with an electromagnetic field?"*

From Maxwell's Equation, we have;

$$\nabla \times \bar{E} = -\frac{\delta \bar{B}}{\delta t}. \tag{10.1}$$

If the permeability of the system is not a function of time,

$$\vec{B} = \mu \vec{H} \tag{10.2}$$

$$\frac{\delta \bar{B}}{\delta t} = -\mu \frac{\delta \bar{H}}{\delta t} \tag{10.3}$$

Therefore;

$$\nabla \times \bar{E} = -\mu \frac{\delta \bar{H}}{\delta t} \tag{10.4}$$

and similarly, if permittivity (ϵ) is not a function of time,

$$\vec{D} = \epsilon \vec{E} \tag{10.5}$$

$$\nabla \times \bar{H} = \bar{J} + \frac{\delta \bar{D}}{\delta t} = J + \epsilon \frac{\delta \bar{E}}{\delta t} \tag{10.6}$$

We start with these two basic equations and then try to investigate the power flow associated with electric and magnetic fields. We essentially make use of the vector identities and then try to find out the meaning of some expressions we would get from using the vector identities.

Recall,

$$\nabla \cdot (\bar{A} \times \bar{C}) = C \cdot (\nabla \times \bar{A}) - \bar{A} \cdot (\nabla \times \bar{C}) \tag{10.7}$$

For arbitrary vectors $\bar{A}$ and $\bar{C}$,

Let $\bar{A}$ be the electric field $\bar{E}$ and $\bar{C}$ be the Magnetic Field $\bar{H}$, substituting for $\bar{A}$ as $\bar{E}$ and $\bar{C}$ as $\bar{H}$ into equation10.7, we have;

$$\nabla \cdot (\bar{E} \times \bar{H}) = \bar{H} \cdot (\nabla \times \bar{E}) - \bar{E} \cdot (\nabla \times \bar{H}) \tag{10.8}$$

Now Substituting for $\nabla \times \bar{E}$ and $\nabla \times \bar{H}$, we have;

$$\nabla \cdot (\bar{E} \times \bar{H}) = \bar{H} \cdot (-\mu \frac{\delta \bar{H}}{\delta t}) - \bar{E} \cdot (\bar{J} + \epsilon \frac{\delta \bar{E}}{\delta t}) \tag{10.9}$$

Recall from vector identities, we have that;

$$\begin{aligned}\frac{\delta}{\delta t}(\bar{A} \cdot \bar{C}) &= \bar{A} \cdot \frac{\delta \bar{C}}{\delta t} + \bar{C} \cdot \frac{\delta \bar{A}}{\delta t} \frac{\delta}{\delta t}(\bar{A} \cdot \bar{A}) \\ &= \bar{A} \cdot \frac{\delta \bar{A}}{\delta t} + \bar{A} \cdot \frac{\delta \bar{A}}{\delta t} \\ &= 2\bar{A} \cdot \frac{\delta A}{\delta t}\end{aligned}$$

Making $\bar{A} \cdot \frac{\delta A}{\delta t}$ the subject of formula, we have;

$$\begin{aligned}\bar{A} \cdot \frac{\delta \bar{A}}{\delta t} &= \frac{1}{2} \frac{\delta}{\delta t}(\bar{A} \cdot \bar{A}) \\ &= \frac{1}{2} \frac{\delta}{\delta t}|A|^2\end{aligned} \tag{10.10}$$

Having refreshed ourselves with these identities, we can deduce from equation10.9, that;

$$\nabla \cdot (\bar{E} \times \bar{H}) = -\mu \left(\bar{H} \cdot \frac{\delta \bar{H}}{\delta t}\right) - \bar{E} \cdot \bar{J} - \epsilon \left(\bar{E} \cdot \frac{\delta \bar{E}}{\delta t}\right)$$

Therefore, relating equation10.10, to equation10.9, we have;

$$\nabla \cdot (\bar{E} \times \bar{H}) = -\frac{\mu}{2} \frac{\delta}{\delta t}|\bar{H}|^2 - \frac{\epsilon}{2} \frac{\delta}{\delta t}|\bar{E}|^2 - \bar{E} \cdot \bar{J} \tag{10.11}$$

We can investigate this expression over a closed surface or volume to have some meaning associated with these quantities.

$$\begin{aligned}\iiint_v \nabla \cdot (\bar{E} \times \bar{H}) dv &= \iiint_v -\frac{\mu}{2} \frac{\delta}{\delta t}|\bar{H}|^2 dv \\ &\quad - \iiint_v \frac{\epsilon}{2} \frac{\delta}{\delta t}|\bar{E}|^2 dv - \iiint_v \bar{E} \cdot \bar{J} dv\end{aligned}$$

If we assume that the volume enclosed is not varying as a function of time i.e. fixed volume, that is to say, that the field varies with time but space does not vary with time. We can take $\frac{\delta}{\delta t}$ out of the integration and apply the divergence theorem on $\int_v \nabla \cdot (\bar{E} \times \bar{H})$ to convert it to a closed surface integral.

Thus,

$$\oint (\bar{E} \times \bar{H}) \cdot \bar{da} = -\frac{\delta}{\delta t} \int_v \frac{\mu}{2} |\bar{H}|^2 dv - \frac{\delta}{\delta t} \int_v \frac{\epsilon}{2} |\bar{E}|^2 dv - \int_v \bar{E} \cdot \bar{J} dv \tag{10.12}$$

We now substitute $\bar{J} = \sigma \bar{E}$, into $\bar{E} \cdot \bar{J}$ to get; $\bar{E} \cdot \bar{J} = \bar{E} \cdot \sigma \bar{E} = \sigma \bar{E} \cdot \bar{E} = \sigma |\bar{E}|^2$.

So,

$$\oint (\bar{E} \times \bar{H}) \cdot \bar{da} = -\frac{\delta}{\delta t} \int_v \frac{\mu}{2} |\bar{H}|^2 dv - \frac{\delta}{\delta t} \int_v \frac{\epsilon}{2} |\bar{E}|^2 dv - \int_v \sigma |\bar{E}|^2 dv \tag{10.13}$$

Each term in the above expression has a physical significance.

$\int_v \frac{\mu}{2} |\bar{H}|^2 dv$ represents the total magnetic energy stored in the volume V hence the term $-\frac{\delta}{\delta t} \int_v \frac{\mu}{2} |\bar{H}|^2 dv$ gives the rate of change of magnetic energy stored in the volume(The negative sign indicates a decrease).

Similarly, $\int_v \frac{\epsilon}{2} |\bar{E}|^2 dv$ represents the total electrical energy stored in volume V hence the term $-\frac{\delta}{\delta t} \int_v \frac{\epsilon}{2} |\bar{E}|^2 dv$ gives the rate of change of electrical energy stored in the volume(Again the negative sign indicates a decrease).

While the term $-\int_v \sigma |\bar{E}|^2 dv$ represents the total ohmic loss in the volume due to the finite conductivity of the medium.

So if we have an amount of energy enclosed in a surface, these 3 total losses combined must be equal to the total outward energy going out from the closed surface (s). Since there is no other mechanism for energy consumption, from the law of conservation of energy, we get that $\oint (\bar{E} \times \bar{H}) \cdot \bar{da}$ must represent the net flow of energy coming out from the closed surface. So $\oint (\bar{E} \times \bar{H}) \cdot \bar{da}$ (the net power outflow from a closed surface) gives the net power flow of the electric and magnetic field from that closed surface. This is called **Poynting theorem**[1]poynting theroem.

10.2 Poynting theorem and Poynting vector

The Poynting theorem states that the cross-product of electric and magnetic fields integrated over a closed surface gives the total power flow from the closed surface. Now we have seen the quantity which represents the total power flow, then $\bar{E} \times \bar{H}$ is essentially the power density on the surface. So that when this power density is integrated over an area, it gives the total power flow from that surface area. It should be kept in mind that, even though $\oint (\bar{E} \times \bar{H}) \cdot \bar{da}$ gives the net power flow from a closed surface, saying that $\bar{E} \times \bar{H}$ should give the power density at any point on the surface, is only an arbitrary definition. The theorem does not state that $\bar{E} \times \bar{H}$ gives the power density (or power flow per unit area) on the surface. Rather it states that the total power flows out of a closed surface is given by $\oint (\bar{E} \times \bar{H}) \cdot \bar{da}$. So saying $\bar{E} \times \bar{H}$ is the power density which is true for every point on the surface of a sphere(due to symmetry), might not be true in some other cases.

Hence, it is arbitrary we take power density from here. It so happens that most of the time in practice, this arbitrary definition gives the power density correctly. But, it should be said that, if knowing $\bar{E} \times \bar{H}$, one already knows the power density at every point on the surface of the volume, this statement is not correct. There are many special cases where $\bar{E} \times \bar{H}$ may give power flow where there is no power flow at that point. So while using $\bar{E} \times \bar{H}$ as power density, one should be very careful. In most practical situations, however, the arbitrary definition that $\bar{E} \times \bar{H}$ gives power density at any location normally is valid.

So we have an important quantity called a power flow density $\bar{P} = \bar{E} \times \bar{H}$ where $\bar{P}$ is called a POYNTING VECTORpoynting vector for these fields. Poynting vector is an important concept in the electromagnetic wave as it tells the power flow associated with any particular point in space and also tells in which direction the power is flowing. The first thing we note here is that if you have an electric and magnetic field, the Poynting vector is the direction perpendicular to both $\bar{E}$ and $\bar{H}$. So if $\bar{P}$ has to be non-zero in terms of power flow, then $\bar{E}$ and $\bar{H}$ should not be parallel to each other. In a case where $\bar{E}$ and $\bar{H}$ are parallel to each other, then $\bar{E} \times \bar{H}$=0 and this connotes that there will be no net power flow associated with this. So only when the component of the electric and magnetic fields are perpendicular to each other would they contribute to power flow and the direction of power flow is perpendicular to $\bar{E}$ and $\bar{H}$ fields. It should be noted that for us to have a power flow, $\bar{E}$ and $\bar{H}$ must cross each other. Whenever $\bar{E}$ and $\bar{H}$ cross each other, there is a possibility of power flow. We use the word possibility here because this quantity $\bar{E} \times \bar{H}$ is telling you the so-called instantaneous power if we know the values of $\bar{E}$ and $\bar{H}$. At that instant of time at some point in space, we can always find out $\bar{E} \times \bar{H}$ and we get $\bar{P}$ which will give the Poynting vector at that instant of time. It is possible that even if $\bar{P}$ is finite at some instance of time, there may be no net power flow over long periods. That is in time average sense, there may not be power flow associated with that system. So Poynting Vector $\bar{P}$ which we defined as $\bar{E} \times \bar{H}$ serves the purpose of defining the power flow. But in Practical Systems, a more useful quantity would be the time average value of the Poynting Vector, because the Poynting vector at some instant of time may be negative and that is like getting negative power. Of course, when dealing with space, we can say negative power flow means direction change. All these complications come in simply if we use $\bar{E} \times \bar{H}$ to get $\bar{P}$ since $\bar{P}$ can be positive or negative and $\bar{P}$ can be a complex quantity too if there is a phase difference between $\bar{E}$ and $\bar{H}$ with time. So what we do is, we try to get the time average value of the Poynting vector and it is a much more useful quantity for finding out if there is a net flow of power associated with the Electric and Magnetic fields.

As we have seen earlier, we are interested only in the analysis of time-harmonic fields, so we assume that the electric and magnetic fields are varying sinusoidally as a function of time. The only thing now is when there is a phase difference between $\bar{E}$ and $\bar{H}$, we then ask what is the average power flow associated with $\bar{E}$ and $\bar{H}$ in that case.

Now we define the general time-varying electric and magnetic fields which vary as a function of space and time as;

$\bar{E} = \bar{E}_0 e^{\jmath\omega t + \jmath\phi_e}$, $\bar{H} = \bar{H}_0 e^{\jmath\omega t + \jmath\phi_h}$. We take the unit vector associated with $\bar{E}_0$ and $\bar{H}_0$ as $\hat{e}$ and $\hat{h}$, so we remember that we

[1]John Henry Poynting (Born 9, September 1852 and died 30, March 1914). He was an English physicist. He was a professor of physics at Mason Science College, from 1880 to 1900, and then at the successor institution, the University of Birmingham until his death.

He is known for the Poynting vector, Poynting Effect, Poynting's theorem Poynting was the youngest son of Thomas Elford Poynting, a Unitarian minister are only taking either the real or the imaginary part of $e^{\jmath\omega t}$ for simple analysis. It doesn't mean $\bar{E}$ cannot be complex. But for simplicity, $\bar{E} = \mathfrak{Re}(E_0 e^{\jmath\omega t} e^{j\phi e}) \cdot \hat{e}$ $\bar{H} = \mathfrak{Re}(H_0 e^{\jmath\omega t} e^{j\phi h}) \cdot \hat{h}$ where $\hat{e}$ and $\hat{h}$ gives the unit vector of the electric and magnetic field, ϕ_e and ϕ_h are the phases of the electric and magnetic fields and E_0 and H_0 are the amplitude of the electric and magnetic fields. So the real parts give an instantaneous value of the electric and magnetic field. Once this is known, we can find the Poynting vector at that instant in time.

$$\bar{E}(t) = E_0 \cos(\omega t + \phi_e)\hat{e}$$
$$\bar{H}(t) = H_0 \cos(\omega t + \phi_h)\hat{h}$$

The power flow density at that instant t is the Poynting vector $\bar{P}$ would be;

$$\bar{P} = \bar{E} \times \bar{H}$$

$$= E_0 H_0 \cos(\omega t + \phi_e) \cos(\omega t + \phi_h)(\hat{e} \times \hat{h})$$

$$= \frac{E_0 H_0}{2}(\cos(\phi_e - \phi_h) + \cos(2\omega t + \phi_e + \phi_h))(\hat{e} \times \hat{h})$$

How did we get to the last expression, we will see below. From compound angle formulas, we have

$$\cos(\omega t + \omega t) = \cos\omega t \cos\omega t - \sin\omega t \sin\omega t$$

$$= \cos^2 \omega t - \sin^2 \omega t$$

$$= \cos^2 \omega t - (1 - \cos^2 \omega t)$$

Simplifying further, we get $2\cos^2 \omega t - 1 = \cos(2\omega t)$ or $\cos^2 \omega t = \frac{\cos^2 \omega t + 1}{2}$.

$$\sin(\omega t + \omega t) = \sin\omega t \cos\omega t + \cos\omega t \sin\omega t$$

Simplifying further, we get $2\sin\omega t \cos\omega t$ or $\sin\omega t \cos\omega t = \frac{\sin 2\omega t}{2}$

$$\begin{aligned}
\bar{P} &= \bar{E} \times \bar{H} \\
&= E_0 H_0 \cos(\omega t + \phi_e) \cos(\omega t + \phi_h) \hat{e} \times \hat{h} \\
&= E_0 H_0 [(\cos\omega t \cos\phi_e - \sin\omega t \sin\phi_e)(\cos\omega t \cos\phi_h \\
&\qquad - \sin\omega t \sin\phi_h)]\hat{e} \times \hat{h} \\
&= E_0 H_0 [\cos^2 \omega t \cos\phi_e \cos\phi_h - \cos\omega t \sin\omega t \cos\phi_e \sin\phi_h \\
&\qquad - \cos\omega t \sin\omega t \sin\phi_e \cos\phi_h + \sin^2 \omega t \sin\phi_e \sin\phi_h]\hat{e} \times \hat{h} \\
&= E_0 H_0 [\cos^2 \omega t \cos\phi_e \cos\phi_h + (1 - \cos^2 \omega t) \sin\phi_e \sin\phi_h \\
&\qquad - \cos\omega t \sin\omega t (cos\phi_e \sin\phi_h + \sin\phi_e \cos\phi_h)]\hat{e} \times \hat{h} \\
&= E_0 H_0 [\cos^2 \omega t (\cos\phi_e \cos\phi_h - \sin\phi_e \sin\phi_h) \\
&\qquad + \sin\phi_e \sin\phi_h \\
&\qquad - \cos\omega t \sin\omega t (\cos\phi_e \sin\phi_h + \sin\phi_e \cos\phi_h)]\hat{e} \times \hat{h} \\
&= E_0 H_0 [\cos^2 \omega t \cos(\phi_e + \phi_h) + \sin\phi_e \sin\phi_h \\
&\qquad - \cos\omega t \sin\omega t \sin(\phi_e + \phi_h)]\hat{e} \times \hat{h} \\
&= E_0 H_0 [(\frac{\cos 2\omega t + 1}{2}) \cos(\phi_e + \phi_h) + \sin\phi_e \sin\phi_h \\
&\qquad - \frac{\sin 2\omega t}{2} \sin(\phi_e + \phi_h)]\hat{e} \times \hat{h} \\
&= E_0 H_0 [\frac{\cos 2\omega t \cos(\phi_e + \phi_h)}{2} - \frac{\sin 2\omega t \sin(\phi_e + \phi_h)}{2} \\
&\qquad + \frac{1}{2}\cos(\phi_e + \phi_h) + \sin\phi_e \sin\phi_h]\hat{e} \times \hat{h} \\
&= E_0 H_0 [\frac{1}{2}\cos(2\omega t + \phi_e + \phi_h) + \frac{1}{2}\cos(\phi_e + \phi_h) \\
&\qquad + \sin\phi_e \sin\phi_h]\hat{e} \times \hat{h} \\
&= E_0 H_0 [\frac{1}{2}\cos(2\omega t + \phi_e + \phi_h) + \frac{1}{2}\cos\phi_e \cos_h \\
&\qquad - \frac{1}{2}\sin\phi_e \sin\phi_h + \sin\phi_e \sin\phi_h]\hat{e} \times \hat{h} \\
&= E_0 H_0 [\frac{1}{2}\cos(2\omega t + \phi_e + \phi_h) + \frac{1}{2}\cos\phi_e \cos\phi_h \\
&\qquad + \frac{1}{2}\sin\phi_e \sin\phi_h]\hat{e} \times \hat{h} \\
&= \frac{E_0 H_0}{2}[\cos(\phi_e - \phi_h) + \cos(2\omega t + \phi_e + \phi_h)]\hat{e} \times \hat{h}
\end{aligned}$$

So when we take an average of this throughout the signal then $\cos(2\omega t + \phi_e + \phi_h)$ will go to zero. Positive area cancels negative area. It corresponds to a waveform having 2f as its frequency. So over a period corresponding to $T = \frac{2\pi}{\omega}$ the time average $\bar{P}_{av}$ is; $\frac{1}{T}\int_0^T \bar{P} dt$ with $T = \frac{2\pi}{\omega}$.

$$\begin{aligned}
\bar{P}_{av} &= \frac{1}{T}\int_0^T \frac{E_0 H_0}{2}\cos(\phi_e - \phi_h) dt (\hat{e} \\
&\qquad \times \hat{h}) \frac{E_0 H_0}{2}\cos(\phi_e - \phi_h)(\hat{e} \times \hat{h}) \\
&= \frac{1}{2}\mathfrak{Re}[(E_0 e^{j\omega t + j\phi_e}\hat{e}) \times (H_0 e^{-j\omega t - j\phi_h}\hat{h})] \\
&= \frac{1}{2}\mathfrak{Re}[(E_0 e^{j\omega t + j\phi_e}\hat{e}) \times (H_0 e^{j\omega t + j\phi_h}\hat{h})^*]
\end{aligned}$$

the * above represents the conjugate.

But Originally; $\bar{E} = \bar{E_0} e^{j\omega t + j\phi_e}$ and $H = \bar{H_0} e^{j\omega t + j\phi_e}$

We now have the average power to be; $\bar{P}_{av} = \mathfrak{Re}(\bar{E} \times \bar{H}^*)$. This means that if we know the electric and magnetic field in complex forms(the electric and magnetic field may not be in the time phase), in general, we can calculate $\bar{E} \times \bar{H}^*$. Taking half of the real part gives the average power flow which is associated with that electromagnetic wave. So $\bar{P}_{av} = \frac{1}{2}\mathfrak{Re}(\bar{E} \times \bar{H}^*)$ is a real quantity. All these problems we had with instantaneous power flow, $\bar{P} = \bar{E} \times \bar{H}$ complex depending on the phase between $\bar{E}$ and $\bar{H}$, have now been taken care of. This also implies that, the overall power flow which is associated with these fields at a particular location. It is possible that, at a particular location, the instantaneous Poynting vector might be negative or positive but the average power Poynting vector will always be positive and that gives the net power flow associated with these electric and magnetic fields.

This is the concept regularly used in finding out the average power flow associated with electromagnetic waves. So in determining average power flow, there are two things now essential. One is, $\bar{E}$ and $\bar{H}$ fields must be perpendicular (or have components perpendicular to each other), only then would we have a cross product which will be non-zero. Secondly, the electric and magnetic fields should not be in TIME QUADRATURE. This means the electric and magnetic fields should not differ by a phase of 90^o. If this happens, $\mathfrak{Re}(\bar{E} \times \bar{H}^*) = 0$ and then we will not have any real power flow associated with the fields at that point. So in general it is possible that if you take the electric and magnetic fields, you get a complex power. The real part of the quantity gives the net power flow at that location and the imaginary part (j conjugate) gives the power oscillating around that point. So at some instant in time, the power might be going in a certain direction, and after some time, the power will be coming back in the same direction. So the imaginary part of $E \times \bar{H}^*$ gives the oscillatory power called the REACTIVE POWER. Whereas the real part gives the net power flow or the RESISTIVE POWER flow at the location. So the concept of the Poynting vector and the average Poynting vector is very important because by using this concept, we can calculate the net power flow at a particular location.

10.3 Uniform Plane wave

uniform plane wave We can then apply this concept to a case of a uniform plane wave. So we may ask, *how much power density does a uniform plane wave carry when it travels in a media?* In Uniform plane waves, electric and magnetic fields are perpendicular to each other. With $\bar{E}$ and $\bar{H}$ shown below

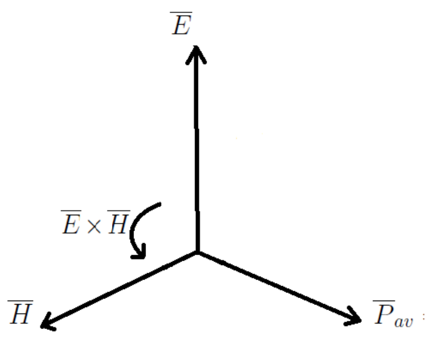


Figure 10.1: Direction of power flow of an electromagnetic wave

The direction of power flow is given by the right-hand rule. $\bar{E} \times \bar{H}$ as we realize has the direction of $\bar{P}_{av}$ which is also the same as the direction of wave propagation. As we have seen in the case of a uniform plane wave, $\bar{E}$, $\bar{H}$ and the direction of propagation are mutually perpendicular to each other. If we say the electric field is $\bar{E} = E_0 e^{-j\beta z}\hat{x}$ having phase variation in the Z direction, therefore; $\bar{H} = H_0 e^{-j\beta z}\hat{y}$,

$$\begin{aligned}\bar{P}_{av} &= \frac{1}{2}\mathfrak{Re}(\bar{E} \times \bar{H}^*)\\ &= \frac{1}{2}\mathfrak{Re}(E_0 e^{-j\beta z} H_0 e^{+j\beta z})(\hat{x} \times \hat{y})\\ &= \frac{1}{2}\mathfrak{Re}(E_0 H_0 \hat{z})\\ &= \frac{1}{2}(E_0 H_0 \hat{z})\end{aligned}$$

So for a uniform plane wave, the average Poynting vector will be half $E_0 H_0$ in direction $\hat{z}$ with $\bar{E}$ in $\hat{x}$ and $\bar{H}$ in $\hat{y}$ direction.

Now we can take specific cases for the uniform plane wave in an unbound medium in which the wave is propagating. For a DIELECTRIC MEDIUM, $\bar{E}$ and $\bar{H}$ are related by the intrinsic impedance of the medium. $\frac{E}{H} = \eta$ = intrinsic impedance.

Assuming H_0 and E_0 are complex quantities, then $\bar{P}_{av} = \frac{1}{2}\mathfrak{Re}(E_0 H_0^* \hat{z})$. Substituting η into $\bar{P}_{av} = \frac{1}{2}\mathfrak{Re}(E_0 H_0^* \hat{z})$ and making it in terms of the electric field only, we have;

$$\bar{P}_{av} = \frac{1}{2}\mathfrak{Re}\left\{E_0(\frac{E_0}{\eta})^*\right\}\hat{z}$$

This can also be written as $\bar{P}_{av} = \frac{1}{2}\mathfrak{Re}\left\{\eta H_0(H_0^*)\right\}\hat{z}$ in terms of magnetic field.

We now have that

$$\bar{P}_{av} = \frac{1}{2}\mathfrak{Re}\left\{\frac{|E_0|^2}{\eta^*}\right\}\hat{z}$$

or

$$\bar{P}_{av} = \frac{1}{2}\mathfrak{Re}\left\{\eta|H_0|^2\right\}\hat{z} \tag{10.14}$$

So from here, we can find the average power flow associated with a uniform plane wave in an unbound medium. For a dielectric medium,for which $\eta = \sqrt{\frac{\mu}{2}}$, that is for an ideal dielectric, η is real. $\bar{P}_{av} = \frac{1}{2}\frac{|E_0|^2}{\eta}$ or $\frac{1}{2}\eta|H_0|^2$. So in a dielectric medium, if we know the peak amplitude of either the electric or magnetic field, the permeability and permittivity of the medium, and then η which is real, we can get the power flow density associated with this uniform plane wave with above equations.

In general, if the medium has a finite conductivity that is non-zero or very large, then we use

$\bar{P}_{av} = \mathfrak{Re}(\frac{|E_0|^2}{2\eta^*})\hat{z}$ or $\bar{P}_{av} = \frac{1}{2}\mathfrak{Re}\left\{\eta H_0(H_0^*)\right\}\hat{z}$ as η is a complex quantity in this case.

For the extreme case of a good conductor, to get the average power, Recall;

$$\eta = \sqrt{\frac{\omega\mu}{2\sigma}} + j\sqrt{\frac{\omega\mu}{2\sigma}}$$

Thus,

$$\bar{P}_{av} = \frac{1}{2}|E_0|^2\sqrt{\frac{\sigma}{2\omega\mu}}$$

that is because,

$$\begin{aligned}\bar{P}_{av} &= \frac{1}{2}\mathfrak{Re}\left\{\frac{|E_0|^2}{\eta^*}\right\}\\ &= \frac{1}{2}\mathfrak{Re}\left\{\frac{|E_0|^2}{\sqrt{\frac{\omega\mu}{2\sigma}} - j\sqrt{\frac{\omega\mu}{2\sigma}}}\right\}\end{aligned}$$

Finding the conjugate

$$\bar{P}_{av} = \frac{1}{2}\mathfrak{Re}\left\{\frac{|E_0|^2\left(\sqrt{\frac{\omega\mu}{2\sigma}} + j\sqrt{\frac{\omega\mu}{2\sigma}}\right)}{\left(\sqrt{\frac{\omega\mu}{2\sigma}} - j\sqrt{\frac{\omega\mu}{2\sigma}}\right)\left(\sqrt{\frac{\omega\mu}{2\sigma}} + j\sqrt{\frac{\omega\mu}{2\sigma}}\right)}\right\}$$

simplifying further,

$$\begin{aligned}\bar{P}_{av} &= \frac{1}{2}\mathfrak{Re}\left\{\frac{|E_0|^2(\sqrt{\frac{\omega\mu}{2\sigma}} + j\sqrt{\frac{\omega\mu}{2\sigma}})}{\frac{2\omega\mu}{2\sigma}}\right\} \\ &= \frac{|E_0|^2}{2}\mathfrak{Re}\left\{\frac{\sigma}{\omega\mu}\left(\sqrt{\frac{\omega\mu}{2\sigma}} + j\sqrt{\frac{\omega\mu}{2\sigma}}\right)\right\} \\ &= \frac{|E_0|}{2}\frac{\sigma}{\omega\mu} \times \sqrt{\frac{\omega\mu}{2\sigma}} \\ &= \frac{1}{2}|E_0|^2\sqrt{\frac{\sigma}{2\omega\mu}}\end{aligned}$$

Hence for a good conductor, we have,

$$\bar{P}_{av} = \frac{1}{2}|E_0|^2\sqrt{\frac{\sigma}{2\omega\mu}} \tag{10.15}$$

So by using the concept of Poynting vector, we can find out the average power flow of any medium and at any particular location in space.

In the case of dielectric, it is straightforward because the intrinsic impedance is real. For a good conductor with finite conductivity, we go to the more general expression as the intrinsic impedance is complex.

10.4 ExerciseList

Ex. 177 — Explain the significance of Maxwell's Equations in the analysis of power flow associated with electromagnetic waves.

Ex. 178 — Derive the expressions for power flow in electromagnetic waves using Maxwell's Equations.

Ex. 179 — Define the Poynting vector and explain its significance in describing power flow associated with electromagnetic fields.

Ex. 180 — Discuss the conditions under which power flow occurs, considering the relationship between electric and magnetic fields.

Ex. 181 — Explain why time average power flow is considered more useful than instantaneous power flow in practical systems

Ex. 182 — Derive the expression for time average power flow associated with electric and magnetic fields in terms of the Poynting vector.

Ex. 183 — Discuss the power flow associated with a uniform plane wave in an unbound medium

Ex. 184 — Calculate the average power flow density for a dielectric medium and a good conductor, considering the intrinsic impedance.

Ex. 185 — Demonstrate how vector identities are applied to arrive at the expression $E \times H = -\mu\left(\frac{\partial}{\partial t}|H|^2\right) - 2\epsilon\left(\frac{\partial}{\partial t}|E|^2\right) - E \cdot J$

Chapter 11

Surface current and power loss in a conductor

11.1 Objective

1. Understand Surface Current Phenomenon.
2. Differentiate between ideal and practical Conductors.
3. Apply mathematical models to calculate surface current density, considering the variation in conductivity and the exponential decay of electric fields within a conductor.
4. Grasp the concept of surface impedance as a crucial boundary parameter, and understand its unit and significance in power calculations on conducting surfaces.
5. Utilize circuit concepts to calculate power loss on the surface of a conductor, finding the relationship between surface current density and resistance in thin slabs.
6. Apply wave propagation concepts, including Poynting vectors, to understand how electromagnetic waves interact with the surface of conductors, leading to power losses

11.1.1

In our previous chapter, we talked about Power Flow in an electromagnetic wave and pointing vectors for electromagnetic waves. The Poynting vector tells us the density of power flow and its direction tells us the direction of the power at any location in space.

In this chapter, we will talk about surface current and power loss in a conductor. We have always wondered how much power is lost in a conducting surface when an electromagnetic wave is incident on the conducting surface, the origin of surface current and if we really have surface current in practice and many more.

Surface Current is essentially a phenomenon which lies on the surface of the medium. This is a phenomenon which is for ideal conducting surfaces.

We will begin this chapter with the volume current density inside a conductor. We know the concept of skin depth in a good conductor, from there we will find out the current which is the surface current and see that, in practical systems the concept of surface current is very useful in finding out how much losses takes place in a conducting surface. First, we will discuss surface current and then go to power loss in a conducting surface.

We established that $\hat{n}\text{x}\hat{H}$ gives the direction of the surface current. $\hat{n}\text{x}\hat{H}$ gives us the direction which goes into the paper as shown in the figure 11.1

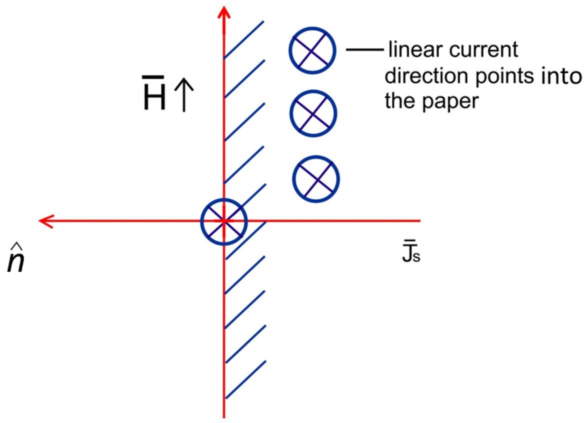


Figure 11.1: Direction of the surface current

This gives the linear surface density $\bar{J}s$ we talked about when we dealt with boundary conditions. Now, if the surface current is flowing here, what is the driving mechanism for this surface current?

It is already established that this surface current phenomenon is associated with infinite conductivity. So if we take an ideal conductor, then maybe at some point if time, instantaneously, some electric fields existed in the medium and since it was a very short-lived phenomenon we had electric and magnetic fields together.

For the momentary existence of the electric field, it would put the charges into motion, which would constitute this current and since the conductivity is infinite, even if the electric field doesn't exist anymore when the charges are put in motion, the charges will keep moving for infinite time which constitutes a current. So when the conductivity is infinite, one may visualize that at some instant of time, some electric field induced motion to these charges. The charges were set in motion and they kept moving which is essentially the surface current and this is balanced by the magnetic field and $\hat{n}\text{x}\hat{H}$ essentially gave us that current density $\bar{J}$.

So basically, there are two situations now; (i) When we have a tangential component of electric field on the conducting surface which is zero but there are no surface currents and (ii) when the tangential component of electric field is zero but there is surface current. So in both situations, we have the tangential component of an electric field to be zero for ideal conductors but in one case, we have surface current and in the other case, we may not have surface current. If we have a surface current, then it must be balanced by the magnetic field. So if we have a magnetic field tangential to the conducting boundary, then we will have surface current otherwise we will not have surface current. This is a hypothetical situation of when surface current is truly flowing on the surface. Now, we may be curious to see what happens if we take a good conductor, whether we can still make use of this concept called surface current or not. It is very clear that if we have a conductivity which is not infinite, then because of the electric field, there will always be a finite conduction current density inside the material of the conductor. Also, the skin depth will be of finite width, which implies that this phenomenon is no more a true surface phenomenon (which requires zero thickness for skin depth). From the diagram of the conductor shown below, let **z** >0 represent the conductor.

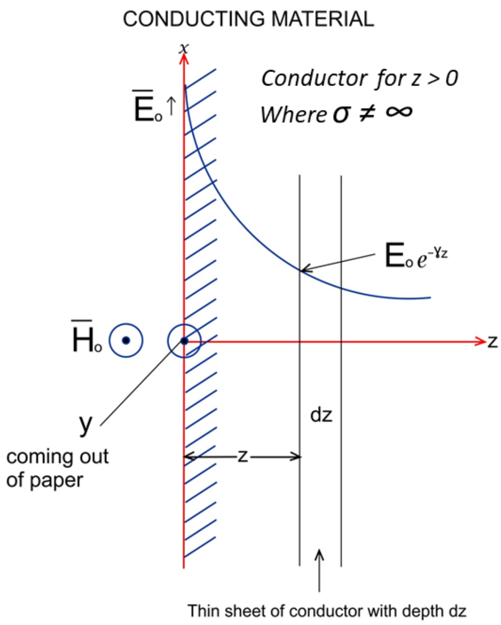


Figure 11.2: Exponential decay of electric field inside a conductor

If we have an electric field in this conductor for a finite σ not equal to ∞, then the tangential component of **E** on this surface is not zero (very small but not zero as for an ideal conductor with ∞ conductivity). Let **E** be in **x** direction and **H** in **y** direction. The electric field **E** dies down exponentially as we go into the conductor as shown in fig11.2. **E** dies down at $e^{-\gamma z}$. Once we know the electric field value at the surface of the conductor, we have

$$\gamma = \sqrt{j\omega\mu\sigma} \tag{11.1}$$

for a good conductor, σ is very large. γ=α+$j\beta$ to get

$$\gamma = \sqrt{\frac{\omega\mu\sigma}{2}} + j\sqrt{\frac{\omega\mu\sigma}{2}} = \alpha + j\beta \tag{11.2}$$

Since the current flow is in the direction of the electric field $\hat{x}$, we have a conduction current density $\bar{J}$=$\sigma\bar{E}_0$ which varies exponentially as we go deeper inside the conductor. We can take a thin sheet with a depth of **dz**. The conduction current is thus, $\bar{J}$=$\sigma\bar{E}_0$e$^{-\gamma z}$. So if we know the current flowing in the sheet per unit area in the xy plane, we can find out the total current flowing through the surface of the xy plane in the x direction. So if we take the conduction current density and integrate it over the depth, we get the total current flowing through the surface xy. If we take that surface as a unit surface, then the current that is flowing is with respect to that unit area. So for the dz region, the conduction current density;

$\bar{J}$(z) = $\sigma\bar{E}$ = $\sigma\bar{E}_0 e^{-\gamma z}\hat{x}$ = $\sigma\bar{E}_0$e$^{-\alpha z}$e$^{-j\beta z}$.

Unit depth in the y direction and width dz represents the area that $\bar{J}$ flows through. For this Area= $1 \times dz = dz$. Since the current flows in $\hat{x}$ direction, the area normal to this flow is in the yz plane and thus we take a unit y depth. So, current in a sheet of unit depth in the y direction along the whole z plane is determined by integrating over the entire z.

$$\begin{aligned} \mathbf{I}(z) &= \bar{J}(z) \times Area \\ &= \bar{J}(z)dz \\ &= \sigma\bar{E}_0 e^{-\gamma z} dz\hat{x} \end{aligned} \tag{11.3}$$

This equation is the current per unit length in the y direction. The total current flowing on the surface is the surface current. Note that there is no true surface current, the current flow into the depth of z but dies down very fast. Hence the little current flow is limited to small dz on the surface of the conductor with large conductivity, we know that the skin depth is very small, so the current flow is confined to the surface, but not truly confined to the surface to a skin depth from the surface; Surface current,

$$\begin{aligned} \bar{J}_s &= \int_0^\infty \sigma\bar{E}_0 e^{-\gamma z} dz\hat{x} \\ &= \frac{\sigma\bar{E}_0}{-\gamma}[e^{-\gamma z}]_0^\infty \hat{x} \\ &= \frac{\sigma}{\gamma}\bar{E}_0\hat{x} \end{aligned} \tag{11.4}$$

Though we call $\bar{J}$s surface current density, it is clear that it is not truly a surface phenomenon but it has all the properties which a surface current has. And one of such properties is that the surface current should be related to a magnetic field, that is, the tangential component of the magnetic field to the surface and it should satisfy the relationship of $\hat{n} \times \hat{H} = \bar{J}_s$.

So if $\bar{J} = \frac{\sigma}{\gamma}\bar{E}_0\hat{x}$ must serve as surface current density, it must be related to the magnetic field and satisfy the relationship $\hat{n}$x$\hat{H}$=$\bar{J}$s.

Since we have an electric field which is $\bar{E}_0$ on the surface and magnetic field $\bar{H}_0$, with the wave travelling inwards in the z-direction, being a transverse electromagnetic wave as the medium is unbound, then $\frac{\bar{E}_0}{\bar{H}_0}$ must satisfy the relationship of being of equal to the intrinsic impedance of the medium. For a good conductor, intrinsic impedance

$$\eta_c = \sqrt{\frac{j\omega\mu}{\sigma}} \tag{11.5}$$

And the electric and magnetic field are related by $H_0 = \frac{E_0}{\eta_c}$ but $\gamma = \sqrt{j\omega\mu\sigma}$ so,

$$\eta_c = \sqrt{\frac{j\omega\mu}{\sigma}} \equiv \sqrt{\frac{j\omega\mu\sigma}{\sigma^2}} = \frac{\gamma}{\sigma} \tag{11.6}$$

Hence,

$$\bar{H}_0 = \frac{\bar{E}_0}{\eta_c} = \frac{\sigma\bar{E}_0}{\gamma} \tag{11.7}$$

But this quantity is the same as what we got for the surface current density $\bar{J}_s$ earlier;
$\bar{J}_s = \frac{\sigma}{\gamma} E_0 \hat{x}$
Thus,

$$|H_0| = \frac{\sigma}{\gamma}|E_0| = |\bar{J}s|. \tag{11.8}$$

So one relationship we wanted was that if $\bar{J}_s$ is surface current density the way we have visualized it, then it must be related to the magnetic field which we have proven as $|H_0| = |\bar{J}_s|$. The second requirement is that $\hat{n} \times \hat{H} = \bar{J}_s$ i.e $\hat{n} \times \hat{H}$ must give surface current direction. $\hat{J}_s$ was in direction of **E** that is $\hat{x}$. So we have a surface current density $\hat{J}_s$ direction that is $\hat{x}$ oriented and a magnetic field **H** that is $\hat{y}$ oriented. The unit normal to the conducting surface is $\hat{n} = -\hat{z}$. Now;

$$\hat{n} \times \hat{H} = -\hat{z} \times H_0\hat{y} = H_0(-\hat{z} \times \hat{y})$$

but $-\hat{z} \times \hat{y}$ gives $\hat{x}$ direction which is the direction of $\bar{J}_s$ thus, $\bar{J}_s = \frac{\sigma}{\gamma} E_0 \hat{x}$ has all the characteristics of a surface current density. The magnitude of the magnetic field should be equal to that of the surface current density. It becomes obvious that, though $\bar{J}$s=$\frac{\sigma}{\gamma} E_0 \hat{x}$ is not a true surface current, we still use it as a surface current nonetheless.

Also when conductivity is very large, this current is effectively confined to a very thin layer called skin depth. As we have seen earlier, if we go to frequencies like 300MHz, skin depth lies in a few micron range. So essentially this is a current flowing through a very thin sheet on the surface of the conductor. Also, it has the characteristics of a true surface current that is, it is related to the magnetic field and $\hat{n}$ x $\bar{H}$ gives its direction. So this quantity can be used as the surface current in a good conductor. Although in the true sense, there is no surface current for a good conductor, in practice, we make use of the quantity $\bar{J}_s$ as surface current in our analysis.

Once we get surface current, we define a quantity called **SURFACE IMPEDANCE** which is a boundary parameter for this boundary. This is useful whenever we do the power calculation on a conducting surface.

$$\text{Surface impedance, } Z_s = \frac{E_{tangetial}}{\bar{J}s} \tag{11.9}$$

$$\bar{J}s = \frac{\sigma}{\gamma} E_0 \hat{x} \tag{11.10}$$

$\bar{J}_s$ has a unit of $\frac{A}{m}$, that is, linear surface current density. To determine the unit of Z_s, we have;

$$Z_s = \frac{V/m}{A/m} = \frac{V}{A} = \Omega \tag{11.11}$$

So as expected, the quantity Z_s is some form of impedance. If we know the tangential component of the electric field, then the surface current can be obtained from what is called the surface impedance or vice versa.

$$Z_s = \frac{E_0}{\frac{\sigma}{\gamma}E_0} = \frac{\gamma}{\sigma} = \frac{\sqrt{j\omega\mu\sigma}}{\sigma} = \sqrt{\frac{j\omega\mu}{\sigma}} = \eta_c \tag{11.12}$$

So for a good conductor, the intrinsic impedance of the medium is the same as the surface impedance. Later we will see that the surface impedance concept is used to calculate the power loss. Now if we know the tangential component of the electric field and know the conductivity, we can calculate the intrinsic impedance of the medium which is the same as the surface impedance. Once we have the surface current density $\bar{J}$s, then we ask how much is the power loss due to this surface current density flowing in this conductor. Considering the surface of the unit area, how much is the power loss in this unit area of the conducting surface? Since the current is flowing into the depth of the conductor, the power loss is not only taking place on the surface, it takes place all along the depth of the conductor. So to determine the total loss, we take the total depth of the conductor, and a unit area of thickness dz as shown below.

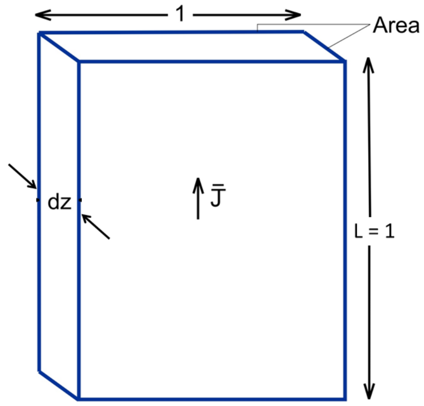


Figure 11.3: A thin slab layer of thickness dz parallel to the surface of the conductor

So if we take a thin slab parallel to the surface of the conductor, since we have a finite conduction current, we ask how much the power loss is in this thin slab. If we integrate over all depth z, we get the power loss per unit area on the surface of the conductor.

All we need to do is to determine the surface current flowing and the resistance of the thin slab.

$\bar{J}_s(1 \times \mathbf{dz})$ is the current flowing in the direction of $\bar{J}_s$

The length L is unity, σ=conductivity, $\frac{1}{\sigma}$= resistivity

$$\begin{aligned} \text{Resistance} &= \frac{\text{Resistivity} \times \text{Length}}{\text{Area}} \\ &= \frac{1}{\sigma} \times \frac{1}{dz} \\ &= \frac{1}{\sigma dz} \end{aligned}$$

current flowing in the area $= Jdz, R = \frac{1}{\sigma dz}, I = Jdz, d\omega =$ ohmic loss in slab
$d\omega = \frac{1}{2} = |I|^2 R$, I and R can be complex hence $\frac{1}{2}|I|^2 R$.

$$I = \bar{J}dz \text{ and } \bar{J} = \sigma E_0 e^{-\gamma z}$$

$$I = \sigma \bar{E}_0 e^{-\gamma z} dz$$

$$d\omega = \frac{1}{2}|\sigma \bar{E}_0 e^{-\alpha z} e^{-j\beta z}|^2 \times \frac{1}{\sigma dz}$$

$$d\omega = \frac{1}{2}\sigma |\bar{E}_0|^2 e^{-2\alpha z} dz$$

The total power loss in the whole depth under the unit area of the surface becomes,

$$\omega = \int_0^\infty \frac{1}{2}\sigma|E_0|^2 e^{-2\alpha z}dz = \frac{1}{2}\sigma \bar{E}_0^2[\frac{e^{-2\alpha z}}{-2\alpha}]_0^\infty \quad (11.13)$$

$$\omega = \frac{1}{2}\frac{\sigma}{2\alpha}|E_0|^2 \quad (11.14)$$

but

$$\gamma = \sqrt{j\omega\mu\sigma} = \alpha + j\beta = \sqrt{\frac{\omega\mu\sigma}{2}} + j\sqrt{\frac{\omega\mu\sigma}{2}} \quad (11.15)$$

So $\alpha = \sqrt{\frac{\omega\mu\sigma}{2}}$

$$w = \frac{1}{2} \times \frac{\sigma}{2\sqrt{\frac{\omega\mu\sigma}{2}}} \times |E_0|^2 \quad (11.16)$$

recall, $\bar{J}_s = \frac{\sigma}{\gamma}E_0\hat{x}$
so, $|E_0|^2 = \frac{|\gamma|^2}{|\sigma|^2}|\bar{J}_s|^2$

$$w = \frac{1}{2} \times \frac{\sigma}{2\sqrt{\frac{\omega\mu\sigma}{2}}} \times \frac{|\gamma|^2}{|\sigma|^2}|\bar{J}_s|^2 \quad (11.17)$$

$$w = \frac{1}{2} \times \frac{1}{2}\sqrt{\frac{2}{\omega\mu\sigma}} \times \omega\mu \times |\bar{J}_s|^2$$

$$w = \frac{1}{2}\sqrt{\frac{\omega\mu}{2\sigma}}|\bar{J}_s|^2 \quad (11.18)$$

$$Z_s = \sqrt{\frac{j\omega\mu}{\sigma}} = \sqrt{\frac{\omega\mu}{2\sigma}} + j\sqrt{\frac{\omega\mu}{2\sigma}} \quad (11.19)$$

Hence we say that the surface impedance has a resistive part called the surface resistance and a reactive part called the surface reactance.

$$Z_s = R_s + jX_s \quad (11.20)$$

$$\omega = \frac{1}{2} \times |\bar{J}s|^2 \times \sqrt{\frac{\omega\mu}{2\sigma}} = \frac{1}{2}|\bar{J}s|^2 R_s \quad (11.21)$$

Thus, the power loss per unit area of conducting surface can be obtained if we know the surface current density and the surface resistance.

Later on, we would see that if we go to structures like waveguides, then the conductor loss is calculated from the surface current density because the conductor surface current density can be obtained from magnetic fields. So if we know the tangential component of the magnetic field on the surface of the conductor, we can find out using $\hat{n} \times \bar{H}$, the linear surface current density with $\bar{J}_s$ known and Z_s(surface impedance) known, from there we can find out the power loss per unit area of the conductor.

Essentially we have calculated the power loss by using the circuit concept that is, we found out the current and the resistance in the slab, then we found out the I^2R loss, and from there we obtained the power loss on the surface of the conductor. We can also determine the power loss by using the wave approach. Once we have an electric and magnetic field, there is

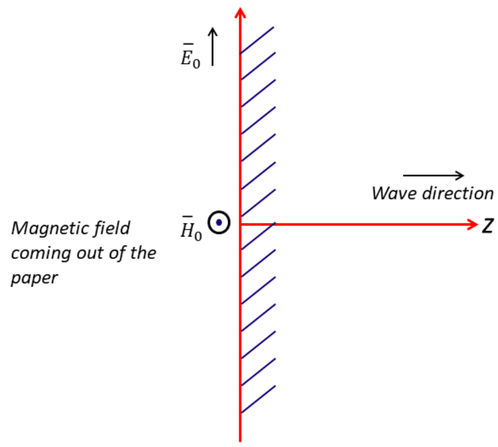


Figure 11.4: Determination of the power loss using wave concept

essentially a wave phenomenon going in the z-direction. So the power going with the wave into the conductor is essentially the power which is lost into the ohmic loss conductor. This power is finite and is decreasing as the wave travels with time as no power is coming back. So whatever the power flow inside the conductor is, is a measure of the power that was lost inside the conductor. So instead of doing the calculation of the power loss from an electrical point of view, by finding out the current and resistance, we can use the wave concept and find out what the power loss is inside the conductor with the interface and **E** and **H** being in the directions shown in the figure below. Since we are asking for power flow per unit area of the surface, we find the Poynting vector at the surface of the conductor. That is the power which is essentially going into the conductor and which is lost as Ohmic losses in the conductor.

Poynting vector,

$$\bar{P} = \bar{E} \times \bar{H} = \frac{1}{2}\mathfrak{Re}(E_0 H_0^* \hat{z}) \quad (11.22)$$

We substitute $\bar{H}_0$=$\frac{E_0}{\eta_c}$ for this medium

$$\bar{P} = \frac{|\bar{E}_0|^2}{2\eta_c^*}, \quad \bar{P} = \frac{1}{2}|\bar{E}_0|^2 \times \frac{1}{\eta_c^*} \quad (11.23)$$

Recall $\eta_c = \sqrt{\frac{j\omega\mu}{\sigma}}$ So that,

$$\eta_c = \sqrt{\frac{\omega\mu}{\sigma}}(\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}})$$

$$\eta_c^* = \sqrt{\frac{\omega\mu}{\sigma}}(\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}})$$

$$\eta_c = \sqrt{\frac{\omega\mu}{\sigma}}(\frac{1}{\sqrt{2}} + j\frac{1}{\sqrt{2}})$$

$$\eta_c^* = \sqrt{\frac{\omega\mu}{\sigma}}(\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}})$$

$$\bar{P} = \frac{1}{2}|\bar{E}_0|^2 \frac{1}{\sqrt{\frac{\omega\mu}{\sigma}}(\frac{1}{\sqrt{2}} - j\frac{1}{\sqrt{2}})} \quad (11.24)$$

$$= \frac{1}{2}|\bar{E}_0|^2 \times \frac{\sqrt{\frac{\omega\mu}{\sigma}}(\frac{1}{\sqrt{2}}+j\frac{1}{\sqrt{2}})}{\sqrt{\frac{\omega\mu}{\sigma}}(\frac{1}{\sqrt{2}}-j\frac{1}{\sqrt{2}}) \times \sqrt{\frac{\omega\mu}{\sigma}}(\frac{1}{\sqrt{2}}+j\frac{1}{\sqrt{2}})}$$

$$\bar{P} = \frac{1}{2}|\bar{E}_0|^2 \mathfrak{Re}\left\{\frac{\sqrt{\frac{\omega\mu}{\sigma}}\left(\frac{1}{\sqrt{2}}+j\frac{1}{\sqrt{2}}\right)}{\frac{\omega\mu}{\sigma}\left(\frac{1}{2}+\frac{1}{2}\right)}\right\}$$

$$\bar{P} = \frac{1}{2}|\bar{E}_0|^2 \mathfrak{Re}\left\{\frac{\sigma}{\omega\mu}\left(\sqrt{\frac{\omega\mu}{\sigma}}\left(\frac{1}{\sqrt{2}}+j\frac{1}{\sqrt{2}}\right)\right)\right\}$$

$$\bar{P} = \frac{1}{2}|\bar{E}_0|^2 \times \sqrt{\frac{\sigma}{\omega\mu}} \times \frac{1}{\sqrt{2}}$$

$$\bar{P} = \frac{1}{2}|\bar{E}_0|^2 \times \sqrt{\frac{\sigma}{2\omega\mu}} \tag{11.25}$$

This expression is the same as we got for power loss using the surface impedance of the conducting medium. So we can calculate the power loss either by using the circuit concept or the wave concept. Using the circuit concept, we find $\bar{J}_s$ and find out the ohmic loss or I^2R loss using the wave concept, we find out what the power going inside the surface of the conductor is and that power should essentially get lost inside the conductor. Since the field at $z = \infty$ goes to zero there is no power flow at $z = \infty$, so whatever power went into the conductor must have been lost in the heating of the conductor. So either the electrical circuit concept or the wave concept can be used to find out the power loss per unit area on the surface of the conductor.

The power that is being lost inside the conductor is proportional to the conductivity from $P = \frac{1}{2}|E_0|^2\sqrt{\frac{\sigma}{2\omega\mu}}$. We are dealing with electromagnetic waves here, SO The higher the conductivity of the medium it propagates in, the more the power loss of this wave so that in a pure dielectric with σ=0, it will have no power loss in that medium. So, large σ means a large part of the electromagnetic wave energy is taken by the medium (or lost to the medium). With high frequency, it losses less of its energy to the medium through which it propagates. So for an ideal conductor with $\sigma \longrightarrow \infty$, the EM wave losses all of its power to this conductor but with skin depth of nearly zero, all of this energy lost flow as pure surface current on this conductor without any Ohmic losses and the entire energy is reflected from the boundary. [1]

For an ideal conductor, there is no power going inside the conductor similarly when the frequency becomes large, there is no power going inside the conductor. What this means is that, for high conductivity or high frequency, there is a resistance to the penetration of the conductor, hence, the power does not go inside the conductor.

This situation is similar to a lossy transmission line but however, in a lossy transmission line, the power gets lost in the heating of the line. In this case, the power is dying down rapidly but the power is not lost in the Ohmic loss because the power is not able to penetrate the layer. So for conductivity = ∞, no power penetrates the conductor so it gets reflected.

In conclusion, when the conductivity is infinite, the wave does not penetrate the medium, there is no power loss and the entire energy is reflected from the boundary.

11.2 ExerciseList

Ex. 186 — Define surface current and discuss its characteristics, particularly in the context of ideal conducting surfaces.

Ex. 187 — Explain the concept of volume current density inside a conductor, and how it relates to the Skin Depth in a good conductor.

Ex. 188 — Describe how the direction of surface current is determined using the nˆ × Hˆ relationship, and illustrate it with a diagram.

Ex. 189 — Discuss the conditions under which surface current is considered in ideal conducting surfaces, and its practical implications.

Ex. 190 — Explain the relationship between the tangential component of the electric field and the existence of surface current.

Ex. 191 — Explore the concept of surface current in good conductors, considering finite conductivity and the impact on conduction current density.

Ex. 192 — Provide a step-by-step calculation of surface current density (J¯s) based on the electric field and conductivity values.

Ex. 193 — Define surface impedance and explain its significance in the context of power calculations on a conducting surface. Also, discuss its unit.

Ex. 194 — Demonstrate how to calculate power loss per unit area on a conducting surface using the circuit concept, considering surface current density and resistance.

Ex. 195 — Explain the alternative method of calculating power loss using the wave concept, particularly focusing on the Poynting vector and wave propagation inside the conductor.

1 

Guglielmo Marconi, 1st Marquis of Marconi (25 April 1874 - 20 July 1937) was an Italian inventor and electrical engineer known for his pioneering work on long-distance radio transmission and for his development of Marconi's law and a radio telegraph system. He is usually credited as the inventor of the radio

Chapter 12

Uniform Plane Wave in an arbitrary direction

12.1 Objective

1. To understand how waves behave when they propagate in different directions and interact with different materials or obstacles.
2. To understand analytical solutions in wave equation facilitating a deeper understanding of wave phenomena.
3. To provide insights and solutions across various scientific and engineering disciplines in studying wave behavior over large distances.

12.2

Prior to now, our discussion on plane wave propagation has been in an unbound medium, orienting our coordinate system such that the wave propagated was in the direction of one of the axes (the z-axis). This was done because an unbound medium is symmetric in all directions, regardless of the direction looked from, the medium appears the same. However, in a bound medium, the choice of the coordinate axis might affect the analysis algebraically. Essentially we select a coordinate system that will enable our analysis to be simple, doing this restricts the choice of coordinate axes. First, we would investigate the propagation of electromagnetic waves in a semi-finite medium (i.e. half of an unbound medium), then move on to a bound medium. However, before this analysis, we require a representation of electromagnetic waves travelling in an arbitrary direction with respect to the coordinate axis.

12.3 Wave in arbitrary direction

Recall when we had a coordinate system such that the wave was propagating in the z-axis, its phase variation was also in the same direction while its electric and magnetic fields were in planes perpendicular to the z-axis(i.e. the x-y plane). Say we have an unbound medium, the wave would still be oriented such that it is moving in an arbitrary direction with respect to the coordinate axis.

In figure 12.1 a wave propagation in an arbitrary direction is shown; a constant phase plane(i.e. the plane in which the electric and magnetic field lie) is shown perpendicular to the wave's direction. ϕ_x,ϕ_y and ϕ_z are the angles which the wave makes with the x, y and z axis respectively. The vector has

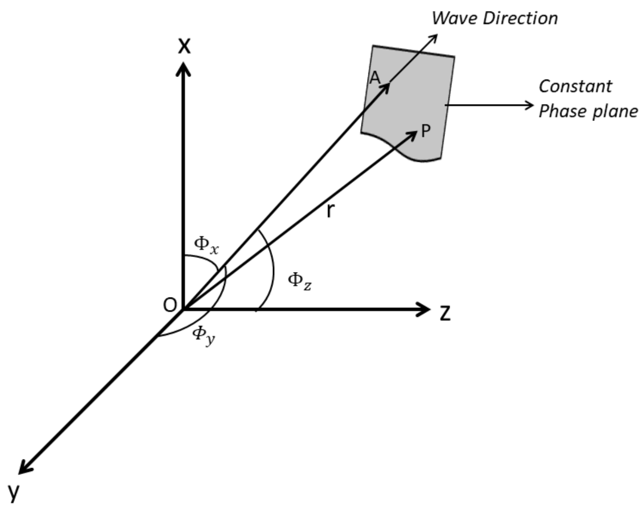


Figure 12.1: wave in arbitrary direction

direction cosines $\cos\phi_x\hat{x}$, $\cos\phi_y\hat{y}$, $\cos\phi_z\hat{z}$. The unit vector:

$$\hat{n} = \cos\phi_x\hat{x} + \cos\phi_y\hat{y} + \cos\phi_z\hat{z} \tag{12.1}$$

Taking two arbitrary points on the wave direction point O and P, where O is the origin and P has coordinates (x,y,z).

Thus $\vec{OP} = x\hat{x} + y\hat{y} + z\hat{z}$

The normal distance to take to any point on the constant phase plane is OA, so if any point is taken on the plane and its projection is found in the direction of the normal, the quantity is fixed regardless of the point it is taking on the phase plane.

The equation of the phase plane will be:

$\hat{n} \cdot \vec{OP}$ (i.e the dot product of $\hat{n}$ and $\vec{OP}$)

$\hat{n} \cdot \vec{OP} = |OA|$ = constant

substituting for $\hat{n}$ and $\vec{OP}$

$x\cos\phi_x + y\cos\phi_y + z\cos\phi_z$ = constant. And this constant is the magnitude of $\vec{OA}$ which gives the distance of the plane from its origin.

If the wave is having a phase constant, β along $\vec{OA}$, then the phase of the plane is β multiplied by the distance travelled which is $|OA|$.

therefore, Phase of the plane = $\beta\ |OA| = \beta\ \hat{n}.\hat{r}$

where $\hat{r} = \vec{OP} = x\hat{x} + y\hat{y} + z\hat{z}$

β = Phase constant, $\hat{n}$ = Unit vector in the direction of the propagated wave, $\hat{r}$ = Position of any Vector on the constant phase plane

Since we now have the phase of the plane, then getting the expression of the electric and magnetic field which corresponds to the wave propagation is straightforward.

$$\text{Electric field, } \bar{E} = \bar{E}_0 e^{-j\beta\hat{n}.\hat{r}}$$

So, $\bar{E}_0.\hat{n} = 0$ since $\bar{E}_0$ is perpendicular to $\hat{n}$.

$\bar{E}_0$ = Magnitude of the Vector, $-j\beta\hat{n}.\hat{r}$ = Phase Variation.

Therefore, $\bar{E} = \bar{E}_0 e^{-j\beta\hat{n}.\hat{r}}$ represents an electromagnetic wave traveling in an arbitrary direction $\hat{n}$. Combining $\beta\hat{n}$ we define another vector called the **wave vector**.

wave vector $\bar{k} \equiv \beta\hat{n} = \beta(\cos\phi_x\hat{x} + \cos\phi_y\hat{y} + \cos\phi_z\hat{z})$

$$\bar{k} = \beta\cos\phi_x\hat{x} + \beta\cos\phi_y\hat{y} + \beta\cos\phi_z\hat{z}$$

Let $K_x = \beta\cos\phi_x$, $K_y = \beta\cos\phi_y$, $K_z = \beta\cos\phi_z$

$\bar{k} = K_x\hat{x} + K_y\hat{y} + K_z\hat{z}$

So if the direction of wave propagation($\hat{n}$) and phase constant(β) which depend on the medium parameters are known, then we can define $\bar{k}$ which completely characterizes the wave in the arbitrary direction.

Recall, $\hat{E}_0.\hat{n} = 0$ since they are perpendicular, then

$$\hat{E}\cdot\hat{k} = 0$$

Therefore, to verify if the wave is travelling in the z-direction then

$$\phi_z = 0,\ \phi_y = 90°,\ \phi_x = 90°.$$
$$\text{Therefore, } \bar{k} = \beta(\cos 90\hat{x} + \cos 90\hat{y} + \cos 0\hat{z})$$
$$\hat{k} = \beta\hat{z}$$
$$\text{Thus, space variation} = e^{-j\beta z}$$

This is the same expression we got for a wave which was travelling in the z-direction.

Therefore, $\hat{E} = \hat{E}_0 e^{-j\beta\hat{n}.\hat{r}}$ is a representation of the electric field for a uniform plane wave travelling in an arbitrary direction making angles ϕ_x, ϕ_y and ϕ_z with the three co-ordinate axes.

To find the magnetic field corresponding to the electric field, we go back to the original Maxwell's equation.

Since we are still dealing with media which does not have conductivity, we say let's have source-free media i.e. no conductivity or currents.

The Maxwell's equations:

$$\bigtriangledown\times\bar{H} = j\omega\epsilon\bar{E} \tag{12.2}$$

and

$$\bigtriangledown\times\bar{E} = -j\omega\mu\bar{H} \tag{12.3}$$

$$\hat{H} = \frac{-1}{j\omega\mu}(\bigtriangledown\times\bar{E}) = \frac{-1}{j\omega\mu}\begin{vmatrix}\hat{x} & \hat{y} & \hat{z}\\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z}\\ E_x & E_y & E_z\end{vmatrix} \tag{12.4}$$

The electric field vector $\hat{E}$ can be determined explicitly in its component as:

$$\bar{E} = [E_{0x}\hat{x} + E_{0y}\hat{y} + E_{0z}\hat{z}]e^{-j\bar{k}.\bar{r}} \tag{12.5}$$

Note that, E_{0x}, E_{0y} and E_{0z} are not functions of the space since the electric field is constant everywhere in the space, hence space variation is only present in $e^{-j\bar{k}.\bar{r}}$. So taking a component and its derivative w.r.t x will be:

$$\frac{\delta}{\delta x}[E_x, E_y, E_z] = -jk_x[E_x, E_y, E_z] \tag{12.6}$$

This means $\frac{\delta}{\delta x} = -jk_x$, $\frac{\delta}{\delta y} = -jk_y$ and $\frac{\delta}{\delta z} = -jk_z$. Therefore, substituting for $\frac{\delta}{\delta x}$, $\frac{\delta}{\delta y}$ and $\frac{\delta}{\delta z}$ into equation(29.4) we get:

$$\hat{H} = \frac{-1}{j\omega\mu}\begin{vmatrix}\hat{x} & \hat{y} & \hat{z}\\ -jk_x & -jk_y & -jk_z\\ E_x & E_y & E_z\end{vmatrix} \tag{12.7}$$

But, since the quantity $\begin{vmatrix}\hat{x} & \hat{y} & \hat{z}\\ -jk_x & -jk_y & -jk_z\\ E_x & E_y & E_z\end{vmatrix}$ is a cross product of $-jk$ and E, $\bar{H}$ can be written as:

$$\frac{-1}{j\omega\mu}[-j\bar{k}\times\bar{E}] = \frac{1}{\omega\mu}\bar{k}\times\bar{E} \tag{12.8}$$

Recall, $\hat{k}$ is in the direction of wave propagation, $\bar{E}$ is in the direction of the electric field and $\bar{H}$ = cross product of $\hat{k}$ and $\bar{E}$, meaning it is perpendicular to both $\hat{k}$ and $\bar{E}$.

Say we have a wave travelling in an arbitrary direction with an electric field:

$$\bar{E} = E_0 e^{-j\bar{k}.\bar{r}} = \bar{E}_0[e^{-j(\beta\cos\phi_x x + \beta\cos\phi_y y + \beta\cos\phi_z z)}] \tag{12.9}$$

Considering only the space variation in the z-direction:

$$\bar{E} = \bar{E}_0 e^{-j\beta(\cos\phi_x x + \cos\phi_y y)}.e^{-j\beta(\cos\phi_z z)} \tag{12.10}$$

Note that if we move in the x-y plane we have a phase variation which means the x-y plane is not a constant phase plane, this is visible from fig 12.1. But there is also a phase variation in the z-direction.

So looking at the wave, what is the velocity with which the phase point moves in the z-direction?

So we take the phase in the direction of z. thus, the effective phase constant that the wave sees in the z-direction is $\beta\cos\phi_z$.

Wave phase constant in z-direction(βz):

$$\beta z = \beta\cos\phi_z \tag{12.11}$$

So phase velocity in the z-direction:

$$v_{pz} = \frac{\omega}{\beta_z} = \frac{\omega}{\beta\cos\phi_z} \tag{12.12}$$

But $\frac{\omega}{\beta}$ is the velocity of the wave in the direction of the wave propagation. Thus,

$$\frac{\omega}{\beta}\times\frac{1}{\cos\phi_z} = \frac{v_0}{\cos\phi_z} \tag{12.13}$$

Therefore, v_0 is the velocity of the wave in the direction perpendicular to the phase i.e. the wave velocity in an unbound medium.

But the phase velocity of the wave in the z-direction is $\frac{v_0}{\cos\phi_z}$.

We can also have the phase velocity velocities in other directions:

$$\beta_x = \beta\cos\phi_x,\ \beta_y = \beta\cos\phi_y$$
$$v_{px} = \frac{\omega}{\beta_x} = \frac{\omega}{\beta\cos\phi_x} = \frac{v_0}{\cos\phi_x}$$
$$v_{py} = \frac{\omega}{\beta_y} = \frac{\omega}{\beta\cos\phi_y} = \frac{v_0}{\cos\phi_y}$$

Since $\cos\phi_x$, $\cos\phi_y$ and $\cos\phi_z$ are always less than 1, v_p is always greater than v_0(intrinsic velocity).

$v_{px}, v_{py}, v_{pz} \geq v_0$ also when $\cos\phi_x$, $\cos\phi_y$, or $\cos\phi_z$ = 0, v_p = ∞ therefore, $\infty \geq v_{px}, v_{py}, v_{pz} \geq v_0$

Since we are talking about velocities greater than the intrinsic velocities, we know light is an electromagnetic wave. And we know the intrinsic velocity of light will be the velocity of light C i.e. $v_0 = C$. But from physics, we know that the velocity of any system cannot be greater than the velocity of light. Does that mean we have found a mechanism that can transfer energy faster than the speed of light? No. This is because this velocity is not the velocity of any energy packet or physical point in space because the wave defines its velocity which is based on the phase fronts. This essentially gives the condition that $v_p < v_0$.

Re-examining it:

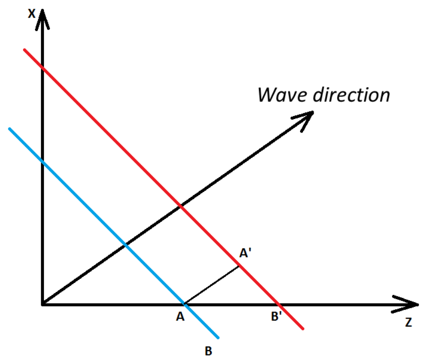


Figure 12.2:

From fig12.2 we have a wave travelling in the x-z plane, with the constant wavefronts shown, by the time the wave travels between two points on its axis, the constant phase points (given by the entire phase fronts) moves between two points(A and B'), unless the wave is parallel to the constant wavefronts the distance moved by the front is always greater than that of the wave, while that happened the point A has not moved to the point B', rather it has measured from A to A', so the point that moved to B' was not A but rather it was from point B which has now moved to B'. So when the wave is moving a defined point (A) on the wavefront, it only moves by a distance AA', however when measuring the phase velocities we simply measure the separation between the points on the z-axis and ask how much time was spent changing positions. So essentially we measure the distance, find how much phase change it has undergone and from there we can get the phase velocity.

So from this, we can see that the phase velocity is not really giving us the velocity of a particular point on the phase line. In fact when we define the phase velocity the entire constant phase plane is behaving like a point, so we just take two points (but both points are representing the same phase) so if we find the same phase we say the distance with a constant phase point has moved from the first to the second point, this is the reason we get the velocity of a particular point on the phase front.

So the velocity, AB (whatever phase velocity we get) is not simply the resolution of the velocity vector in 3-directions, because resolving the velocity in 3-directions will always make it less than the actual velocity but in this case, we see that the components of the phase velocity in three directions v_{px}, v_{py} and v_{pz} are always greater than the velocity vector, so this is not a simple vector resolution of the velocity of the wave, in fact, the phase velocities are calculated from the distance travelled by the constant phase point along the z-axis and that gives the velocity:

$$v_p = \frac{\text{Intrinsic Velocity}}{\text{Direction Cosines}}$$

So as the wave becomes more and more perpendicular to the z-axis: meaning moving in the x direction, the phase front becomes parallel to the z-axis, at point A a small tilt shows that the point is moving very rapidly. So by a small tilt, it would have moved by a large distance in the z-direction, that is the point has moved a small distance in the x-axis when the phase front is almost parallel to the z-axis, what we find is that if the wave was moving along the x-axis and if the phase fronts were parallel to the z-axis, then for a small movement of the wave, the point moves by a very large distance. And if the wave was perfectly parallel to the z-axis the point moves from $-\infty$ to $+\infty$ even for infinitesimal movement of the wave in the x-direction.

So when the phase velocity approaches infinity the wave move in the x-direction, so that the point vector for the wave is in the x-direction, and there is no point moving in the z-direction. As the phase velocity approaches infinity we might ask with what velocity does the energy travel in the z-direction. So in general we might ask a question *if the wave is moving with a velocity v_p in the z-direction, with what velocity is the energy moving in the z-direction?* And we say that the velocity in the AB' direction which is v_0 will be $v_0 \times \cos\phi_z$, so the velocity with which the point A moves in the z-direction will be $v_0 \cos\phi_z$ and if it is in the x-direction it will be $v_0 \cos\phi_x$. So the velocity with which a given point on the phase front moves in the z-direction is called the **group velocity**(v_g)

Group velocity in the z-direction $v_{gz} = v_0 \cos\phi_z$

So since $\cos\phi_z < 1$, v_g(which is the velocity of a particular point)$\leq v_0$

Group velocity in the following:

(i) x-direction: $v_{gx} = v_0 \cos\phi_x$

(ii) y-direction: $v_{gy} = v_0 \cos\phi_y$

(iii) z-direction: $v_{gz} = v_0 \cos\phi_z$

So $0 \leq v_{gx}, v_{gy}, v_{gz} \leq v_0$ but recall, $\infty \geq v_{px}, v_{py}, v_{pz} \geq v_0$.

From this we can see that v_p never goes below v_0 and v_g never goes above v_0, So we have a kind of dividing line:

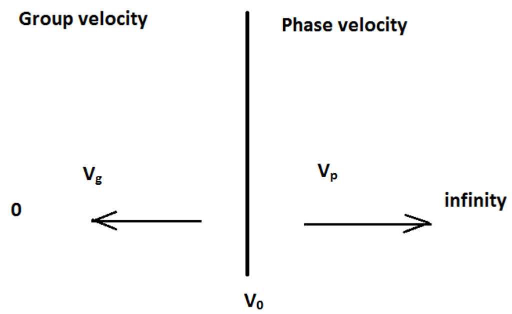


Figure 12.3:

Only in a situation where the wave is travelling in the x,y,z direction or if we find the velocity of the wave both v_p and v_g

in the direction of wave motion then both v_p and v_p will be v_0^2.

$$v_p v_g = v_0^2 \tag{12.14}$$

So taking v_p and v_g in the same direction their product is v_0^2, this means that when v_p approaches infinity the product of v_p and v_g is constant i.e v_0^2.

So as v_p increases v_g reduces. So when calculating the speed of energy flow we find v_g, but when we talk of the movement of phase in the medium, the velocity is given as the phase velocity, v_p.

Wavelength in the following:

(i) x-direction:

$$\begin{aligned}\frac{v_{px}}{f} &= \frac{v_0/\cos\phi_x}{f}\\ &= \frac{\lambda_0}{\cos\phi_x}\end{aligned}$$

(ii) y-direction:

$$\begin{aligned}\frac{v_{py}}{f} &= \frac{v_0/\cos\phi_y}{f}\\ &= \frac{\lambda_0}{\cos\phi_y}\end{aligned}$$

(iii) z-direction:

$$\begin{aligned}\frac{v_{pz}}{f} &= \frac{v_0/\cos\phi_z}{f}\\ &= \frac{\lambda_0}{\cos\phi_z}\end{aligned}$$

12.4 ExerciseList

Ex. 196 — How does wave polarization impact the behavior of a uniform plane wave in transmission and reception?

Ex. 197 — A uniform plane wave exhibiting a reference polarization of z propagates away from the z-axis. Develop representations of this wave in ray-fixed and global cartesian coordinates?

Ex. 198 — Compare and contrast phase velocity and group velocity in the context of uniform plane waves?

Ex. 200 — How does the propagation of a uniform plane wave differ when it travels in different directions?

Chapter 13

Wave propagating in a bound medium

13.1 Learning Outcomes:

i. Understand the complexities of wave propagation within bound mediums with varying permittivity and permeability.

ii. Analyze incident waves at media interfaces and predict their intricate behaviors upon interaction, considering diverse scenarios.

iii. Explore the nuanced concept of phase gradients and their profound impact on altering wave propagation directions.

iv. Apply the laws of reflection and Snell's law to electromagnetic waves in multifaceted boundary conditions.

v. Define and intricately calculate reflection and transmission coefficients for electric fields, considering the complexities of different mediums.

vi. Investigate the profound implications of intrinsic impedance on the nuanced behavior of electric and magnetic fields within complex media interfaces.

vii. Analyze and differentiate between the subtleties of parallel and perpendicular polarization of electric fields concerning the intricate dynamics of the plane of incidence.

viii. Examine and evaluate the characteristics of waves polarized at arbitrary angles, utilizing comprehensive analysis of reflection and transmission coefficients within various mediums.

13.2 Introduction

The wave propagated in some arbitrary direction with respect to the coordinate axes has been studied and investigated in the last chapter. Henceforth, we will study wave propagation in a bound medium. As earlier stated, when we have a bound medium, we do not have the freedom to choose a coordinate axis. Hence, in a bound medium, the choice of the coordinate axis is that which when chosen, will align along the axis, and then the wave propagates in an arbitrary direction with respect to the coordinate axis because the wave is now incident on the boundary from some arbitrary angle.

13.3 Plane wave at media interface

Instead of having a finite medium, we can divide the medium into two semi-infinite media thereby producing what is called a media interface between the two semi-infinite media. The medium at the left of the media interface is infinite while that at the right of the media interface is infinite. Assuming on the left side we have medium 1 and on the right side we have medium 2 and also assuming that the conductivity (σ) for both of the mediums is still zero, that is the media is still lossless but the permittivity (ε) and permeability (μ) for both media are different.

Therefore the permeability for medium 1 is μ_1, the permeability for medium 2 is μ_2, the permittivity for medium 1 is ε_1, and the permittivity for medium 2 is ε_2 as shown in figure 13.1.

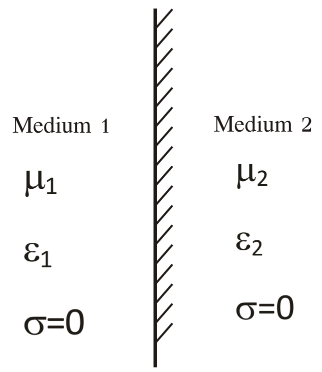


Figure 13.1:

13.4 Incident wave at a media interface

We assume that the coordinate system is oriented such that the media interface is in the XY plane, the x-axis is the vertical axis, the z-axis is the perpendicular axis, that is, the axis perpendicular to the x-axis (90^0 to the x-axis) and the y axis is pointing outward. This is shown in figure 13.2.

Given a wave that is incident on the XY plane at some arbitrary angle with respect to the coordinate axis as illustrated in Fig 13.2, We may be wondering what will happen to the wave when incident on the surface. Some part of the energy of the wave will be transferred to the other medium (medium 2) which constitutes some kind of wave propagation in medium 2. A phenomenon that is also observed is that some part of the wave will get reflected from the boundary interface back into medium 1.

So if the wave is incident on the media interface, part of the wave's energy is transmitted to the second medium and continues with its propagation in medium 2, and the other part gets reflected into the first medium. When the wave is incident on the media interface, two kinds of fields thus will exist. One will be existing in medium 2 and also the field in the first medium must be modified to satisfy the boundary conditions. Part of the energy goes into the second medium and part of the energy gets reflected, that is comes back to medium 1.

As a result of this, *what happens to the plane wave nature? In what direction will the energy be going? What is the magnitude of the field going into the second medium and how much power is transferred into the second medium? How much power gets back from the interface to the first medium and what happens to the direction of the electric field?* These are some of the issues that relate to the propagation of uniform plane waves at the media interface. The media interface becomes a dielectric media interface if conductivity is zero as in this case.

Let θ_i be the angle of incidence at the media interface (the angle between the plane wave and the normal to the XY plane) as shown in fig13.2. The media is uniform in the XY direction and the sudden change in uniformity is at $z = 0$. The z-direction as we have seen is perpendicular to the media interface and hence it is normal to the media interface. So what happens to the wave when it is incident to the media interface at this angle θ_i? The wave is first represented in a form that is a phase, and amplitude, without specifically saying whether we are referring to an electric or a magnetic field. So the wave has some field vectors with the definite phase function since it is travelling at an angle θ_i with respect to the coordinate axis.

Hence, this vector makes three angles θ_x, θ_y,θ_z with the respective coordinate axis. θ_x is the angle which the wave makes with the x axis and it is $\frac{\pi}{2}$-θ_i. The angle θ_y is the angle it makes with the y axis which is $\frac{\pi}{2}$ and θ_z=θ_i. With these angles known, the direction cosines can be written down and then the phase function can be determined. Let β_1 and β_2 be the phase for medium 1 and medium 2 respectively.

$$\beta_1 = \sqrt{\mu_1\varepsilon_1}$$
$$\beta_2 = \sqrt{\mu_2\varepsilon_2}$$

let $\bar{F_i}$ represent this wave incident on the interface (it can be electric or magnetic field) with magnitude$\bar{F_{oi}}$, then

$$\bar{F_i} = \bar{F_{oi}}e^{-j\beta_1(\cos\theta_x x+\cos\theta_y y+\cos\theta_z z)}$$

Since

$$\theta_x = \frac{\pi}{2} - \theta_i$$
$$\cos\theta_x = \sin\theta_i$$
$$\cos\theta_y = 0$$
$$\cos\theta_z = \cos\theta_i$$

Substituting in the incident wave vector,

$$\bar{F_i} = \bar{F_{oi}}e^{-j\beta_1(\cos\theta_x+\cos\theta_y+\cos\theta_z)}$$

So the wave which is travelling at an angle θ_i with respect to the normal to the interface can be represented by a field

$$\bar{F_i} = \bar{F_{oi}}e^{j\beta_1(x\sin\theta_i+z\cos\theta_i)}$$

At some instant of time, the variation of the field amplitude as a function of x,y, and z is obtained by taking the real part

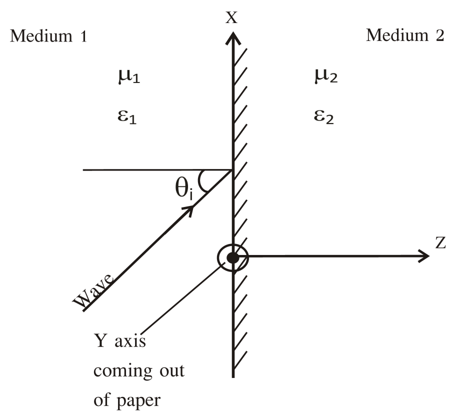


Figure 13.2:

of $\bar{F_i}$ magnitude of the field on the XY plane at $z = 0$, $\mathfrak{Re}\{\bar{F_i}\}$=$\bar{F_{oi}}\cos(\beta_1 x\sin\theta_i)$

So the field is having a variation which is cosine variation in the x direction and it does not have any variation in the Y direction. If the field were to be plotted, the amplitude will vary along the X axis and it will be constant in the y direction as shown in figure 13.3.

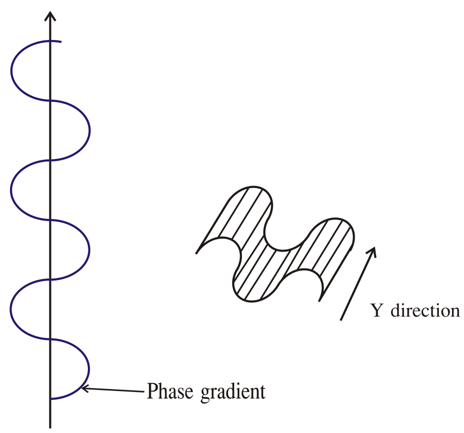


Figure 13.3:

Essentially, it is more like a corrugated surface on the XY plane(think of a zinc sheet for roofing). It has the asbestos sheet plane of undulation as shown in figure 13.3.

It has been established that if the wave is incident in the direction shown, the amplitude variation on the XY media interface appears like a corrugated surface. So we can visualize the field as an amplitude variation in the x direction or it has a phase variation $\beta_1\sin\theta_i$ or phase gradient -$\beta_1\sin\theta_i$. The phase gradient will be equal to zero if $\theta_i = 0$.

Thus, the phase becomes constant and this is what happens when the wave is moving perpendicular to the XY plane and therefore will not have any gradient in that plane. So the direction of the wave, as it changes, affects the phase gradient in the plane.

Creating a phase gradient and changing the direction of the wave are similar. This simply means that changing the wave direction leads to a particular gradient and also creating a phase gradient leads to a wave that orients in the same direction to satisfy the phase gradient. If the phase gradient at the interface is given and this value of phase gradient is equated to $-\beta_1\sin\theta_i$, we get some quantity θ_1 and this gives us the effective direction in which the wave is travelling. In essence, this simply means that if a wave is given in some arbitrary direction, a phase gradient is created but if the phase gradient is given, we get a wave travelling at an angle that will satisfy the phase gradient created.

We have seen the wave travelling from medium 1 which was making an angle θ_i with respect to the media interface and have also seen that the wave creates a phase gradient. If the wave was incident from medium 2 in the direction of θ_i, it creates exactly the same phase gradient on the XY plane. So irrespective of whether this incident wave is coming from medium 1 or 2 both are going to create exactly the same phase gradient on the media interface.

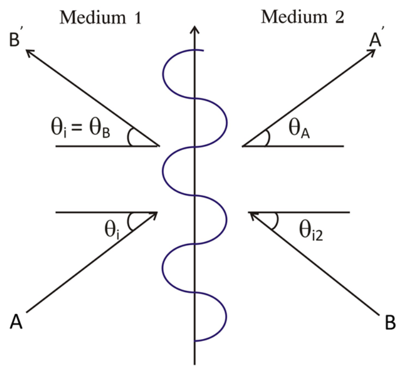


Figure 13.4:

From the diagram in figure 13.4, waves A and B are incident waves from medium 1 and medium 2 respectively. These two waves will create the same phase gradient respectively. B in medium 2 corresponds to a wave B' at an angle θ_B travelling away from the interface in medium 1. A in medium 1 corresponds to a wave A' at an angle θ_A travelling away from the interface in medium 2.

A wave that is approaching the interface and a wave that is going away from the interface at the same angle made with respect to the normal creates the same phase gradient. Two other waves and their direction that can result in some phase gradient are B' and A. B' is a result of the reflection of incident wave A and A' is a result of the wave transmitted into medium 2 as a result of incident wave A. All these waves result in a constant phase gradient at the XY plane when z=0 (A' is called the transmitted wave or refracted wave) and B' is called the reflected wave and wave A that brought energy to the media interface is called the incident wave as shown in figure 13.5.

In general, the wave vector for incident wave, reflected wave, and transmitted wave all lie on the same plane called the plane of incidence. The y-axis is perpendicular to the plane of incidence. The important conclusion is that when a wave was incident on a dielectric interface (in this case) at an angle, it induces two waves called the reflected wave going from the interface through medium 1 and the refracted wave going away from the interface to medium 2 as illustrated in figure 13.5.

All these three waves have the same phase variation and all lie in the same plane called the plane of incidence. If **E** is the magnitude of the electric field and H is the magnitude of the magnetic field, then $\dfrac{E}{H} = \eta$ holds for medium 1 and medium 2 as the wave always sees an infinite medium ahead of its media interface. This incident wave always excites the electric and magnetic field in medium 2. This excitation creates **E** and **H** that will satisfy $\dfrac{E}{H} = \eta_2$. η is known as the intrinsic impedance. We can't satisfy boundary conditions by having $\dfrac{E}{H} = \eta_1$ in medium 1 and $\dfrac{E}{H} = \eta_2$ in medium 2 for both electric and magnetic fields. This means that a certain field has to be induced in the first medium. In other words, we have to modify the fields in the first medium so that the boundary condition can be satisfied for electric and magnetic fields at the interface. So there will be fields induced in both medium 1 and medium 2 due to the incident wave in order to satisfy boundary conditions. The reason is that the incident wave leads to both refracted waves in medium 2 and reflected waves in medium 1 as shown in figure 13.5. Since the inducing phenomenon is due to phase variation, the two other waves induced in the interface will have the same phase variation as the original incident wave. Since the fields are time-varying, they constitute a wave phenomenon.

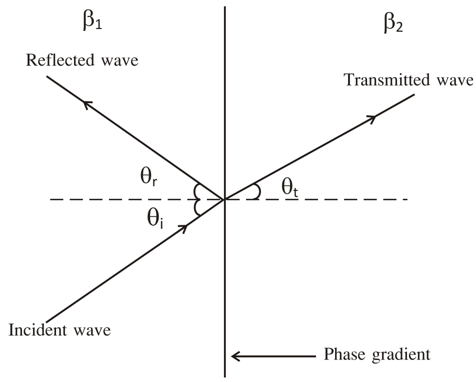


Figure 13.5:

We recall that light is a transverse electromagnetic wave and if we can recall the various laws of reflection for light, the first law states that the incident ray, reflected ray, and transmitted ray all lie on the same plane called the plane of incidence. So the first law of reflection comes from satisfying the phase condition on the interface. The plane of incidence is that which contains the wave vector and the normal to the interface. Thus we get the first law of reflection established by the phase condition.

If -$\beta_1 \sin\theta_1$ is the phase gradient created because the wave moved in medium 1, for medium 2, θ_2 is obtained by equating the phase gradient in medium 1 and medium 2 (that is -$\beta_1 \sin\theta_1 = -\beta_2 \sin\theta_t$, where θ_t is the angle the transmitted wave makes with the normal). In general, the angle that the waves make with the normal is given by the formula below

$$\theta = \frac{\text{phase gradient}}{\text{phase constant}}$$

Let θ_r and θ_t denote the angle of reflection and the angle of refraction respectively. The diagram below clearly explains θ_r and θ_t. From

$$-\beta_1 \sin\theta_i = -\beta_1 \sin\theta_r$$

$$\sin\theta_r = -\frac{-\beta_1 \sin\theta_1}{\beta_1}$$

From the above formula we get

$$\theta_i = \theta_r$$

This is called the second law of reflection. The angle of incidence and the angle of refraction are equal. This law is familiar in optics, that if a light ray is incident at an angle and the refractive index of the two mediums are given as m_1 and m_2 respectively, then the transmitted or reflected angle θ_t is given by

$$\sin\theta_t = \frac{m_1 \sin\theta_i}{m_2}$$

This is a special case of Snells law.

So whatever angle of incidence with respect to the normal it is the same angle with which the wave is reflected. similarly

$-\beta_1 \sin\theta_1 = -\beta_2 \sin\theta_2$[1]

or

$$\frac{\sin\theta_i}{\sin\theta_t} = \frac{\beta_2}{\beta_1}$$

But $\beta_1 = \omega\sqrt{\mu_1\varepsilon_1}$ and also $\beta_2 = \omega\sqrt{\mu_2\varepsilon_2}$. Substituting, we have

$$\omega\sqrt{\mu_1\varepsilon_1}\sin\theta_1 = \omega\sqrt{\mu_2\varepsilon_2}\sin\theta_2$$

ω cancels out and the resulting equation gives Snell's law. This is the important law that tells the direction of transmitted electromagnetic waves with respect to the direction of incident waves. We call this the generalized Snell's law. So for pure dielectric with μ negligible, Snell's law reduces to

$$\sqrt{\varepsilon_1}\sin\theta_1 = \sqrt{\varepsilon_2}\sin\theta_2$$

If ε_{r1} and ε_{r2} is the refractive index of medium 1 and medium 2 respectively and $\varepsilon_{r1} = n_1$ and $\varepsilon_{r2} = n_2$, then

$$\sqrt{\varepsilon_o\varepsilon_{r1}}\sin\theta_i = \sqrt{\varepsilon_o\varepsilon_{r2}}\sin\theta_2$$

reduces to $n_1 \sin\theta_i = n_2 \sin\theta_T$.

Therefore, in summary, we started the wave propagation in some arbitrary direction. In this arbitrary direction, a phase gradient is created at the media interface. This phase gradient is the same for the wave which is induced on both sides of the media called the transmitted wave and reflected wave. We found that changing the direction of the wave and changing the phase gradient is essentially equivalent. Hence the induced field has the same gradient in the interface and there essentially we can find out the direction in which the wave will be travelling. From there, the laws of reflection and Snell's law were established. Let us consider the diagram in figure 13.6 where we have the magnitude of the electric field. The arrows show the wave direction of propagation meaning that the electric field is oriented in some other direction. So with amplitude, $\bar{E}$, $\bar{E}_r$,and $\bar{E}_T$, We want to find out the relationship between $\bar{E}_t$ and $\bar{E}_i$ which is $\frac{\bar{E}_r}{\bar{E}_i}$.

This quantity is what we call the TRANSMISSION COEFFICIENT, the ratio of the interface. The ratio of E_r to E_i is what we call the reflection coefficient. We have some situations where the wave sees a change in intrinsic impedance which causes impedance discontinuity, a part of the energy is thus reflected and the reflection coefficient is $\frac{\bar{E}_r}{\bar{E}_i}$. The reflection coefficient $\Gamma = \frac{\bar{E}_r}{\bar{E}_i}$, transmission coefficient $\tau = \frac{\bar{E}_r}{\bar{E}_i}$. So we have Γ and τ as the electric field reflection coefficient and electric field transmission coefficient. These quantities can be defined for magnetic fields also. As we already know, the $\bar{E}$ and $\bar{H}$ fields are related by the intrinsic impedance of the medium. If we know $\bar{E}$ then $\bar{H}$ can be found out since $\frac{\bar{E}}{\bar{H}} = \eta$ where

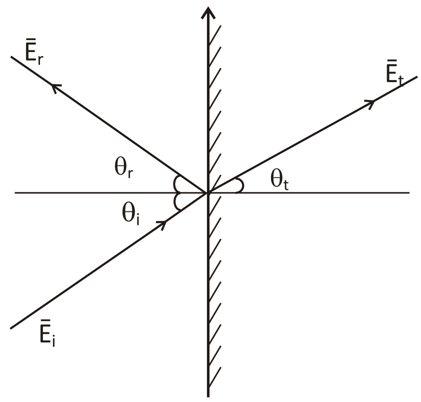


Figure 13.6:

η is the intrinsic impedance. So for medium 1, $\frac{E_r}{H_r} = \eta_1$ and $\frac{\bar{E}_t}{\bar{H}_t} = \eta_1$. For medium 2, $\frac{\bar{E}_t}{\bar{H}_t} = \eta_2$. η_1 and η_2 are the intrinsic impedance for medium one and medium 2 respectively. The electric field will be in a direction with an angle to the plane of incidence in which the wave vector lies. These electric fields can be resolved into two components. These are parallel polarization where the electric field lies in the plane of incidence and perpendicular polarization in which the electric field lies perpendicular to the plane of incidence. Γ and τ have to be calculated in order to analyse these two cases. Once this is done, the total characteristics of the wave which is polarized at an arbitrary angle can be examined.

Questions

1. Elaborate on the intricacies of wave propagation within bound mediums with disparate permittivity and permeability, highlighting the implications on incident, reflected, and refracted waves at media interfaces.
2. Discuss the interrelation between phase gradients and alterations in wave direction, explaining their profound impact on incident waves at interfaces and their resultant behaviors.
3. Utilizing Snell's law and the laws of reflection, derive and elucidate the generalized form of Snell's law for electromagnetic waves in diverse mediums. Analyze its implications on waves encountering interfaces with varying intrinsic impedances.
4. Define and intricately calculate the reflection and transmission coefficients for electric fields within bound mediums exhibiting complex boundary conditions and varying intrinsic impedances.
5. Analyze and differentiate between parallel and perpendicular polarization of electric fields concerning the nuanced dynamics of the plane of incidence within diverse media interfaces.
6. Explore the complexities of wave behavior within dielectric interfaces, considering the intricate relationships between incident, reflected, and transmitted waves in contexts with non-negligible magnetic permeability.
7. Discuss the implications of altering the incident angle on the reflection and transmission coefficients in electromag-

netic wave propagation within bound mediums with diverse intrinsic impedances.

8. Elaborate on the behavior of incident waves polarized at arbitrary angles, discussing the resultant complexities in reflection and transmission coefficients within media interfaces with varying permittivity and permeability.

9. Analyze the nuanced behavior of electric fields within media interfaces experiencing impedance discontinuities, elaborating on the effects of altered intrinsic impedances on wave polarization and propagation.

10. Investigate the intricate relationships between wave direction, intrinsic impedance alterations, and the resultant behaviors of incident waves encountering diverse boundary conditions in media interfaces.

Chapter 14

Total Internal Reflection

14.1 Learning Outcomes:

i. Understand the conditions and implications of total internal reflection.
ii. Analyze reflection coefficients and their relevance in different polarizations during total internal reflection.
iii. Explain the behavior of fields in medium 2 and their role in total internal reflection.
iv. Define reactive fields and their significance in the context of power flow during total internal reflection.
v. Describe the relationship between total internal reflection and guiding electromagnetic energy in wave guides.

14.2 Introduction

In this chapter, we will discuss a special case of reflection across a dielectric boundary called Total Internal Reflection. We studied total internal reflection at the high school level where it was understood that, if a ray is incident at an angle greater than the critical angle, the ray is completely reflected back into the same medium. However, this understanding is a rather superficial one. Here we do a detailed study of the phenomenon of TOTAL INTERNAL REFLECTION.

14.3 Snell's Law

From Snell's Law,

$$\beta_1 \sin\theta_i = \beta_2 \sin\theta_t \tag{14.1}$$

where $\beta_1 and \beta_2$ are the phase constants in the two media. Recall;

$$\beta_1 = \omega\sqrt{\mu_1\epsilon_1}$$
$$\beta_2 = \omega\sqrt{\mu_2\epsilon_2}$$

$$\omega\sqrt{\mu_1\epsilon_1}\sin\theta_i = \omega\sqrt{\mu_2\epsilon_2}\sin\theta_t$$
or
$$\sqrt{\mu_1\epsilon_1}\sin\theta_i = \sqrt{\mu_2\epsilon_2}\sin\theta_t$$

If the medium parameters are such that,

$$\frac{\beta_1}{\beta_2}\sin\theta_i > 1 \tag{14.2}$$

Then there is no wave nature in medium 2 for which the angle $\sin\theta_t$ can propagate. It is more of an imaginary angle which lacks direction. And even though phase conditions in equation 14.1 is satisfied, there is no physical angle θ_t for which

$$\sin\theta_t > 1 \tag{14.3}$$

This poses a very interesting case. Now we require in our reflection co-efficient expression that the quantity

$$\cos\theta_t = \sqrt{1-\sin^2\theta_t} \tag{14.4}$$

Substituting the expression for $\sin^2\theta_t$,

$$\cos\theta_t = \sqrt{1-(\frac{\beta_1}{\beta_2}\sin\theta_i)^2}$$

this leaves the quantity $1-(\frac{\beta_1}{\beta_2}\sin\theta_i)^2$ always negative and thus, the quantity $\sqrt{1-(\frac{\beta_1}{\beta_2}\sin\theta_i)^2}$ imaginary. So we make it positive to satisfy the medium parameter in which $\frac{\beta_1}{\beta_2}\sin\theta_i > 1$

$$\cos\theta_t = \jmath\sqrt{(\frac{\beta_1}{\beta_2}\sin\theta_i)^2-1} \tag{14.5}$$

Now again, we see that from equation 14.3 and equation 14.5 that there is no physical angle θ_t for which the wave would be propagated.

However for the angle θ_t and θ_i we would still have reflection and transmission coefficients, so we go back and substitute $\cos\theta_t$ in the reflection co-efficient expression.

Recall we had two cases of reflection coefficients.

For the Perpendicular case:

$$\Gamma_\perp = \frac{\eta_2\cos\theta_i - \eta_1\cos\theta_t}{\eta_2\cos\theta_i + \eta_1\cos\theta_t} \tag{14.6}$$

For the Parallel case:

$$\Gamma_\parallel = \frac{\eta_1\cos\theta_i - \eta_2\cos\theta_t}{\eta_1\cos\theta_i + \eta_2\cos\theta_t} \tag{14.7}$$

Note: η_1 and η_2 are interchanged for both cases otherwise they are similar.

substituting equation 14.5 in equation 14.7,

$$\Gamma_\parallel = \frac{\eta_1\cos\theta_i - \jmath\eta_2\sqrt{(\frac{\beta_1}{\beta_2}\sin\theta_i)^2-1}}{\eta_1\cos\theta_i + \jmath\eta_2\sqrt{(\frac{\beta_1}{\beta_2}\sin\theta_i)^2-1}} \tag{14.8}$$

Note: we would get the same expression for the perpendicular polarization case but with η_1 and η_2 interchanged.

What's important to note here is that the expression in equation 14.8 is of the form

$$\frac{a - \jmath b}{a + \jmath b} \tag{14.9}$$

and as we know, expressions of the form always have a magnitude equal to 1 at an angle $\angle - 2\tan^{-}1(\frac{b}{a})$ that is:

$$\frac{a - \jmath b}{a + \jmath b} = 1\angle - 2\tan^{-1}(\frac{b}{a}) \tag{14.10}$$

And since the modulus(or magnitude) of the reflection coefficient is 1, whatever power is incident on the media interface is totally reflected. So in this case, even though we have a dielectric boundary, the entire power is reflected back into medium 1 with a phase change given by $\angle - 2\tan^{-}1(\frac{b}{a})$ and no power is transmitted into medium 2. Hence, this phenomenon is referred to as **Total internal reflection**.

14.4 Conditions For Total Internal Reflection

(i) For total internal reflection to occur, then the condition below has to be satisfied.

$$\frac{\beta_1}{\beta_2}\sin\theta_i > 1$$

in terms of medium parameters

$$\frac{\sqrt{\mu_1\epsilon_1}}{\sqrt{\mu_2\epsilon_2}}\sin\theta_i > 1$$

θ_i can attain a maximum value of 90^o, therefore the maximum value of $\sin\theta_i$ = sin90 = 1, this means, for the above condition to be meant,

$$\sqrt{\mu_1\epsilon_1} > \sqrt{\mu_2\epsilon_2} \tag{14.11}$$

So if $\sqrt{\mu_1\epsilon_1} > \sqrt{\mu_2\epsilon_2}$, then there is a possibility of total internal reflection.

$\sqrt{\mu_1\epsilon_1} > \sqrt{\mu_2\epsilon_2}$ implies that medium 1 is a denser medium compared to medium 2. So essentially, when the wave goes from a denser medium to a less dense medium, then there is a possibility of total internal reflection.

For a pure dielectric(or non magnetic material) where, μ_1 and μ_2 are equal,

$$\sqrt{\epsilon_1} > \sqrt{\epsilon_2}$$

Recall, refractive index, $n = \sqrt{\epsilon}$, thus

$$n_1 > n_2$$

Where n_1 and n_2 are the refractive indexes of medium 1 and medium 2 respectively.

Now the angle at which $\sin\theta_t = 1$ is called the CRITICAL ANGLE, beyond which if an incident wave is launched, there would be no transmitted wave in the second medium.

This phenomenon is shown in Figure 14.1.

$$\sqrt{\mu_1\epsilon_1}\sin\theta_c = \sqrt{\mu_2\epsilon_2}$$

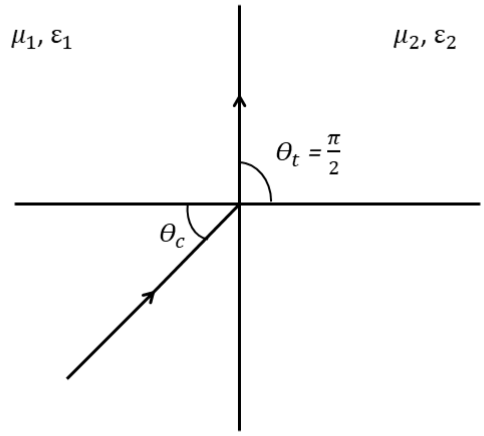


Figure 14.1: Critical angle

similarly for pure dielectric media,

$$n_1\sin\theta_c = n_2$$

Or the critical angle

$$\theta_c = \sin^{-1}(\frac{n_2}{n_1})$$

Now what we have not discussed, however, is that, when the total internal reflection takes place, the wave which is reflected, undergoes a phase change at the interface. This phase change is a function of the medium parameters and the angle of incidence. So in general, we can say that by choosing appropriate parameters for the media, and the angle of incidence, we can generate an arbitrary phase difference between the incidence and the reflected wave. This would happen for both parallel and perpendicular polarization. However, the parallel and perpendicular polarization have η_1 and η_2 interchanged between them. So a and b in $\frac{a-\jmath b}{b+\jmath b}$ are different for parallel and perpendicular polarizations. This means that the phase changes which the wave undergoes for the two polarizations are different.

(ii) The wave undergoes a phase change at Total Internal Reflection the phase change is different for parallel and perpendicular polarizations.

If we write out explicitly, the phase changes for parallel and perpendicular polarization, we have;

$$\phi_{\parallel} = -2\tan^{-1}\frac{\eta_2\sqrt{(\frac{\beta_1}{\beta_2}\sin\theta_i)^2 - 1}}{\eta_1\cos\theta_i} \tag{14.12}$$

$$\phi_{\perp} = -2\tan^{-1}\frac{\eta_1\sqrt{(\frac{\beta_1}{\beta_2}\sin\theta_i)^2 - 1}}{\eta_2\cos\theta_i} \tag{14.13}$$

From here we see that the two polarizations do not undergo the same phase change. But later we'll see the implication of the phase change for the different kinds of polarization when Total Internal Reflection takes place in the dielectric boundary.

(iii) For fields in medium 2

Recall:

The phase expression for a transmitted wave in medium 2 is

$$\bar{E}_t = E_t e^{-j\beta_2(x\sin\theta_t + z\cos\theta_t)}$$

From Snell's Law and keeping phase gradient same $\bar{E}_t$ is written out explicitly as:

$$\bar{E}_t = E_t e^{-jx\beta_2\sin\theta_t - jz\beta_2\cos\theta_t} \tag{14.14}$$

substituting for $\cos\theta_t$ from equation 14.5 and $\sin\theta_t$ from equation 14.1

$$\bar{E}_t = E_t e^{-jx\beta_1\sin\theta_i - z\beta_2\sqrt{\left(\frac{\beta_1}{\beta_2}\sin\theta_i\right)^2 - 1}}$$

$$\bar{E}_t = E_t e^{-jx\beta_1\sin\theta_i - z\sqrt{(\beta_1\sin\theta_i)^2 - \beta_2^2}} \tag{14.15}$$

From the above equation,

The amplitude term is:

$$e^{-z\sqrt{\beta_1^2\sin^2\theta_i - \beta_2^2}} \tag{14.16}$$

and the phase term:

$$e^{-jx\beta_1\sin\theta_i} \tag{14.17}$$

From these equations, we see the field is exponentially dying down as a function of z and the phase varies as a function of x in medium 2.

Now with total internal reflection, the travelling wave nature is only in the x-direction as the phase variation only occurs in the x-direction. So the fields are no longer travelling in the z-direction.

Essentially, equation 14.16 gives decaying fields in the z direction and equation 14.17 gives a travelling wave in the x direction. This is depicted in figure 14.2

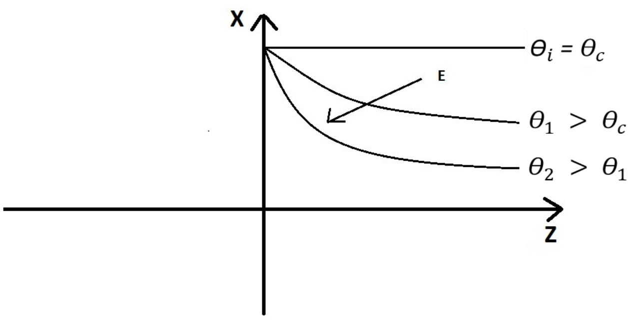


Figure 14.2: Amplitude decay of fields in medium 2

At angle $\theta_i = \theta_c$, the term $e^{-z\sqrt{\beta_1^2\sin^2\theta_i - \beta_2^2}}$ reduces to 1 so the amplitude of the field remains constant. Hence the field does not decay in medium 2. However, as θ_i increases more and more beyond θ_c, the decay becomes faster as shown in figure 14.2. So at the critical angle, the field essentially extends up to infinity at a constant amplitude and as the angle of incidence increases more and more beyond the critical angle, the field gets more and more confined towards the dielectric interface. The important observation is that the field never goes to zero in medium 2 under any condition. The field is always non-zero in medium 2. Theoretically, no matter how small this amplitude is, it will extend up to infinity in the z-direction as it gets closer and closer to zero but never reaches zero. So when total internal reflection takes place at the media interface, from what we learnt in high school, it was like the total power got reflected in medium 1 and nothing existed in medium 2. There was no discussion at all as to what was the behaviour of the fields in medium 2 or the role medium 2 needed to play in total internal reflection. Once we had the critical condition for total internal reflection satisfied, we didn't bother about the point beyond the boundary into medium 2. But now we understand that we cannot ignore the region of medium 2 beyond the boundary. The fields have to exist in the form shown in medium 2 for total internal reflection to take place, and they are as important as the fields in medium 1. The fields in medium 2 are required for us to satisfy the continuity equation at the boundary. It is these fields in medium 2 that support the total internal reflection phenomenon. If any disturbance is created to these fields in medium 2, then the total internal reflection phenomenon is also disturbed and we cannot have a total internal reflection. So to have a good total internal reflection phenomenon, we must have the field in medium 2 as confined to the interface as much as possible. That means we must launch a wave with an angle larger than the critical angle for as much as possible so that the field dies down rapidly in medium 2 as fast as possible and thus confined to a thickness z in medium 2 that is extremely small.

Secondly, we must provide a certain region from the dielectric interface in medium 2 of a thickness where the fields in medium 2 are protected. If these two conditions are guaranteed, then we will have a total internal reflection phenomenon. Theoretically, unless we provide an infinite region in medium 2, and the fields properly protected, then we cannot have an ideal total internal reflection phenomenon.

14.5 Reactive Fields

We conclude that when total internal reflection takes place, the fields in medium 2 exist, and the transmission coefficient is not zero. When the total power is reflected, $|\Gamma| = 1$ but this doesn't mean that the transmission coefficient is zero because the transmission coefficient doesn't always mean that there is a power flow in medium 2. So as we have seen, the fields which are dying down in medium 2 are the ones which do not constitute power flow as the wave is not travelling inside the medium rather the wave is still travelling in the x direction along the dielectric interface. Thus there is no power flow in the z direction, but the fields exist. So we should clearly understand the difference between having fields and no power flow and having fields with power flow. You may have electric and magnetic fields and they may constitute power flow, as we have seen in earlier cases. But however, we might have a situation like this where we have electric and magnetic fields but there is no power flow. In the electrical circuit terminology, we can call these fields **REACTIVE FIELDS**, as they have the energy stored in the transient phase when the field is being set up. We visualize it as when the wave is incident on the interface or boundary as the wave is being set up, there is some power flow into the second medium. That power flow gets stored in the second medium in the fields. Once these fields are set up in the steady state, no net power flows in the second medium and the power flow is essentially in medium 1. These fields in medium 2 are referred to as EVENESCENT FIELDS (reactive fields). They will exist in medium 2 without constituting any power flow. The power flow will exist in the interface in the direction of x. Now for medium 1, we had a wave incident on it and another wave which is reflected from the interface into medium 1. There is an arbitrary phase change at the point of incidence and reflection as shown below.

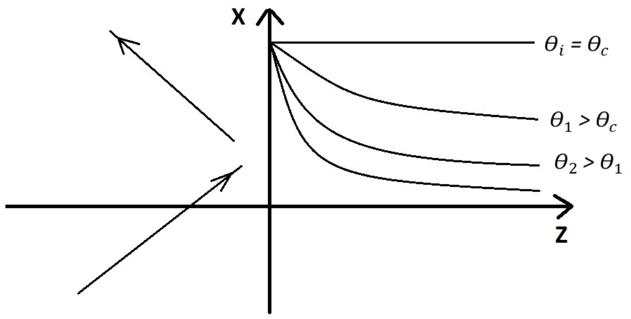


Figure 14.3: Graph of incidence and reflection

For fields in medium 1;

$$E_i = E_i e^{-\jmath\beta_1(x\sin\theta_i + z\cos\theta_i)} \tag{14.18}$$

$$E_r = E_r e^{-\jmath\beta_1(x\sin\theta_i - z\cos\theta_i)} \tag{14.19}$$

So the total electric field in medium 1 is given as;

$$E = E_i + E_r$$

$$E = E_i e^{-\jmath\beta_1(x\sin\theta_i + z\cos\theta_i)} + \mathbf{E}^{\cdot}\mathbf{r} e^{-\jmath\beta_1(x\sin\theta_i - z\cos\theta_i)}$$

Since the magnitude of the reflection coefficient is 1, $|E_i| = |E_r|$ but with a phase difference. So we can say that, $E_r = E_i e^{j\phi}$. Hence,

$$E = E_i e^{-\jmath\beta_1 x\sin\theta_i}\{e^{-\jmath\beta_1 z\cos\theta_i} + e^{\jmath\phi}e^{\jmath\beta_1 z\cos\theta_i}\}$$

Depending on the phase difference ϕ, the term $e^{-\jmath\beta_1 z\cos\theta_i} + e^{\jmath\phi}e^{\jmath\beta_1 z\cos\theta_i}$ is always going to create a constructive or destructive interference in the z-direction. While the term $e^{-\jmath\beta_1 x\sin\theta_i}$ gives a travelling wave in the x direction.

So in medium 1 from the equation, we have a fully developed standing wave in the z direction but a travelling wave in the x direction. In medium 2 it was an exponentially decaying field and a travelling wave in the x direction with the same phase constant as in medium 1 travelling wave in the x direction. Now we can then plot the fields in the two media depending upon the phase of reflection ϕ, so the field distribution is like a corrugated surface in medium 1 and a decaying field in medium 2. The constant phase plane is determined by $e^{-\jmath\beta_1 x\sin\theta_i}$. So the constant phase plane is the yz plane. The constant amplitude plane is in the xy plane. So this complex amplitude distribution of the wave travels in the x direction with a phase constant $\beta_1 \sin\theta_i$.

So whenever total internal reflection occurs, you have this complex distribution of amplitude in space, one side exponentially decaying fields, the other side standing wave kind of fields. This amplitude distribution travels in the x direction with a phase velocity in the x direction $v_{px} = \frac{\omega}{\beta_1 \sin\theta_i}$.

Now we may ask, *what is the direction of net power flow?* The fields in medium 2 do not constitute any wave, so no power flow in the z-direction. With no power flow in the z direction in medium 2, there can be no power flow in the z direction in medium 1 also. Since if the power had come from medium 1 in the z direction, and the power can't flow in medium 2, the interface does not consume power since we are talking of a completely lossless medium. Hence whatever power came from the medium 1 incident on the boundary, essentially gets reflected back to give the standing wave shown in medium 1. This implies that there is no power flow perpendicular to the interface but there is a power flow along the interface since we have a travelling wave in the x direction. So when the total internal reflection phenomenon takes place, the net power flow is along the interface or in other words, the dielectric interface is capable of guiding electromagnetic energy.

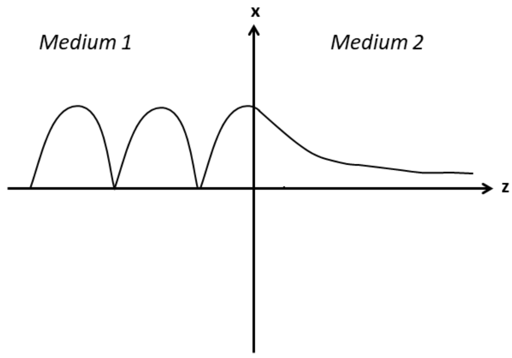


Figure 14.4: Graph of field distribution

14.6 Wave Guides

So we can make use of a dielectric interface for guiding electromagnetic energy along a path. Precisely this is what is used in structures called WAVEGUIDES, and especially when we talk about dielectric media this is what happens in optical fibres. An optical fibre is nothing but a structure having a dielectric interface. So with a wave launched at an angle $\theta_i > \theta_c$, then there is a total internal reflection in medium 1. The field in medium 2 will exponentially decay. Power flow will be along the interface.

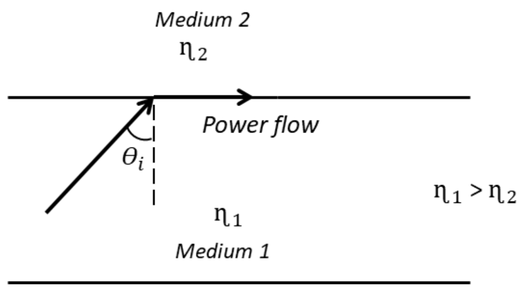


Figure 14.5: Power flow in optical fibre structure

So if we create some structure like this, it can carry electromagnetic energy over a long distance without any loss as such. Energy is confined in the region of n_1 and it will propagate along the interface as a result of total internal reflection. So total internal reflection has played an important role in modern optical communication and helped in the design of structures called optical fibres.

Questions

1. Describe the mathematical conditions necessary for total internal reflection and explain the critical angle's role.
2. Derive the reflection coefficients for both perpendicular and parallel polarizations in total internal reflection.
3. How do phase changes differ between parallel and perpendicular polarizations during total internal reflection? Explain with equations.
4. Analyze the behavior of fields in medium 2 during total internal reflection and discuss their impact on the phenomenon.
5. Define reactive fields and their relationship to power flow in the context of total internal reflection. Provide examples.
6. Discuss the role of Snell's Law in determining the possibility of total internal reflection and its significance.
7. Compare and contrast the amplitude and phase distributions of waves in medium 1 and medium 2 during total internal reflection.
8. Explain the critical angle's significance in determining the behavior of transmitted waves during total internal reflection.
9. How does total internal reflection facilitate the guiding of electromagnetic energy in wave guides such as optical fibers? Discuss its practical applications.
10. Analyze the phase velocity in the x-direction and power flow along the interface during total internal reflection, elaborating on their relationships and implications.

Chapter 15

Polarization across Dielectric Interface

15.1 Learning Outcomes:

i. Understanding how arbitrary polarization of an electromagnetic wave across a dielectric interface can be decomposed into orthogonal polarizations, parallel and perpendicular to the plane of incidence.
ii. Exploring how incident waves decomposed into parallel and perpendicular components reflect and transmit differently at dielectric boundaries.
iii. Analyzing the effects of normal reflection and total internal reflection on the polarization state, particularly with linearly and circularly polarized incident waves.
iv. Understanding Brewster's angle for parallel and perpendicular polarization, how it defines angles where reflected waves become linearly polarized or transmission occurs without reflection.
v. Utilizing dielectric interfaces to manipulate and change the state of polarization in electromagnetic waves, especially achieving linear polarization from arbitrarily polarized waves.

15.2 Introduction

In the previous chapters, we have been investigating the behaviour of uniform plane waves across a dielectric media interface. We studied two cases which are the parallel and perpendicular polarization, and we saw the transmission and reflection across the dielectric boundary. We also investigated a special case called total internal reflection across a dielectric boundary. Now we are going to discuss another important aspect of electromagnetic waves called the **Polarization across dielectric interface**.

We consider a wave which is arbitrarily polarized i.e. the electric field vector makes an arbitrary angle with the plane of incidence and might also be varying as a function of time. We then see how the polarization changes when the wave is reflected from the dielectric interface. So what we are going to investigate now is that, if the incoming wave is coming with a certain polarization, what will be the polarization of the reflected wave and what will also be the polarization of the transmitted wave?

As we have seen earlier, any arbitrary state of polarization can be decomposed into two orthogonal polarizations. Similarly, in this case, we take two linear orthogonal polarizations. One is parallel to the plane of incidence and the other is perpendicular to the plane of incidence. This implies that when the wave is incident on the media interface, the electric field $\bar{E}$ can be decomposed into $\mathbf{E}_{\|}$[1] and $\mathbf{E}_{\perp}$[2] with a phase difference of $e^{j\phi}$ between $\mathbf{E}_{\|}$ and $\mathbf{E}_{\perp}$.

So the polarized wave will be;

$$\mathbf{E} = \mathbf{E}_{\|} + \mathbf{E}_{\perp}.e^{j\phi}$$

15.3 Decomposition of Polarized wave

When an incident polarized wave $\mathbf{E}_i$ hits a dielectric media interface, part of it gets transmitted through the media($\mathbf{E}_t$), while the other part gets reflected($\mathbf{E}_r$). Each of the fields of the wave i.e. $\mathbf{E}_i$, $\mathbf{E}_t$ and $\mathbf{E}_r$ can be decomposed into their orthogonal components as explained above.

15.3.1 Incident wave

An incident polarized wave $\mathbf{E}_i$ can be decomposed into two orthogonal components as shown in equation 15.1

$$\mathbf{E}_i = \mathbf{E}_{i\|} + \mathbf{E}_{i\perp}e^{j\phi} \tag{15.1}$$

figure 15.1 shows the two orthogonal components.

15.3.2 Reflected Wave

Similarly, a reflected polarized wave $\mathbf{E}_r$ can be decomposed into two orthogonal components as shown in figure 15.2. It can be represented in terms of reflection coefficient as shown in figure 15.3

$$\mathbf{E}_r = \mathbf{E}_{r\|} + \mathbf{E}_{r\perp}e^{j\phi} \tag{15.2}$$

$$\mathbf{E}_r = \Gamma_{\|}\mathbf{E}_{i\|} + \Gamma_{\perp}\mathbf{E}_{i\perp}e^{j\phi} \tag{15.3}$$

figure 15.1 shows the two orthogonal components.

15.3.3 Transmitted Wave

A transmitted polarized wave $\mathbf{E}_t$ can be decomposed into two orthogonal components as shown in equation 15.4

$$\mathbf{E}_t = \mathbf{E}_{t\|} + \mathbf{E}_{t\perp}e^{j\phi} \tag{15.4}$$

It can be represented in terms of the transmission coefficient;

$$\mathbf{E}_t = \tau_{\|}\mathbf{E}_{i\|} + \tau_{\perp}\mathbf{E}_{i\perp}e^{j\phi} \tag{15.5}$$

figure 15.1 shows the two orthogonal components.

[1] Parallel component of the polarized wave
[2] The vertical component of the polarized wave

15.3.4 Combination of $\mathbf{E}_i$, $\mathbf{E}_r$ and $\mathbf{E}_t$

In combining the incident wave $\mathbf{E}_i$, reflected wave $\mathbf{E}_r$ and transmitted wave $\mathbf{E}_t$, we will get a wave structure as that shown in figure 15.1

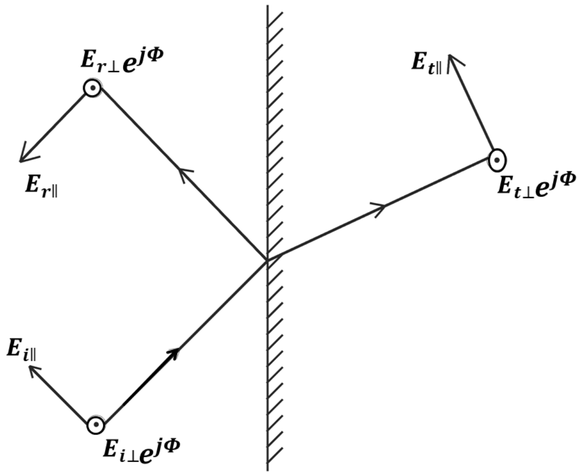


Figure 15.1: Combination of $\mathbf{E}_i$, $\mathbf{E}_r$ and $\mathbf{E}_t$

Depending on whether it is a normal reflection or total internal reflection, τ and Γ could be real or complex quantities. For normal reflection, τ and Γ are real quantities that can have positive or negative values. However, for total internal reflection, τ and Γ become complex with a magnitude of one. Now we have various possibilities for special cases which we shall discuss in the next section as we may be curious to know what happens to the polarization if the wave was linearly polarized and the reflections were normal or a case of total internal reflection, or perhaps what happens if the wave was circularly polarized.

15.4 Linearly Polarized Wave

In a linearly polarized wave, the electric field vector does not change direction as a function of time. That is, the incident electric field vector makes an angle with the plane of incidence, but its direction remains the same irrespective of time. And as we know, a linearly polarized wave can be decomposed into two components that are orthogonal with no phase difference. So if the wave is linearly polarized, the phase difference between $\mathbf{E}_{i\|}$ and $\mathbf{E}_{i\perp}$ is zero. Hence for this case $\phi = 0$. We shall now investigate what will happen in the case of normal reflection and also in the case of total internal reflection.

15.4.1 Normal Reflection

Recall that for normal reflection, $\Gamma_{\|}$ and $\Gamma_{\perp}$ are real. So if $\Gamma_{\|}$ and $\Gamma_{\perp}$ are real and $\phi = 0$, then $\mathbf{E}_{i\|}$ and $\mathbf{E}_{i\perp}$ will have a phase of 0 or π radians between them, depending upon the direction of the electric field. In both cases of 0 or π phase difference, the polarization remains linear since π radians phase shift means a reversal of the electric field. So the orientation of the electric field will change but the polarization will remain linear. So in general, $\Gamma_{\|} \neq \Gamma_{\perp}$. Hence, a linearly polarized wave will remain linearly polarized since the phase difference is unchanged, but the relative amplitude for the reflected waves might change depending upon whether the condition $\Gamma_{\|} \neq \Gamma_{\perp}$ is meant or in some special cases where $\Gamma_{\|} = \Gamma_{\perp}$. This implies that the ratio of both components of the reflected wave will not be the same as it was for the incident wave thus, the direction of the electric field vector is changed for the reflected wave but the wave will remain linearly polarized. So the plane of polarization might change depending upon whether $\Gamma_{\|}$ and $\Gamma_{\perp}$ are equal or not, but the linearly polarized nature of the wave will be maintained.

15.4.2 Total Internal Reflection

On the other hand, if we have total internal reflection, then $\Gamma_{\|}$ and $\Gamma_{\perp}$ become complex, but $|\Gamma_{\|}| = |\Gamma_{\perp}| = 1$. So the phase difference between $\Gamma_{\|}$ and $\Gamma_{\perp}$ component in the reflected wave will change as $\Gamma_{\|}\ \mathbf{E}_{i\|}$ will have its phase and $\Gamma_{\perp}\ \mathbf{E}_{i\perp}\ e^{j\phi}$ will have its own phase. So here the phase difference could be arbitrary which means we would have a linear polarization changing to an elliptic polarization, for the case of total internal reflection. The same thing also applies to the transmitted wave.

Hence for normal reflection, a linearly polarized wave remains linearly polarized whereas for total internal reflection, a linearly polarized wave becomes elliptically polarized.

15.5 Circular Polarization

Now if the incident wave is circularly polarized, what happens to the transmitted and reflected waves at normal reflections and total internal reflection?

For circular polarization, $|\mathbf{E}_{r\|}| = |\mathbf{E}_{r\perp}|$ and the phase difference between them is $\pm\frac{\pi}{2}$. So $\phi = \pm\frac{\pi}{2}$.

15.5.1 Normal Reflection

Again under normal reflection, $|\ \Gamma_{\|}\ | \neq |\ \Gamma_{\perp}\ |$ but $\Gamma_{\|}$ and $\Gamma_{\perp}$ are real quantities. That means the phase difference between the reflected wave remains $\phi = \pm\frac{\pi}{2}$ but the amplitude ratio for the reflected wave for perpendicular and parallel are not equal because $|\ \Gamma_{\|}\ | \neq |\ \Gamma_{\perp}\ |$. So $|\ \mathbf{E}_{r\|}\ | \neq |\ \mathbf{E}_{r\perp|}$ and the phase difference between the two remains $\pm\frac{\pi}{2}$. Now, the vector sum in this case with varying amplitude gives an elliptical polarization but, the major and minor axis align with the coordinate axis for the ellipse. So the incidence wave was circularly polarized but the reflected wave became elliptically polarized. Since the phase difference between these reflected waves remains $\pm\frac{\pi}{2}$, the major and minor axis of the reflected wave ellipse will lie in the plane of incidence and perpendicular to the plane of incidence respectively or vice versa but the angle which the major axis makes with the plane of incidence could be either zero or $\frac{\pi}{2}$ so the tilt angle for the ellipse with respect to the plane of incidence is either 0 or $\frac{\pi}{2}$.

15.5.2 Total Internal Reflection

With total internal reflection, $|\ \Gamma_{\|}\ | = |\ \Gamma_{\perp}\ | = 1$ but $\Gamma_{\|}$ and $\Gamma_{\perp}$ are complex quantities, meaning $\mathbf{E}_{r\|}$ and $\mathbf{E}_{r\perp}$ will have same amplitude with $\mathbf{E}_{i\|}$ and $\mathbf{E}_{i\perp}$ respectively but with different phase which is arbitrary. So we get a wave once again that is elliptically polarized.

In summary, a circularly polarized wave under normal reflection gives a reflected wave with elliptical polarization with its major axis at 0 or $\dfrac{\pi}{2}$ to the plane of incidence (tilt angle) and under total internal reflection, a circularly polarized incident wave, in general, gives an elliptically polarized reflected wave.

So in general, we can draw the conclusion that, for an incident wave which is elliptically polarized, the state of polarization of the reflected wave will change depending upon the medium parameters. Hence, an elliptically polarized wave can become circularly polarized or linearly polarized depending upon the medium.

Essentially the important point we have to note from this discussion is that the media interface can be used to alter the state of polarization of an electromagnetic wave. So if we have a certain state of polarization for the incoming wave, then by launching the wave at an arbitrary angle, we should be able to change the polarization to a desired state.

One of the special cases of this is that if the incident wave is some arbitrarily polarized wave, it is possible to make the reflected wave linearly polarized. There are many applications where we require a linearly polarized wave. The source we might have may not generate physically a linearly polarized wave. So we look for some mechanism by which an arbitrarily polarized wave can be converted into a linearly polarized wave as we saw that the dielectric interface can change the polarization of the electromagnetic wave.

15.6 Brewster Angle

We recall from earlier sections that

$$\Gamma_{\|} = \frac{\eta_1 \cos\theta i - \eta_2 \cos\theta t}{\eta_1 \cos\theta i - \eta_2 \cos\theta t} \tag{15.6}$$

$$\tau_{\|} = \frac{2\eta_2 \cos\theta i}{\eta_1 \cos\theta i + \eta_2 \cos\theta t} \tag{15.7}$$

and,

$$\Gamma_{\perp} = \frac{\eta_2 \cos\theta i - \eta_1 \cos\theta t}{\eta_2 \cos\theta i - \eta_1 \cos\theta t} \tag{15.8}$$

$$\tau_{\perp} = \frac{2\eta_2 \cos\theta i}{\eta_1 \cos\theta i + \eta_2 \cos\theta t} \tag{15.9}$$

If $\eta_1 \cos\theta_i$ - $\eta_2 \cos\theta_t$ = 0 for parallel or $\eta_2 \cos\theta_i$ - $\eta_1 \cos\theta_t$ = 0 for perpendicular polarization i.e Γ = 0, then the wave incident on the dielectric interface is completely transmitted to the second medium. That means if we put a perpendicularly polarized wave at an angle $\Gamma_\perp$ = 0, then the perpendicularly polarized wave goes through without reflection. Similarly a parallel polarized wave incident on the medium with the angle at which $\Gamma_\|$ = 0, the parallel polarized wave passes through and there is no reflection. Now if we have an arbitrarily polarized wave and launch at an angle which satisfies $\Gamma_\perp$ = 0, only the perpendicularly polarized component goes through. If it launched at an angle at which $\Gamma_\|$ = 0 only the parallel component of polarization passes through the media interface. In either case, the reflected wave would have parallel and horizontal polarization for $\Gamma_\|$ = 0 and $\Gamma_\perp$ = 0 respectively. In both situations what is transmitted becomes essentially linearly polarized. We have now the interesting case that for a given parameter, we look for the angle at which $\Gamma_\|$ or $\Gamma_\perp$ goes to zero. At this angle, one of the polarization is reflected, and a complimentary polarization is transmitted to the second medium. So an arbitrary state of polarization can be converted to linear polarization.

For parallel polarization, $\Gamma_\|$ = 0 when $\eta_1 \cos\theta$i - $\eta_2 \cos\theta$t = 0 or $\eta_1 \cos\theta$i = $\eta_2 \cos\theta$t. For perpendicular polarization, $\Gamma_\perp$ = 0 i.e $\eta_2 \cos\theta$i - $\eta_1 \cos\theta$t = 0 or $\eta_2 \cos\theta$i = $\eta_1 \cos\theta$t. Where η_1 and η_2 are the intrinsic impedances of medium 1 and 2 respectively. From Snell's law $\beta_1 \sin\theta$i = $\beta_2 \sin\theta$t. By doing some algebra, we can solve for the incident angle θ_i in the case of parallel and perpendicular polarization and the expressions are shown in equation 15.10 and 15.11.

$$\tan\theta_{B\|} = \frac{\beta_2}{\eta_2}\left\{\frac{\eta_1^2 - \eta_2^2}{\beta_2^2 - \beta_1^2}\right\}^{\frac{1}{2}} \tag{15.10}$$

$$\tan\theta_{B\perp} = \frac{\beta_2}{\eta_1}\left\{\frac{\eta_2^2 - \eta_1^2}{\beta_2^2 - \beta_1^2}\right\}^{\frac{1}{2}} \tag{15.11}$$

These angles at which the reflection coefficient goes to zero, that is, $\Gamma_\|$ = $\Gamma_\perp$ = 0 are called the BREWSTER ANGLE, therefore θ_B is the Brewster angle. So we have a Brewster angle for parallel and perpendicular polarizations. In general, if we take a medium which could be magnetic, then the Brewster angle for both polarization exists. So at the Brewster angle for parallel polarization, the reflected wave will have only a perpendicularly polarized wave and at the Brewster angle for perpendicular polarization, the reflected wave has only a parallel wave. The Brewster angles for parallel and perpendicular polarization can be derived as follows.

15.6.1 Parallel Polarization

Previously, we derived the relationships below

$$\begin{aligned}\eta_1 \cos\theta_i &= \eta_2 \cos\theta_t \\ &= \eta_2\sqrt{1 - \sin^2\theta_t} \\ &= \eta_2\sqrt{1 - \left\{\frac{\beta_1}{\beta_2}\right\}^2 \sin^2\theta_i}\end{aligned} \tag{15.12}$$

because $\beta_1 \sin\theta_i$ = $\beta_2 \sin\theta_t$ from snell's law. so

$$\beta_2\eta_1 \cos\theta_i = \eta_2\sqrt{\beta_2^2 - \beta_1^2 \sin^2\theta_i}$$

or

$$\cos\theta_i \frac{\beta_2\eta_1}{\eta_2} = \sqrt{\beta_2^2 - \beta_1^2 \sin^2\theta_i} \tag{15.13}$$

if we square both sides of equation 15.13, we get

$$\cos^2\theta_i \frac{\beta_2^2\eta_1^2}{\eta_2^2} = \beta_2^2 - \beta_1^2 \sin^2\theta_i$$

or

$$\frac{\beta_2^2\eta_1^2}{\eta_2^2}(1 - \sin^2\theta_i) = \beta_2^2 - \beta_1^2 \sin^2\theta_i \tag{15.14}$$

$$\frac{\beta_2^2\eta_1^2 - \beta_2^2\eta_2^2}{\eta_2^2} = \left(\frac{\beta_2^2\eta_1^2}{\eta_2^2} - \beta_1^2\right)\sin^2\theta_i$$

$$\beta_2^2\eta_1^2 - \beta_2^2\eta_2^2 = (\beta_2^2\eta_1^2 - \beta_1^2\eta_2^2)\sin^2\theta_i$$

$$\sin^2\theta_i = \frac{\beta_2^2\eta_1^2 - \beta_2^2\eta_2^2}{\beta_2^2\eta_1^2 - \beta_1^2\eta_2^2} \tag{15.15}$$

from trigonometry, we have that,

$$\cos^2\theta_i = 1 - \sin^2\theta_{ii}$$

so substituting that in equation 15.15 we get,

$$\cos^2\theta_i = 1 - \left\{ \frac{\beta_2^2\eta_1^2 - \beta_2^2\eta_2^2}{\beta_2^2\eta_1^2 - \beta_1^2\eta_2^2} \right\} \tag{15.16}$$

also from trigonometry, we know that

$$\tan^2\theta_i = \frac{\sin^2\theta_i}{\cos^2\theta_i} = \left\{ \frac{\beta_2^2\eta_1^2 - \beta_2^2\eta_2^2}{\beta_2^2\eta_1^2 - \beta_1^2\eta_2^2} \right\} = \left\{ \frac{\beta_2^2\eta_2^2 - \beta_2^2\eta_1^2}{\beta_2^2\eta_1^2 - \beta_1^2\eta_2^2} \right\}$$

replacing θ_i with $\theta_{B\|}$

$$\tan\theta_{B\|} = \sqrt{\frac{\beta_2^2(\eta_1^2 - \eta_2^2)}{\eta_2^2(\beta_2^2 - \beta_1^2)}} = \frac{\beta_2}{\eta_2}\left\{ \frac{\eta_1^2 - \eta_2^2}{\beta_2^2 - \beta_1^2} \right\}^{\frac{1}{2}} \tag{15.17}$$

15.6.2 Perpendicular Polarization

Let's assume that the electric field was perpendicular to the plane of incidence as shown in figure 15.2 (Pointing out of the plane of the paper as indicated by the dot). Irrespective of variation in θ_i, $\mathbf{E}_i$ will remain perpendicular to the plane of incidence and its amplitude is fixed. So when the continuity of the electric field is applied, we get a reflected and transmitted field. The sum of $\mathbf{E}_i$ and $\mathbf{E}_r$ must be equal to $\mathbf{E}_t$ i.e. tangential component of the electric field must be continuous across the boundary. The amplitude of $\mathbf{E}_i$, $\mathbf{E}_r$ and $\mathbf{E}_t$ don't depend on lunching angle θ_i. Hence the reflected field $\mathbf{E}_r$ will always be needed to satisfy the boundary condition at the interface.

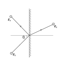


Figure 15.2: Perpendicular polarization

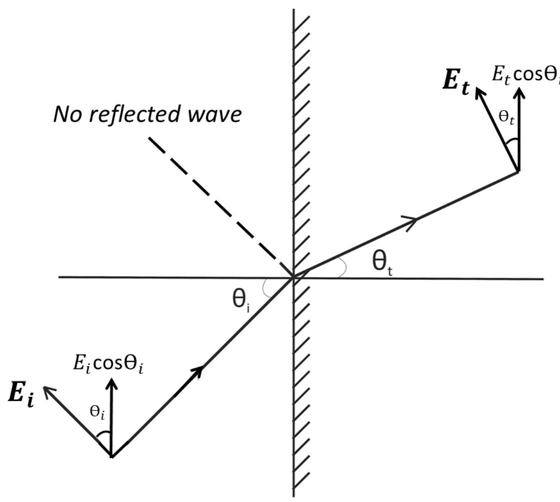


Figure 15.3: Parallel polarization

However, for the parallel polarization, variation of θ_i affects the tangential component of $\mathbf{E}_i$ and $\mathbf{E}_t$ and they vary with θ_i as shown in figure 15.3. So there is an angle at which $\mathbf{E}_t$ tangential equals $\mathbf{E}_i$ tangential i.e $\mathbf{E}_i\cos\theta_i = \mathbf{E}_t\cos\theta_t$, and then, the electric field due to reflection is not required any more to satisfy the boundary condition. That means $\mathbf{E}_r$ will not be needed anymore to satisfy the boundary condition at the interface hence, the reflected wave is absent. So at $\mathbf{E}_i\cos\theta_i = \mathbf{E}_t\cos\theta_t$, the wave incident is completely transmitted. Since the tangential component varies with the angle of incidence, at some point, there is no reflected electric field required to satisfy the boundary condition.

For perpendicular polarization, this is not the case because the electric field perpendicular to the plane of the paper does not depend on θ_i and so $\mathbf{E}_r$ is always required, so that $\mathbf{E}_i + \mathbf{E}_r = \mathbf{E}_t$. If $\mu_1 \neq \mu_2$ i.e. the media were magnetic, so whatever applies to the electric field applies to the magnetic fields and then we have Brewster angle for perpendicular polarization also. However, for a media which is a pure dielectric, ie $\mu_1 = \mu_2$, then the Brewster angle will only exist for parallel polarization and will not exist for perpendicular polarization.

$\theta_{B\|} = \tan^{-1}\left(\sqrt{\frac{\epsilon_2}{\epsilon_1}}\right) = \tan^{-1}\left(\frac{\eta_2}{\eta_1}\right)$ were η_2 and η_1 are the refractive index of medium 2 and medium 1 respectively because $\frac{\eta_2}{\eta_1} = \sqrt{\frac{\epsilon_2}{\epsilon_1}}$.

Generally, for the dielectric media, we have the Brewster angle for parallel polarization. One can make use of this Brewster angle for converting an arbitrarily polarized wave to a linearly polarized wave. From the diagram below, the incident wave is arbitrarily polarized hence it would have perpendicular and parallel polarized components. At an incidence angle = $\theta_{B\|}$, the reflected wave is made out of a purely perpendicular polarized field $\mathbf{E}_{r\perp}$ as the parallel polarized reflected wave will not exist. For the transmitted wave , the $\mathbf{E}_{i\|}$ passes through since incidence angle is $\theta_{B\|}$. Some of the perpendicular components will pass through also as $\theta_{B\|}$ is not its polarizing angle of $\theta_{B\perp}$.

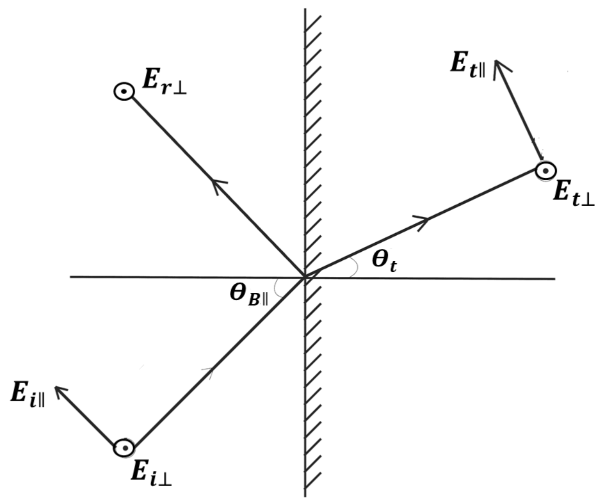


Figure 15.4:

Hence it is just like a normal wave incident at an angle. So the transmitted wave can in general be elliptically polarized since it has both parallel and perpendicular polarized components. So without worrying about the state of polarization of the incoming wave, whether linear, circular or elliptically polarized, if the wave is launched at the Brewster angle, then the reflected wave will have a polarization which is linear polarization and its orientation will be perpendicular to the plane of incidence. Hence for any arbitrarily polarized wave incident at the Brewster angle $\theta_{B\|}$, the reflected wave $\mathbf{E}_{r\perp}$ points outward in the manner shown perpendicular to the plane of incidence. For this reason, $\theta_{B\|}$ is also called the **polarizing angle**. Since a wave launched at that angle has a reflected wave which is linearly polarized. The polarizing angle finds application in polarizing the light beams in lasers, i.e. if the polarization was arbitrary, we can always use the concept of Brewster angle to make sure the reflected wave is linearly polarized.

So our conclusion from the discussion is that, by using the dielectric interfaces, one can change the state of polarization of electromagnetic waves. Though this has been shown here for optical beams, for other radio frequencies, one can also make use of this phenomenon for changing the state of polarization of electromagnetic waves. Later when we talk about reflection and refraction from conducting boundaries, we will consider how the state of polarization would be affected from the incident to the reflected wave. We have seen for the light incident at Brewster angle, an arbitrary polarization gives a reflected light that is linearly polarized. In conducting boundaries, however, the case is a little simpler and we will observe an interesting phenomenon then. So this essentially gives us an idea of how to use dielectric boundaries to change the state of polarization of electromagnetic waves.

Now suppose we have a general media boundary for which $\sigma_1 \neq \sigma_2 \neq 0$, we may ask, *what happens to the reflected and transmitted wave in this general media?* To solve this problem we first satisfy the boundary conditions at the interface that is, we have to satisfy the boundary condition for the tangential components of the electric and magnetic fields. Even though conductivity is not zero, it is not infinity either so we still do not have a surface current and because of that, we can still apply the boundary conditions which we applied earlier for the dielectric media interface hence, the expressions which we got for $\Gamma_{\|}$, $\Gamma_{\perp}$, $\tau_{\|}$ and $\tau_{\perp}$ are still valid. But the expressions for the intrinsic impedances in both cases η_1 and η_2 now becomes $\eta_1 = \sqrt{\dfrac{j\omega\mu_1}{\sigma_1 + j\omega\epsilon_1}}$ and $\eta_2 = \sqrt{\dfrac{j\omega\mu_2}{\sigma_2 + j\omega\epsilon_2}}$ as we have taken the conductivity into consideration. And because of this, the transmission and reflection coefficients become complex quantities.

So for a general lossy media like this, the analysis is the same with η_1 and η_2 appropriately chosen. But with $\gamma = \alpha + j\beta$ being complex in the presence of loss, indicating that there is some attenuation in the system.

Earlier in our discussion, when a plane wave was incident on a dielectric interface, it did not matter from where the wave originated since the wave amplitude doesn't change even if we travel an infinite distance for a lossless medium. However for a lossy media if we assume that this wave was incident on the boundary at an infinite distance from the surface and γ_1 = $\alpha_1 + j\beta_1$ ie complex, then the wave amplitude decreases exponentially as it travels. So after travelling an infinite distance, the amplitude which reaches the boundary interface of the lossy media will be zero. Similarly, if we had considered that the wave had a finite amplitude at the interface, coming from an infinite distance as a plane wave, then it should have had an infinite amplitude from the point where the wave originated. So in the case of a lossy medium, it would look as if we require an infinite amount of energy to start with if the wave travels from infinity. This is a hypothetical situation. Without getting into the issue of where the wave originated from, we can assume that when the wave reaches the media interface, it will have a certain amplitude and then the amplitude of the reflected wave just before the interface and the amplitude of the transmitted wave just after the media interface will be given by the reflection and transmission coefficient respectively.

Now when a wave travels backwards from medium 2 to medium 1, the incident wave amplitude is going to grow exponentially as you move away from the interface further into medium 1. So in medium 2, we have a travelling wave with propagation constant γ_2 and whose amplitude is exponentially decaying as we move in the direction of travel further into medium 2.

In medium 1, the fields seen are a superposition of the incident and reflected waves. The incidence wave exponentially grows and the reflected wave exponentially decays as we go away from the interface. This case is identical to a lossy transmission line. As we go away from the interface or termination point in a transmission line, the reflected wave becomes weaker and weaker, the transmitted wave grows larger. So essentially as you go away from the interface into medium 1, we see a phenomenon equivalent to that of a travelling wave in a transmission line and when we come close to the boundary, we see some kind of standing wave behaviour. This is identical to what we had seen for a lossy transmission line.

So in conclusion, the analysis of a lossy media interface where the conductivity is finite(it is not 0 or ∞), i.e. non of the media is an ideal conductor or dielectric, the problem can still be handled exactly in the same way as we did for the dielectric media interface. We can apply the same boundary conditions and the same expressions for the reflection and transmission coefficients. Then substituting appropriately for the intrinsic impedances for the two media, we can get the expressions for the reflection and transmission coefficients for a lossy media interface.

Questions

1. How does the decomposition of an arbitrary polarization into parallel and perpendicular components assist in understanding the behavior of waves at dielectric interfaces?
2. Explain how the phase difference between the parallel and perpendicular components of a polarized wave affects its polarization after reflection and transmission across a dielectric boundary.
3. In what scenarios does a linearly polarized incident wave maintain its polarization after reflection? Contrast this with the outcomes during total internal reflection.
4. Describe the transformation of polarization for circularly polarized incident waves in normal reflection versus total internal reflection across a dielectric interface.
5. What role does Brewster's angle play in altering the state of polarization? Explain how it impacts incident waves and the resulting reflected waves at dielectric boundaries.
6. How does the Brewster angle aid in converting arbitrarily polarized waves into linearly polarized waves at a dielectric interface?
7. Discuss the significance of Brewster's angle in practical applications involving lasers and polarization of light beams.
8. Compare the polarization outcomes for dielectric interfaces with those for conducting boundaries. What notable differences exist in the behavior of the reflected waves?
9. Explain the influence of conductivity differences between two media on the reflected and transmitted waves at a general media boundary.

10. Why is it possible to apply boundary conditions for tangential components of electric and magnetic fields even when conductivity is neither zero nor infinity at a dielectric interface?

Chapter 16

Reflection from a Conducting Boundary

16.1 Learning Outcomes:

i. Explain the phenomenon of reflection from a conducting boundary and its implications in guiding electromagnetic waves.
ii. Apply boundary conditions for perpendicular polarization to derive the reflection coefficient for a conducting boundary.
iii. Analyze the behavior of electric and magnetic fields in a dielectric medium upon reflection from a conducting boundary.
iv. Interpret the concept of complex waves involving standing waves and traveling waves resulting from reflections.
v. Evaluate the utility of conducting boundaries in guiding electromagnetic energy and the principles underlying waveguides.

16.2 Introduction

In this chapter, we will discuss the reflection of a uniform plane wave from a conducting boundary. This will lay a foundation for a structure called a *Waveguide*. A waveguide is basically a linear structure that conveys or guides electromagnetic waves between its endpoints.

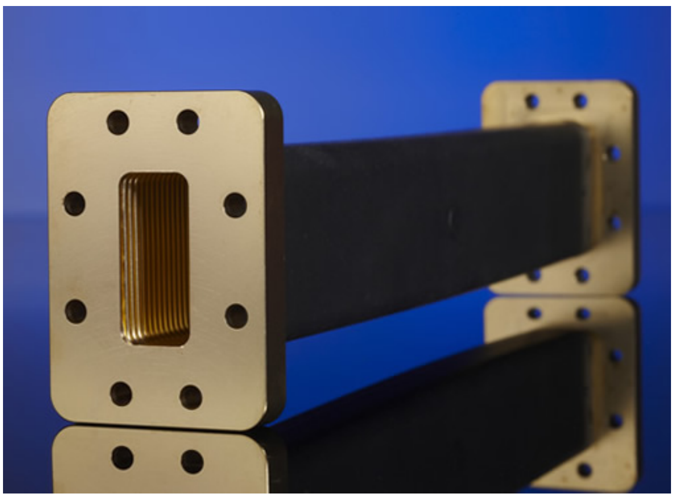


Figure 16.1: A waveguide

The phenomenon of reflection from a conducting surface is considered in terms of the solutions of Maxwell equations where amplitudes of reflected and transmitted waves are found as functions of an incident wave and the conductivity of the plane. This conducting plane lies between two distinct media:

(i) An ideal dielectric medium.

(ii) An ideal conductive medium.

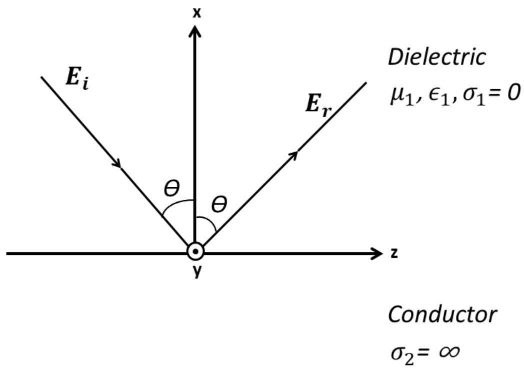


Figure 16.2: Conducting plane between two media

With reference to figure 16.2, we have a horizontal conductive plane which divides the medium into two parts. Below the plane is an ideal conductor with an infinite conductivity ($\sigma_2 = \infty$) and above the plane is an ideal dielectric with conductivity $\sigma_1 = 0$, permeability μ_1 and permittivity ε_1. This plane or boundary can therefore be referred to as a **Dielectric conductor boundary**.

There can be no wave propagation in the conductive medium as time-varying fields in a conductor are zero due to its infinite conductivity. So the wave can only be propagated from the dielectric medium and then incident on the conductive plane boundary.

As shown in figure 16.2, the wave $\boldsymbol{E_i}$ is incident from the dielectric side making an angle θ with the plane but it is not propagated to the conductor side because of the infinite conductivity of the conductor, hence no wave is transmitted. The wave E_r is however reflected at an angle θ with the normal of the plane[1].

Again we can consider the case of perpendicular and parallel polarizations and carry out the analysis of the boundary as we have done for a dielectric media interface. However, this is a simpler case as we only have two waves(incident and reflected waves) to match the boundary conditions.

[1]From the law of reflection, the angle of incidence is equal to the angle of reflection.

16.3 Perpendicular Polarization

Taking the perpendicular polarization, E_i and E_r are oriented in the y direction and by using a Poynting vector, we get the direction for the magnetic fields H_i (incident magnetic field) and H_r (reflected magnetic field). We can resolve the magnetic field into two components which are either **parallel** or **perpendicular** to the plane:

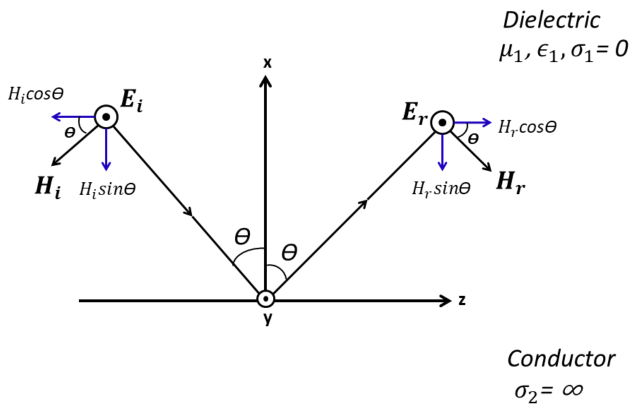


Figure 16.3: Magnetic Field

(i) $H_i sin\theta$ and $H_r sin\theta$ are the normal (perpendicular) components to the plane.

(ii) $H_i cos\theta$ and $H_r cos\theta$ are the tangential (parallel) components of the plane.

Since we are having a conducting boundary, we will have the presence of surface current, hence, we can either use the boundary conditions with appropriate surface current and the tangential component of the magnetic fields or use only the boundary condition for the perpendicular components of the magnetic fields without worrying about the surface current. Since there is no wave propagation in the conducting medium, the time-varying fields E_t and H_t are zero, and the boundary conditions now have to be satisfied by only the incident and reflected waves.

Writing the expressions for the incident and reflected Electric and Magnetic fields and taking the appropriate components of these fields to satisfy the boundary conditions, we have that the ratio of the electric to the magnetic field for both the incident and reflected waves is equal to the intrinsic impedance η. That is η is the intrinsic impedance given as $\eta = \sqrt{\frac{\mu}{\varepsilon}} = \dfrac{E_r}{H_r} = \dfrac{E_i}{H_i}$.

Note that we are not using subscript 1 for the dielectric medium(medium 1) or subscript 2 for medium 2 as there is no wave propagation in the conducting medium(medium 2).

16.3.1 Electric Field

$$\bar{E}_i = E_i e^{-j\beta(-x\cos\theta + z\sin\theta)}\hat{y} \tag{16.1}$$

$$\bar{E}_r = E_r e^{-j\beta(x\cos\theta + z\sin\theta)}\hat{y}^2 \tag{16.2}$$

Where $\bar{E}_i$ is the incident electric field and $\bar{E}_r$ is the reflected electric field. E_i and E_r are the amplitudes of the electric fields, $e^{-j\beta}$ is the phase function and β is the propagation constant given as $\omega\sqrt{\mu\varepsilon}$

[2] $\hat{y}$ signifies the field is in the y direction

16.3.2 Magnetic Field

$$\bar{H}_i = (-H_i sin\theta\hat{x} - H_i cos\theta\hat{z})e^{-j\beta(-xcos\theta + zsin\theta)} \tag{16.3}$$

$$\bar{H}_r = (-H_r sin\theta\hat{x} + H_r cos\theta\hat{z})e^{-j\beta(xcos\theta + zsin\theta)} \tag{16.4}$$

[3] where $\bar{H}_i$ is the incident magnetic field and $\bar{H}_r$ is the reflected magnetic field. H_i and H_r are the amplitudes of the magnetic fields. As stated earlier $\frac{E_r}{H_r} = \eta = \frac{E_i}{H_i}$, hence,

$$H_i = \frac{E_i}{\eta} \tag{16.5}$$

$$H_r = \frac{E_r}{\eta} \tag{16.6}$$

From equations (16.5) and (16.6), we can rewrite the expressions for H_i and H_r in equations (16.3) and (16.4) in terms of E_i and E_r respectively;

$$\bar{H}_i = -\frac{E_i}{\eta}(sin\theta\hat{x} + cos\theta\hat{z})e^{-j\beta(-xcos\theta + zsin\theta)} \tag{16.7}$$

$$\bar{H}_r = \frac{E_r}{\eta}(-sin\theta\hat{x} + cos\theta\hat{z})e^{-j\beta(xcos\theta + zsin\theta)} \tag{16.8}$$

So now that we have the expressions for the electric and magnetic fields of the incident and the reflected waves, we need to obtain the boundary conditions which are appropriate for the plane. The boundary conditions are:

(i) The tangential component of the electric field should be continuous across the plane i.e. $\bar{E}_{tangential}$ must be continuous.

(ii) The normal component of the magnetic field should also be continuous across the plane i.e. $\bar{H}_{normal}$ must be continuous.

In figure 16.3, we can see that the tangential component to the plane is the electric field. At this point, if we equate $x = 0$ in equations 16.1 and 16.2, the sum of the two electric fields should be zero because the fields are continuous and there are no fields in the conductive medium.

$$\bar{E} = (E_i + E_r)e^{-j\beta(zsin\theta)}\hat{y} = 0 \tag{16.9}$$

Such that it simplifies to $E_i = -E_r$ i.e, $\frac{E_r}{E_i}$ = -1 **(reflection coefficient)**.

So in this case, the electric field reflection coefficient is always equal to -1. Similarly, the sum of the two normal components of the magnetic fields at x equal to 0 should be 0 because the normal component is also continuous at the boundary.

Recall that, in transmission line analysis, the reflection coefficient is -1 for a **short-circuited load**. This means that the conducting boundary is identical to the short circuit condition on a transmission line. Therefore, the dielectric medium on which the wave is propagated is analogous to a transmission line. When the electric wave reaches this ideal conductive boundary, it is completely reflected from the boundary with a phase

[3] $\hat{x}$ signifies it's in the x direction and $\hat{z}$ signifies it's in the z direction

reversal (phase difference of 180 degrees). So this boundary essentially behaves like a short circuit in the transmission line terminology.

Now, we want to know the kind of field patterns that are created when the electric field is completely reflected in the dielectric medium.

16.4 Electric fields in the dielectric medium

Recall that E_i = -E_r

By superimposition of the two waves,
$\bar{E} = \bar{E}_i + \bar{E}_r$

Substitute E_i = - E_r and add equations (16.1) and (16.2)

$$\bar{E} = \bar{E}_i e^{-j\beta z \sin\theta}(e^{j\beta x \cos\theta} - e^{-j\beta x \cos\theta})\hat{y} \tag{16.10}$$

Recall that $e^{jx} - e^{-jx} = 2jsinx$
From that analogy,
$e^{j\beta xcos\theta} - e^{-j\beta xcos\theta} = 2jsin(\beta xcos\theta)$

Substituting into equation 16.10

$$\bar{E} = 2j\bar{E}_i \sin(\beta x \cos\theta)e^{-j\beta z \sin\theta}\hat{y} \tag{16.11}$$

From equation 16.11, we have the electric fields with a sinusoidal variation in the x-direction and a phase term which is in the z-direction. This means that the electric field has a **standing wave** behaviour in the x direction because the function sin($\beta xcos\theta$) does not have a phase but it has an amplitude variation which is the nature of a complete standing wave. So equation 16.11 represents something like a standing wave which is in the x direction and a travelling wave which is given by this term $e^{-j\beta zsin\theta}$ in the z-direction. The equation thus gives us a **standing wave** in the x-direction and a **travelling wave** in the z-direction. Therefore, the wave is a composite phenomenon of the incident and the reflected wave making it a **complex wave**.

A complex wave is the combination of a standing wave which is in a direction perpendicular to the plane or boundary and a travelling wave which is in the direction of the plane.

The same can be obtained for the magnetic fields when we combine the incident and reflected magnetic fields. We therefore would have an expression for standing and travelling magnetic waves.

16.5 Magnetic fields in the dielectric medium

Taking the two components of the magnetic fields in the x and z direction $\boldsymbol{H_x}$ and $\boldsymbol{H_z}$ respectively, and adding equations 16.3 and 16.4, we obtain

$$H_x = \frac{-E_i}{\eta}(sin\theta\hat{x} + cos\theta\hat{z})e^{-j\beta(-xcos\theta+zsin\theta)} + \frac{E_r}{\eta}\left(-sin\theta\hat{x} + cos\theta\hat{z}\right)e^{-j\beta(xcos\theta+zsin\theta)}$$

$$H_x = \frac{-E_i}{\eta}(sin\theta e^{j\beta(xcos\theta)}) - \frac{E_r}{\eta}\sin\theta e^{-j\beta(xcos\theta)}(e^{-j\beta zsin\theta}) \tag{16.12}$$

Recall that E_i = -E_r and substituting into equation 16.12

$$Hx = \frac{-E_i}{\eta}(sin\theta)(e^{j\beta(x\cos\theta)} - e^{-j\beta x\cos\theta})(e^{-j\beta zsin\theta}) \tag{16.13}$$

Recall that $e^{jx} - e^{-jx} = 2jsinx$. From that analogy, $e^{j\beta xcos\theta} - e^{-j\beta xcos\theta} = 2jsin\beta xcos\theta$ and substitute into equation (16.13)

$$H_x = -2j\frac{E_i}{\eta}sin\theta sin(\beta xcos\theta)(e^{-j\beta zsin\theta}) \tag{16.14}$$

The x component H_x has similar behaviour with the electric field. That is, it has a standing wave component $(sin\beta xcos\theta)$ which is in the x-direction and a travelling wave component $(e^{j\beta zsin\theta})$ in the z-direction.

The same can be done for the z component of the magnetic field H_z.

$$H_z = \frac{-E_i}{\eta}cos\theta(e^{j\beta xcos\theta} + e^{-j\beta xcos\theta})(e^{-j\beta zsin\theta}) \tag{16.15}$$

Recall that $e^{jx} + e^{-jx} = 2cosx$
From that analogy, $e^{j\beta xcos\theta} - e^{-j\beta xcos\theta} = 2cos(\beta xcos\theta)$ and substitute into equation (16.15)

$$H_z = \frac{-2E_i}{\eta}cos\theta cos(\beta xcos\theta)(e^{-j\beta zsin\theta}) \tag{16.16}$$

From equation (16.16), we see that the z component of the magnetic field Hz has a standing wave component $(cos\beta xcos\theta)$ and a travelling wave component $(e^{-j\beta zsin\theta})$. So in general, the fields in the dielectric medium are a combination of travelling and standing waves.

All of these fields travel in the positive z direction (they travel along the plane). However, note that the standing waves are in a direction perpendicular to the interface.

Now we can make some observations from the expressions which we got for the electric and magnetic fields. First, if we plot the amplitude of the electric field as a function of x when x is equal to 0 then the amplitude is zero, therefore, the electric field is zero. The same happens for the x component of the magnetic field H_x, when x is equal to zero the magnetic field will be 0. So whenever the electric field is zero, the x component of the magnetic field H_x is also zero. It can therefore be deduced that the amplitude behaviour of H_x and electric field E_y is identical as a function of x. And the magnetic field component of H_z is a cos function. That means it is shifted by a quarter cycle in the x direction. So whenever H_x is zero, H_z will be maximum and whenever H_x is maximum, H_z will be zero. The plot of the amplitude of the electric and magnetic fields is shown in figure 16.4.

Referring to the figure 16.4, we have a boundary consisting of an electric field E_y, a magnetic field H_x for the x component and a magnetic field on the z component H_z. We can see that both fieldsE_y and H_x are exactly identical in behaviour whereas the z component of the magnetic field H_z is shifted by a quarter cycle. So wherever H_x is maximum, H_z is 0 and vice versa. Since H_z which is the tangential component (along the plane) is not zero, there will therefore be surface currents on the plane and the magnitude of the surface current will be equal to the tangential component of the magnetic field. When the wave is incident on the conducting boundary, the surface currents are

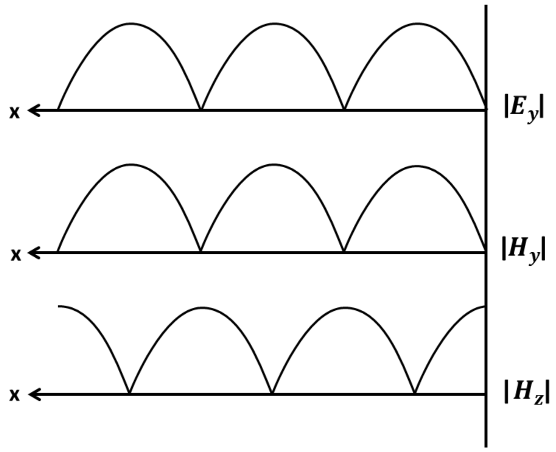


Figure 16.4: Amplitude of electric and magnetic fields

going to get induced on the surface which is due to the tangential component of the magnetic field. Also, the normal component of the magnetic field and the tangential component of the electric field will be zero.

Now referring back to the expression for the electric field in equation 16.11, the electric field is zero when x is equal to zero and it will also be zero whenever the quantity $(\beta x cos\theta)$ is a multiple of π. This means that E_y will be 0 when $(\beta x cos\theta)$ is a multiple of π. i.e E_y will be zero when:
$(\beta x cos\theta) = m\pi$ where m is an integer (0,1,2, 3 ...)
Hence,
$x = \frac{m\pi}{\beta cos\theta}$
but $\beta = \frac{2\pi}{\lambda}$
where λ is the wavelength in the dielectric medium.

$$x = \frac{m\pi}{\frac{2\pi}{\lambda}\cos\theta}$$

Thus,

$$x = \frac{m\lambda}{2cos\theta} \tag{16.17}$$

So at this distance x from the plane, the electric field will be 0. This distance is however dependent on the angle at which the wave is launched on the plane. There are multiple planes here in which the electric field is zero and since the electric field and the x component of the magnetic field H_x have the same behaviour in those planes, both quantities will be zero but H_z will be maximum in those planes. If we go backwards or forwards by a quarter cycle away, we will see that H_z will be 0 and the quantities E_y and H_x will then be maximum.

Now, we know that the wave is traveling along the z direction so there must be a power flow in that direction. This essentially gives us the pointing vector in the z-direction. Note that there is no net power flow in the x direction. Hence, given that we have a conducting boundary, there will be no power flow through the boundary. So whatever power is incident on the boundary is reflected back to the source medium. The wave which is incident in a particular direction has components of propagation in the direction normal and parallel to the interface. The wave which is incident normal to the interface is completely reflected, hence, a standing wave is created. So there is no net power flow in the x direction. This implies that, given a conducting boundary, then the boundary can be used to guide a wave of energy along it, hence, the conducting boundary has the capability of guiding electromagnetic energy. So when a wave is launched at an arbitrary angle, the net power flow is always along the surface of the boundary interface. This is the principle used in making a waveguide.

So in a wave-guiding structure, the conducting boundaries are used so that the electromagnetic energy is guided along these boundaries. Later in this book, we will see that, in those planes where the electric field was 0(given by location $x = \frac{m\lambda}{2cos\theta}$ away from the boundary), we can insert another conducting boundary without affecting the field distribution and thus create a structure that is bounded from both sides. Within this structure, electromagnetic energy can be trapped and hence we can have a net propagation of electromagnetic energy along the structure. This structure is called the **parallel plane waveguide** and will be considered in the next section.

Questions:

1. Explain the role of the infinite conductivity of a conductor in the reflection of electromagnetic waves from a conducting boundary.
2. How do the boundary conditions for perpendicular polarization differ between a dielectric-dielectric boundary and a dielectric-conductor boundary?
3. Discuss the significance of the reflection coefficient being -1 for a conducting boundary and its analogy to a short-circuited load in transmission line analysis.
4. Describe the field patterns generated when an electric field is completely reflected in a dielectric medium from a conducting boundary.
5. Illustrate the characteristics of the composite wave resulting from the incident and reflected waves at a conducting boundary.
6. Explain the behavior of electric and magnetic fields in the dielectric medium concerning standing and traveling wave components after reflection from a conducting boundary.
7. What role does the surface current play in the context of the tangential component of the magnetic field upon reflection from a conducting boundary?
8. Discuss the conditions where the electric field becomes zero and the implications for the magnetic field components in a dielectric medium after reflection.
9. Explain the absence of net power flow in the x-direction and the implications for power transmission in structures involving conducting boundaries.
10. Describe the principle behind creating a wave-guiding structure using conducting boundaries and its application in waveguide design.

Index

Anisotropic medium, 98
Resistance, 98
electric field, 97
magnetic field, 97

adaptive array antenna, 5
admittance, 37
amperes law, 100
attenuation, 117
attenuation constant, 17
azimuth resolution, 7

balanced parallel wire transmission line, 11
balanced structure, 84
balun, 71, 84
Boundary Condition, 107
brewster angle, 153

characteristic impedance, 20, 25–27, 47
characteristic length, 18
characteristics admittance, 48, 49
circle of constant reactance, 44
circle of constant resistance, 44
coaxial cable, 4, 11
complex co-efficient plane, 43
conduction current density, 98
constant vswr circle, 48
coordinate system, 87
coordinate systems, 87
corrugated surface, 140
coupled equations, 15
critical angle, 145, 146
current carrying element, 97
current density, 111

De Moivre's Theorem, 117
del operator, 91
dielectric boundary, 145, 151
dielectric interface, 151
dilectric interface, 141
displacement vector, 99
distributed parameters, 13
divergence, 87
double stub matching technique, 74

electric field, 98
electromagnetic interference, 4, 9
electromagnetic wave, 137
emf, 101
evenscent fields, 148

fading, 6
faraday's law, 102
Faraday's law of electromagnetic induction, 105
fibre optic cable, 6
finite conductivity, 112
flat ribbon cable, 82, 84
frequency response of the circuit, 67
frequency selectivity, 69

gamma plane, 43
gauss law, 99
gradient, 87

half wave dipole, 82

impedance discontinuity, 142
impedance matching, 71
Impedance Transformation, 51
impedance transformation, 47, 51
intrinsic, 115
intrinsic impedance, 141, 142
Isotropic medium, 111

limit circle, 45, 47
lossless transmission line, 26
low-loss transmission line, 27
lumped circuit analysis, 13

magenetic flux density, 97
magneto-motive force (mmf), 97
matched condition, 23, 25
matching condition, 37, 71, 84
matching power transfer condition, 37
matching transformer, 71
maximum normalized resistance, 52
maximum value of resistance, 52
maxwell's equation, 95
microstrip line, 12, 83
minimum and maximum voltage and current, 52
minimum voltage, 52
momentary, 129

net power flow, 148
normalized admittances, 49
normalized impedance, 42

Ohm's law, 98
optic fibre, 4

parallel plane wave guide, 160
parallel polarization, 142
permeability, 97, 98
Perpendicular Polarization, 158
perpendicular polarization, 142
phase constant, 17
phase gradient, 140
plane of undulation, 140
Point Relation, 107
polarizing angle, 154
Poynting, 124

poynting theroem, 124
poynting vector, 124
primary constants, 13
propagation, 126
propagation constant, 15, 26, 27

quality factor, 67
quarter wavelength transformer technique, 71

range resolution, 7
reactive fields, 147
reflection coefficient, 21, 84
right-handed coordinate system, 87

satellite communication, 6
scalar, 87
scalar product, 90
sectioned radiation patterns, 6
slotted transmission line, 59
smith chart, 42, 46
snell's law, 142
snells law, 142
standing wave, 30
standing wave pattern, 56
standing wave patterns, 57
stub, 73
stub matching technique, 73
Surface Charge Density, 107
surface charge density, 107
surface charges, 108
surface density, 129
swath width, 8
switched beam antenna, 5
synthetic aperture radar, 8

telegraphers equations, 15
The characteristic admittance, 48
the minimum and maximum voltage and current, 52
The Smith chart, 46
three stub matching, 75
total internal reflection, 145, 151
transit time, 12, 84
transmission coefficient, 142
traveling wave phenomenon, 15
Types of Load, 56
types of load, 56

unbalanced parallel wire transmission line, 11
unbalanced structure, 84
uniform plane wave, 126
unknown impedance, 59

vector product, 90
vector products, 91
vector quantity, 98
vectors, 87, 88
voltage maximum, 52
voltage standing wave ratio, 29, 32, 33, 36
volume charge density, 107
VSWR circles, 48

wave guides, 148
waveguide, 4, 5

yagi antenna, 82, 84

Answer (Ex. 12) — 2.68cm.

Answer (Ex. 29) — 2.45 + 105137.8j

Answer (Ex. 30) — 2000Hz

Answer (Ex. 45) — (a) $Z_0 = 600\angle-1.2^o\Omega$ (b) $\gamma = 1.048 \times 10^{-4}\angle 88.8^o$/m (c) $v = 2.998 \times 10^8$m/s and (d) $\lambda = 59.98$km.

Answer (Ex. 47) — Reflection coefficient= 0.2
VSWR= 1.5

Answer (Ex. 48) — 7.55ohms per meters

Answer (Ex. 50) — (a) $V_{max} = 100V$
(b) $I_{max} = 2A$
(c) $I_{min} = -2A$
(d) $VSWR = -1$

Answer (Ex. 51) — (a) 4.01
(b) $R_{max} = 28.84\Omega$ and $R_{min} = 17.44\Omega$

Answer (Ex. 52) — $v_p = 5 \times 10^7 m/s$

Answer (Ex. 53) — $R = 20.8\Omega$/mi, $L = 630\mu H$/mi, $G = 0$, and $C = 0.103\mu F$/mi.

Answer (Ex. 54) — (a) $\Gamma_L = -0.2+j0.4$ (b) $\Gamma(L) = -0.09+j0.20$ (c) $Z(L) = 502.7\angle 22.8^o\Omega$ (d) $V(L) = 75.0\angle 167.3^o$V (e) $|V|_{\max}$= 85.5V

Answer (Ex. 63) — (a) $P_L = 9.55W$ (b) $|V_L|$= 30.9V

Answer (Ex. 64) — The power delivered to each antenna is 76.68W

Answer (Ex. 77) — solution for exercise 77

Answer (Ex. 86) — Impedance, $Z_{in} = 300\Omega$

Answer (Ex. 95) — Solution to exercise 95

Answer (Ex. 96) — Solution to exercise 96

Answer (Ex. 97) — Solution to exercise 97

Answer (Ex. 98) — Solution to exercise 98

Answer (Ex. 99) — Solution to exercise 99

Answer (Ex. 100) — (a) The wavelength is 20πm, (b) the velocity of propagation is $20\pi f$ and (c) the loss is $-0.086l$dB/m.

Answer (Ex. 101) — Solution to exercise 101.

Answer (Ex. 102) — (a) $\Gamma_L = 0.45\angle 116.6°$, $\Gamma(L) = 0.22\angle 113.6°$, $Z(L) = 502.7\angle 22.8°\Omega$; (b) $V(L) = 75.0\angle 167.3°$V, $|V|_{\max}$= 85.5V at $x = 0.162\lambda$.

Answer (Ex. 103) — Solution to exercise 102.

Answer (Ex. 110) — (i). $3\hat{x}$+6$\hat{y}$-2$\hat{z}$ (ii).-3$\hat{x}$-6$\hat{y}$+2$\hat{z}$

Answer (Ex. 111) — $\nabla F = \frac{-5\sin(5x)}{y}\hat{x} - \frac{\cos(5x)}{y^2}\hat{y}$

Answer (Ex. 112) — $\nabla T = (2xyz^3 + y^2z^2)\hat{x} + (x^2z^3 + 2xyz^2)\hat{y} + (3x^2yz^2 + 2xy^2z)\hat{z}$

Answer (Ex. 113) — (i).$\nabla\dot{\bar{P}}$=$\sin(2y)\hat{z}$ (ii).$\nabla\times\bar{P}$= $2x\cos(2y)\hat{z}$ (iii).$\nabla\dot{\bar{Q}}$ = 0 (iv).$\nabla\times\bar{Q}$= -$e^y\sin(z)\hat{y}$ + ($\cosh(x)$ - $ye^y\cos(z))\hat{z}$

Answer (Ex. 125) —

$$u = 2.28 \times 10^{-19}$$

Answer (Ex. 127) — 6.0 x 10^3 v/m

Answer (Ex. 128) — V˙a - V˙b = -1

Answer (Ex. 129) — 1v/m.

Answer (Ex. 130) —

$$E = \frac{\rho_0}{3\varepsilon_0}\left(R - \frac{r^2}{R}\right)$$

Answer (Ex. 130) —

$$B = \frac{\mu_0 \cdot 5}{2\pi \cdot 0.05}\,\text{T}$$

Answer (Ex. 199) — 13.33m, 23.10m, 20m

3.Find the phase velocity in the x-direction, y-direction and in the z-direction

Ex. 199 — Discuss the factors affecting the propagation of a uniform plane wave in different mediums? Consider a 30MHz uniform wave propagating in free space and given by the **E** field vector; $E = 5(ax + \sqrt{3}ay)\cos\left[6\pi \times 10^7 t - 0.05\pi(3x - \sqrt{3}y + 2z)\right]$

1. Calculate for the phase constant
2. Find the wavelength in x-direction, y-direction and z-direction