NATIONAL OPEN UNIVERSITY OF NIGERIA

SCHOOL OF SCIENCE AND TECHNOLOGY

COURSE CODE: PHY 206

COURSE TITLE: NETWORK ANALYSIS AND DEVICES

Contents

INTRODUCTION

In PHY 122 which serves as prerequisite to this course, you would have become familiar with basic electrical network concepts and were encouraged to develop an enquiring attitude towards electrical devices which abound and which you interact with directly and indirectly every single day.

It is the objectives of this course to build upon the lessons learnt in the prerequisite course, and formally to introduce to you the underlying principles of electrical network analysis with the view to greater strengthening your understanding of the underlying concepts upon which developmental work and research on electrical networks and discrete component are based,

THE COURSE

PHY 206 Network Analysis and Devices

This course comprises a total of fourteen Units distributed across four modules as follows”

Module 1 is composed of 4 Units
Module 2 is composed of 4 Units
Module 3 is composed of 4 Units
Module 4 is composed of 2 Units

Module 1 will introduce idealised circuit elements to you in Unit 1 with emphasis laid on the fact that the millions of real worlds components you encounter daily in the real world are actually made up of no more than nine classes of idealized circuit element. This we will follow up with a brief introduction to Kirchhoff’s circuit laws in Unit 2 where through these laws, it will be emphasized to you that the laws of conservation of energy and of electrical charge are inviolable.

Unit 3 will explain to you that complex impedances are electrical vector quantities which have magnitude and phase angle. You will also be encouraged to distinguish between impedance and reactance and the conditions when impedance and the reactance of electrical networks become identical. We shall be occupied with the subject of Current and Voltage Source Transformations in the 4^{th} Unit of module 1 where the transformations will be made clear to you in a brief, simple straightforward and efficient manner.

Module 2 will expose you to key circuit theorems upon which most network analytical work is predicated. Those theorems you will cover in Unit 1 are Equivalent Impedance Transforms, Equivalent Circuit, Extra Element Theorem, Felici’s Law and Foster’s Reactance Theorem.

In Unit 2, you will meet Kirchhoff's Voltage Law and Kirchhoff's Current Law again; however, there shall be additional emphasis. Having progressed this far in your coursework you will begin to perceive that many of the network theorems are either derived from, or are very closely related to Kirchhoff's laws. Maximum Power Transfer Theorem will show you with proof that the conditions for maximum power transfer are not the same as the conditions for maximum efficiency. Finally in this Unit your coursework will guide you through Miller Theorem for voltages and the Dual miller theorem for current.

You will progress to Millman's Theorem, Norton's Theorem, Ohm's Law and Reciprocity in Unit 3 while with Superposition Theorem, Tellegen's Theorem, Thevenin's Theorem and Star – Delta Transformation we shall finally conclude our study of network theorems.

Your knowledge of electronic devices, particularly semiconductor devices, will be augmented as you work through module 3as you shall visit Vacuum Tubes in Unit 1, Semiconductor Materials in Unit 2, P-N Junction Diodesin and Transistors in Unit 4.By the time you have completed this module you will have been adequately equipped to speak with confidence on contemporary active electronics devices.

The final module in this coursework is intended as part of our objective, to establish a relationship between passive fitters and resonant circuits on the one hand, and Attenuators and Impedance Matching on the other.

COURSE AIMS AND OBJECTIVES

The aim of PHY 206 is to further intimate you with idealised circuit elements and their parametric characteristics, to acquaint you with the electrical network laws and theorems let establish their indispensability in network analysis, as illustration, to accustom you to simple but universal electronic circuits and to describe to you the evolution, processing, application and operation of vacuum tube and solid state devices.

Specifically, after working through this course diligently, upon completing it you should be able to:

- List the nine basic network elements
- Categorise the nine basic network elements into three
- Distinguish between linear and non linear network elements
- Identify the number of ports of different types of network elements

- Know the energy exchange mechanisms for lossless network elements
- Differentiate between dissipative and lossless network elements
- Appreciate that real world components (network Elements) are not perfect
- Understand the differences between dependent and independent sources
- State Kirchhoff’s first and second circuit laws
- Relate Kirchhoff’s current laws to the laws of conservation of charge.
- Relate Kirchhoff’s voltage laws to the laws of conservation of energy
- Understand The Meaning Of Complex Impedance
- Work with the Complex Impedance Plane
- Add Complex Impedance Vectors
- Subtract Complex Impedance Vectors
- Describe and Work With Phasors
- Derive Expressions for Device Specific Impedances
- Relate Resistance and Reactance to Impedance Phase
- Combine Impedances in Series and Parallel Configurations
- Solve Problems Involving Complex Impedance
- Transform a current Source into a Voltage source
- Transform a voltage source into a current source
- Know the relationship with Thevenin’s and Norton’s Theorems
- Use Ohm’s law to perform Transformations
- Explain Equivalent Impedance Transforms
- Identify Two and Three Element Networks
- Apply Network Transform Equations to Networks

- Describe Equivalent circuits

- Relate Thevenin’s and Norton’s Theorems to Equivalent Circuit Analysis
- Understand the similarities between Extra Element Theorem and Thevenin’s Theorem

- Use Simple RC circuit to demonstrate the Extra Element Theorem

- State Felici’s Law

- Use Felici’s Law to calculate net charge in a given period using initial and final flux

- State Foster’s Reactance Theorem

- Describe Admittance, Susceptance and Immittance

- Plot Reactance Curves for Capacitance, Inductance and Resonant Circuits

- Understand Foster’s First and Second Form Driving Point Impedance

- State Kirchhoff’s Voltage and Current Laws

- Understand How Kirchhoff’s Voltage Law based On the Law of Conservation Of Energy

- Understand How Kirchhoff’s Current Law based On the Law of Conservation Of Charge

- State the Maximum Power Transfer Theorem

- Understand why Maximum Power Transfer (MPT) conditions do not result in maximum efficiency.

- Prove that source and load impedances should be complex conjugates for reactive MPT

- State Miller Theorem
- Understand how the two versions of Miller Theorem are based on the two Kirchhoff's circuit laws

- State Millman’s Theorem

- Understand How Millman's Theorem Is Used To Compute Parallel Branch Voltage

- Establish That Millman's Theorem Is Derived from Ohm's And Kirchhoff's Laws
- Work with Supernode
- State Norton's Theorem
- Use To Calculate Equivalent Circuits
- Understand That Norton's Theorem Is an Extension of Thevenin's Theorem
- State Ohm's Law
- Apply Ohm's Law to Purely Resistive Networks
- Use Ohm's Law to Calculate Current in Reactive Circuits
- Explain Why Ohm's Law Does Not Apply To P-N Junction Devices
- Recognise the Other Versions of Ohm's Law
- Understand the Principles of Reciprocity
- Identify Practical Applications of Reciprocity in Spectral Radiators and Absorbers
- State the Superposition Theorem
- Apply superposition in converting circuits into Norton or Thevenin's equivalent
- State Tellegen's Theorem
- Establish the relationship between Tellegen's theorem and Kirchhoff's Laws
- State Thevenin's Theorem
- Derive the Thevenin's equivalent circuit for a Black Box Circuit
- Know when not to apply the Thevenin's Equivalent method
- Understand the procedure for Star-Delta Transformation
- Apply Star Delta transformations to passive linear networks

- Describe Vacuum Tubes
- Explain the historical progression of vacuum tube development
- Understand The Functional Sub Units Of The Vacuum Tube
- Distinguish Between Diode and Triode Vacuum Tubes
- Explain the Action of the Screen Grid
- Justify the Construction of Multi Section Vacuum Tube Design
- Describe the Functioning Of Different Special; Purpose Vacuum Tube
- Explain the Functional Similarities between the Vacuum Tube Triode and the Bipolar Transistor
- Compare and Contrast Vacuum Tube Technology with Semiconductor Technology
- Describe the Electrical Properties of Semiconductor Materials
- List Different Semiconductor Devices
- Distinguish Between Electrons and Holes as Carriers
- Understand the Process of Semiconductor Doping
- Explain Charge Depletion in Semiconductor Junction
- Describe the Flattening Of the Fermi-Dirac Distribution with Temperature
- Explain Band Diagram of P-N Junction Operation
- Discuss the Czochralski Process of Semiconductor Purification
- List Different Semiconductor Materials
- Explain How a P-N Junction Functions
- Describe the Depletion Region of A P-N Junction
- Appreciate Forward, Reverse and Zero Voltage Effect on Depletion

- Understand Reverse Breakdown on the P-N Junction
- Describe the Devices Whose Operation Depend On P-N Junction
- Know That Not All P-N Junctions Rectify Current
- Describe the Process for Manufacturing P-N Junctions
- Describe the similarities and the differences between transistor and vacuum tube triode
- Distinguish between Bipolar Junction Transistors and Field Effect Transistors
- Describe the operation of a transistor
- Calculate the current gain of a Bipolar Junction Transistor
- Draw and label the terminals of BJT and FET transistors
- List major advantages and limitations of FET over BJT
- Sketch simple diagrams of transistor as a switch and as an amplifier
- Describe an LC Circuit
- Explain Resonance
- Understand the Energy Exchange Mechanism In LC Circuits
- Calculate the Frequency Of Oscillation Of A Resonant Circuit
- Relate the “Q” Factor to Frequency Selectivity
- Know Why Unsustained Oscillations Die Down Asymptotically
- Work with both Parallel and Series LC Circuits
- Use LC Circuits As Frequency Band Pass or Rejector Filters
- Distinguish Between Active and Passive Filters
- Identify Filter Topology
- Calculate Filter Response To Frequency.

- Understand what attenuators are and what they do

- Be able to sketch the basic attenuator topologies

- Be able to distinguish between balanced and unbalanced attenuators

- Know how to qualify attenuators by their specifications

- Easily explain the meaning of impedance matching

- Know the significance of complex conjugate in complex load matching

- Be able to explain reflectionless impedance matching and maximum power transfer

- Easily identify impedance matching devices in the real world

- Recognise the role of the Smith Chart in Transmission Line matching networks

WORKING THROUGH THE COURSE

This course requires you to spend quality time to read. Whereas the content of this course is quite comprehensive, it is presented in clear language with lots of illustrations that you can easily relate to. The presentation style might appear rather qualitative and descriptive. This is deliberate and it is to ensure that your attention in the course content is sustained as a terser approach can easily "frighten" particularly when new concepts are being introduced.

You should take full advantage of the tutorial sessions because this is a veritable forum for you to "rub minds" with your peers – which provides you valuable feedback as you have the opportunity of comparing knowledge with your course mates.

COURSE MATERIAL

You will be provided course material prior to commencement of this course, which will comprise your Course Guide as well as your Study Units. You will receive a list of recommended textbooks which shall be an invaluable asset for your course material. These textbooks are however not compulsory.

STUDY UNITS

You will find listed below the study units which are contained in this course and you will observe that there are four modules. Each module comprises four Units each, except for module 4 which has two Units.

Module 1 CIRCUIT ANALYSIS

Unit 1 Circuit Elements
Unit 2 Kirchhoff's Circuit Laws
Unit 3 Complex Impedances
Unit 4 Current-Voltage Source Transformations

Module 2 CIRCUIT THEOREMS

Unit 1 Equivalent Impedance Transforms
Equivalent Circuit, Extra Element Theorem,
Felici's Law, Foster's Reactance Theorem

Unit 2 Kirchhoff's Voltage Law, Kirchhoff's
Current Law, Maximum Power Transfer
Theorem, Miller Theorem

Unit 3 Millman's Theorem, Norton's Theorem,
Ohm's Law, Reciprocity

Unit 4 Superposition Theorem, Tellegen's
Theorem, Thevenin's Theorem, Star – Delta
Transformation

Module 3 ELECTRONIC DEVICES

Unit 1 Vacuum Tubes
Unit 2 Semiconductor Materials
Unit 3 P-N Junction Diodes
Unit 4 Transistors

Module 4 ELECTRONIC CIRCUITS

Unit 1 Resonant Circuits and Passive Filters
Unit 2 Attenuators and Impedance Matching

TEXTBOOKS

There are more recent editions of some of the recommended textbooks and you are advised to consult the newer editions for your further reading.

Electronic Devices and Circuit Theory 7^{th} Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Semiconductor Device Fundamentals
By Robert F. Pierret Published by Prentice Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

Analog Filter Design
By M. E. Van Valkenburg Published by Holt, Rinehart and Winston

ASSESSMENT

Assessment of your performance is partly trough Tutor Marked Assessment which you can refer to as TMA, and partly through the End of Course Examinations.

TUTOR MARKED ASSIGNMENT

This is basically Continuous Assessment which accounts for 30% of your total score. During this course you will be given 4 Tutor Marked Assignments and you must answer three of them to qualify to sit for the end of year examinations. Tutor Marked Assignments are provided by your Course Facilitator and you must return the answered Tutor Marked Assignments back to your Course Facilitator within the stipulated period.

END OF COURSE EXAMINATION

You must sit for the End of Course Examination which accounts for 70% of your score upon completion of this course. You will be notified in advance of the date, time and the venue for the examinations which may, or may not coincide with National Open University of Nigeria semester examination.

SUMMARY

Each of the four modules of this course has been designed to stimulate your interest in network concepts and both electrical and electronic devices and components.

Module 2 in particular is specifically tailored to provide you a sound understanding of the laws and the theorems which will enable you to analyse circuits with relative ease and facilitate the translation of abstract theoretical concepts to real world devices subsystems and systems which you readily around you.

Module 3 provides you invaluable insight into the discovery, development and the functioning of conceptually simple components which perhaps have been most catalytic in transforming the world to that which we live in today.

Module 1 Unit 1 has provided you with a technical “mnemonic” with which you can easily classify the millions of components which constitute modern networks into just nine types of idealised elements. It also emphasises that all real world components – which are far from ideal components – can be modeled by the appropriate combination of just the nine ideal circuit elements. Unit 1 also shows you that an exotic component, the memristance ought to exist if all combinations of state variables (V, I, Q and Φ) are satisfied.

You will upon completion of this course confidently proffer realistic solutions and answers to everyday questions that arise such as:

- Is a fuse wire a resistance?

- Why are the most successful electrical products single component appliances such as the Electric Iron, the Incandescent Light Bulb, the Emersion Heater and incidentally, the electric Fuse?

- Why can an “L” topology filter only be a high pass or low pass but never a band pass filter?

- What happens when you connect two Zener diodes in series?

It will be an understatement to say that this course will change the way you see the world around you in more ways than one. You just make sure that you have enough referential and study material available and at your disposal. On this note;

We wish you the very best as you pursue knowledge.

Course Code	PHY 206
Course Title	Network Analysis and Devices

Writer

ESURUOSO Adedayo
Independent Researcher
And
Dr. AJIBOLA Saheed O.
School of Science and Technology
National Open University of Nigeria

Course Editing Team	Dr. Ajibola S. O	NOUN.
	Dr. Fayose	FUTA
	Mr. Adeola	FUTA

NATIONAL OPEN UNIVERSITY OF NIGERIA

National Open University of Nigeria
Headquarters
14/16 Ahmadu Bello Way
Victoria Island
Lagos

Abuja Annex
245 Samuel Adesujo Ademulegun Street
Central Business District
Opposite Arewa Suites
Abuja

e-mail: @nou.edu.ng
URL .nou.edu.ng

National Open University of Nigeria 2011

First Printed 2011
:
ISBN:

Printed and Bound in

Table of Content **Page**

UNIT 1 CIRCUIT ELEMENTS

CONTENTS

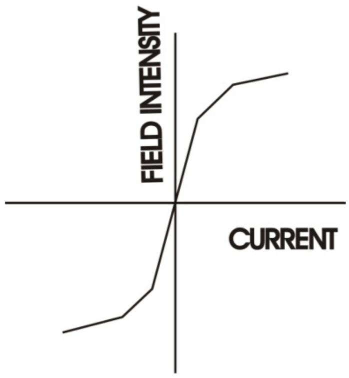
1.0 INTRODUCTION

I intend to teach you the basic circuit elements in this unit and I will explain to you that circuit elements are really abstract representations of ideal electrical components. We already know some common examples of these components; we know resistors, capacitors, and inductors and all of these are used in analysis of electrical networks.

If I show you a schematic diagram, which is really a graphical representation of a physical circuit, I can easily explain, and you also can easily understand that the schematic diagram is comprised of interconnections of components, and that these components affect the current and the voltage at different parts of the network. The components in a circuit diagram (schematic diagram) are represented by ideal components which do not exist in the real world. This is because uncertainties exist in the values and specifications of real world components; however, despite these uncertainties and non-linearity, they are a close approximation to the idealized elements and the lower the component tolerance, the closer it represents the idealized electrical element.

I will emphasize to you that you can only represented Electrical components by circuit elements over a limited range of physical parameters such as temperature, pressure, electric field potential, magnetic field intensity, high energy radiation and current densities. This is due to non-linearity which you will observe as each of these parameters are driven over an extended range.

Let me illustrate this; as you increase the current through the primary windings of a transformer, the magnetic flux in the core will increase proportionately until the core becomes magnetically saturated, and increase in current will produce very little increment in magnetic flux. A plot of current against magnetic flux will produce an “S” curve clearly indicating the non-linearly due to magnetic saturation – this particular curve which you should have encountered before now is called the Hysteresis curve.



Let me also emphasize to you that every physical electrical component is also often a combination of more than one ideal circuit element. You can therefore simulate a physical resistor with an ideal resistive element in series with an ideal inductive element in combination with distributed ideal capacitive elements. Similarly you can represent a physical capacitor by an ideal capacitive element in series with an ideal inductive element and with both series and parallel ideal resistive elements to take account of such imperfections as dielectric leakage amongst others.

It is essential that you take note of the usage of the word “physical” and the usage of the word “ideal” in the two preceding paragraphs.

By using electrical elements for electric circuit analysis, it will be possible for you to understand electrical networks which use real world components and easy for you to predict how a real electrical circuit behaves based on the effect each component has on the network.

2.0 OBJECTIVES

After reading through this unit, you will be able to

1 List the nine basic network elements
2 Categorise the nine basic network elements into three
3 Distinguish between linear and non linear network elements
4 Identify the number of ports of different types of network elements
5 Know the energy exchange mechanisms for lossless network elements
6 Differentiate between dissipative and lossless network elements
7 Appreciate that real world components (network Elements) are not perfect
8 Understand the differences between dependent and independent sources

3.0 MAIN CONTENT

3.1 Electrical Circuit Elements

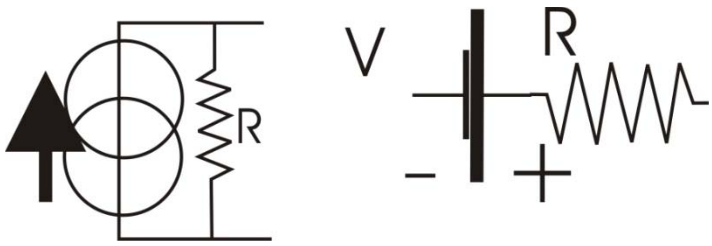
Most of us are not aware that on a daily basis, we interact directly and indirectly with thousands of electrical and electronic devices which often subtly affect our lives. Can you tell me how calculators, GSM phones, computers, DVD players, Plasma and LCD television sets, light bulbs and simple electrical fuses affect your life every day? Good. Now have you seen the circuitry inside some of these appliances and have you observed that all these electrical circuits are made up of interconnections of electrical and electronic components?

Often you will be amazed at the myriad of components which constitute these circuits and which appear almost limitless in variety. In actual fact all circuits and electrical networks comprise very few component types – some of which you can recognize as capacitors, resistors and transistors, and at this point I want to emphasize to you that there are only nine types of circuit elements in all, and that these nine types are all that is required to synthesize any given real component. These nine type of components comprise four active elements and five passive elements which are defined by their relation to the four state variables as follows:

- Voltage (V)
- Current (I)
- Charge (Q) and
- Magnetic flux (Φ)

Let us classify and learn more about the first group of circuit elements. You should always refer to this group as Active Circuit Elements; there are two types and they are both sources of energy:

- Voltage source
- Current source



Have you seen a voltage source before? no doubt you have. We see dry cell batteries, lead acid car batteries and electricity generators every day.

But have you seen a current source before, more likely than not, when you come across them you do not recognize them. They are most commonly encountered in electronic circuitry as transistor current sources and deliver current through high output impedance.

We have a second group, which also comprises of sources, however, these sources are controlled by either current or voltage and they are referred to as the Abstract Active Circuit Elements; they are:

- Voltage controlled voltage source
- Voltage controlled current source
- Current controlled voltage source
- Current controlled current source

The final group of circuit elements which we shall treat in this unit is known as Passive Circuit Elements and this group comprises:

- Resistance
- Capacitance
- Inductance

You should inspect the preceding groupings of circuit elements and always, you should remember that:

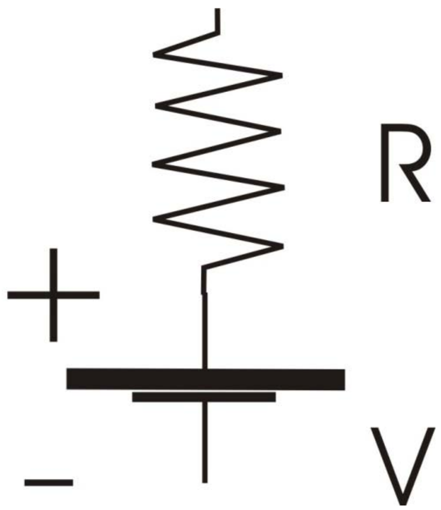
- there are a total of nine circuit elements
- each circuit element is classified as active or passive
- any single circuit element belongs to one of three categories; active source element, abstract active source element and passive circuit element

It is important that you note that in the real physical world, the combinations of these nine circuit elements are all that is required to describe any electrical circuit. We shall now treat the characteristics of each member of the three categories of circuit elements.

3.2 Source Elements

We have learnt that there are two non abstract active circuit elements and both of them are sources. They are both two terminal circuit elements which mean that they are single port (output port) elements. These elements can provide a constant output, or a time varying output of which the sinusoidal output is very frequently encountered.

Voltage source



A voltage source produces a potential difference between two terminals which is related to magnetic flux by the equation.

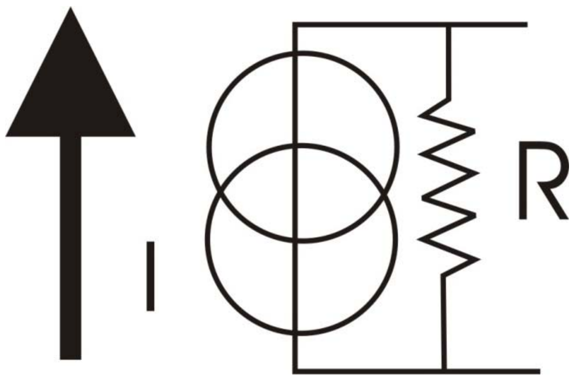
$$d\Phi = Vdt$$

The unit of measurement of voltage is the Volt designated (V). You will recognize in the equation above that Φ is the time integral of voltage. It may, or may not represent a physical quantity subject to the nature of the source and it is only meaningful if the voltage source is generated by electromagnetic induction.

Voltage may also be generated by a variety of means which include electrochemical, photovoltaic, thermoelectric or piezoelectric to mention a few. In all of these cases; the quantity Φ in the equation above does not apply.

An ideal voltage source is a two terminal circuit element whose terminal voltage is independent of the current drawn from it. The ideal voltage source has zero internal resistance. In reality, voltage sourced do possess internal resistance and the lower the internal resistance of a voltage source, the closer it resembles an ideal voltage source.

Current source



A current source produces current in a conductor which is related to electric charge by the equation

$$dQ = -\ Idt$$

You will also recognize in the equation above that the electric charge Q is the time integral of electric current and it represents the quantity of electric charge which is delivered by the current source. The unit of measurement of current is the Ampere designated (A).

In similarity with the ideal voltage source, an ideal current source is also a two terminal circuit element which delivers a constant current irrespective of the voltage across its terminals. The ideal current source would appear to possess infinitely high value of internal resistance which in the physical, real world is not true of current sources. A high but finite internal resistance exists across the terminals of real world current sources and the higher this internal resistance, the closer a current source emulates the ideal current source.

3.3 Abstract Active Element

We can easily see why there are only four abstract active elements when you combine voltage or current input with voltage or current output for a source which has an input and an output. They are all four terminal circuit elements. They possess input terminals (input port) and output terminals (output port) and are therefore classified as two port elements. Because the magnitude of the output state variables (V and I) depends on the input variables they are referred to as dependent sources.

We now discuss each of these dependent sources in relation to their input and output state variable of voltage and current, as well as their input and output impedance.

Voltage controlled voltage source

A voltage source which generates a voltage based on another voltage and which output voltage is related to its input voltage by a gain factor is known as a Voltage Controlled Voltage Source (VCVS). Further to this, the ideal Voltage Controlled Voltage Sources possess zero output impedance and infinite input impedance. (VCVS) is a dependent network element and is a four terminal (2 port) element.

Voltage controlled current source

A current source which generates a current based on another voltage and which output current is related to its input voltage by a gain factor is known as a Voltage Controlled Current Source (VCCS). Ideally, Voltage Controlled Current Sources possess infinite output impedance and infinite input impedance. (VCCS) is also a dependent four terminal (2 port) network element.

Current controlled voltage source

A voltage source which generates a voltage based on another current and which output voltage is related to its input current by a gain factor is known as a Current Controlled Voltage Source (CCVS). The ideal Current Controlled Voltage Source has zero output impedance and zero input impedance. By virtue of the fact that its output depends on its input, the (CCVS) is a dependent four terminal (2 port) network element.

Current controlled current source

The Current Controlled Current Source (CCCS) is the fourth four terminal (2 port) dependent network element being treated. It generates an output current which is a function of an input current; the two currents being related though a gain f actor. The ideal (CCCS) possesses zero input impedance and infinite output impedance.

3.4 Passive Elements

Apart from the active circuit elements, we regularly come across passive circuit elements - these are resistance, inductance and capacitance often referred to as the conventional circuit components. They are common in radio and television receivers amongst a host of other consumer electronics. We will discuss the properties of these passive circuit elements below:

Resistance

Resistance is a circuit element which produces a voltage across its terminals which is proportional to the current which flows through it. It is measured in Ohms symbolized by Ω . Do you remember this symbol? Yes you do; but do you remember the relationship

$$V = IR$$

And how it is derived?

The relationship which resistance defines between voltage and current is expressed as

$$dV = RdI$$

Where R, V and I are the resistance, the voltage across the terminals of and the current which flows through resistance element respectively.

When electrons flow through materials, they collide in-elastically with the particles in the material and lose energy. While the time rate of flow of electrons through the cross sectional area of the material is the current, electron energy loss per unit charge is drop in potential.

If E is the electric field intensity, J the electric current density and the electrical conductivity of the material

Then for a material of cross sectional area A and length L

$$J = \sigma E$$

$$I = JA$$

$$V = E/L$$

$$V/I = EL/JA = L/\sigma A$$

$$R = L/\sigma A$$

The quantity $L/\sigma A$ is constant for a given cross sectional geometry of a material. This constant is the resistance.

Thus $R = L/\sigma A$

The unit of resistance is the Ohm (Ω) after the name of the discoverer of the relation above.

Capacitance

When voltage changes across its terminals, capacitance produces a current which is proportional to the rate of voltage change. It is related to electric charge and terminal voltage by the equation

$$dQ = CdV$$

Do you recall $C = Q/V$

Capacitance is measured in farads

An electric field is set up which creates a force between two parallel metallic plates if one plate is positively charged and the other negatively charged.

If D is the electric field flux density, A the area of the plates, E the electric field intensity, q the electric charge and d the distance between the two metallic plates

Then for parallel conducting plates of area A

$$D = q/A$$

$$D = \epsilon E$$

$$E = q/\epsilon A = V/d$$

The quantity $\epsilon A/d$ is constant for a given cross sectional area of parallel conducting plates. This constant is the capacitance.

Thus $C = \epsilon A/d$

The unit of capacitance is the farad

Inductance

When you make a current flow through an inductor it produces a magnetic flux which is proportional to the rate of current change. It is related to magnetic flux and the inducing current by the equation

$$d\Phi = LdI$$

Inductance is measured in henries

A magnetic field is set up when a current flows through inductance which creates a magnetic force which is detectable with a magnetic compass. For a general geometry conductor

If Φ is magnetic flux surrounding the conductor, N the number of turns, f the total flux linkage,

Then for a conducting coil with N turns

$$\phi = \int_s B.ds$$

$$\Psi = N\phi = (N\int(\int dB)ds)i$$

The quantity (N∫(∫dB)ds) is the inductance.

$$\text{Thus} \quad L = (N\int(\int dB)ds)I$$

The unit of inductance is henry.

3.5 Non linear circuit elements

As I earlier told you, all real world circuit elements approximately exhibit parametric linearity over only a limited range and for you to obtain optimal description of the passive circuit elements; you will require adopting their constitutive relation instead of mere proportionality.

You can form six constitutive relations for any two of the state variables which are voltage (V), current (I), charge (Q), and magnetic flux (Φ). These relations when placed side by side with the five elements found in linear networks suggest the existence of a theoretical fourth passive circuit element (in addition to resistance, capacitance and inductance). This element called memristance is a non linear circuit element which reduces to resistance when considered as a linear circuit element.

For the purpose of circuit analysis, you will be introduced to two additional non linear circuit elements. These circuit elements which are not the ideal counterpart of any real component are:

- Nullator: That is a circuit element which restricts the value of voltage and current to zero

- Norrator: That is a circuit element which places no restriction on voltage or current.

4.0 CONCLUSION

We have learnt that there are there are nine types of network elements which are defined by their relationship to the state variables; Voltage (V), Current (I), Charge (Q) and Magnetic flux (Φ). E have also learnt that these nine network elements are are grouped into three broad types; two types are sources while one is passive elements. We have also learnt that of the passive elements, one is dissipative while two are not dissipative. Finally, we have learnt thet by inference and for completeness, a fourth passive network element was only recently discovered – which is memristance.

5.0 SUMMARY

- Only nine network elements exist
- These nine elements fall into two groups of source elements and passive elements
- Source elements are further grouped into two which are active elements and abstractive active elements
- Passive network elements fall into the two groups of dissipative and non dissipative elements
- Real world components are far from ideal network elements but can be simulated by combinations of appropriate aggregates of idealized elements.
- Real world components are generally non linear and only exhibit parametric linearity over a limited range of physical variables.
- Circuit analysis is facilitated by the introduction of two hypothetical elements called the Nullator and the Norrator.

6.0 TUTOR MARKED ASSIGNMENTS

1. List the nine basic network elements
2. Network elements are broadly classified into two groups, Discuss.
3. What are, and how many are abstractive source network elements?
4. Which amongst the following idealized elements are dissipative network elements?

 6 volt battery

 10 kilo Ohm resistor

10 Micro farad capacitor

15 micro Henry Inductor

1 Amp Constant Current Source

5. Explain how a memristance works.

6. With what ideal network elements can you simulate a real world inductor?

7. Describe the hysteresis process and highlight why it is a non linear process.

8. Name two lossless network elements.

9. State the four state variables which govern the operation of electrical network elements.

10. What is the expected value of internal impedance of an idealized current source?

11. Describe the relationship between the input and the output of a voltage controlled current source.

12. What is the unit of capacitance? And what is the unit of inductance?

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

UNIT 2 KIRCHOFF'S CIRCUIT LAWS

CONTENTS

1.0 INTRODUCTION

In this unit you will learn about Kirchhoff's circuit laws also known as Kirchhoff's current law (KCL) and Kirchhoff's voltage law (KVL).While it is useful to be able to reduce series and parallel resistors in a circuit, circuits are however not always composed exclusively of serial and parallel combinations of resistors. You will recognize that such circuits include star and delta configurations. In such cases you will find it expedient to utilize the powerful set of relations called Kirchhoff's laws which will enable you to analyze arbitrary circuits.

Do you know that the formulas that are used for star to delta and delta to star conversions are derived from Kirchhoff's laws? where the resistances in the three terminal networks are equivalent to the other because they have equivalent resistances across any one pair of terminals.

Kirchhoff's Laws provides the practical means for you to solve for unknowns in a circuit and it makes it possible for you to take a circuit with two or more loops and several power sources and determine loop equations, solve loop currents, and solve individual element currents as Kirchhoff's two laws reveal a unique relationship between current, voltage, and resistance in electrical circuits that is vital to performing and understanding electrical circuit analysis.

You are reminded at this point that Kirchhoff's laws only re-affirm the laws governing energy and charge conservation since all of the power provided from the source is consumed by the load. Energy and charge are conserved since voltage and current can be related to energy and charge.

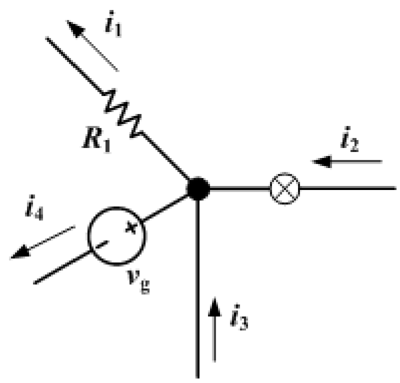
2.0 OBJECTIVES

After reading through this unit, you will be able to

1. State Kirchhoff's first and second circuit laws
2. Relate Kirchhoff's current laws to the laws of conservation of charge.
3. Relate Kirchhoff's voltage laws to the laws of conservation of energy

3.0 MAIN CONTENT

3.1 Kirchhoff's Current Law



Kirchhoff's current law states that "The current arriving at any junction point in a circuit is equal to the current leaving that junction" This law is also known as Kirchhoff's point rule,

Kirchhoff's junction rule (or nodal rule), and Kirchhoff's first rule and you are advised to take note that they all refer to Kirchhoff's current law

At any given node of an electrical circuit, the principle of conservation of charge stipulates that the sum of current which flows into the node should equal the sum of current flowing out of it as there can be no accumulation of current. In other words, the algebraic sum of currents in a network of conductors meeting at a point is zero. (Assuming that current entering the junction is taken as positive and current leaving the junction is taken as negative).

Mathematically, this can be stated as

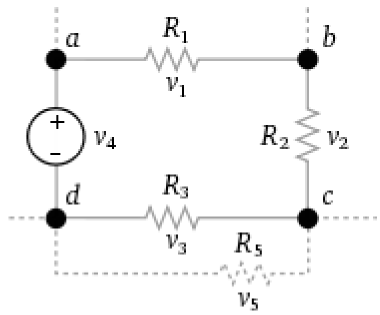
$$\sum_{k=1}^{n} I_k = 0$$

where n is the total number of branches with currents flowing towards or away from the node.

You will see that this relationship is also valid for complex currents and is signed which takes into account currents flowing towards the node (positive) and current flowing away from the node (negative).

It is easy to see how charge is conserved by Kirchhoff's current law when you recall that charge (measured in coulombs) is the product of the current (in amperes) and the time (which is measured in seconds).

3.2 Kirchhoff's Voltage Law



Kirchhoff's voltage law states that that "The sum of the voltage drops around a closed loop is equal to the sum of the voltage sources of that loop". This law is also known as Kirchhoff's second law, Kirchhoff's loop (or mesh) rule, and Kirchhoff's second rule and you are again advised to note that they all refer to Kirchhoff's voltage law.

Mathematically Kirchhoff's voltage law can be stated as

$$\sum_{k=1}^{n} V_k = 0$$

where n is the total number of voltages measured.

Once again, this relationship is also valid for complex currents and is signed which takes into account voltage polarities along the loop trajectory.

Kirchhoff's voltage law is based on the conservation of energy given/taken or energy taken by potential field excluding energy taken by dissipation. A charge which has completed a closed loop does not gain or lose energy for a given voltage potential. It simply goes back to initial potential level.

The law is valid even with energy dissipating resistance in the circuit as electrical charges do not return to their starting potential due to energy dissipation but just terminate at the negative terminal instead of positive terminal. This means all the energy given by the potential difference is been fully dissipated by resistance in the form of heat.

Kirchhoff's voltage law is a law relating to potential generated by voltage sources regardless of the electronic components which are present in the circuit whereby the gain or loss in "energy given by the potential field" must be zero when a charge completes a closed loop.

4.0 CONCLUSION

We learnt in Unit 2 that Kirchhoff's circuit laws are as a consequence of the laws of conservation of energy and electrical charge, and that many of the other specialized theorems and laws of electrical network analysis are derived from them.

5.0 SUMMARY

- Kirchhoff's Current Law can be stated mathematically as

$$\sum_{k=1}^{n} I_k = 0$$

- Kirchhoff's Voltage Law can be stated mathematically as

$$\sum_{k=1}^{n} V_k = 0$$

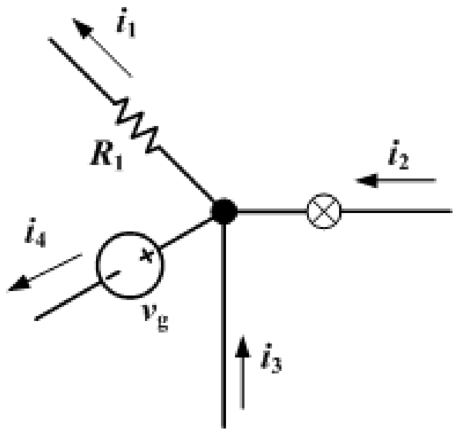
- Star and delta electrical networks configurations are best analysed by using Kirchhoff's laws

6.0 TUTOR MARKED ASSIGNMENTS

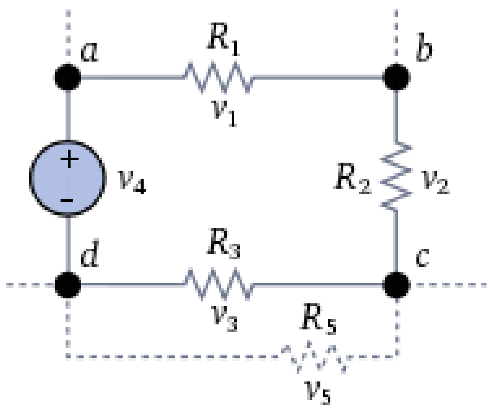
1. State Kirchhoff's Current law

2. Give two other common names by which Kirchhoff's Voltage law is known?

3. In the diagram below, what is the value of the current i_4 if i_1, i_2 and i_3 are 7, 15 and -18 milliamps respectively, and taking current directed towards the node as being positive in value?



4. What is the voltage across R_2 if the voltages across R_1 and R_3 are 3 volts and 5 volts respectively? The source voltage is 12 volts.



7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

UNIT 3 COMPLEX IMPEDANCES

1.0 INTRODUCTION

Electrical impedance which you may simply refer to as impedance, describes a measure of opposition to alternating current (AC) and electrical impedance extends the concept of resistance to AC circuits, describing not only the relative amplitude units of the voltage and current, but also the relative
Phases. When a circuit is driven with direct current (DC) there is no distinction between impedance and resistance; you may therefore imagine it as impedance with zero phase angle.

The symbol Z is often used for impedance and it may be represented by writing its magnitude and phase in the form $|Z|\angle\theta$. You will however discover that, complex number representation of electrical impedance is more powerful for circuit analysis purposes. The term impedance was first used by Oliver Heaviside in July 1886 while Arthur Kennelly was the first to represent impedance with complex numbers in 1893.

Impedance is defined as the frequency domain ratio of the voltage to the current. In other words, it is the voltage–current ratio for a single complex exponential at a particular frequency ω. In general, impedance will be a complex number, with the same units as resistance, for which the SI unit is the ohm (Ω). For a sinusoidal current or voltage input, the polar form of the complex impedance relates the amplitude and phase of the voltage and current. In particular, The magnitude of the complex impedance is the ratio of the voltage amplitude to the current amplitude

- The phase of the complex impedance is the phase shift by which the current is ahead of the voltage.

- The reciprocal of impedance is admittance. Admittance is the current-to-voltage ratio, and it conventionally carries units of Siemens which was formerly called mhos.

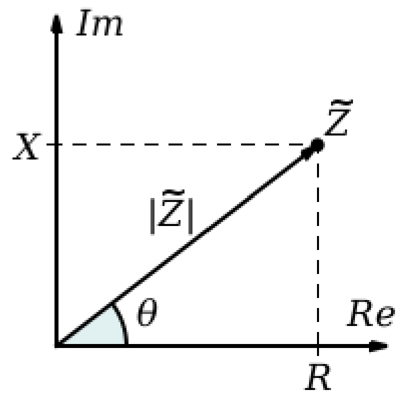
2.0 OBJECTIVES

After reading through this unit, you should be able to

1. Understand The Meaning Of Complex Impedance
2. Work with the Complex Impedance Plane
3. Add Complex Impedance Vectors
4. Subtract Complex Impedance Vectors
5. Describe and Work With Phasors
6. Derive Expressions for Device Specific Impedances
7. Relate Resistance and Reactance to Impedance Phase
8. Combine Impedances in Series and Parallel Configurations
9. Solve Problems Involving Complex Impedance

3.0 MAIN CONTENT

3.1 Equivalent Impedance Transform



The Complex Impedance Plane

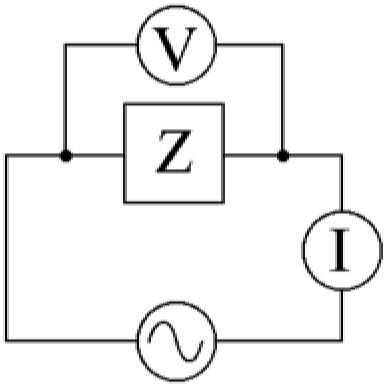
Impedance is represented as a complex quantity Z and the term complex impedance may be used interchangeably; the polar form conveniently captures both magnitude and phase characteristics,

$$Z = |Z|e^{j\theta}$$

where the magnitude $|Z|$ represents the ratio of the voltage difference amplitude to the current amplitude, while the argument θ gives the phase difference between voltage and current and j is the imaginary unit. In Cartesian form, where the real part of impedance is the resistance R and the imaginary part is the reactance.

where it is required for you to add or subtract impedances the cartesian form is more convenient, but when quantities are multiplied or divided the calculation becomes simpler if the polar form is used. A circuit calculation, such as finding the total impedance of two impedances in parallel, may require conversion between forms several times during the calculation. If you wish to convert between the forms you must follow the normal conversion rules of complex numbers.

Ohm's law



An AC supply applying a voltage *V* across a load Z, driving a current *I*.

The meaning of electrical impedance can be understood by substituting it into Ohm's law.

$$V = I\,Z = I\,|Z|\,e^{J\theta}$$

The magnitude of the impedance $|Z|$ acts just like resistance, giving the drop in voltage amplitude across an impedance Z for a given current I.

The phase factor tells us that the current lags the voltage by a phase of θ (i.e. in the time domain, the current signal is shifted $\frac{\theta}{2\pi}$ to the right with respect to the voltage signal).

Just as impedance extends Ohm's law to cover AC circuits, other results from DC circuit analysis such as voltage division, current division, Thevenin's theorem, and Norton's theorem, can also be extended to AC circuits by replacing resistance with impedance.

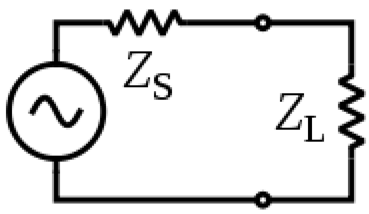
3.2 Complex Voltage and Current

We can draw generalized impedances in a circuit with the same symbol as a resistor or with a labelled box.

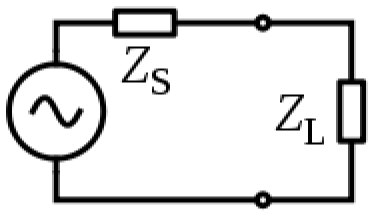
In order to simplify your calculations, it will be easier for you to represent sinusoidal voltage and current waves as complex-valued functions of time denoted as V and I.

Let us see what exactly we mean by these below.

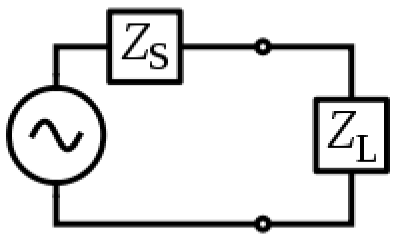
Resistor Symbol



Box Symbol



Labelled box symbol



$$V = |V|e^{j(\omega t+\phi_V)}$$

$$I = |I|e^{j(\omega t+\phi_I)}$$

When we define impedance as the ratio of these quantities

$$Z = \frac{V}{I}$$

And we substitute these into Ohm's law we have

$$\begin{aligned} |V|e^{j(\omega t+\phi_V)} &= |I|e^{j(\omega t+\phi_I)}|Z|e^{j\theta} \\ &= |I||Z|e^{j(\omega t+\phi_I+\theta)} \end{aligned}$$

And we note that this must hold for all values of t, we may equate the magnitudes and phases to obtain

$$|V| = |I||Z|$$

$$\phi_V = \phi_I + \theta$$

The magnitude equation is the familiar Ohm's law applied to the voltage and current amplitudes, while the second equation defines the phase relationship.

Validity of complex representation

We can justify this representation using complex exponentials by noting that

$$\cos(\omega t + \phi) = \frac{1}{2}\left[e^{j(\omega t+\phi)} + e^{-j(\omega t+\phi)}\right]$$

a real-valued sinusoidal function which represents our voltage or current waveform can be broken into two complex-valued functions. By the principle of superposition, we may analyze the behaviour of the sinusoid on the left-hand side by analysing the behaviour of the two complex terms on the right-hand side. Given the symmetry, we only need to perform the analysis for one right-hand term; the results will be identical for the other.

At the end of any calculation, we may return to real-valued sinusoids by further noting that

$$\cos(\omega t + \phi) = \Re\left\{e^{j(\omega t+\phi)}\right\}$$

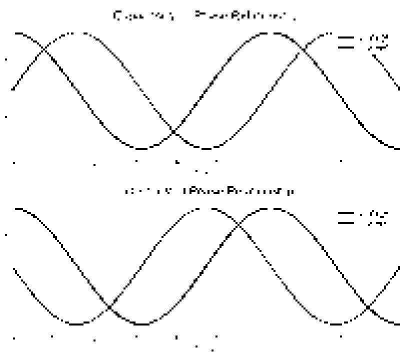
In other words, we simply take the real part of the result.

Phasors

A phasor is a constant complex number which you can usually expressed in exponential form, representing the complex amplitude (magnitude and phase) of a sinusoidal function of time. Phasors are used in the field of electrical engineering to simplify computations involving sinusoids, where they can often reduce a differential equation problem to an algebraic one.

The impedance of a circuit element can be defined as the ratio of the phasor voltage across the element to the phasor current through the element, as determined by the relative amplitudes and phases of the voltage and current. This is identical to the definition from Ohm's law given above, recognizing that the factors of $e^{j\omega t}$ cancel.

Let us look at these two examples in the diagram below



The phase angles in the equations for the impedance of inductors and capacitors indicate that the voltage across a capacitor lags the current through it by a phase of 2π while the voltage across an inductor leads the current through it by 2π

The identical voltage and current amplitudes tell us that the magnitude of the impedance is equal to one.

The impedance of an ideal resistor is purely real and is referred to as resistive impedance:

$$Z_R = R.$$

Ideal inductors and capacitors have a purely imaginary reactive impedance:

$$Z_L = j\omega L,$$

$$Z_C = \frac{1}{j\omega C}\ .$$

You should note the following identities for the imaginary unit and its reciprocal:

$$j = \cos\left(\frac{\pi}{2}\right) + j\sin\left(\frac{\pi}{2}\right) = e^{j\frac{\pi}{2}},$$

$$\frac{1}{j} = -j = \cos\left(-\frac{\pi}{2}\right) + j\sin\left(-\frac{\pi}{2}\right) = e^{j(-\frac{\pi}{2})}.$$

Thus we can rewrite the inductor and capacitor impedance equations in polar form:

$$Z_L = \omega L e^{j\frac{\pi}{2}},$$

$$Z_C = \frac{1}{\omega C} e^{j(-\frac{\pi}{2})}.$$

The magnitude tells us the change in voltage amplitude for given current amplitude through the impedance, while the exponential factors give the phase relationship.

3.3 Device Specific Impedances

What follows below is a derivation of impedance for each of the three basic circuit elements, the resistor, the capacitor, and the inductor. Although the idea can be extended to define the relationship between the voltage and current of any arbitrary signal, these derivations will assume sinusoidal signals, since any arbitrary signal can be approximated as a sum of sinusoids through Fourier analysis.

Resistor

For a resistor, we have the relation:

$$v_{\mathrm{R}}(t) = i_{\mathrm{R}}(t)R.$$

This is simply a statement of Ohm's Law.

Considering the voltage signal to be

$$v_{\mathrm{R}}(t) = V_p \sin(\omega t)\,,$$

it follows that

$$\frac{v_{\mathrm{R}}(t)}{i_{\mathrm{R}}(t)} = \frac{V_p \sin(\omega t)}{I_p \sin(\omega t)} = R.$$

This tells us that the ratio of AC voltage amplitude to AC current amplitude across a resistor is R, and that the AC voltage leads the AC current across a resistor by 0 degrees.

This result is commonly expressed as

$$Z_{\mathrm{resistor}} = R\,.$$

Capacitor

For a capacitor, we have the relation:

$$i_{\mathrm{C}}(t) = C\frac{\mathrm{d}v_{\mathrm{C}}(t)}{\mathrm{d}t}.$$

Considering the voltage signal to be

$$v_{\mathrm{C}}(t) = V_p \sin(\omega t)$$

it follows that

$$\frac{\mathrm{d}v_{\mathrm{C}}(t)}{\mathrm{d}t} = \omega V_p \cos(\omega t)\,.$$

And thus

$$\frac{v_{\mathrm{C}}(t)}{i_{\mathrm{C}}(t)} = \frac{V_p \sin(\omega t)}{\omega V_p C \cos(\omega t)} = \frac{\sin(\omega t)}{\omega C \sin\left(\omega t + \frac{\pi}{2}\right)}.$$

This tells us that the ratio of AC voltage amplitude to AC current amplitude across a capacitor is $\frac{1}{\omega C}$, and that the AC voltage leads the AC current across a capacitor by -90 degrees (or the AC current leads the AC voltage across a capacitor by 90 degrees).

This result is commonly expressed in polar form, as

$$Z_{\mathrm{capacitor}} = \frac{1}{\omega C} e^{-j\frac{\pi}{2}}$$

or, by applying Euler's formula, as

$$Z_{\mathrm{capacitor}} = -j\frac{1}{\omega C} = \frac{1}{j\omega C}\,.$$

Inductor

For the inductor, we have the relation:

$$v_{\mathrm{L}}(t) = L\frac{\mathrm{d}i_{\mathrm{L}}(t)}{\mathrm{d}t}.$$

This time, considering the current signal to be

$$i_{\mathrm{L}}(t) = I_p \sin(\omega t)\,,$$

it follows that

$$\frac{\mathrm{d}i_{\mathrm{L}}(t)}{\mathrm{d}t} = \omega I_p \cos(\omega t)\,.$$

And thus

$$\frac{v_{\mathrm{L}}(t)}{i_{\mathrm{L}}(t)} = \frac{\omega I_p L \cos(\omega t)}{I_p \sin(\omega t)} = \frac{\omega L \sin\left(\omega t + \frac{\pi}{2}\right)}{\sin(\omega t)}.$$

This tells us that the ratio of AC voltage amplitude to AC current amplitude across an inductor is ωL, and that the AC voltage leads the AC current across an inductor by 90 degrees.

This result is commonly expressed in polar form, as

$$Z_{\text{inductor}} = \omega L e^{j\frac{\pi}{2}}.$$

Or, more simply, using Euler's formula, as

$$Z_{\text{inductor}} = j\omega L.$$

S-plane impedance

Impedance defined in terms of jω can strictly only be applied to circuits which are energised with a steady-state AC signal. The concept of impedance can be extended to a circuit energised with any arbitrary signal by using complex frequency instead of jω. Complex frequency is given the symbol s and is, in general, a complex number. Signals are expressed in terms of complex frequency by taking the Laplace transform of the time domain expression of the signal. The impedance of the basic circuit elements in this more general notation is as follows:

Element	**Impedance expression**
Resistor	*R*
Inductor	*sL*
Capacitor	*1/sC*

For a DC circuit you can simplify this to s = 0. For a steady-state sinusoidal AC signal s = jω.

Resistance and reactance

Resistance and reactance together determine the magnitude and phase of the impedance through the following relations:

$$|Z| = \sqrt{ZZ^*} = \sqrt{R^2 + X^2}$$

$$\theta = \arctan\left(\frac{X}{R}\right)$$

In many applications the relative phase of the voltage and current is not critical so only the magnitude of the impedance is significant.

Resistance

Resistance $R$is the real part of impedance; a device with purely resistive impedance exhibits no phase shift between the voltage and current.

$$R = |Z| \cos\theta$$

Reactance

Reactance $X$is the imaginary part of the impedance; a component with a finite reactance induces a phase shift θbetween the voltage across it and the current through it.

$$X = |Z| \sin\theta$$

A purely reactive component is distinguished by the fact that the sinusoidal voltage across the component is in quadrature with the sinusoidal current through the component. This implies that the component alternately absorbs energy from the circuit and then returns energy to the circuit. A pure reactance will not dissipate any power.

Capacitive reactance

A capacitor has purely reactive impedance which is inversely proportional to the signal frequency. A capacitor consists of two conductors separated by an insulator, also known as a dielectric.

At low frequencies a capacitor is open circuit, as no charge flows in the dielectric. A DC voltage applied across a capacitor causes charge to accumulate on one side; the electric field due to the accumulated charge is the source of the opposition to the current. When the potential associated with the charge exactly balances the applied voltage, the current goes to zero.

Driven by an AC supply, a capacitor will only accumulate a limited amount of charge before the potential difference changes sign and the charge dissipates. The higher the frequency, the less charge will accumulate and the smaller the opposition to the current.

Inductive reactance

Inductive reactance X_L is proportional to the signal frequency f and the inductance L.

$$X_L = \omega L = 2\pi f L$$

An inductor consists of a coiled conductor. Faraday's law of electromagnetic induction gives the back emf $\mathcal{E}$(voltage opposing current) due to a rate-of-change of magnetic flux density B through a current loop.

$$\mathcal{E} = -\frac{d\Phi_B}{dt}$$

For an inductor consisting of a coil with N loops this gives.

$$\mathcal{E} = -N\frac{d\Phi_B}{dt}$$

The back-emf is the source of the opposition to current flow. A constant direct current has a zero rate-of-change, and sees an inductor as a short-circuit usually made from a material with a low resistivity. An alternating current has a time-averaged rate-of-change that is proportional to frequency; this causes the increase in inductive reactance with frequency.

3.4 Combination of Impedances

The total impedance of many simple networks of components can be calculated using the rules for combining impedances in series and parallel.

The rules are identical to those used for combining resistances, except that the numbers in general will be complex numbers. In the general case however, equivalent impedance transforms in addition to series and parallel will be required.

Series combination

When we connect components in series, the current through each circuit element is the same; the total impedance is simply the sum of the component impedances. It is easier for you to visualize this by looking at this sketch.

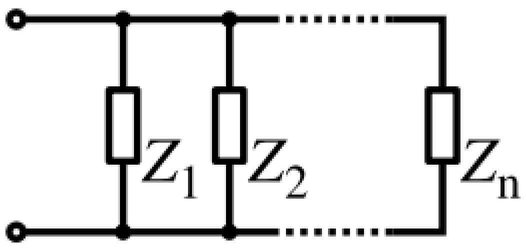
$$Z_{\text{eq}} = Z_1 + Z_2 + \cdots + Z_n$$

Or explicitly in real and imaginary terms:

$$Z_{\text{eq}} = R + jX = (R_1 + R_2 + \cdots + R_n) + j(X_1 + X_2 + \cdots + X_n)$$

Parallel combination

For components connected in parallel, the voltage across each circuit element is the same; the ratio of currents through any two elements is the inverse ratio of their impedances. You can similarly visualize this by also consulting this sketch below



Hence the inverse total impedance is the sum of the inverses of the component impedances:

$$\frac{1}{Z_{\text{eq}}} = \frac{1}{Z_1} + \frac{1}{Z_2} + \cdots + \frac{1}{Z_n}$$

or, when n = 2:

$$Z_{\text{eq}} = Z_1 \| Z_2 \stackrel{\text{def}}{=} \frac{Z_1 Z_2}{Z_1 + Z_2}$$

The equivalent impedance Z_{eq}can be calculated in terms of the equivalent resistance $R_{\text{eq}}$and reactance X_{eq}.

$$Z_{\text{eq}} = R_{\text{eq}} + jX_{\text{eq}}$$

$$R_{\text{eq}} = \frac{(X_1R_2 + X_2R_1)(X_1 + X_2) + (R_1R_2 - X_1X_2)(R_1 + R_2)}{(R_1 + R_2)^2 + (X_1 + X_2)^2}$$

$$X_{\text{eq}} = \frac{(X_1R_2 + X_2R_1)(R_1 + R_2) - (R_1R_2 - X_1X_2)(X_1 + X_2)}{(R_1 + R_2)^2 + (X_1 + X_2)^2}$$

4.0 CONCLUSION

We will recall in this Unit that Electrical impedance describes the measure of opposition to alternating current extends the concept of resistance to AC circuits. We will also remember that impedance is a complex quantity which has a magnitude and an angle between the voltage and the current when visualized in polar coordinates. We learnt about Phasors which represent magnitude and phase of a sinusoidal function of time and how Phasors can often reduce a differential equation problem to an algebraic one. Also visited are the expressions of impedance for resistor, capacitor and inductor and distinguished impedance from reactance. We learnt how to combine impedances in series and in parallel to arrive at equivalent impedance and finally saw how a phase lead or a phase lag between voltage and current results when there is impedance in the network.

5.0 SUMMARY

- Impedance is a complex quantity where a phase angle exists between the driving voltage and the current.

- Reactance is the impedance of a purely lossless network element and has a phase angle of 90 degrees between driving voltage and current.

- Phasors are constant complex numbers which can be expressed in exponential form and which represent the complex amplitude of a sinusoidal function of time. They simplify computations involving sinusoids and can reduce differential equation problems to algebraic form.

- Device specific impedances are as detailed beow for risitance, inductor and capacitor

Resistor	*R*
Inductor	*sL*
Capacitor	*1/sC*

- The expression for series combination of impedances is

$$Z_{\mathrm{eq}} = Z_1 + Z_2 + \cdots + Z_n$$

- While the expression for parallel combination of impedances is

$$\frac{1}{Z_{\mathrm{eq}}} = \frac{1}{Z_1} + \frac{1}{Z_2} + \cdots + \frac{1}{Z_n}$$

6.0 TUTOR MARKED ASSIGNMENTS

1. What is the value of inductive impedance of a fluorescent choke when connected directly across the 220 volts 50 hertz mains voltage when a current of 0.3 Amp is observed to flow through its windings?

2. Distinguish between impedance and reactance and state the condition when both are equal.

3. Which network component does not introduce a phase shift between its driving voltage and current?.

4. State the expressions for reactive and capacitive reactances respectively.

5. Describe the S plane.

6. What is a phasor and how does it function?

7. Determine the Impedance and the phase angle of a serial resistor and capacitor when driven by a 1 Kilohertz 20 volt signal if the resistor is 10 Kilo ohm and the capacitor 2 micro Farad.

8. At what frequency will the reactances of a 1.5 microfarad capacitor and a 50 milli henry inductor become equal? And what is the name given to this frequency?

9. What is the resultant impedance of 30 micro Farad capacitor in parallel with 2200 Ohm resistor?

10. Which of the following is Admittance?

    a. voltage -to- current ratio
    b. current-to-voltage ratio
    c. current-to- current ratio
    d. voltage -to-voltage ratio

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

UNIT 4 CURRENT-VOLTAGE SOURCE TRANSFORMATIONS

1.0 INTRODUCTION

Finding a solution to a circuit can be somewhat difficult without employing methods that make the circuit appear simpler. Circuit solutions are often simplified, especially with mixed sources, by transforming a voltage into a current source, and vice versa. This process is known as a source transformation, and is an application of Thevenin's theorem and Norton's theorem which you shall treat in greater detail in module 2. This Unit therefore shall only serve as a brief introduction to the subject.

2.0 OBJECTIVES

After reading through this unit, you should be able to

1 Transform a current Source into a Voltage source
2 Transform a voltage source into a current source
3 Know the relationship with Thevenin’s and Norton’s Theorems
4 Use Ohm’s law to perform Transformations

3.0 MAIN CONTENT

3.1 Source Transformations

Performing a source transformation is the process of using Ohms Law to take an existing voltage source in series with a resistance, and replace it with a current source in parallel with the same resistance. Remember that Ohms law states that a voltage in a material is equal to the material's resistance times the amount of current through it. Since source transformations are bilateral, one can be derived from the other.

Source transformations are not limited to resistive circuits however. They can be performed on a circuit involving capacitors and inductors, as long as the circuit is first put into the frequency domain. In general, the concept of source transformation is an application of Thevenin's theorem to a current source, or Norton's theorem to a voltage source.

Specifically, source transformations are used to exploit the equivalence of a real current source and a real voltage source, such as a battery. Application of Thevenin's theorem and Norton's theorem gives the quantities associated with the equivalence. Specifically, suppose we have a real current source I, which is an ideal current source in parallel with an impedance.

If the ideal current source is rated at I amperes, and the parallel resistor has an impedance Z, then applying a source transformation gives an equivalent real voltage source, which is ideal, and in series with the impedance. This new voltage source V, has a value equal to the ideal current source's value times the resistance contained in the real current source. The impedance component of the real voltage source retains its real current source value.

In general you can summarize source transformations by keeping Ohms Law in mind and always remembering that impedances remain the same. Therefore source transformations are very easy for you to perform as long as you are familiar with Ohms Law and basically, if there is a voltage source in series with an impedance, it is possible to find the value of the equivalent current source in parallel with the impedance by dividing the value of the voltage source by the value of the impedance.

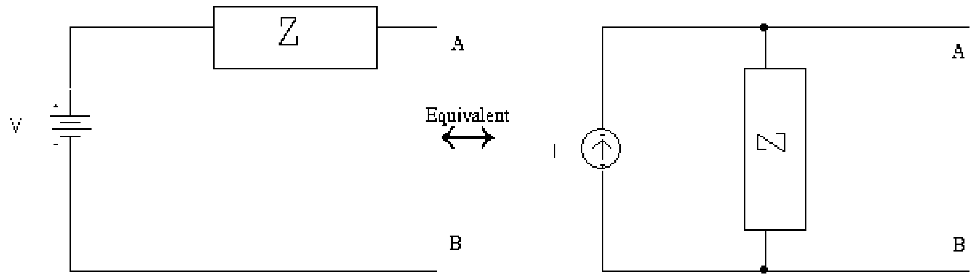
We may also apply the converse as it also applies. Initially, if a current source in parallel with an impedance is present, multiplying the value of the current source with the value of the impedance will result in the equivalent voltage source in series with the impedance.

We can illustrate what happens during a source transformation as follows:

You must remember that:

$$V = I \cdot Z$$

$$I = \frac{V}{Z}$$



DC source transformation. You should note that the impedance Z is the same in both configurations.

4.0 CONCLUSION

We have learnt that transforming a voltage into a current source and vice versa can greatly simplify the derivation of circuit solutions. We have learnt in addition that the two most powerful tools in the transformation of sources are Thevenin's theorem and Norton's theorem. Whilst transforming sources, we learnt that the source impedances remain the same.

5.0 SUMMARY

- Circuit solutions involving mixed sources are often simplified by source transformation.

- Source transformation involves the application of Thevenin's theorem and Norton's theorem.

6.0 TUTOR MARKED ASSIGNMENTS

1. Describe the process of source transformation

2. When is source transformation useful

3. An analog voltmeter is a high input impedance device that measures voltage. If a shunt resistor of very low value is connected in parallel, then the meter can measure the current through the resistance as a function of the voltage across it. Is this a good example of source transformation?

4. A voltage source comprises a12 volt source with internal impedance of 12 ohms. What are the electrical parameters of the voltage source when transformed into a current source?

5. What two theorems must one bear in mind when carrying out source transformation?

6. Is Ohms law useful in source transformation, If yes, how?

7. Can source transformation be applied to reactive circuits?

8. Explain the meaning of “source transformations are bilateral”

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

UNIT 1 EQUIVALENT IMPEDANCE TRANSFORMS, EQUIVALENT CIRCUIT, EXTRA ELEMENT THEOREM, FELICI'S LAW, FOSTER'S REACTANCE THEOREM

CONTENTS

1.0 INTRODUCTION

In this Unit, we shall study Equivalent Impedance Transforms which will enable us to derive an equivalent circuit of impedances which presents impedance between all pairs of terminals as did the given network, Equivalent Circuit with which we can calculate the simplest form of circuits that retains all of the electrical characteristics of the original, we shall learn to use the Extra Element Theorem to break down complicated problems into several simpler ones, you will be taught to use Felici's Law to calculate the net charge through a circuit and Foster's Reactance Theorem to analyze and synthesize electrical networks.

2.0 OBJECTIVES

After reading through this unit, you should be able to

1. Explain Equivalent Impedance Transforms
2. Identify Two and Three Element Networks
3. Apply Network Transform Equations to Networks
4. Describe Equivalent circuits
5. Relate Thevenin's and Norton's Theorems to Equivalent Circuit Analysis
6. Understand the similarities between Extra Element Theorem and Thevenin's Theorem
7. Use Simple RC circuit to demonstrate the Extra Element Theorem
8. State Felici's Law
9. Use Felici's Law to calculate net charge in a given period using initial and final flux
10. State Foster's Reactance Theorem
11. Describe Admittance, Susceptance and Immittance
12. Plot Reactance Curves for Capacitance, Inductance and Resonant Circuits
13. Understand Foster's First and Second Form Driving Point Impedance

3.0 MAIN CONTENT

3.1 Equivalent Impedance Transforms

You can say that equivalent impedance is an equivalent circuit of an electrical network of impedance elements which presents the same impedance between all pairs of terminals as did the given network and we shall now discuss the mathematical transformations between some passive, linear impedance networks commonly found in electronic circuits.

As you will discover, some equivalent circuits in linear network analysis include resistors in series, resistors in parallel and the extension to series and parallel circuits for capacitors, inductors and general impedances,

We shall highlight Norton and Thevenin's equivalent current generator and voltage generator circuits respectively, and a good example is the Y-Δ transform.

That the number of equivalent circuits that a linear network can be transformed into is unlimited; even with the most trivial cases is true. For instance, you might ask how many different combinations of resistors in parallel are equivalent to a given combined resistor and your answer is that it is virtually limitless. This Unit will could never hope to be comprehensive, but there are some generalisations possible.

For your information, a man named Wilhelm Cauer found a transformation that could generate all possible equivalents of a given rational, passive, linear one-port, or in other words, any given two-terminal impedance. Transformations of 4-terminal, especially 2-port, networks are also commonly found and transformations of yet more complex networks are possible.

The extensive nature of the subject of equivalent circuits is exemplified by Ronald Foster and George Campbell's 1920 paper on non-dissipative four-ports. In the course of their work they looked at the ways four ports could be interconnected with ideal transformers and found a number of combinations which might have practical applications. When they tried to calculate every single one of them, they arrived at an enormous total of 83,539 equivalents.

2-Terminal, 2-Element Networks

A single impedance has two terminals to connect to the outside world, hence you can describe it as a 2-terminal, or a one-port, network. Despite this simple description, it might surprise you that there is no limit to the number of meshes, the complexity and number of elements that the impedance network may have.

2-element networks are common in circuit design and you will discover that filters, for instance, are often LC networks. Practically, circuit designers prefer RC networks because inductors of the LC networks are more difficult to manufacture that RC networks.

Transformations of 2-element networks are simpler and are easier to find than for 3-element networks. Take note that One-element networks can be thought of as a special case of two-element network.

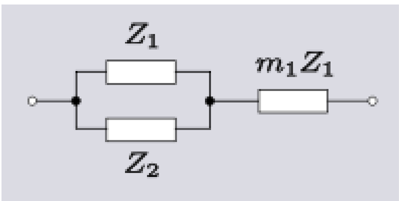
While it is possible for you to use the transformations in this Unit on certain 3-element networks by substituting a network of elements for element Z_n, you will be limited to the substitution of a maximum of two impedances.

3-Element Networks

Whereas one-element networks, and two-element, two-terminal networks which comprise either two elements in series or two elements in parallel are trivial, the smallest number of elements that a network that is non-trivial may have is three. and there are two 2-element-kind non-trivial transformations possible, one being both the reverse transformation and the topological dual, of the other.

Now let us take a look at this 3-element network, its topological dual and this network's transform equations.

Network



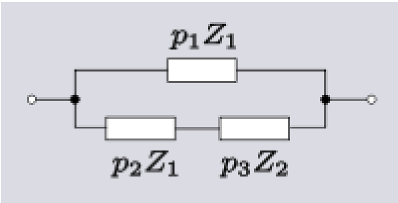
Transform Equations

$$p_1 = 1 + m_1 \ ,$$

$$p_2 = m_1(1 + m_1) \ ,$$

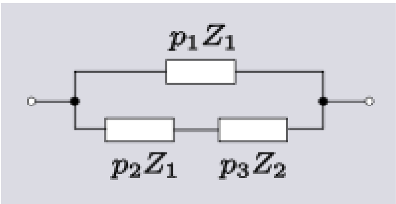
$$p_3 = (1 + m_1)^2 \ .$$

Transformed Network (Topological Dual)



Let us now take a look at how you can start from this topological dual network and derive its reverse transform: follow carefully:

Network



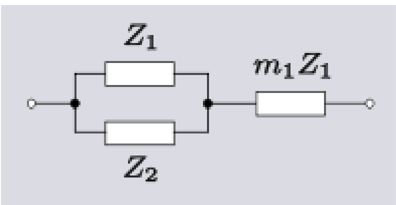
Transform Equations

$$p_1 = \frac{{m_1}^2}{1+m_1} \ ,$$

$$p_2 = \frac{m_1}{1+m_1} \ ,$$

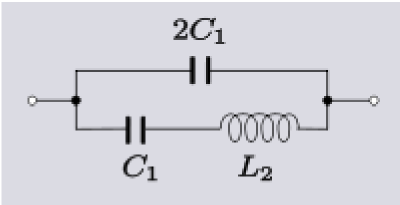
$$p_3 = \left(\frac{m_1}{1+m_1}\right)^2 \ .$$

Transformed Network



We will now use a practical example to illustrate the utility of transforms. Inductors in circuits can have impractically high values. By utilizing network transformation, you can derive a practical value of inductance which will be much lower in value and is practically achievable in reality. Let us see how.

Network



Transform Equations

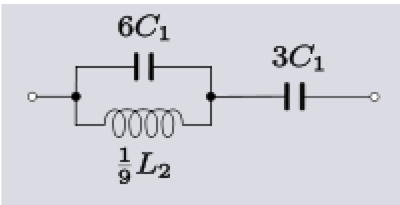
$m_1 = 0.5$,

$p_1 = \frac{1}{6}$,

$p_2 = \frac{1}{3}$,

$p_3 = \frac{1}{9}$.

Transformed Network



We can see that the value of the transformed inductance is only $1/9^{th}$ the original value, which is both smaller in physical size and therefore cheaper.

3.2 Equivalent Circuit

You can say without contradiction that an equivalent circuit is the simplest form of a circuit that retains all of the electrical characteristics of the original (and more complex) circuit.

In its most common form, an equivalent circuit is made up of linear, passive elements. However, more complex equivalent circuits are used that approximate the nonlinear behaviour of the original circuit as well.

These more complex circuits are often called macro models of the original circuit. Have you seen an operational amplifier before? Yes, an operational amplifier is a very good example of a macro model.

There are two very important two-terminal equivalent circuits you should know about:

> Thevenin's equivalent - Which reduces a two-terminal circuit to a single voltage source and a series Thevenin's impedance
>
> Norton equivalent – Which reduces a two terminal circuit to a current source and a parallel Norton impedance

For a restricted set of linear four-terminal circuits, you can set up equivalent two-port networks. The restriction upon a two-port representation is that of a port: the current entering each port must be the same as the current leaving that port By linearizing a nonlinear circuit about its operating point, you can derive such a two-port representation for transistors. We have two good examples of this representation in hybrid pi and h-parameter transistor equivalent circuits.

By way of analogy we can also make equivalent circuits which describe and model the electrical properties of materials or biological systems. A biological cell membrane for instance can be modelled as a lipid bi-layer capacitor in parallel with resistance-battery combinations, which represent ion channels powered by an ion gradient across the membrane. Quite an interesting analogy would you say.

3.3 Extra Element Theorem

The Extra Element Theorem (EET) is an analytic technique for simplifying the process of deriving driving point and transfer functions for linear electronic circuits and as you will see, it functions quite like Thevenin's theorem in that the extra element theorem breaks down one complicated problem into several simpler ones.

We can generally find Driving point and transfer functions using the methods of Kirchhoff's Voltage Law and Kirchhoff's Current Law. We may however discover that several complicated equations can result that offer little insight into a circuit's behaviour. By utilizing the extra element theorem, you can remove a circuit element such as a resistor from a circuit and then you can find the desired driving point or transfer function.

By removing the element that most complicates the circuit, like elements that creates feedback, you can obtain the desired function more easily. Then two correctional factors must be found and combined with the previously derived function to find the exact expression.

The general form of the extra element theorem is called the N-extra element theorem and allows multiple circuit elements to be removed at once. Let us see how this is done.

We will use driving point impedances to illustrate a special case of the Extra Element Theorem to find the input impedance of a network.

Let us write the Extra Element Theorem as shown below:

$$Z_{in} = Z_{in}^{\infty}\left(\frac{1+\frac{Z_e^0}{Z}}{1+\frac{Z_e^{\infty}}{Z}}\right)$$

Where

Z is the impedance chosen as the extra element

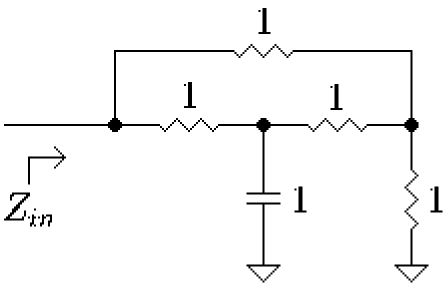
Z_{in}^{∞} is the input impedance with Z removed (or made infinite)

Z_e^0 is the impedance seen by the extra element Z with the input shorted (or made zero)

Z_e^{∞} is the impedance seen by the extra element Z with the input open (or made infinite)

When we compute these three terms, it may seem like extra effort, but be rest assured that they are often easier to compute than the overall input impedance.

Let us use this simple RC circuit to demonstrate the extra element theorem where we denote the capacitor as the extra element.



Consider the problem of finding Z_{in} for the circuit above using the extra element theorem. We make all component values unity for simplicity.

$$Z = \frac{1}{s}$$

Removing this capacitor from the circuit we find

$$Z_{in}^{\infty} = 2\|1 + 1 = \frac{5}{3}$$

Calculating the impedance seen by the capacitor with the input shorted we find

$$Z_e^0 = 1\|(1 + 1\|1) = \frac{3}{5}$$

Calculating the impedance seen by the capacitor with the input open we find

$$Z_e^{\infty} = 2\|1 + 1 = \frac{5}{3}$$

Therefore using the Extra Element Theorem, we find

$$Z_{in} = \frac{5}{3}\left(\frac{1 + \frac{3}{5}s}{1 + \frac{5}{3}s}\right)$$

It is essential for you to observe that this problem was solved by calculating three simple driving point impedances by inspection.

3.4 Felici's Law

You can always use Felici's law whenever we want to calculate the net charge through a circuit configuration in which there is a current induced by a variable magnetic field. With Felici's law it is possible to calculate net charge in a period using initial and final flux for a conductor coil immersed in a variable magnetic field.

$$q(t) = \frac{1}{\mathrm{R}}[\Phi(0) - \Phi(t)]$$

Let us demonstrate this as follows

$$q(t) = \int_0^t i(\tau)d\tau = \frac{1}{\mathrm{R}}\int_0^t f_{em}(\tau)d\tau = \frac{1}{\mathrm{R}}[\Phi(0) - \Phi(t)]$$

This will be easy for you to understand if you recall that

$$f_{em}(t) = -\frac{d\Phi(t)}{dt}$$

3.5 Foster's Reactance Theorem

Foster's reactance theorem is important for us in electrical network analysis and synthesis. This theorem states that the reactance of a passive, lossless two-terminal (one-port) network always monotonically increases with frequency.

As strange as the statement above might appear to you, the proof of it was first presented in 1924 by Ronald Martin Foster; from which the law derives its name.

Reactance is the imaginary part of the complex electrical impedance. The specification that the network must be passive and lossless implies that there are no resistors (lossless), or amplifiers or energy sources (passive) in the network. The network consequently must consist entirely of inductors and capacitors and the impedance will be purely an imaginary number with zero real part. Other than that, the theorem is quite general; in particular, it applies to distributed element circuits although Foster formulated it in terms of discrete inductors and capacitors. Foster's theorem applies equally to the admittance of a network that is the susceptance (imaginary part of admittance) of a passive, lossless one-port monotonically increases with frequency.

Once again, this result may seem counterintuitive to you since admittance is the reciprocal of impedance, but it is easily proved.

Now let us be attentive as we prove it together.

If an impedance

$$Z = iX$$

where,

Z is impedance

X is reactance

i is the imaginary unit

then the admittance is given by

$$Y = \frac{1}{iX} = -i\frac{1}{X} = iB$$

Where,

Y is admittance

B is susceptance

If X is monotonically increasing with frequency then 1/X must be monotonically decreasing. −1/ X must consequently be monotonically increasing and hence it is proved that B is also increasing.

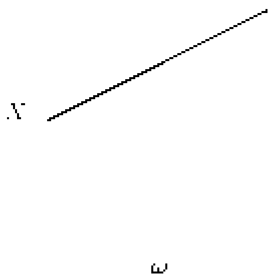
It is often the case in network theory that a principle or procedure apply equally to impedance or admittance as they do here. It is convenient in these circumstances to use the concept of immittance which can mean either impedance or admittance.

The mathematics are carried out without stating which it is or specifying units until it is desired to calculate a specific example. Foster's theorem can thus be stated in a more general form as,

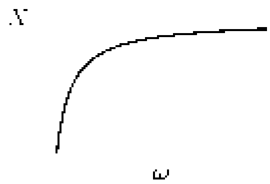
Foster's theorem (immittance form)

The imaginary immittance of a passive, lossless one-port monotonically increases with frequency

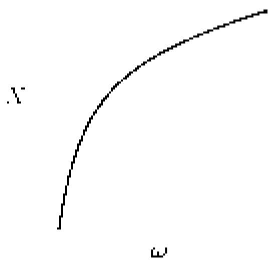
Now study these four examples



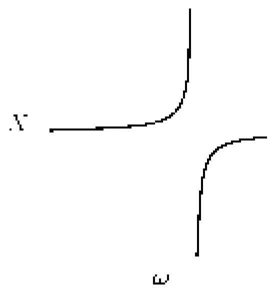
Plot of the reactance of an inductor against frequency



Plot of the reactance of a capacitor against frequency



Plot of the reactance of a series LC circuit against frequency



Plot of the reactance of a parallel LC circuit against frequency

With the following we illustrate the theorem in a number of simple circuits.

Inductor

The impedance of an inductor is given by $Z = i\omega L$

L is inductance

ω is angular frequency

Therefore the reactance is, $X = \omega L$

by inspection you can see that the reactance is monotonically, and linearly increasing with frequency

Capacitor

The impedance of a capacitor is given by,

$$Z = \frac{1}{i\omega C}$$

C is capacitance

Therefore the reactance is,

$$X = -\frac{1}{\omega C}$$

which again as you can see is monotonically increasing with frequency.

The impedance function of the capacitor is identical to the admittance function of the inductor and vice versa. Always remember that it is a general result that the dual of any immittance function that obeys Foster's theorem will also follow Foster's theorem.

Series resonant circuit

A series LC circuit has an impedance that is the sum of the impedances of an inductor and capacitor,

$$Z = i\omega L + \frac{1}{i\omega C} = i\left(\omega L - \frac{1}{\omega C}\right)$$

At low frequencies the reactance is dominated by the capacitor and so is large and negative. This monotonically increases towards zero and the magnitude of the capacitor reactance becomes smaller. The reactance passes through zero at the point where the magnitudes of the capacitor and inductor reactances are equal. This occurs at the resonant frequency after which it then continues to monotonically increase as the inductor reactance becomes progressively dominant.

Parallel resonant circuit

A parallel LC circuit is the dual of the series circuit and hence its admittance function is the same form as the impedance function of the series circuit,

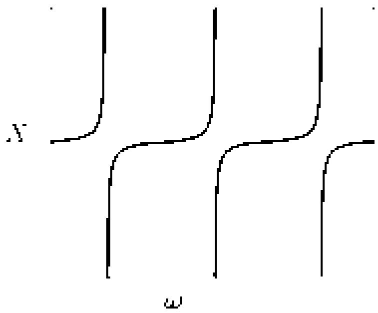
$$Y = i\omega C + \frac{1}{i\omega L}$$

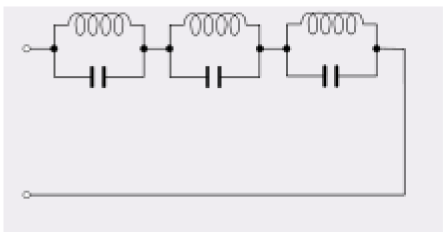
The impedance function is,

$$Z = i\left(\frac{\omega L}{1 - \omega^2 LC}\right)$$

At low frequencies the reactance is dominated by the inductor and is small and positive. This monotonically increases towards a pole at the anti-resonant frequency where the susceptance of the inductor and capacitor are equal and opposite and cancel. Past the pole the reactance is large and negative and increasing towards zero where it is dominated by the capacitance.

Poles and zeroes





Take a look at this plot of the reactance of Foster's first form of canonical driving point impedance which shows the pattern of alternating poles and zeroes. You might intuitively have worked it out that three anti-resonators are required to realise this impedance function.

A consequence of Foster's theorem is that the poles and zeroes of any passive immittance function must alternate with increasing frequency.

After passing through a pole the function will be negative and is obliged to pass through zero before reaching the next pole if it is to be monotonically increasing.

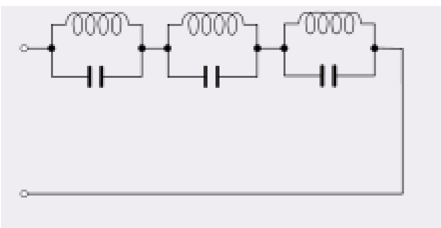
Take cognizance of the fact that not all networks obey Foster's theorem. You can refer to those that do as Foster networks and to those that do not as non-Foster networks.

A circuit containing an amplifier may or may not be not be a Foster network and in particular, it is possible to simulate negative capacitors and inductors with negative impedance conversion circuits. These circuits will have an immittance function of frequency with a negative slope.

When a scaling factor is added, the poles and zeroes of an immittance function completely determine the frequency characteristics of a Foster network. Two Foster networks that have identical poles and zeroes will be equivalent circuits in the sense that their immittance functions will be identical.

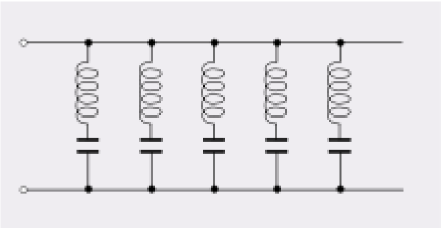
Also, another consequence of Foster's theorem is that the plot of a Foster immittance function on a Smith chart must always travel around the chart in a clockwise direction with increasing frequency.

This is Foster's first form of canonical driving point impedance realization in the network diagram below.



If the polynomial function has a pole at ω=0 one of the LC sections will reduce to a single capacitor. If the polynomial function has a pole at ω=∞ one of the LC sections will reduce to a single inductor. If both poles are present then two sections reduce to a series LC circuit.

The network below is Foster's second form of canonical driving point impedance realization. Can you distinguished it from Foster's first form?



If the polynomial function has a zero at ω=0 one of the LC sections will reduce to a single capacitor. If the polynomial function has a zero at ω=∞ one of the LC sections will reduce to a single inductor. If both zeroes are present then two sections reduce to a parallel LC circuit.

A one-port passive immittance consisting of discrete elements (that is, not a distributed element circuit) is described as rational in that in can be represented as a rational function of s,

$$Z(s) = \frac{P(s)}{Q(s)}$$

where,

$Z(s)$ is immittance

$P(s),\ Q(s)$ are polynomials with real, positive coefficients

s is the Laplace operator, which can be replaced with $i\omega$ when dealing with steady-state AC signals.

This is sometimes referred to as the driving point impedance because it is the impedance at the place in the network at which the external circuit is connected and "drives" it with a signal.

Foster described how such a lossless rational function may be realised in two ways. Foster's first form consists of a number of series connected parallel LC circuits. Foster's second form of driving point impedance consists of a number of parallel connected series LC circuits. The realization of the driving point impedance is by no means unique. Foster's realization has the advantage that the poles and/or zeroes are directly associated with a particular resonant circuit.

4.0 CONCLUSION

In this unit we have learnt about Equivalent Impedance Transforms, Equivalent Circuit, Extra Element Theorem, Felici's Law and Foster's Reactance Theorem.

An equivalent circuit of an electrical network of impedance elements which presents the same impedance between all pairs of terminals as did the given network we now know is the equivalent impedance. We visited 2-Terminal, 2-Element Networks, 3-Element Networks and have calculated topological duals along with their transform equations. We also saw how to overcome some physical device constraints through equivalent impedance transforms

We studied the two very important two-terminal equivalent circuits; Thevenin's equivalent - Which reduces a two-terminal circuit to a single voltage source and a series Thevenin's impedance, and Norton equivalent – Which reduces a two terminal circuit to a current source and a parallel Norton impedance.

We have seen how with the Extra Element Theorem, it is possible to simplify a network by removing the element that most complicates the circuit such as feedback inducing elements.

Finally, we learnt Felici's Law and Foster's Reactance Theorem as applied to reactive networks.

5.0 SUMMARY

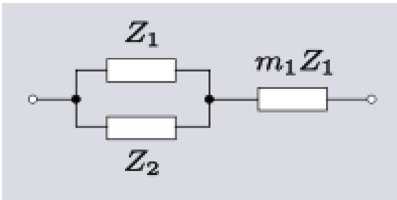
- Equivalent impedance is an equivalent circuit of an electrical network of impedance elements which presents the same impedance between all pairs of terminals as did the given network

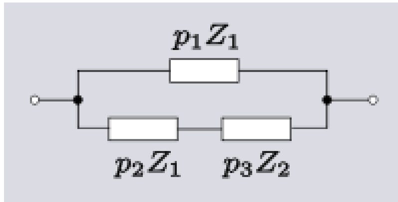
- Some equivalent circuits in linear network analysis include resistors in series, resistors in parallel and circuits of capacitors, inductors and general impedances,

- 2-Terminal, 2-Element Networks represents a single impedance which has two terminals to connect to the outside world, hence can be described as a one-port, network.

- One-element networks can be thought of as a special case of two-element network.

- Topological dualism is is exemplified by the two diagrams below





- Equivalent circuit is the simplest form of a circuit that retains all of the electrical characteristics of the original (and more complex) circuit.

- Thevenin's equivalent and Norton equivalent are two very important two-terminal equivalent circuits you should know about:

- By linearizing a nonlinear circuit about its operating point, it is possible to derive two-port representation for transistors with examples of this representation being the hybrid pi and h-parameter transistor equivalent circuits.

- Extra Element Theorem presents an analytic technique for simplifying the process of deriving driving point and transfer functions for linear electronic circuits.

- Extra Element Theorem functions like Thevenin's theorem in that the extra element theorem breaks down one complicated problem into several simpler ones.

- In order to calculate the net charge through a circuit configuration in which there is a current induced by a variable magnetic field, we can employ Felici's law as it enables us to calculate net charge in a period using initial and final flux for a conductor coil immersed in a variable magnetic field.

- Foster's reactance theorem states that the reactance of passive, lossless two-terminal network always monotonically increases with frequency.

6.0 TUTOR MARKED ASSIGNMENTS

1. Describe Equivalent Impedance Transforms.

2. What is an Equivalent Circuit.

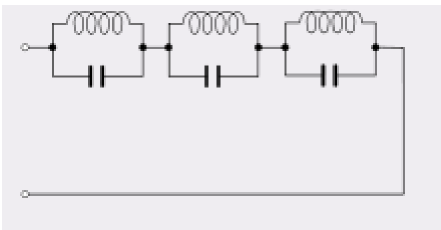
3. List two important Equivalent circuits in electrical network analysis

4. State the Extra Element Theorem

5. Can Felici's Law be applied to purely resistive networks? Why.

6. According to Foster's Reactance Theorem the impedance function of a capacitor is identical to the admittance function of an inductor and vice versa. Prove this statement

7. In the diagram below, the poles and zeroes of the passive immittance function must alternate with increasing frequency as a consequence of Foster's theorem. Sketch a frequency domain graph to illustrate this.



7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

Analog Filter Design
By M. E. Van Valkenburg Published by Holt, Rinehart and Winston

UNIT 2 KIRCHHOFF'S VOLTAGE LAW, KIRCHHOFF'S CURRENT LAW, MAXIMUM POWER TRANSFER THEOREM, MILLER THEOREM

CONTENTS

1.0 INTRODUCTION

In this Unit, we shall study Kirchhoff's Voltage Law which relates to the principle of conservation of energy and Kirchhoff's Current Law which obeys the principle of conservation of electric charge.

You will be instructed on the Theorem that states the necessary conditions for Maximum Power to be transferred from a source to a load and that maximum power transfer is not synonymous with maximum efficiency. Finally in this Unit, we shall take a close look at Miller Theorem for voltages and for current and you shall see that the two Miller Theorems are actually derived from Kirchhoff's laws.

It is safe for you to presume that most of the work in this unit centres on Kirchhoff's laws and their derivatives. You therefore have to pay special attention to the underlying concepts of Kirchhoff's laws.

2.0 OBJECTIVES

After reading through this unit, you should be able to

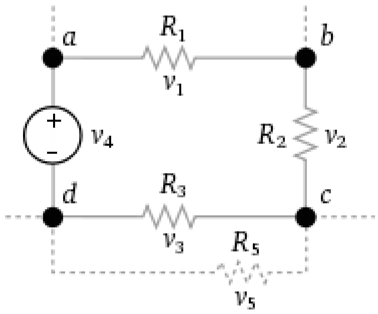
1 State Kirchhoff's Voltage and Current Laws

2 Understand How Kirchhoff's Voltage Law based On the Law of Conservation Of Energy
3 Understand How Kirchhoff's Current Law based On the Law of Conservation Of Charge
4 State the Maximum Power Transfer Theorem
5 Understand why Maximum Power Transfer (MPT) conditions do not result in maximum efficiency.
6 Prove that source and load impedances should be complex conjugates for reactive MPT
7 State Miller Theorem
8 Understand how the two versions of Miller Theorem are based on the two Kirchhoff's circuit laws

3.0 MAIN CONTENT

3.1 Kirchhoff's Voltage Law

Kirchhoff's circuit laws are two equalities that deal with the conservation of charge and energy in electrical circuits, and were first described in 1845 by Gustav Kirchhoff and they are widely used in electrical engineering. They are also called Kirchhoff's rules or simply Kirchhoff's laws. Both circuit rules can be directly derived from Maxwell's equations,



Kirchhoff's Voltage Law may be stated thus: The sum of all the voltages around a loop is equal to zero. $v_1 + v_2 + v_3 - v_4 = 0$

This law is also called Kirchhoff's second law, Kirchhoff's loop (or mesh) rule, and Kirchhoff's second rule and the principle of conservation of energy implies that the directed sum of the electrical potential differences (voltage) around any closed circuit is zero.

or more simply; The sum of the emf in any closed loop is equivalent to the sum of the potential drops in that loop.

or

The algebraic sum of the products of the resistances of the conductors and the currents in them in a closed loop is equal to the total emf available in that loop.

Similarly to Kirchhoff's Current Law, it can be stated as:

$$\sum_{k=1}^{n} V_k = 0$$

Here, n is the total number of voltages measured. The voltages may also be complex:

$$\sum_{k=1}^{n} \tilde{V}_k = 0$$

This law is based on the conservation of "energy given/taken by potential field" (not including energy taken by dissipation). Given a voltage potential, a charge which has completed a closed loop doesn't gain or lose energy as it has gone back to initial potential level.

You will discover that this law holds true even when resistance (which causes dissipation of energy) is present in a circuit. The validity of this law in this case can be understood if one realizes that a charge in fact doesn't go back to its starting point, due to dissipation of energy. A charge will just terminate at the negative terminal, instead of positive terminal. This means all the energy given by the potential difference has been fully consumed by resistance which in turn loses the energy as heat dissipation.

We summarise: Kirchhoff's voltage law has nothing to do with gain or loss of energy by electronic components (resistors, capacitors, etc). It is a law referring to the potential field generated by voltage sources. In this potential field, regardless of what electronic components are present, the gain or loss in "energy given by the potential field" must be zero when a charge completes a closed loop.

Electric field and electric potential

You can safely assume that Kirchhoff's voltage law as a consequence of the principle of conservation of energy. Otherwise, it would be possible to build a perpetual motion machine that passed a current in a circle around the circuit.

Considering that electric potential is defined as a line integral over an electric field, Kirchhoff's voltage law can be expressed equivalently as

$$\oint_C \mathbf{E} \cdot d\mathbf{l} = 0,$$

Do you see the significance of this equation; that this line integral of the electric field around closed loop C is zero?

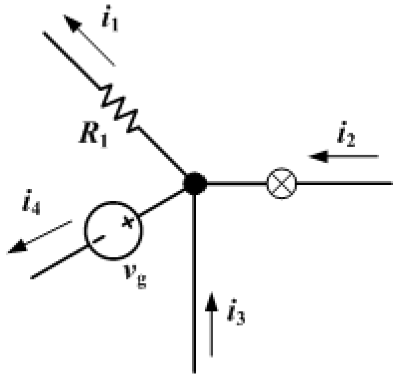
In order to return to the more special form, you can "cut in pieces" this integral can be in order to get the voltage at specific components.

Limitations

Do you know that Kirchhoff's voltage law is a simplification of Faraday's law of induction for the special case where there is no fluctuating magnetic field linking the closed loop? Now you do and for this reason it practically suffices for explaining circuits containing only resistors and capacitors.

In the presence of a changing magnetic field the electric field is not conservative and it cannot therefore define a pure scalar potential—the line integral of the electric field around the circuit is not zero. This is because energy is being transferred from the magnetic field to the current (or vice versa). In order to "fix" Kirchhoff's voltage law for circuits containing inductors, an effective potential drop, or electromotive force (emf), is associated with each inductance of the circuit, exactly equal to the amount by which the line integral of the electric field is not zero by Faraday's law of induction.

3.2 Kirchhoff's Current Law



The current entering any junction is equal to the current leaving that junction. $i_1 + i_4 = i_2 + i_3$

This law is also called Kirchhoff's point rule, Kirchhoff's junction rule (or nodal rule), and Kirchhoff's first rule.

The principle of conservation of electric charge implies that: At any node (junction) in an electrical circuit, the sum of currents flowing into that node is equal to the sum of currents flowing out of that node.

or

The algebraic sum of currents in a network of conductors meeting at a point is zero. (Assuming that current entering the junction is taken as positive and current leaving the junction is taken as negative).

If you will recall that current is a signed (positive or negative) quantity reflecting direction towards or away from a node, then you will see thay this principle can be stated as:

$$\sum_{k=1}^{n} I_k = 0$$

n is the total number of branches with currents flowing towards or away from the node.

This formula is also valid for complex currents:

$$\sum_{k=1}^{n} \tilde{I}_k = 0$$

The law is based on the conservation of charge whereby the charge (measured in coulombs) is the product of the current (in amperes) and the time (which is measured in seconds).

Changing charge density

The restriction regarding capacitor plates means that Kirchhoff's current law is only valid if the charge density remains constant in the point that it is applied to. This is normally not a problem because of the strength of electrostatic forces: the charge build-up would cause repulsive forces to disperse the charges.

However, a charge build-up can occur in a capacitor, where the charge is typically spread over wide parallel plates, with a physical break in the circuit that prevents the positive and negative charge accumulations over the two plates from coming together and cancelling. In this case, the sum of the currents flowing into one plate of the capacitor is not zero, but rather is equal to the rate of charge accumulation. However, if the displacement current dD/dt is included, Kirchhoff's current law once again holds. (This is really only required if one wants to apply the current law to a point on a capacitor plate. In circuit analyses, however, the capacitor as a whole is typically treated as a unit, in which case the ordinary current law holds since exactly the current that enters the capacitor on the one side leaves it on the other side.)

More technically, Kirchhoff's current law can be found by taking the divergence of Ampere’s law with Maxwell's correction and combining with Gauss's law, yielding:

$$\nabla \cdot \mathbf{J} = -\nabla \cdot \frac{\partial \mathbf{D}}{\partial t} = -\frac{\partial \rho}{\partial t}$$

This is simply the charge conservation equation (in integral form, it says that the current flowing out of a closed surface is equal to the rate of loss of charge within the enclosed volume (Divergence theorem). Kirchhoff's current law is equivalent to the statement that the divergence of the current is zero, true for time-invariant ρ, or always true if the displacement current is included with J.

3.3 Maximum Power Transfer Theorem

Maximum power transfer theorem states that, to obtain maximum external power from a source with a finite internal resistance, the resistance of the load must be equal to the resistance of the source as viewing from the output terminals. Moritz von Jacobi published the maximum power (transfer) theorem around 1840, which is also referred to as "Jacobi's law"

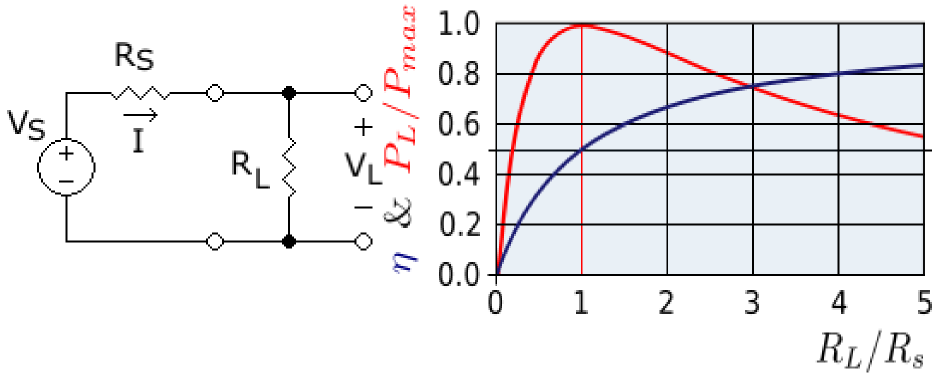
You must note that this theorem results in maximum power transfer, and not maximum efficiency. If the resistance of the load is made larger than the resistance of the source, then efficiency is higher, since a higher percentage of the source power is transferred to the load, but the magnitude of the load power is lower since the total circuit resistance goes up.

If you make the load resistance smaller than the source resistance, then most of the power ends up being dissipated in the source, and although the total power dissipated is higher, due to a lower total resistance, it turns out that the amount dissipated in the load is reduced.

You can extend the theorem to AC circuits which include reactance, and state that maximum power transfer occurs when the load impedance is equal to the complex conjugate of the source impedance.

Maximizing power transfer versus power efficiency

This theorem was originally misunderstood (notably by Joule) to imply that a system consisting of an electric motor driven by a battery could not be more than 50% efficient since, when the impedances were matched, the power lost as heat in the battery would always be equal to the power delivered to the motor. In 1880 this assumption was shown to be false by either Edison or his colleague Francis Robbins Upton, who realized that maximum efficiency was not the same as maximum power transfer. To achieve maximum efficiency, the resistance of the source (whether a battery or a dynamo) could be made close to zero. Using this new understanding, they obtained an efficiency of about 90%, and proved that the electric motor was a practical alternative to the heat engine.



The condition of maximum power transfer does not result in maximum efficiency. If we define the efficiency η as the ratio of power dissipated by the load to power developed by the source, then it is straightforward to calculate from the above circuit diagram that

$$\eta = \frac{R_{\mathrm{load}}}{R_{\mathrm{load}} + R_{\mathrm{source}}} = \frac{1}{1 + \frac{R_{\mathrm{source}}}{R_{\mathrm{load}}}}$$

Consider three particular cases:

If $R_{\mathrm{load}} = R_{\mathrm{source}}$, then $\eta = 0.5$.

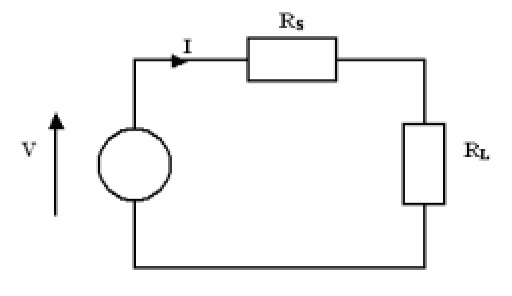
If $R_{\mathrm{load}} = \infty$or $R_{\mathrm{source}} = 0$, then $\eta = 1$.

If $R_{\mathrm{load}} = 0$, then $\eta = 0$.

The efficiency is only 50% when maximum power transfer is achieved, but approaches 100% as the load resistance approaches infinity, though the total power level tends towards zero. Efficiency also approaches 100% if the source resistance can be made close to zero. When the load resistance is zero, all the power is consumed inside the source (the power dissipated in a short circuit is zero) so the efficiency is zero.

Note that for a time varying voltage and current, the above applies only to the resistive component of the power, not the reactive component. As shown below, a proper pure reactive match can theoretically achieve 100% power transfer AND 100% efficiency simultaneously in a perfect resistance free reactive circuit. In an imperfect reactive circuit (i.e. one with both reactance and resistance), the efficiency will be the power dissipated in the load divided by the power provided by the source. When the resistance is a small fraction of the reactance, the efficiency will then be very high even at the maximum power transfer point. This is a key reason why power grids use AC versus DC power, and why radio frequency (RF) circuits can operate with relatively high efficiency.

This is impedance matching which you should apply to purely resistive circuits. Sketch this circuit and label it off head.



In the diagram, power is being transferred from the source, with voltage V and fixed source resistance R_{S}, to a load with resistance R_{L}, resulting in a current I. By Ohm's law, I is simply the source voltage divided by the total circuit resistance:

$$I = \frac{V}{R_{\mathrm{S}} + R_{\mathrm{L}}}$$

The power P_{L} dissipated in the load is the square of the current multiplied by the resistance:

$$P_{\mathrm{L}} = I^2 R_{\mathrm{L}} = \left(\frac{V}{R_{\mathrm{S}} + R_{\mathrm{L}}}\right)^2 R_{\mathrm{L}} = \frac{V^2}{R_{\mathrm{S}}^2/R_{\mathrm{L}} + 2R_{\mathrm{S}} + R_{\mathrm{L}}}$$

The value of R_{L} for which this expression is a maximum could be calculated by differentiating it, but it is easier to calculate the value of R_{L} for which the denominator

$$R_{\mathrm{S}}^2/R_{\mathrm{L}} + 2R_{\mathrm{S}} + R_{\mathrm{L}}$$

is a minimum. The result will be the same in either case. Differentiating the denominator with respect to R_{L}:

$$\frac{d}{dR_{\mathrm{L}}}\left(R_{\mathrm{S}}^2/R_{\mathrm{L}} + 2R_{\mathrm{S}} + R_{\mathrm{L}}\right) = -R_{\mathrm{S}}^2/R_{\mathrm{L}}^2 + 1$$

For a maximum or minimum, the first derivative is zero, so

$$R_{\mathrm{S}}^2/R_{\mathrm{L}}^2 = 1$$

Or

$$R_{\mathrm{L}} = \pm R_{\mathrm{S}}$$

In practical resistive circuits, R_{S} and R_{L} are both positive, so the positive sign in the above is the correct solution. To find out whether this solution is a minimum or a maximum, the denominator expression is differentiated again:

$$\frac{d^2}{dR_{\mathrm{L}}^2}\left(R_{\mathrm{S}}^2/R_{\mathrm{L}} + 2R_{\mathrm{S}} + R_{\mathrm{L}}\right) = 2R_{\mathrm{S}}^2/R_{\mathrm{L}}^3$$

This is always positive for positive values of R_{S} and R_{L}, showing that the denominator is a minimum, and the power is therefore a maximum, when

$$R_{\mathrm{S}} = R_{\mathrm{L}}$$

A note of caution to you is in order here. This last statement as written implies to many people that for a given load, the source resistance must be set equal to the load resistance for maximum power transfer. However, this equation only applies if the source resistance cannot be adjusted, e.g. with antennas (see the first line in the proof stating "fixed source resistance"). For any given load resistance a source resistance of zero is the way to transfer maximum power to the load. As an example, a 100 volt source with an internal resistance of 10 ohms connected to a 10 ohm load will deliver 250 watts to that load. Make the source resistance zero ohms and the load power jumps to 1000 watts.

Impedance matching in reactive circuits

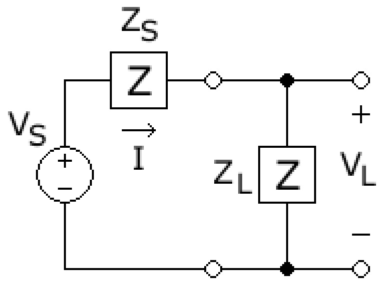
The theorem also applies where the source and/or load are not totally resistive. This invokes a refinement of the maximum power theorem which says that any reactive components of source and load should be of equal magnitude but opposite phase. (See below for a derivation.) This means that the source and load impedances should be complex conjugates of each other. In the case of purely resistive circuits, the two concepts are identical. However, physically realizable sources and loads are not usually totally resistive; having some inductive or capacitive components, and so practical applications of this theorem, under the name of complex conjugate impedance matching, do, in fact, exist.

If the source is totally inductive (capacitive), then a totally capacitive (inductive) load, in the absence of resistive losses, would receive 100% of the energy from the source but send it back after a quarter cycle. The resultant circuit is nothing other than a resonant LC circuit in which the energy continues to oscillate to and fro. This is called reactive power. Power factor correction (where an inductive reactance is used to "balance out" a capacitive one), is essentially the same idea as complex conjugate impedance matching although it is done for entirely different reasons.

For a fixed reactive source, the maximum power theorem maximizes the real power (P) delivered to the load by complex conjugate matching the load to the source.

For a fixed reactive load, power factor correction minimizes the apparent power (S) (and unnecessary current) conducted by the transmission lines, while maintaining the same amount of real power transfer. This is done by adding a reactance to the load to balance out the load's own reactance, changing the reactive load impedance into resistive load impedance.

Let us prove this: now, follow me



In this diagram, AC power is being transferred from the source, with phasor magnitude voltage $|V_S|$ (peak voltage) and fixed source impedance Z_S, to a load with impedance Z_L, resulting in a phasor magnitude current $|I|$. $|I|$ is simply the source voltage divided by the total circuit impedance:

$$|I| = \frac{|V_S|}{|Z_S + Z_L|}.$$

The average power P_L dissipated in the load is the square of the current multiplied by the resistive portion (the real part) R_L of the load impedance:

$$P_L = I_{rms}^2 R_L = \frac{1}{2}|I|^2 R_L = \frac{1}{2}\left(\frac{|V_S|}{|Z_S + Z_L|}\right)^2 R_L$$
$$= \frac{1}{2}\frac{|V_S|^2 R_L}{(R_S + R_L)^2 + (X_S + X_L)^2},$$

where the resistance R_S and reactance X_S are the real and imaginary parts of Z_S, and X_L is the imaginary part of Z_L.

In order to determine the values of R_L and X_L (since V_S, R_S, and X_S are fixed) for which this expression is a maximum, we first find, for each fixed positive value of R_L, the value of the reactive term X_L for which the denominator

$$(R_S + R_L)^2 + (X_S + X_L)^2$$

is a minimum. Since reactances can be negative, this denominator is easily minimized by making

$$X_L = -X_S.$$

The power equation is now reduced to:

$$P_L = \frac{1}{2}\frac{|V_S|^2 R_L}{(R_S + R_L)^2}$$

and it remains to find the value of R_L which maximizes this expression. However, this maximization problem has exactly the same form as in the purely resistive case, and the maximizing condition $R_L = R_S$ can be found in the same way.

The combination of conditions

$$R_L = R_S$$

$$X_L = -X_S$$

can be concisely written with a complex conjugate (the *) as:

$$Z_S = Z_L^*.$$

3.4 Miller Theorem

Miller theorem refers to the process of creating equivalent circuits. The general circuit theorem asserts that a floating impedance element supplied by two connected in series voltage sources may be split into two grounded elements with corresponding impedances. There is also a dual Miller theorem with regards to impedance supplied by two connected in parallel current sources. The two versions are based on the two Kirchhoff's circuit laws.

Miller theorems are not only pure mathematical expressions. These arrangements explain important circuit phenomena about modifying impedance (Miller effect, virtual ground, bootstrapping, negative impedance, etc.) and help designing and understanding various popular circuits (feedback amplifiers, resistive and time-dependent converters, negative impedance converters, etc.) They are useful in the area of circuit analysis especially for analyzing circuits with feedback and certain transistor amplifiers at high frequencies.

There is a close relationship between Miller theorem and Miller effect: the theorem may be considered as a generalization of the effect and the effect may be thought as a special case of the theorem.

Miller theorem (for voltages)

Miller theorem establishes that in a linear circuit, if there exists a branch with impedance Z, connecting two nodes with nodal voltages V_1 and V_2, we can replace this branch by two branches connecting the corresponding nodes to ground by impedances respectively Z/(1 - K) and KZ/(K - 1), where K = V_2/V_1. Miller theorem may be proved by using the equivalent two-port network technique to replace the two-port to its equivalent and by applying the source absorption theorem. This version of Miller theorem is based on Kirchhoff's voltage law; for that reason, it is named also Miller theorem for voltages.

Miller theorem implies that an impedance element is supplied by two arbitrary (not necessarily dependent) voltage sources that are connected in series through the common ground. In practice, one of them acts as a main (independent) voltage source with voltage V_1 and the other - as an additional (linearly dependent) voltage source with voltage $V_2 = KV_1$. The idea of Miller theorem (modifying circuit impedances seen from the sides of the input and output sources) is revealed below by comparing the two situations - without and with connected an additional voltage source V_2.

If V_2 was zero (there was not a second voltage source or the right end of the element with impedance Z was just grounded), the input current flowing through the element would be determined, according to Ohm's law, only by V_1

$$I_{in0} = \frac{V_1}{Z}$$

and the input impedance of the circuit would be

$$Z_{in0} = \frac{V_1}{I_{in0}} = Z.$$

If you were to include a second voltage source, the input current will depend on both the voltages. According to its polarity, V_2 is subtracted or added from/to V_1; so, the input current decreases/increases

$$I_{in} = \frac{V_1 - V_2}{Z} = \frac{(1-K)}{Z}V_1 = (1-K)I_{in0}$$

and the input impedance of the circuit seen from the side of the input source accordingly increases/decreases

$$Z_{in} = \frac{V_1}{I_{in}} = \frac{Z}{1-K}.$$

So, Miller theorem expresses the fact that connecting a second voltage source with proportional voltage $V_2 = KV_1$ in series with the input voltage source changes the effective voltage, the current and respectively, the circuit impedance seen from the side of the input source. Depending on the polarity, V_2 acts as a supplemental voltage source helping or opposing the main voltage source to pass the current through the impedance.

Besides by presenting the combination of the two voltage sources as a new composed voltage source, you may explain the theorem by combining the actual element and the second voltage source into a new virtual element with dynamically modified impedance. From this viewpoint, V_2 is an additional voltage that artificially increases/decreases the voltage drop V_z across the impedance Z thus decreasing/increasing the current. The proportion between the voltages determines the value of the obtained impedance (see the tables below) and gives in total six groups of typical applications.

The original Miller effect is implemented by capacitive impedance connected between the two nodes. Miller theorem generalizes Miller effect as it implies arbitrary impedance Z connected between the nodes. It is supposed also a constant coefficient K; then the expressions above are valid. But modifying properties of Miller theorem exist even when these requirements are violated and this arrangement can be generalized further by dynamizing the impedance and the coefficient.

Non-linear element

Besides impedance, Miller arrangement can modify the IV characteristic of an arbitrary element. The circuit of a diode log converter is an example of a non-linear virtually zeroed resistance where the logarithmic forward IV curve of a diode is transformed to a vertical straight line overlapping the Y axis.

Not constant coefficient

If the coefficient K varies, some exotic virtual elements can be obtained. A gyrator circuit is an example of such a virtual element where the resistance R_L is modified so that to mimic inductance, capacitance or inversed resistance.

Dual Miller theorem (for currents)

There is also a dual version of Miller theorem that is based on Kirchhoff's current law (Miller theorem for currents): if there is a branch in a circuit with impedance Z connecting a node, where two currents I_1 and I_2 converge to ground, we can replace this branch by two conducting the referred currents, with impedances respectively equal to $(1 + \alpha)Z$ and (1 +

α)Z/α, where $\alpha = I_2/I_1$. The dual theorem may be proved by replacing the two-port network by its equivalent and by applying the source absorption theorem.

Dual Miller theorem actually expresses the fact that connecting a second current source producing proportional current $I_2 = KI_1$ in parallel with the main input source and the impedance element changes the current flowing through it, the voltage and accordingly, the circuit impedance seen from the side of the input source. Depending on the direction, I_2 acts as a supplemental current source helping or opposing the main current source I_1 to create voltage across the impedance. The combination of the actual element and the second current source may be thought as of a new virtual element with dynamically modified impedance.

Dual Miller theorem is usually implemented by an arrangement consisting of two voltage sources supplying the grounded impedance Z through floating impedances The combinations of the voltage sources and belonging impedances form the two current sources - the main and the auxiliary one. As in the case of the main Miller theorem, the second voltage is usually produced by a voltage amplifier. Depending on the kind of the amplifier (inverting, non-inverting or differential) and the gain, the circuit input impedance may be virtually increased, infinite, decreased, zero or negative.

As the main Miller theorem, besides helping circuit analysis process, the dual version is a powerful tool for designing and understanding circuits based on modifying impedance by additional current. Typical applications are some exotic circuits with negative impedance as load cancellers, capacitance neutralizers,

The Howland current source consists of an input voltage source V_{IN}, a positive resistor R, a load (the capacitor C acting as impedance Z) and a negative impedance converter INIC ($R_1 = R_2 = R_3 = R$ and the op-amp). The input voltage source and the resistor R constitute an imperfect current source passing current I_R through the load The INIC acts as a second current source passing "helping" current I_{-R} through the load. As a result, the total current flowing through the load is constant and the circuit impedance seen by the input source is increased. As a comparison, in a load canceller, the INIC passes all the required current through the load; the circuit impedance seen from the side of the input source (the load impedance) is almost infinite.

4.0 CONCLUSION

You have just covered the subjects of Kirchhoff's Voltage Law, Kirchhoff's Current Law, Maximum Power Transfer Theorem and Miller Theorem.

While treating Kirchhoff's laws as applied to electrical networks, you learnt that they are merely a restatement of the laws of conservation of charge and energy in electrical circuits and that both circuit rules can be directly derived from Maxwell's equations, You will do well to enquire what happens when Kirchhoff's Voltage Law is applied to a network which include magnetically coupled inductors and have you wondered what happens when Kirchhoff's Current Law is applied to the interface between the two plates of a capacitor in c network which includes capacitors?

When studying maximum Power Transfer, you learnt that Maximum Power transfer does not imply Maximum Efficiency as half of the energy is dissipated by the load. When treating the Miller theorem you observed that it has broad practical applications in modern electronics and that the two versions of Miller theorem are based on the two Kirchhoff's circuit laws.

5.0 SUMMARY

- Kirchhoff's Voltage Law which relates to the principle of conservation of energy while Kirchhoff's Current Law results from the principle of conservation of electric charge.

- Kirchhoff's Voltage Law states that the sum of all the voltages around a loop is equal to zero. $v_1 + v_2 + v_3 - v_4 = 0$

- A limitation in applying Kirchhoff's Voltage Law is that, in the presence of a changing magnetic field the electric field is not conservative and it cannot therefore define a pure scalar potential—the line integral of the electric field around the circuit is not zero because energy is being transferred from the magnetic field to the current (or vice versa).

- Kirchhoff's Current Law states that the current entering any junction is equal to the current leaving that junction. $i_1 + i_4 = i_2 + i_3$

- A restriction regarding capacitor plates infers that Kirchhoff's current law is only valid if the charge density remains constant at the point that it is applied to.

- In order to obtain maximum external power from a source with a finite internal resistance, the resistance of the load must be equal to the resistance of the source as viewing from the output terminals.

- Maximum power transfer is not synonymous with maximum efficiency. The higher the resistance of the load compared with the internal resistance of the source, the higher the efficiency.

- Miller theorem is the process of creating equivalent circuits where a floating impedance element supplied by two connected in series voltage sources may be split into two grounded elements with corresponding impedances.

- The Dual Miller theorem refers to impedance supplied by two connected in parallel current sources.

- Both Miller theorems are based on the two Kirchhoff's circuit laws.

6.0 TUTOR MARKED ASSIGNMENTS

1. State Kirchhoff's Voltage Law

2. Under what condition does Kirchhoff's Voltage Law encounter difficulties?

3. State Kirchhoff's Current Law

4. What happens when you apply Kirchhoff's Current Law to the region between the plates of a capacitor?

5. Explain why Kirchhoff's Circuit Laws are said to obey the laws of conservation of energy and electric charge?

6. Which laws are stated by the expressions below?

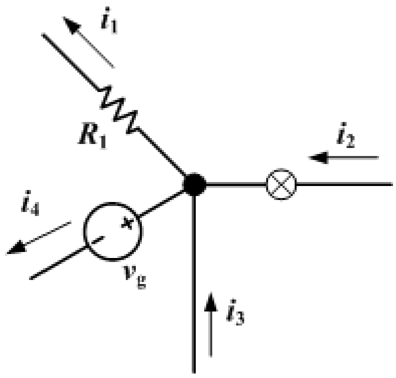
$$\sum_{k=1}^{n} V_k = 0 \qquad\qquad \sum_{k=1}^{n} \tilde{V}_k = 0$$

7. Prove that the maximum efficiency derivable from a source is 50%

8. Maximum power transfer does not imply maximum efficiency. Explain this statement.

9. Which circuit law is the dual Miller Theorem derived from?

10. In the diagram below, if i1, i2 and 13 are 10, 15 and -5 milli Amps respectively, then what is the current i4?



7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

UNIT 3 MILLMAN'S THEOREM, NORTON'S THEOREM, OHM'S LAW, RECIPROCITY

CONTENTS

1.0 INTRODUCTION

You will learn that parallel generator theorem is the same as Millman's theorem and that it is a method to simplify the solution of a circuit, In this Unit, we will study the Mayer–Norton theorem for collection of voltage sources, current sources and resistors. You will re-discover Ohms law and the derivation of the law, its various equivalent expressions and its applications. Finally we shall treat the concept of reciprocity in this Unit.

2.0 OBJECTIVES

After reading through this unit, you should be able to

1 State Millman's Theorem
2 Understand How Millman's Theorem Is Used To Compute Parallel Branch Voltage
3 Establish That Millman's Theorem Is Derived from Ohm's And Kirchhoff's Laws
4 Work with Supernode
5 State Norton's Theorem
6 Use To Calculate Equivalent Circuits
7 Understand That Norton's Theorem Is an Extension of Thevenin's Theorem

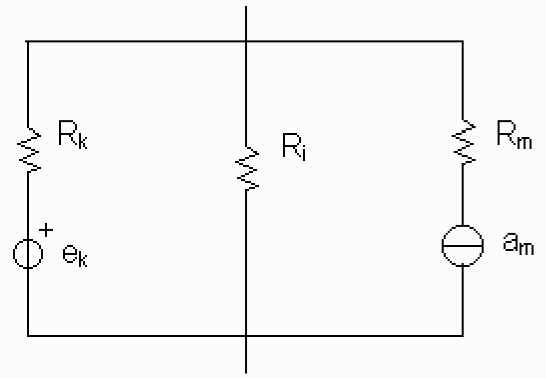
8 State Ohm's Law
9 Apply Ohm's Law to Purely Resistive Networks
10 Use Ohm's Law to Calculate Current in Reactive Circuits
11 Explain Why Ohm's Law Does Not Apply To P-N Junction Devices
12 Recognise the Other Versions of Ohm's Law
13 Understand the Principles of Reciprocity
14 Identify Practical Applications of Reciprocity in Spectral Radiators and Absorbers

3.0 MAIN CONTENT

3.1 Millman's Theorem

Millman's theorem (or the parallel generator theorem) is a method to simplify the solution of a circuit. Specifically, Millman's theorem is used to compute the voltage at the ends of a circuit made up of only branches in parallel. For emphasis, you should underscore "circuit made up of only branches in parallel"

It is named after Jacob Millman, who proved the theorem.



Let e_k be the voltage generators and a_m the current generators.

Let R_i be the resistances on the branches with no generator.

Let R_k be the resistances on the branches with voltage generators.

Let R_m be the resistances on the branches with current generators.

Then Millman states that the voltage at the ends of the circuit is given by:

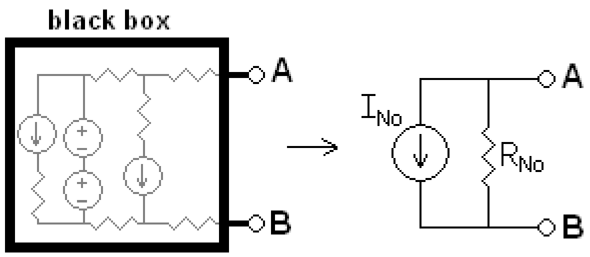
$$v = \frac{\sum \frac{\pm e_k}{R_k} + \sum \pm a_m}{\sum \frac{1}{R_k} + \sum \frac{1}{R_i}}$$

You can prove this by considering the circuit as a single super node. Then, according to Ohm’s and Kirchhoff’s laws, is the voltage between the ends of the circuit not equal to the total current entering the supernode divided by the total equivalent conductance of the Supernode? Yes it is equal to it. The total current is the sum of the currents flowing in each branch. The total equivalent conductance of the supernode is the sum of the conductance of each branch, since all the branches are in parallel. When computing the equivalent conductance all the generators have to be switched off, so all voltage generators become short circuits and all current generators become open circuits. That's why the resistances on the branches with current generators do not appear in the expression of the total equivalent conductance.

3.2 Norton’s Theorem

Norton's theorem for linear electrical networks which you can also refer to as the Mayer–Norton theorem, states that any collection of voltage sources, current sources, and resistors with two terminals is electrically equivalent to an ideal current source, I, in parallel with a single resistor, R. For single-frequency AC systems the theorem can also be applied to general impedances, not just resistors. The Norton equivalent is used to represent any network of linear sources and impedances, at a given frequency. The circuit consists of an ideal current source in parallel with an ideal impedance (or resistor for non-reactive circuits).

For your information, Norton's theorem is an extension of Thévenin's theorem and was introduced in 1926 separately by two people: Hause-Siemens researcher Hans Ferdinand Mayer (1895–1980) and Bell Labs engineer Edward Lawry Norton (1898–1983). Only Mayer actually published on this topic, but Norton made known his finding through an internal technical report at Bell Labs.



Any black box containing only voltage sources, current sources, and resistors can be converted to a Norton equivalent circuit.

Calculation of a Norton equivalent circuit

The Norton equivalent circuit is a current source with current I_{No} in parallel with a resistance R_{No} and for you to find the equivalent you must do the following:

- find the Norton current I_{No}. Calculate the output current, I_{AB}, with a short circuit as the load (meaning 0 resistance between A and B). This is I_{No}.

- Find the Norton resistance R_{No}. When there are no dependent sources (i.e., all current and voltage sources are independent), there are two methods of determining the Norton impedance R_{No}.

- Calculate the output voltage, V_{AB}, when in open circuit condition (i.e., no load resistor — meaning infinite load resistance). R_{No} equals this V_{AB} divided by I_{No}.

or you can:

- Replace independent voltage sources with short circuits and independent current sources with open circuits. The total resistance across the output port is the Norton impedance R_{No}.

or alternatively you:

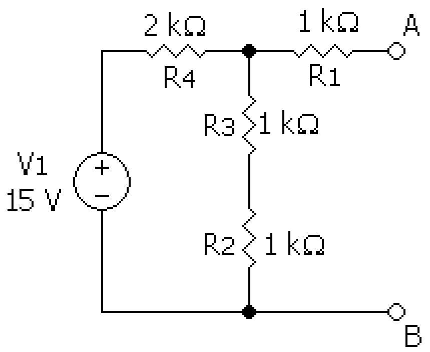
Use a given Thevenin's resistance: as the two are equal.

However, when there are dependent sources, then you should use the general method.

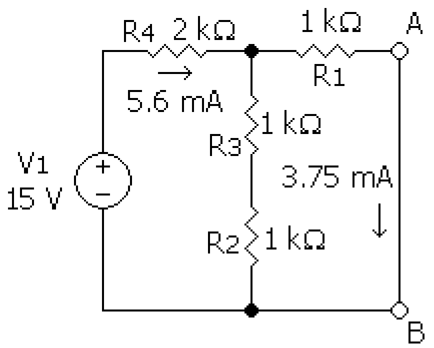
- Connect a constant current source at the output terminals of the circuit with a value of 1 Ampere and calculate the voltage at its terminals. The quotient of this voltage divided by the 1 A current is the Norton impedance R_{No}. This method must be used if the circuit contains dependent sources, but it

Let us consider the following illustrative examples

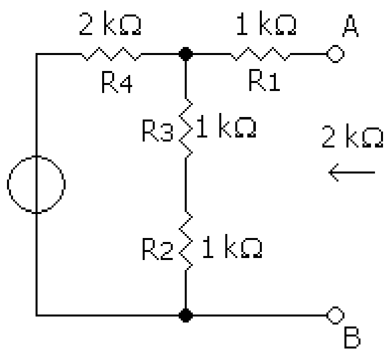
Step 0: The original circuit



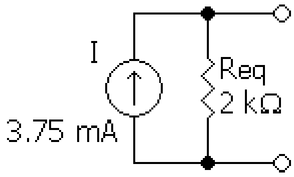
Step 1: Calculating the equivalent output current



Step 2: Calculating the equivalent resistance



Step 3: The equivalent circuit



In the example, the total current I_{total} is given by:

$$I_{\text{total}} = \frac{15\text{V}}{2\,\text{k}\Omega + 1\,\text{k}\Omega \| (1\,\text{k}\Omega + 1\,\text{k}\Omega)} = 5.625\text{mA}$$

The current through the load is then, using the current divider rule:

$$I = \frac{1\,\text{k}\Omega + 1\,\text{k}\Omega}{(1\,\text{k}\Omega + 1\,\text{k}\Omega + 1\,\text{k}\Omega)} \cdot I_{\text{total}}$$

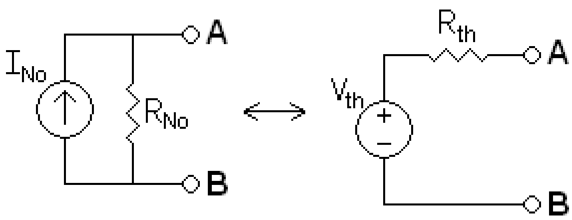
$$= 2/3 \cdot 5.625\text{mA} = 3.75\text{mA}$$

And the equivalent resistance looking back into the circuit is:

$$R = 1\,\text{k}\Omega + 2\,\text{k}\Omega \| (1\,\text{k}\Omega + 1\,\text{k}\Omega) = 2\,\text{k}\Omega$$

So the equivalent circuit is a 3.75 mA current source in parallel with a 2 kΩ resistor.

Let us now see how we can convert to the Thevenin equivalent circuit



A Norton equivalent circuit is related to the Thevenin equivalent by the following equations:

$$R_{Th} = R_{No}$$

$$V_{Th} = I_{No} R_{No}$$

$$V_{Th} / R_{Th} = I_{No}$$

3.3 Ohm’s Law

Ohm's law states that the current through a conductor between two points is directly proportional to the potential difference or voltage across the two points, and inversely proportional to the resistance between them.

The mathematical equation that describes this relationship is:

$$I = \frac{V}{R}$$

where I is the current through the conductor in units of amperes, V is the potential difference measured across the conductor in units of volts, and R is the resistance of the conductor in units of ohms. More specifically, Ohm's law states that the R in this relation is constant, independent of the current.

The law was named after the German physicist Georg Ohm, who, in a treatise published in 1827, described measurements of applied voltage and current through simple electrical circuits containing various lengths of wire. He presented a slightly more complex equation than the one above to explain his experimental results.

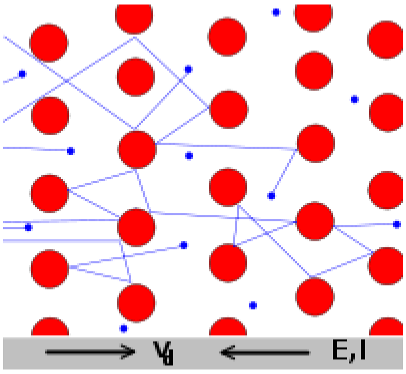
$I = \frac{V}{R}$ is the modern form of Ohm's law.

The term Ohm's law is also used to refer to various generalizations of the law originally formulated by Ohm and the simplest example of this is:

$$\boldsymbol{J} = \sigma \boldsymbol{E},$$

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. This reformulation of Ohm's law is due to Gustav Kirchhoff.

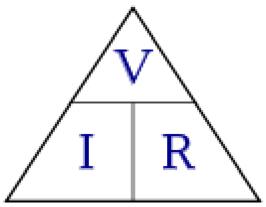
I want you to take a close look at this model description of Ohms Law developed by Paul Drude.



The Drude model developed by Paul Drude in 1900 treats electrons (or other charge carriers) like pinballs bouncing between the ions that make up the structure of the material. Electrons will be accelerated in the opposite direction to the electric field by the average electric field at their location. With each collision, though, the electron is deflected in a random direction with a velocity that is much larger than the velocity gained by the electric field. The net result is that electrons take a tortuous path due to the collisions, but generally drift in a direction opposing the electric field.

The drift velocity then determines the electric current density and its relationship to **E** and is independent of the collisions. Drude calculated the average drift velocity from $\mathbf{p} = -e\mathbf{E}\tau$ where **p** is the average momentum, $-e$ is the charge of the electron and τ is the average time between the collisions. Since both the momentum and the current density are proportional to the drift velocity, the current density becomes proportional to the applied electric field; this leads to Ohm's law.

Circuit analysis



Study the triangle above. Do you know what it is called? It is known as the Ohm's law triangle which graphically illustrates the relationship between voltage, current and resistance. I will ask you to recall it later in this unit – so – remember it.

In circuit analysis, three equivalent expressions of Ohm's law are used interchangeably:

$$I = \frac{V}{R} \quad \text{or} \quad V = IR \quad \text{or} \quad R = \frac{V}{I}.$$

Each equation is quoted by some sources as the defining relationship of Ohm's law, or all three are quoted, or derived from a proportional form, or even just the two that do not correspond to Ohm's original statement may sometimes be given.

The interchangeable nature of the equation may be represented by a triangle, where V (voltage) is placed on the top section, the I (current) is placed to the left section, and the R (resistance) is placed to the right. The line that divides the left and right sections indicate multiplication, and the divider between the top and bottom sections indicates division (hence the division bar).

Resistive circuits

Resistors are circuit elements that impede the passage of electric charge in agreement with Ohm's law, and are designed to have a specific resistance value R. In a schematic diagram the resistor is shown as a zigzag symbol. An element (resistor or conductor) that behaves according to Ohm's law over some operating range is referred to as an ohmic device (or an ohmic resistor) because Ohm's law and a single value for the resistance suffice to describe the behaviour of the device over that range.

Ohm's law holds for circuits containing only resistive elements and you should remember that no capacitances or inductances are allowed for all forms of driving voltage or current, regardless of whether the driving voltage or current is constant (DC) or time-varying such as AC. At any instant of time Ohm's law is valid for such circuits.

Resistors which are in series or in parallel may be grouped together into a single "equivalent resistance" in order to apply Ohm's law in analyzing the circuit.

Reactive circuits with time-varying signals

When reactive elements such as capacitors, inductors, or transmission lines are involved in a circuit to which AC or time-varying voltage or current is applied, the relationship between voltage and current becomes the solution to a differential equation, so Ohm's law (as defined above) does not directly apply since that form contains only resistances having value R, not complex impedances which may contain capacitance ("C") or inductance ("L").

Equations for time-invariant AC circuits take the same form as Ohm's law, however, the variables are generalized to complex numbers and the current and voltage waveforms are complex exponentials.

In this approach, a voltage or current waveform takes the form Ae^{st}, where t is time, s is a complex parameter, and A is a complex scalar. In any linear time-invariant system, all of the currents and voltages can be expressed with the same s parameter as the input to the system, allowing the time-varying complex exponential term to be canceled out and the system described algebraically in terms of the complex scalars in the current and voltage waveforms.

The complex generalization of resistance is impedance, usually denoted Z; it can be shown that for an inductor,

$$Z = sL$$

and for a capacitor,

$$Z = \frac{1}{sC}.$$

We can now write,

$$\boldsymbol{V} = \boldsymbol{I} \cdot \boldsymbol{Z}$$

where **V** and **I** are the complex scalars in the voltage and current respectively and **Z** is the complex impedance.

This form of Ohm's law, with Z taking the place of R, generalizes the simpler form. When Z is complex, only the real part is responsible for dissipating heat.

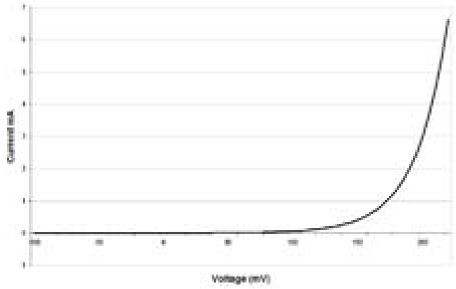
In the general AC circuit, Z varies strongly with the frequency parameter s, and so also will the relationship between voltage and current.

For the common case of a steady sinusoid, the s parameter is taken to be jω, corresponding to a complex sinusoid $Ae^{j\omega t}$. The real parts of such complex current and voltage waveforms describe the actual sinusoidal currents and voltages in a circuit, which can be in different phases due to the different complex scalars.

Linear approximations

Ohm's law is one of the basic equations used in the analysis of electrical circuits. It applies to both metal conductors and circuit components (resistors) specifically made for this behaviour. Both are ubiquitous in electrical engineering. Materials and components that obey Ohm's law are described as "ohmic" which means they produce the same value for resistance (R = V/I) regardless of the value of V or I which is applied and whether the applied voltage or current is DC (direct current) of either positive or negative polarity or AC (alternating current).

In a true ohmic device, the same value of resistance will be calculated from R = V/I regardless of the value of the applied voltage V. That is, the ratio of V/I is constant, and when current is plotted as a function of voltage the curve is linear (a straight line). If voltage is forced to some value V, then that voltage V divided by measured current I will equal R. Or if the current is forced to some value I, then the measured voltage V divided by that current I is also R. Since the plot of I versus V is a straight line, then it is also true that for any set of two different voltages V_1 and V_2 applied across a given device of resistance R, producing currents $I_1 = V_1/R$ and $I_2 = V_2/R$, that the ratio $(V_1-V_2)/(I_1-I_2)$ is also a constant equal to R. The operator "delta" (Δ) is used to represent a difference in a quantity, so we can write $\Delta V = V_1-V_2$ and $\Delta I = I_1-I_2$. Summarizing, for any truly ohmic device having resistance R, $V/I = \Delta V/\Delta I = R$ for any applied voltage or current or for the difference between any set of applied voltages or currents.



Look at this plot above. It is a plot of the i–v curve of an ideal p-n junction diode and does not follow ohm's law as it is not a straight line.

There are, however, components of electrical circuits which do not obey Ohm's law; that is, their relationship between current and voltage (their I–V curve) is nonlinear. An example is the p-n junction diode (curve at right). As seen in the figure, the current does not increase linearly with applied voltage for a diode. One can determine a value of current (I) for a given value of applied voltage (V) from the curve, but not from Ohm's law, since the value of "resistance" is not constant as a function of applied voltage. Further, the current only increases significantly if the applied voltage is positive, not negative. The ratio V/I for some point along the nonlinear curve is sometimes called the static, or chordal, or DC, resistance, but as seen in the figure the value of total V over total I varies depending on the particular point along the nonlinear curve which is chosen. This means the "DC resistance" V/I at some point on the curve is not the same as what would be determined by applying an AC signal having peak amplitude ΔV volts or ΔI amps centred at that same point along the curve and measuring $\Delta V/\Delta I$. However, in some diode applications, the AC signal applied to the device is small and it is possible to analyze the circuit in terms of the dynamic, small-signal, or incremental resistance, defined as the one over the slope of the V–I curve at the average value (DC operating point) of the voltage (that is, one over the derivative of current with respect to voltage). For sufficiently small signals, the dynamic resistance allows the Ohm's law small signal resistance to be calculated as approximately one over the slope of a line drawn tangentially to the V-I curve at the DC operating point.

Temperature effects

Ohm's law has sometimes been stated as, "for a conductor in a given state, the electromotive force is proportional to the current produced." That is, that the resistance, the ratio of the applied electromotive force (or voltage) to the current, "does not vary with the current strength ." The qualifier "in a given state" is usually interpreted as meaning "at a constant temperature," since the resistivity of materials is usually temperature dependent. Because the conduction of current is related to Joule heating of the conducting body, according to Joule's first law, the temperature of a conducting body may change when it carries a current. The dependence of resistance on temperature therefore makes resistance depend upon the current in a typical experimental setup, making the law in this form difficult to directly verify. Maxwell and others worked out several methods to test the law experimentally in 1876, controlling for heating effects.

Relation to heat conductions

Ohm's principle predicts the flow of electrical charge (i.e. current) in electrical conductors when subjected to the influence of voltage differences; Jean-Baptiste-Joseph Fourier's principle predicts the flow of heat in heat conductors when subjected to the influence of temperature differences.

The same equation describes both phenomena, the equation's variables taking on different meanings in the two cases. Specifically, solving a heat conduction (Fourier) problem with temperature (the driving "force") and flux of heat (the rate of flow of the driven "quantity", i.e. heat energy) variables also solves an analogous electrical conduction (Ohm) problem having electric potential (the driving "force") and electric current (the rate of flow of the driven "quantity", i.e. charge) variables.

The basis of Fourier's work was his clear conception and definition of thermal conductivity. He assumed that, all else being the same, the flux of heat is strictly proportional to the gradient of temperature. Although undoubtedly true for small temperature gradients, strictly proportional behaviour will be lost when real materials (e.g. ones having a thermal conductivity that is a function of temperature) are subjected to large temperature gradients.

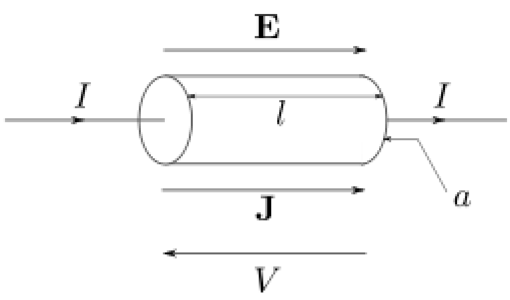
A similar assumption is made in the statement of Ohm's law: other things being alike, the strength of the current at each point is proportional to the gradient of electric potential. The accuracy of the assumption that flow is proportional to the gradient is more readily tested, using modern measurement methods, for the electrical case than for the heat case.

Other versions of Ohm's law

Ohm's law, in the form above, is an extremely useful equation in the field of electrical/electronic engineering because it describes how voltage, current and resistance are interrelated on a "macroscopic" level, that is, commonly, as circuit elements in an electrical circuit. Physicists who study the electrical properties of matter at the microscopic level use a closely related and more general vector equation, sometimes also referred to as Ohm's law, having variables that are closely related to the V, I, and R scalar variables of Ohm's law, but are each functions of position within the conductor. Physicists often use this continuum form of Ohm's Law:

$$\mathbf{E} = \rho \mathbf{J}$$

where "**E**" is the electric field vector with units of volts per meter (analogous to "V" of Ohm's law which has units of volts), "**J**" is the current density vector with units of amperes per unit area (analogous to "I" of Ohm's law which has units of amperes), and "ρ" (Greek "rho") is the resistivity with units of ohmmeters (analogous to "R" of Ohm's law which has units of ohms). The above equation is sometimes written as **J** = σ**E** where "σ" is the conductivity which is the reciprocal of ρ.



Current flowing through a uniform cylindrical conductor (such as a round wire) with a uniform field applied.

The potential difference between two points is defined as.

$$\Delta V = -\int \mathbf{E} \cdot d\mathbf{l}$$

with $d\mathbf{l}$ the element of path along the integration of electric field vector **E**. If the applied **E** field is uniform and oriented along the length of the conductor as shown in the figure, then defining the voltage V in the usual convention of being opposite in direction to the field and with the understanding that the voltage V is measured differentially across the length of the conductor allowing us to drop the Δ symbol, the above vector equation reduces to the scalar equation:

$$V = El \quad \text{or} \quad E = \frac{V}{l}.$$

Since the **E** field is uniform in the direction of wire length, for a conductor having uniformly consistent resistivity ρ, the current density **J** will also be uniform in any cross-sectional area and oriented in the direction of wire length, so we may write:

$$J = \frac{I}{a}.$$

Substituting the above 2 results (for E and J respectively) into the continuum form shown at the beginning of this section:

$$\frac{V}{l} = \frac{I}{a}\rho \qquad \text{or} \qquad V = I\rho\frac{l}{a}.$$

The electrical resistance of a uniform conductor is given in terms of resistivity by:

$$R = \rho\frac{l}{a}$$

where l is the length of the conductor in SI units of meters, a is the cross-sectional area (for a round wire $a = \pi r^2$ if r is radius) in units of meters squared, and ρ is the resistivity in units of ohmmeters.

After substitution of R from the above equation into the equation preceding it, the continuum form of Ohm's law for a uniform field (and uniform current density) oriented along the length of the conductor reduces to the more familiar form:

$$V = IR.$$

A perfect crystal lattice, with low enough thermal motion and no deviations from periodic structure, would have no resistivity, but a real metal has crystallographic defects, impurities, multiple isotopes, and thermal motion of the atoms. Electrons scatter from all of these, resulting in resistance to their flow.

The more complex generalized forms of Ohm's law are important to condensed matter physics, which studies the properties of matter and, in particular, its electronic structure. In broad terms, they fall under the topic of constitutive equations and the theory of transport coefficients.

Magnetic effects

The continuum form of the equation is only valid in the reference frame of the conducting material. If the material is moving at velocity **v** relative to a magnetic field **B**, a term must be added as follows:

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = \mathbf{J}\rho.$$

Lorentz force and Hall effect are some other implications of a magnetic field and I want you to find out more about these phenomenon. This equation is not a modification to Ohm's law. Rather, it is analogous in circuit analysis terms to taking into account inductance as well as resistance.

3.4 Reciprocity

Reciprocity refers to a variety of related theorems involving the interchange of time-harmonic electric current densities (sources) and the resulting electromagnetic fields in Maxwell's equations for time-invariant linear media under certain constraints. Reciprocity is closely related to the concept of Hermitian operators from linear algebra, applied to electromagnetism.

Perhaps the most common and general such theorem is Lorentz reciprocity (and its various special cases such as Rayleigh-Carson reciprocity), named after work by Hendrik Lorentz in 1896 following analogous results regarding sound by Lord Rayleigh and Helmholtz (Potton, 2004). Loosely, it states that the relationship between an oscillating current and the resulting electric field is unchanged if one interchanges the points where the current is placed and where the field is measured. For the specific case of an electrical network, it is sometimes phrased as the statement that voltages and currents at different points in the network can be interchanged. More technically, it follows that the mutual impedance of a first circuit due to a second is the same as the mutual impedance of the second circuit due to the first.

There is also an analogous theorem in electrostatics, known as Green's reciprocity, relating the interchange of electric potential and electric charge density.

Forms of the reciprocity theorems are used in many electromagnetic applications, such as analyzing electrical networks and antenna systems. For example, reciprocity implies that antennas work equally well as transmitters or receivers, and specifically that an antenna's radiation and receiving patterns are identical. Reciprocity is also a basic lemma that is used to prove other theorems about electromagnetic systems, such as the symmetry of the impedance matrix and scattering matrix, symmetries of Green's functions for use in boundary-element and transfer-matrix computational methods, as well as orthogonality properties of harmonic modes in waveguide systems (as an alternative to proving those properties directly from the symmetries of the Eigen-operators).

Lorentz reciprocity

Specifically, suppose that one has a current density $\mathbf{J}_1$ that produces an electric field $\mathbf{E}_1$ and a magnetic field $\mathbf{H}_1$, where all three are periodic functions of time with angular frequency ω, and in particular they have time-dependence exp(− iωt). Suppose that we similarly have a second current $\mathbf{J}_2$ at the same frequency ω which (by itself) produces fields $\mathbf{E}_2$

and $\mathbf{H}_2$. The Lorentz reciprocity theorem then states, under certain simple conditions on the materials of the medium described below, that for an arbitrary surface S enclosing a volume V:

$$\int_V [\mathbf{J}_1 \cdot \mathbf{E}_2 - \mathbf{E}_1 \cdot \mathbf{J}_2]\, dV = \oint_S [\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1] \cdot \mathbf{dA}.$$

Equivalently, in differential form (by the divergence theorem):

$$\mathbf{J}_1 \cdot \mathbf{E}_2 - \mathbf{E}_1 \cdot \mathbf{J}_2 = \nabla \cdot [\mathbf{E}_1 \times \mathbf{H}_2 - \mathbf{E}_2 \times \mathbf{H}_1]\,.$$

This general form is commonly simplified for a number of special cases. In particular, one usually assumes that $\mathbf{J}_1$ and $\mathbf{J}_2$ are localized (i.e. have compact support), and that there are no incoming waves from infinitely far away. In this case, if one integrates over all space then the surface-integral terms cancel (see below) and one obtains:

$$\int \mathbf{J}_1 \cdot \mathbf{E}_2\, dV = \int \mathbf{E}_1 \cdot \mathbf{J}_2\, dV$$

This result (along with the following simplifications) is sometimes called the Rayleigh-Carson reciprocity theorem, after Lord Rayleigh's work on sound waves and an extension by John R. Carson (1924; 1930) to applications for radio frequency antennas. Often, one further simplifies this relation by considering point-like dipole sources, in which case the integrals disappear and one simply has the product of the electric field with the corresponding dipole moments of the currents. Or, for wires of negligible thickness, one obtains the applied current in one wire multiplied by the resulting voltage across another and vice versa.

Another special case of the Lorentz reciprocity theorem applies when the volume V entirely contains both of the localized sources (or alternatively if V intersects neither of the sources). In this case:

$$\oint_S (\mathbf{E}_1 \times \mathbf{H}_2) \cdot \mathbf{dA} = \oint_S (\mathbf{E}_2 \times \mathbf{H}_1) \cdot \mathbf{dA}$$

Reciprocity for electrical networks

Above, Lorentz reciprocity was phrased in terms of an externally applied current source and the resulting field. Often, especially for electrical networks, one instead prefers to think of an externally applied voltage and the resulting currents. The Lorentz reciprocity theorem describes this case as well, assuming ohmic materials (i.e. currents that respond linearly to the applied field) with a 3×3 conductivity matrix σ that is required to be symmetric, which is implied by the other conditions below. In order to properly describe this situation, one must carefully distinguish between the externally applied fields (from the driving voltages) and the total fields that result (King, 1963).

More specifically, the $\mathbf{J}$ above only consisted of external "source" terms introduced into Maxwell's equations. We now denote this by $\mathbf{J}^{(e)}$ to distinguish it from the total current produced by both the external source and by the resulting electric fields in the materials. If this external current is in a material with a conductivity σ, then it corresponds to an externally applied electric field $\mathbf{E}^{(e)}$ where, by definition of σ:

$$\mathbf{J}^{(e)} = \sigma \mathbf{E}^{(e)}$$

Moreover, the electric field $\mathbf{E}$ above only consisted of the response to this current, and did not include the "external" field $\mathbf{E}^{(e)}$. Therefore, we now denote the field from before as $\mathbf{E}^{(r)}$, where the total field is given by $\mathbf{E} = \mathbf{E}^{(e)} + \mathbf{E}^{(r)}$.

Now, the equation on the left-hand side of the Lorentz reciprocity theorem can be rewritten by moving the σ from the external current term $\mathbf{J}^{(e)}$ to the response field terms $\mathbf{E}^{(r)}$, and also adding and subtracting a $\sigma \mathbf{E}_1^{(e)} \mathbf{E}_2^{(e)}$ term, to obtain the external field multiplied by the total current $\mathbf{J} = \sigma \mathbf{E}$:

$$\int_V \left[\mathbf{J}_1^{(e)} \cdot \mathbf{E}_2^{(r)} - \mathbf{E}_1^{(r)} \cdot \mathbf{J}_2^{(e)}\right] dV = \int_V \left[\sigma \mathbf{E}_1^{(e)} \cdot (\mathbf{E}_2^{(r)} + \mathbf{E}_2^{(e)}) - (\mathbf{E}_1^{(r)} + \mathbf{E}_1^{(e)}) \cdot \sigma \mathbf{E}_2^{(e)}\right] dV$$

$$= \int_V \left[\mathbf{E}_1^{(e)} \cdot \mathbf{J}_2 - \mathbf{J}_1 \cdot \mathbf{E}_2^{(e)}\right] dV$$

For the limit of thin wires, this gives the product of the externally applied voltage (1) multiplied by the resulting total current (2) and vice versa. In particular, the Rayleigh-Carson reciprocity theorem becomes a simple summation:

$$\sum_n V_1^{(n)} I_2^{(n)} = \sum_n V_2^{(n)} I_1^{(n)}$$

where V and I denote the (complex) amplitudes of the AC applied voltages and the resulting currents, respectively, in a set of circuit elements (indexed by n) for two possible sets of voltages V_1 and V_2.

Most commonly, this is simplified further to the case where each system has a single voltage source V, at $V_1^{(1)} = V$ and $V_2^{(2)} = V$.

Then the theorem becomes simply $I_1^{(2)} = I_2^{(1)}$: the current at position (1) from a voltage at (2) is identical to the current at (2) from the same voltage at (1).

Proof of Lorentz reciprocity

The Lorentz reciprocity theorem is simply a reflection of the fact that the linear operator $\hat{O}$ relating $\mathbf{J}$ and $\mathbf{E}$ at a fixed frequency (in linear media):

$$\mathbf{J} = \frac{1}{i\omega}\left[\left(\nabla \times \frac{1}{\mu}\nabla\times\right) - \omega^2\varepsilon\right]\mathbf{E} \equiv \hat{O}\mathbf{E}$$

$$(\mathbf{F}, \mathbf{G}) = \int \mathbf{F} \cdot \mathbf{G}\, dV$$

is usually a Hermitian operator under the inner product for vector fields $\mathbf{F}$ and $\mathbf{G}$. (Technically, this unconjugated form is not a true inner product because it is not real-valued for complex-valued fields, but that is not a problem here. In this sense, the operator is not truly Hermitian but is rather complex-symmetric.) This is true whenever the permittivity ε and the magnetic permeability μ, at the given ω, are symmetric 3×3 matrices (symmetric rank-2 tensors) — this includes the common case where they are scalars (for isotropic media), of course. They need not be real—complex values correspond to materials with losses, such as conductors with finite conductivity σ (which is included in ε via $\varepsilon \rightarrow \varepsilon + i\sigma/\omega$)—and because of this the reciprocity theorem does not require time reversal invariance. The condition of symmetric ε and μ matrices is almost always satisfied; see below for an exception.

For any Hermitian operator $\hat{O}$ under an inner product (f, g), we have $(f, \hat{O}g) = (\hat{O}f, g)$ by definition, and the Rayleigh-Carson reciprocity theorem is merely the vectorial version of this statement for this particular operator $\mathbf{J} = \hat{O}\mathbf{E}$: that is, $(\mathbf{E}_1, \hat{O}\mathbf{E}_2) = (\hat{O}\mathbf{E}_1, \mathbf{E}_2)$.

The Hermitian property of the operator here can be derived by integration by parts. For a finite integration volume, the surface terms from this integration by parts yield the more-general surface-integral theorem above. In particular, the key fact is that, for vector fields $\mathbf{F}$ and $\mathbf{G}$, integration by parts (or the divergence theorem) over a volume V enclosed by a surface S gives the identity:

$$\int_V \mathbf{F} \cdot (\nabla \times \mathbf{G})\, dV = \int_V (\nabla \times \mathbf{F}) \cdot \mathbf{G}\, dV - \oint_S (\mathbf{F} \times \mathbf{G}) \cdot \mathbf{dA}$$

This identity is then applied twice to $(\mathbf{E}_1, \hat{O}\mathbf{E}_2)$ to yield $(\hat{O}\mathbf{E}_1, \mathbf{E}_2)$ plus the surface term, giving the Lorentz reciprocity relation.

Surface-term

The cancellation of the surface terms on the right-hand side of the Lorentz reciprocity theorem, for an integration over all space, is not entirely obvious but can be derived in a number of ways.

The simplest argument would be that the fields goes to zero at infinity for a localized source, but this argument fails in the case of lossless media: in the absence of absorption, radiated fields decay inversely with distance, but the surface area of the integral increases with the square of distance, so the two rates balance one another in the integral.

Instead, it is common (e.g. King, 1963) to assume that the medium is homogeneous and isotropic sufficiently far away. In this case, the radiated field asymptotically takes the form of planewaves propagating radially outward (in the $\hat{\mathbf{r}}$ direction) with $\hat{\mathbf{r}} \cdot \mathbf{E} = 0$ and $\mathbf{H} = \hat{\mathbf{r}} \times \mathbf{E}/Z$ where Z is the impedance $\sqrt{\mu/\epsilon}$ of the surrounding medium. Then it follows that $\mathbf{E}_1 \times \mathbf{H}_2 = \mathbf{E}_1 \times \hat{\mathbf{r}} \times \mathbf{E}_2/Z$, which by a simple vector identity equals $\hat{\mathbf{r}}(\mathbf{E}_1 \cdot \mathbf{E}_2)/Z$. Similarly, $\mathbf{E}_2 \times \mathbf{H}_1 = \hat{\mathbf{r}}(\mathbf{E}_2 \cdot \mathbf{E}_1)/Z$ and the two terms cancel one another.

The above argument shows explicitly why the surface terms can cancel, but lacks generality. Alternatively, one can treat the case of lossless surrounding media by taking the limit as the losses (the imaginary part of ε) go to zero. For any nonzero loss, the fields decay exponentially with distance and the surface integral vanishes, regardless of whether the medium is homogeneous. Since the left-hand side of the Lorentz reciprocity theorem vanishes for integration over all space with any non-zero losses, it must also vanish in the limit as the losses go to zero. (Note that we implicitly assumed the standard boundary condition of zero incoming waves from infinity, because otherwise even an infinitesimal loss would eliminate the incoming waves and the limit would not give the lossless solution.)

Green's function

The inverse of the operator $\hat{O}$, i.e. in $\mathbf{E} = \hat{O}^{-1}\mathbf{J}$ (which requires a specification of the boundary conditions at infinity in a lossless system), has the same symmetry as $\hat{O}$ and is essentially a Green's function convolution. So, another perspective on Lorentz reciprocity is that it reflects the fact that convolution with the electromagnetic Green's function is a complex-symmetric (or anti-Hermitian, below) linear operation under the appropriate conditions on ε and μ. More specifically, the Green's function can be written as $G_{nm}(\mathbf{x}', \mathbf{x})$ giving the n-th component of $\mathbf{E}$

at $\mathbf{x}'$ from a point dipole current in the m-th direction at $\mathbf{x}$(essentially, G gives the matrix elements of $\hat{O}^{-1}$), and Rayleigh-Carson reciprocity is equivalent to the statement that $G_{nm}(\mathbf{x}', \mathbf{x}) = G_{mn}(\mathbf{x}, \mathbf{x}')$. Unlike $\hat{O}$, it is not generally possible to give an explicit formula for the Green's function (except in special cases such as homogeneous media), but it is routinely computed by numerical methods.

One case in which ε is not a symmetric matrix is for magneto-optic materials, in which case the usual statement of Lorentz reciprocity does not hold (see below for a generalization, however). If we allow magneto-optic materials, but restrict ourselves to the situation where material absorption is negligible, then ε and μ are in general 3×3 complex Hermitian matrices. In this case the operator

$$\nabla \times \frac{1}{\mu}\nabla \times -(\omega^2/c^2)\varepsilon$$

is Hermitian under the conjugated inner product $(\mathbf{F}, \mathbf{G}) = \int \mathbf{F}^* \cdot \mathbf{G}\, dV$, and a variant of the reciprocity theorem still holds:

$$-\int_V [\mathbf{J}_1^* \cdot \mathbf{E}_2 + \mathbf{E}_1^* \cdot \mathbf{J}_2]\, dV = \oint_S [\mathbf{E}_1^* \times \mathbf{H}_2 + \mathbf{E}_2 \times \mathbf{H}_1^*] \cdot \mathbf{dA}$$

where the sign changes come from the 1 / iω in the equation above, which makes the operator $\hat{O}$anti-Hermitian (neglecting surface terms). For the special case of $\mathbf{J}_1 = \mathbf{J}_2$, this gives a re-statement of conservation of energy or Poynting's theorem (since here we have assumed lossless materials, unlike above): the time-average rate of work done by the current (given by the real part of $-\mathbf{J}^* \cdot \mathbf{E}$) is equal to the time-average outward flux of power (the integral of the Poynting vector). By the same token, however, the surface terms do not in general vanish if one integrates over all space for this reciprocity variant, so a Rayleigh-Carson form does not hold without additional assumptions.

The fact that magneto-optic materials break Rayleigh-Carson reciprocity is the key to devices such as Faraday isolators and circulators. A current on one side of a Faraday isolator produces a field on the other side but not vice-versa.

For a combination of lossy and magneto-optic materials, and in general when the ε and μ tensors are neither symmetric nor Hermitian matrices, one can still obtain a generalized version of Lorentz reciprocity by considering $(\mathbf{J}_1, \mathbf{E}_1)$ and $(\mathbf{J}_2, \mathbf{E}_2)$ to exist in different systems.

In particular, if $(\mathbf{J}_1, \mathbf{E}_1)$ satisfy Maxwell's equations at ω for a system with materials (ε_1, μ_1), and $(\mathbf{J}_2, \mathbf{E}_2)$ satisfy Maxwell's equations at ω for a system with materials $(\varepsilon_1^T, \mu_1^T)$, where T denotes the transpose, then the equation of Lorentz reciprocity holds.

Exceptions

For nonlinear media, no reciprocity theorem generally holds. Reciprocity also does not generally apply for time-varying ("active") media; for example, when ε is modulated in time by some external process. (In both of these cases, the frequency ω is not generally a conserved quantity.)

Feld-Tai reciprocity

A closely related reciprocity theorem was articulated independently by Y. A. Feld and C. T. Tai in 1992 and is known as **Feld-Tai reciprocity** or the **Feld-Tai lemma**. It relates two time-harmonic localized current sources and the resulting magnetic fields:

$$\int \mathbf{J}_1 \cdot \mathbf{H}_2 dV = \int \mathbf{H}_1 \cdot \mathbf{J}_2 dV.$$

However, the Feld-Tai lemma is only valid under much more restrictive conditions than Lorentz reciprocity. It generally requires time-invariant linear media with isotropic homogeneous impedance, i.e. a constant scalar μ/ε ratio, with the possible exception of regions of perfectly conducting material.

Green's reciprocity

Whereas the above reciprocity theorems were for oscillating fields, Green's reciprocity is an analogous theorem for electrostatics with a fixed distribution of electric charge (Panofsky and Phillips, 1962).

In particular, let φ_1 denote the electric potential resulting from a total charge density ρ_1. The electric potential satisfies Poisson's equation, $-\nabla^2\phi_1 = \rho_1/\varepsilon_0$, where $\varepsilon_0$is the vacuum permittivity. Similarly, let φ_2 denote the electric potential resulting from a total charge density ρ_2, satisfying $-\nabla^2\phi_2 = \rho_2/\varepsilon_0$. In both cases, we assume that the charge distributions are localized, so that the potentials can be chosen to go to zero at infinity. Then, Green's reciprocity theorem states that, for integrals over all space:

$$\int \rho_1\phi_2 dV = \int \rho_2\phi_1 dV.$$

This theorem is easily proven from Green's second identity. Equivalently, it is the statement that

$\int \phi_2(\nabla^2\phi_1)dV = \int \phi_1(\nabla^2\phi_2)dV$, i.e. that $\nabla^2$is a Hermitian operator (as follows by integrating by parts twice).

4.0 CONCLUSION

In section 2 unit 3 we took a look at Millman's Theorem, Norton's Theorem, Ohm's Law and Reciprocity. We learnt that the parallel generator theorem is the same as Millman's theorem and that it is a method to simplify the solution of a circuit made up of only branches.

We saw that any collection of voltage sources, current sources, and resistors with two terminals is electrically equivalent to an ideal current source, in parallel with a single resistor through Norton's theorem.

We were thrilled to re-discover Ohm's law and the underlying concepts behind its derivation – also we saw the various forms of Ohms law as well as the Ohms law triangle which graphically illustrates the relationship between voltage, current and resistance. We were also intimated with the concept of Ohmic and non Ohmic devices and elements – which present linear and non-linear V-I curves respectively.
We concluded this Unit with Reciprocity which interestingly refers to a variety of theorems involving the interchange of time-harmonic electric current densities. We saw the application of this in the illustration that a good transmitter is also a good receiver and a transmitting antenna

designed for optimum performance at a given frequency will also perform the reciprocal function of an optimum receiver at that same frequency.

5.0 SUMMARY

- Millman's theorem also called the parallel generator theorem simplifies the solution of a circuit made up of only branches.

- Onm's and Kirchhoff's laws serve as the basis upon which Millman's theorem is derived.

- The total equivalent conductance of a supernode is the sum of the conductance of each branch according to Millman's theorem.

- Norton's theorem which is an extension of thevenin's theorem states Norton theorem, states that any collection of voltage sources, current sources, and resistors with two terminals is electrically equivalent to an ideal current source in parallel with a single resistor.

- Norton's theorem is applicable not only to Direct Current circuit analysis, but also to single-frequency Alternating Current systems and it can also be applied to general impedances and not just resistors.

- Norton equivalent can be used to represent any network of linear sources and impedances, at a given frequency.

- Ohms law relates the voltage across a network element to the current flowing through it by the expression V -= I R

- In the Drude model, electrons are treated like pinballs creating average momentum and current density proportional to drift velocity. The current density becomes proportional to the applied electric field which leads to Ohm's law

- Ohm's law is applicable to resistive circuits for both Direct and Alternating current and is applicable to reactive circuits for alternating current where impedance substitutes resistance.
- Ohm's law is linear and circuit elements that do not have a linear Voltage – Current curve are referred to as non-Ohmic

- Electrical conduction in all materials is affected by temperature and Ohm's law applies when temperature is constant.

- Reciprocity refers to a variety of theorems involving the interchange of time-harmonic electric current densities and the resulting electromagnetic fields in Maxwell's equations for time-invariant linear media under certain constraints.

- Reciprocity is closely related to the concept of Hermitian operators from linear algebra, applied to electromagnetism.

6.0 TUTOR MARKED ASSIGNMENTS

1. State Millman's theorem.

2. What is the relationship between Kirchhoff's law and Millman's theorem?

3. State Norton's theorem

4. What relationship exists between Thevenin's theorem and Norton's theorem?

5. State Ohm's law and mention three forms of Ohm's law.

6. What is an Ohm's law triangle and how does it function

7. What concept is Reciprocity closely related to in linear algebra?

8. Millman's theorem can be used to simplify the solution to circuits on one condition. What is that condition?

9. State those things which you must do to find a Norton's equivalent.

10. Describe the Drude model.

11. How does temperature affect Ohm's law?

12. Can you discuss four examples of Reciprocity which we see around us every day?

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

Analog Filter Design
By M. E. Van Valkenburg Published by Holt, Rinehart and Winston

UNIT 4 SUPERPOSITION THEOREM, TELLEGEN'S THEOREM, THEVENIN'S THEOREM, STAR – DELTA TRANSFORMATION

CONTENTS

1.0 INTRODUCTION

In this Unit I am going to acquaint you with the Superposition theorem which aggregates the contribution of individual sources to the overall voltage or current response of any branch of a bilateral linear network. After this, we shall visit Tellegen's and Thevenin's Theorems which are two very powerful theorems for interrogating electrical networks by reducing networks to easily digestible equivalents – incidentally both are very closely related to Kirchhoff's laws. We shall round up this Unit by learning how to transform a star to a delta network and vice versa.

You are expected to pay close attention; particularly to the derivation of equations and transforms as well as their application to problem solving as they are of great practical significance in the real world.

2.0 OBJECTIVES

After reading through this unit, you should be able to

1 State the Superposition Theorem
2 Apply superposition in converting circuits into Norton or Thevenin's equivalent
3 State Tellegen's Theorem

4 Establish the relationship between Tellegen's theorem and Kirchhoff's Laws
5 State Thevenin's Theorem
6 Derive the Thevenin's equivalent circuit for a Black Box Circuit
7 Know when not to apply the Thevenin's Equivalent method
8 Understand the procedure for Star-Delta Transformation
9 Apply Star Delta transformations to passive linear networks

3.0 MAIN CONTENT

3.1 Superposition Theorem

The superposition theorem for electrical circuits states that the response (Voltage or Current) in any branch of a bilateral linear circuit having more than one independent source equals the algebraic sum of the responses caused by each independent source acting alone, while all other independent sources are replaced by their internal impedances.

To ascertain the contribution of each individual source, all of the other sources first must be "turned off" (set to zero) by:

Replacing all other independent voltage sources with a short circuit (thereby eliminating difference of potential. i.e. V=0, internal impedance of ideal voltage source is zero (short circuit)).

Replacing all other independent current sources with an open circuit (thereby eliminating current. i.e. I=0, internal impedance of ideal current source is infinite (open circuit).

This procedure is followed for each source in turn, then the resultant responses are added to determine the true operation of the circuit. The resultant circuit operation is the superposition of the various voltage and current sources.

The superposition theorem is very important in circuit analysis. It is used in converting any circuit into its Norton equivalent or Thevenin equivalent.

Superposition theorem is applicable to linear networks (time varying or time invariant) consisting of independent sources, linear dependent sources, linear passive elements Resistors, Inductors, Capacitors and linear transformers.

3.2 Tellegen's Theorem

Tellegen's theorem is one of the most powerful theorems in network theory. Most of the energy distribution theorems and extremum principles in network theory can be derived from it. It was published in 1952 by Bernard Tellegen. Fundamentally, Tellegen's theorem gives a simple relation between magnitudes that satisfy the Kirchhoff's laws of electrical circuit theory.

The Tellegen theorem is applicable to a multitude of network systems. The basic assumptions for the systems are the conservation of flow of extensive quantities (Kirchhoff's current law, KCL) and the uniqueness of the potentials at the network nodes (Kirchhoff's voltage law, KVL). The Tellegen theorem provides a useful tool to analyze complex network systems among them electrical circuits, biological and metabolic networks, pipeline flow networks, and chemical process networks.

Consider an arbitrary lumped network whose graph G has b branches and n_t nodes. In an electrical network, the branches are two-terminal components and the nodes are points of interconnection. Suppose that to each branch of the graph we assign arbitrarily a branch potential difference W_k and a branch current F_k for $k = 1, 2, \ldots, b$, and suppose that they are measured with respect to arbitrarily picked associated reference directions. If the branch potential differences $W_1, W_2, \ldots, W_b$ satisfy all the constraints imposed by KVL and if the branch currents $F_1, F_2, \ldots, F_b$ satisfy all the constraints imposed by KCL, then

$$\sum_{k=1}^{b} W_k F_k = 0.$$

Tellegen's theorem is extremely general; it is valid for any lumped network that contains any elements, linear or nonlinear, passive or active, time-varying or time-invariant. The generality is extended when W_k and F_k are linear operations on the set of potential differences and on the set of branch currents (respectively) since linear operations don't affect KVL and KCL. For instance, the linear operation may be the average or the Laplace transform. Another extension is when the set of potential differences W_k is from one network and the set of currents F_k is from an entirely different network, so long as the two networks have the same topology (same incidence matrix). This extension of Tellegen's Theorem leads to many theorems relating to two-port networks.

Definitions

Incidence matrix: The $n_t \times n_f$ matrix $\mathbf{A_a}$ is called node-to-branch incidence matrix for the matrix elements a_{ij} being

$$a_{ij} = \begin{cases} 1, & \text{if flow } j \text{ leaves node } i \\ -1, & \text{if flow } j \text{ enters node } i \\ 0, & \text{if flow } j \text{ is not incident with node } i \end{cases}$$

A reference or datum node P_0 is introduced to represent the environment and connected to all dynamic nodes and terminals.

$(n_t - 1) \times n_f$

The matrix $\mathbf{A}$ above, where the row that contains the elements a_{0j} of the reference node P_0 is eliminated is called reduced incidence matrix.

The conservation laws (KCL) in vector-matrix form:

$$\mathbf{AF} = \mathbf{0}$$

The uniqueness condition for the potentials (KVL) in vector-matrix form:

$$\mathbf{W} = \mathbf{A^T w}$$

where w_k are the absolute potentials at the nodes to the reference node P_0.

Using KVL:

$$\mathbf{W^T F} = (\mathbf{A^T w})^\mathbf{T} \mathbf{F} = (\mathbf{w^T A})\mathbf{F} = \mathbf{w^T AF} = \mathbf{0}$$

because $\mathbf{AF} = \mathbf{0}$ by KCL. So:

$$\sum_{k=1}^{b} W_k F_k = \mathbf{W^T F} = 0$$

Applications

Network analogs have been constructed for a wide variety of physical systems, and have proven extremely useful in analyzing their dynamic behaviour. The classical application area for network theory and Tellegen's theorem is electrical circuit theory. It is mainly in use to design filters in signal processing applications.

A more recent application of Tellegen's theorem is in the area of chemical and biological processes. The assumptions for electrical circuits (Kirchhoff laws) are generalized for dynamic systems obeying the laws of irreversible thermodynamics. Topology and structure of reaction networks (reaction mechanisms, metabolic networks) can be analyzed using the Tellegen theorem.

A formulation for Tellegen's theorem of process systems: Another application of Tellegen's theorem is to determine stability and optimality of complex process systems such as chemical plants or oil production systems. The Tellegen theorem can be formulated for process systems using process nodes, terminals, flow connections and allowing sinks and sources for production or destruction of extensive quantities.

$$\sum_{j=1}^{np} W_j \frac{\mathrm{d}Z_j}{\mathrm{d}t} = \sum_{k=1}^{nf} W_k f_k + \sum_{j=1}^{np} w_j p_j + \sum_{j=1}^{n_t} w_j t_j, \quad j = 1, \ldots, n_p + n_t$$

where p_j are the production terms, t_j are the terminal connections.

3.3 Thevenin’s Theorem

Thevenin’s theorem for linear electrical networks states that any combination of voltage sources, current sources, and resistors with two terminals is electrically equivalent to a single voltage source V and a single series resistor R. For single frequency AC systems the theorem can also be applied to general impedances, not just resistors. The theorem was first discovered by German scientist Hermann von Helmholtz in 1853, but was then rediscovered in 1883 by French telegraph engineer Léon Charles Thevenin (1857–1926).

This theorem states that a circuit of voltage sources and resistors can be converted into a Thevenin equivalent, which is a simplification technique used in circuit analysis. The Thevenin equivalent can be used as a good model for a power supply or battery (with the resistor representing the internal impedance and the source representing the electromotive force). The circuit consists of an ideal voltage source in series with an ideal resistor.

Any black box containing only voltage sources, current sources, and other resistors can be converted to a Thevenin equivalent circuit, comprising exactly one voltage source and one resistor.

The Thevenin's equivalent

To calculate the equivalent circuit, the resistance and voltage are needed, so two equations are required. These two equations are usually obtained by using the following steps, but any conditions placed on the terminals of the circuit should also work:

Calculate the output voltage, V_{AB}, when in open circuit condition (no load resistor—meaning infinite resistance). This is V_{Th}.

(1) Calculate the output current, I_{AB}, when the output terminals are short circuited (load resistance is 0). R_{Th} equals V_{Th} divided by this I_{AB}.

The equivalent circuit is a voltage source with voltage V_{Th} in series with a resistance R_{Th}.

(2) Replace voltage sources with short circuits, and current sources with open circuits.

Calculate the resistance between terminals A and B. This is R_{Th}.

The Thevenin's-equivalent voltage is the voltage at the output terminals of the original circuit. When calculating a Thevenin's-equivalent voltage, the voltage divider principle is often useful, by declaring one terminal to be V_{out} and the other terminal to be at the ground point.

The Thevenin's-equivalent resistance is the resistance measured across points A and B "looking back" into the circuit. It is important to first replace all voltage- and current-sources with their internal resistances. For an ideal voltage source, this means replace the voltage source with a short circuit. For an ideal current source, this means replace the current source with an open circuit. Resistance can then be calculated across the terminals using the formulae for series and parallel circuits. This method is valid only for circuits with independent sources. If there are dependent sources in the circuit, another method must be used such as connecting a test source across A and B and calculating the voltage (notice that R_1 is not taken into consideration, as above calculations are done in an open circuit condition between A and B, therefore no current flows through this part which means there is no current through R_1 and therefore no voltage drop along this part)

Limitations

Many, if not most circuits are only linear over a certain range of values, thus the Thevenin's equivalent is valid only within this linear range and may not be valid outside the range.

The Thevenin equivalent has an equivalent I-V characteristic only from the point of view of the load.

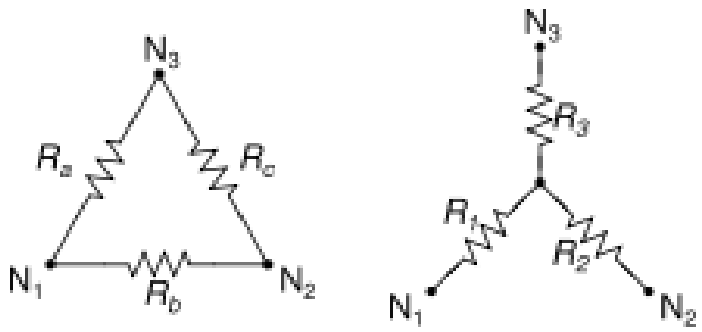
The power dissipation of the Thevenin equivalent is not necessarily identical to the power dissipation of the real system. However, the power dissipated by an external resistor between the two output terminals is the same however the internal circuit is represented.

3.4 Star – Delta Transformation

The Y-Δ transform, also referred to as Y-delta, Wye-delta, Kennelly's delta-star transformation, star-mesh transformation, T-Π or T-pi transform is a mathematical technique to simplify the analysis of an electrical network.

The transformation is used to establish equivalence for networks with three terminals. Where three elements terminate at a common node and none are sources, the node is eliminated by transforming the impedances. For equivalence, the impedance between any pair of terminals must be the same for both networks. The equations given here are valid for complex as well as real impedances.

Basic Y-Δ transformation



Look at the diagram above and you will see a Delta (**Δ)** circuit on the left and a star network (**Y)** on the right. Can you see why they are called delta and Star networks?

Equations for the transformation from Δ-load to Y-load 3-phase circuit

Our objective is to compute the impedance R_y at a terminal node of the Y circuit with impedances R', R'' to adjacent nodes in the Δ circuit by

$$R_y = \frac{R'R''}{\sum R_\Delta}$$

where R_Δ are all impedances in the Δ circuit. This yields the specific formulae

$$R_1 = \frac{R_a R_b}{R_a + R_b + R_c},$$

$$R_3 = \frac{R_a R_c}{R_a + R_b + R_c}.$$

$$R_3 = \frac{R_a R_c}{R_a + R_b + R_c}.$$

Equations for the transformation from Y-load to Δ-load 3-phase circuit

As a general rule, we compute an impedance R_Δ in the Δ circuit by

$$R_\Delta = \frac{R_P}{R_{\text{opposite}}}$$

where $R_P = R_1R_2 + R_2R_3 + R_3R_1$ is the sum of the products of all pairs of impedances in the Y circuit and R_{opposite} is the impedance of the node in the Y circuit which is opposite the edge with R_Δ. The formula for the individual edges are thus

$$R_a = \frac{R_1R_2 + R_2R_3 + R_3R_1}{R_2},$$

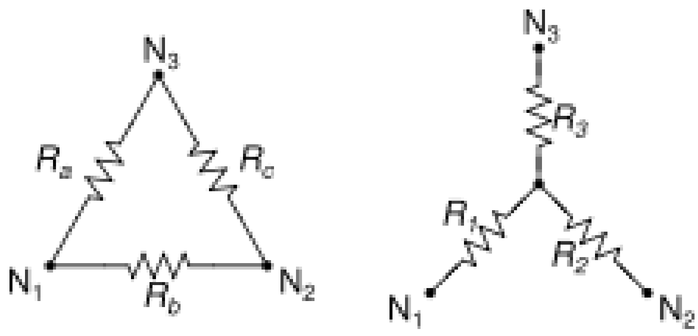
$$R_b = \frac{R_1R_2 + R_2R_3 + R_3R_1}{R_3},$$

$$R_c = \frac{R_1R_2 + R_2R_3 + R_3R_1}{R_1}.$$

Graph theory

In graph theory, the Y-Δ transform means replacing a Y sub graph of a graph with the equivalent Δ subgraph. The transform preserves the number of edges in a graph, but not the number of vertices or the number of cycles. Two graphs are said to be **Y-Δ equivalent** if one can be obtained from the other by a series of Y-Δ transforms in either direction. For example, the Petersen graph family is a Y-Δ equivalence class.

Δ-load to Y-load transformation equations



These sketches above are Δ and Y circuits with the labels that are used in this Unit.

To relate $\{R_a, R_b, R_c\}$ from Δ to $\{R_1, R_2, R_3\}$ from Y, the impedance between two corresponding nodes is compared. The impedance in either configuration is determined as if one of the nodes is disconnected from the circuit.

The impedance between N_1 and N_2 with N_3 disconnected in Δ:

$$R_\Delta (N1, N2) = Rb \,||(Ra + Rc)$$

$$= \frac{R_b(R_a + R_c)}{R_a + R_b + R_c}$$

To simplify, let R_T be the sum of $\{R_a, R_b, R_c\}$.

$$R_T = R_a + R_b + R_c$$

Thus,

$$R_\Delta (N_1, N_2) = \frac{R_b(R_a + R_c)}{R_T}$$

The corresponding impedance between N_1 and N_2 in Y is simple:

$$R_Y(N_1, N_2) = R_1 + R_2$$

hence:

$$R_1 + R_2 = \frac{R_b(R_a + R_c)}{R_T} \qquad 1$$

Repeating for $R(N_2, N_3)$:

$$R_2 + R_3 = \frac{R_c(R_a + R_b)}{R_T} \qquad 2$$

and for $R(N_1, N_3)$:

$$R_1 + R_3 = \frac{R_a(R_b + R_c)}{R_T} \qquad 3$$

From here, the values of $\{R_1, R_2, R_3\}$ can be determined by linear combination (addition and/or subtraction).

For example, adding (1) and (3), then subtracting (2) yields

$$R_1 + R_2 + R_1 + R_3 - R_2 - R_3 = \frac{R_b(R_a + R_c)}{R_T} + \frac{R_a(R_b + R_c)}{R_T} - \frac{R_c(R_a + R_b)}{R_T}$$

$$2R_1 = \frac{2R_b R_a}{R_T}$$

thus,

$$R_1 = \frac{R_b R_a}{R_T}$$

where $R_T = R_a + R_b + R_c$

For completeness:

$$R_1 = \frac{R_b R_a}{R_T} \qquad 4$$

$$R_2 = \frac{R_b R_c}{R_T} \qquad 5$$

$$R_3 = \frac{R_a R_c}{R_T} \qquad 6$$

Y-load to Δ-load transformation equations

Let

$$R_T = R_a + R_b + R_c.$$

We can write the Δ to Y equations as

$$R_1 = \frac{R_a R_b}{R_T} \quad 1$$

$$R_2 = \frac{R_b R_c}{R_T} \quad 2$$

$$R_3 = \frac{R_a R_c}{R_T} \quad 3$$

Multiplying the pairs of equations yields

$$R_1 R_2 = \frac{R_a R_b^2 R_c}{R_T^2} \quad 4$$

$$R_1 R_3 = \frac{R_a^2 R_b R_c}{R_T^2} \quad 5$$

$$R_2 R_3 = \frac{R_a R_b R_c^2}{R_T^2} \quad 6$$

and the sum of these equations is

$$R_1 R_2 + R_1 R_3 + R_2 R_3 = \frac{R_a R_b^2 R_c + R_a^2 R_b R_c + R_a R_b R_c^2}{R_T^2} \quad 7$$

When we factor $R_aR_bR_c$ from the right side, leaving R_T in the numerator, canceling with an R_T in the denominator.

$$R_1 R_2 + R_1 R_3 + R_2 R_3 = \frac{(R_a R_b R_c)(R_a + R_b + R_c)}{R_T^2}$$

$$R_1 R_2 + R_1 R_3 + R_2 R_3 = \frac{R_a R_b R_c}{R_T} \qquad 8$$

You are to note the similarity between (8) and {(1),(2),(3)}

Divide (8) by (1)

$$\frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_1} = \frac{R_a R_b}{R_T} \cdot \frac{R_c R_T}{R_a R_b}$$

$$\frac{R_1 R_2 + R_1 R_3 + R_2 R_3}{R_1} = R_c$$

which is the equation for R_c. Dividing (8) by R_2 or R_3 gives the other equations.

4.0 CONCLUSION

In this Unit you have been able to cover the last four of the circuit the circuit theorems namely; Superposition Theorem, Tellegen's Theorem, Thevenin's Theorem and the very interesting Star – Delta Transformation

Superposition theorem showed you how to aggregate the contribution of individual sources to the overall voltage or current response of any branch of a bilateral linear network by explaining the steps you must take. It also taught you how to convert any circuit into its Norton equivalent or Thevenin equivalent.

,

You learnt that most energy distribution theorems and extremum principles in network theory can be derived from Tellegen's theory which makes it one of the most powerful theorems in network theory.

With Thevenin's theorem firmly in your grasp, you can now combine voltage sources, current sources, and resistors with two terminals and find out the electrically equivalent circuit which comprise a single voltage source and a single series resistor. You also now know that for single frequency Alternating Current systems, Thevenin's theorem can be applied to general impedances and not just resistors.

Finally, you learnt how to convert a star network into a delta equivalent and vice versa. You know the transformation equations, their derivation and the mnemonics which make you remember them easily.

5.0 SUMMARY

- Superposition theorem for electrical circuits states that the response in any branch of a bilateral linear circuit having more than one independent source equals the algebraic sum of the responses caused by each independent source acting alone, while all other independent sources are replaced by their internal impedances.

- To ascertain the contribution of each individual source with the superposition theorem, we must first turn off all voltage and current sources by replacing independent voltage sources with a short circuit and replacing all other independent current sources with an open circuit.

- Superposition theorem in circuit analysis is used in converting any circuit into its Norton equivalent or Thevenin equivalent and is applicable to linear networks consisting of independent sources, linear dependent sources, linear passive elements Resistors, Inductors, Capacitors and linear transformers.

- Tellegen's theorem is one of the most powerful theorems in network theory and most of the energy distribution theorems and extremum principles in network theory can be derived from it. It is closely related to Kirchhoff’s circuit laws

- Very few circuit theorems are as general as Tellegen's theorem which is valid for any lumped network that contains any linear or nonlinear, passive or active, time-varying or time-invariant elements.

- Thevenin’s equivalent which is a single source and a single impedance simulation of a more complex network of elements and sources derives fro Thevenin’s theorem. It is simple to apply yet it is powerful.

- Star to Delta transformation and vice versa establishes a set of transform equations which facilitate the transformation of composite three terminal networks. There are short cut methods for remembering the transform equations.

6.0 TUTOR MARKED ASSIGNMENTS

1. State the Superposition Theorem.

2. State Tellegen's Theorem.

3. State Thevenin's Theorem.

4. Derive the transform equations for a Star to Delta Transformation,

5. List the steps in the process of using Superposition theorem.

6. What is a node-to-branch incidence matrix? Explain in detail.

7. What limitation does Thevenin's equivalent circuit possess?

8. Explain the relevance of graph theory to Star to Delta transformation

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

Analog Filter Design
By M. E. Van Valkenburg Published by Holt, Rinehart and Winston

UNIT 1 VACUUM TUBES

CONTENTS

1.0 INTRODUCTION

Let us go back into history; to be precise; 1883. That is when Thomas Edison discovered that electric current flowed from the hot filament to a metal plate at the bottom of a light bulb. This discovery was called the "Edison effect" at that time and it showed that electrical current did not need a physical conductor, and that indeed, it was possible to make current flow through a vacuum.

As you might expect, this discovery that current can travel through a vacuum at that time appeared illogical to a number of people and was not put into practical use until 1904 when a British scientist named John A. Fleming made a vacuum tube called the diode. This acted as a valve because it forced current in the tube to travel in one direction. This unidirectional flow is very important in turning alternating current into direct current.

The vacuum tube underwent a major transformation when Lee De Forest invented a vacuum device which not only forced current to move in a single direction, but also increased the current as it passed through the vacuum tube. Lee De Forest placed a metal grid in the middle of the vacuum tube and by using a small input current to change the voltage on the grid, Lee De Forest could control the flow of a more powerful current, through the tube. The strength of two currents was not necessarily related as a weak current might be applied to the tube's grid, but a much amplified current result at the main electrodes of the tube.

Turning weak currents into strong currents was crucial for a number of new technologies at that time and Bell Laboratories made instrumental in the development of this technology which found applications in everything from hearing aids to radios to televisions.

Vacuum tubes are constructed of a glass tube surrounding a vacuum with embedded electrical contacts at the ends; Vacuum tubes are also referred to as electron tubes and thermionic valve. They are devices used to amplify, to switch otherwise modify, or create an electrical signal controlling the movement of electrons a low-pressure space.

If I told you that some vacuum tubes are filled with gas, you would argue that they were no longer “vacuum” tubes – an apparent misnomer, but some special function vacuum tubes are actually filled with low-pressure gas and are called soft tubes as distinct from the hard vacuum type which have the internal gas pressure reduced as much as physically possible and as a rule, virtually all vacuum tubes depend on the thermionic emission of electrons. You should take note of the facts above as we will revisit them later in your assignments.

2.0 OBJECTIVES

After reading through this unit, you should be able to

1. Describe Vacuum Tubes
2. Explain the historical progression of vacuum tube development
3. Understand The Functional Sub Units Of The Vacuum Tube
4. Distinguish Between Diode and Triode Vacuum Tubes
5. Explain the Action of the Screen Grid
6. Justify the Construction of Multi Section Vacuum Tube Design

7 Describe the Functioning Of Different Special; Purpose Vacuum Tube
8 Explain the Functional Similarities between the Vacuum Tube Triode and the Bipolar Transistor
9 Compare and Contrast Vacuum Tube Technology with Semiconductor Technology

3.0 MAIN CONTENT

3.1 Description Of Vacuum Tubes

Vacuum tubes have been critical to the development of electronic technology, which has driven the expansion and commercialization of radio broadcasting, television, radar, sound reproduction, huge telephone networks, analog and digital computing, and industrial process control. Some of these applications pre-dated electronics, but the vacuum tube made them widespread and practical.

While for most purposes, the vacuum tube has been replaced by solid state devices such as transistors and solid-state diodes which last much longer are smaller, more efficient, more reliable, and cheaper than equivalent vacuum tube devices. Vacuum tubes still find wide applications in specialized areas such as in high-power radio frequency transmissions; distinct sound signature, and as cathode ray tubes in oscilloscopes. A specialized form of the electron tube, the magnetron, is the primary source of microwave energy in microwave ovens and radar systems. The klystron, which is a powerful but narrow-band radio-frequency amplifier, is commonly deployed by broadcasters as high-power UHF television transmitters.

Vacuum tubes consist of electrodes in a vacuum all within in an insulating heat-resistant envelope. Because this envelope is often in the shape of a tube, early references to this device referred to it as vacuum tube

Many vacuum tubes have glass envelopes, though some types such as power tubes may be constructed with ceramic or metal envelopes. The electrodes are attached to leads which pass through the envelope via airtight seal. Most tubes, almost without exception feature socket pins through which the leads are designed to plug into tube sockets for easy placement and replacement.

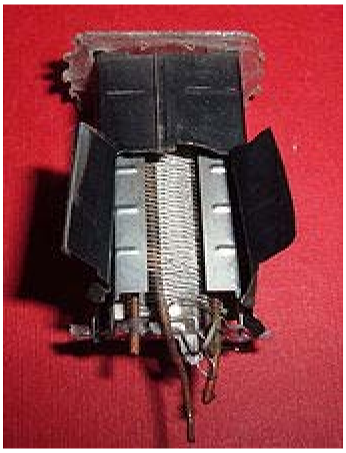
The simplest vacuum tubes have a filament called the cathode which is housed an evacuated glass envelope. When hot, the filament through thermionic emission releases electrons into the vacuum. These electrons result is a negatively charged electron cloud called space charge. The space charge electrons are drawn to a metal plate anode inside the envelope when the anode is positively charged relative to the filament. This translates to a net flow of electrons from the cathode to the anode. This implies that conventional current flows from the anode to the cathode.

Vacuum tubes need a considerable temperature differential between the hot cathode and the cold anode. Because of this, vacuum tubes are relatively power-inefficient as the heating of the filament consumes energy. If the vacuum tube is encased within a heat-retaining envelope of insulation, then the entire tube would reach the same temperature. This results in electron emission from the anode that would counter the normal one-way current. Because the tube requires a vacuum to operate, convection cooling of the anode is not generally possible unless the anode forms a part of the vacuum envelope.

Anode cooling occurs in most tubes through black body radiation and conduction of heat to the outer glass envelope via the anode mounting frame. In cold cathode tubes some form of gas discharge underlies the operation as they do not rely on thermionic emission at the cathode. They are usually applicable in lighting such as neon bulbs, and in voltage regulation.

If a control grid, is added between the cathode and the anode the vacuum tube is called a triode because it now has three electrodes. A triode is voltage-controlled in that a voltage applied as an input to the grid can be used to control the flow of electrons between cathode and anode. The relationship between this input voltage and the output current is determined by transconductance. Control grid current is practically negligible in most circuits.

3.2 History And Development Of Vacuum Tubes



Let us take a look at the picture above, It shows us a vacuum tube with its plate cut open so that we can see the inside of the tube.

Let us step back into history again. In the 19th century, scientists experimented with such tubes mostly for specialized scientific applications or novelties with the exception of the light bulb, and the foundation laid by 19^{th} century scientists and inventors was critical to the development of vacuum tube technology.

Though reported earlier, thermionic emission is more often credited to Thomas Edison as he patented his Edison effect even though he did not understand the underlying physics, or the potential value of the discovery. It wasn't until the early 20th century that this effect was put to use by John Ambrose Fleming and Lee De Forest as the diode and the triode respectively.

The development of the thermionic diode and the triode led to great improvements in telecommunications technology, particularly the birth of broadcast radio.

Diodes and triodes

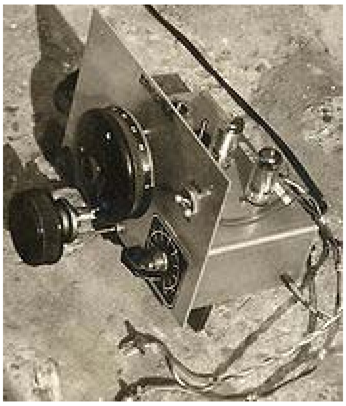
The physicist John Ambrose Fleming who worked as an engineering consultant for such communications firms as Edison Telephone and the Marconi Company in 1904 developed the Fleming valve which was used as a rectifier for alternating current and as a radio wave detector. All this was as a consequence of experiments he conducted on Edison effect bulbs imported from the USA.

Within two years Robert Von Lieben filed for a patent on a three electrode vacuum tube capable of amplification but it was Lee De Forest who in 1907 placed a bent wire to serve as a screen between the filament and anode that brought the triode to the forefront in radio communications application.

Lee De Forest cautioned against operation which might cause the vacuum to become too hard however in 1915, Langmuir was one of the first scientists to realize that a harder vacuum would improve the amplifying characteristics of the triode.

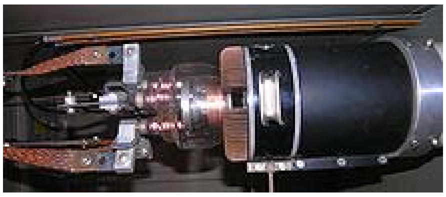
The non-linear characteristic of the triode caused harmonic distortions at low volumes in early vacuum tube audio amplifiers. This non-linearity is remedied by applying a grid bias negative voltage.

Tetrodes and pentodes



Take a look at this relic in the picture. It is a very old vacuum tube radio – one of the earliest which uses only two tubes.

Early application of triodes in radio transmission and reception was plagued by uncontrollable oscillations which resulted from parasitic anode-to-grid capacitance. Several efforts to reduce these parasitic capacitances were unsatisfactory over a broad spectrum of frequencies until it was discovered that the addition of a second grid, located between the control grid and the plate solved the problem. This grid is called the screen grid and a positive voltage slightly lower than the plate voltage applied to it completely eliminating the oscillation problem. A consequence of the screen grid is that the Miller capacitance is reduced with improvement in gain at high frequency. This two-grid tube is called a Tetrode because it has four active electrodes.



This picture shows you a high-power vacuum tube for a specialised application.

In all vacuum tubes, electrons strike the anode hard enough to knock out secondary electrons and in triodes these electrons being less energetic cannot reach the grid or cathode – they are re-captured by the anode. However in a tetrode, they can be captured by the second grid, reducing the plate current and the amplification of the circuit. Since secondary electrons can outnumber the primary electrons, when anode voltage falls below the screen voltage, the valve can exhibit negative-resistance known as the tetrode kink. This can also overload the screen grid and can cause it to overheat and melt, destroying the tube.

This problem is solved by introducing the suppressor grid biased at either ground or cathode voltage. Its negative voltage relative to the anode voltage electrostatically suppresses the secondary electrons by repelling them back toward the anode. This is the five electrode vacuum tube which is also known as the Pentode. The pentode was invented in 1928 by Bernard Tellegen.

3.3 Improvements In Vacuum Tubes

Early vacuum tubes resembled incandescent light bulbs and were made by lamp manufacturers, who had equipment to manufacture glass envelopes and powerful vacuum pumps required to evacuate the enclosures. Later, specialized manufacturers using more economical construction methods were set up to fill growing demand for broadcast receivers and bare tungsten filaments operated at a temperature of around 2200 ^{o}C.

The development of oxide-coated filaments reduced filament temperatures to around 700 ^{o}C which in turn reduced the thermal distortion of the tube structure and allowed closer spacing of tube elements. This improved tube gain, since the gain of a triode is inversely proportional to the spacing between grid and cathode. Development of the indirectly-heated cathode, with the filament inside a cylinder of oxide-coated nickel, further reduced distortion of the tube elements and also allowed the cathode heaters to be run from an AC supply without the super imposition of the mains signal on the output.

Heat Transfer and Appearance of Tubes

Take a look at this power tube in the picture. Do you know that the anode of this transmitting triode has been designed to dissipate up to 500W of heat? Let us learn more about heat generation and heating in tubes.

Since considerable amount of heat is produced when vacuum tubes operate, they are typically about 30-60% efficient which means that 40-70 % of input power to an amplifier stage is lost as heat. The requirements for heat removal significantly changed the appearance of high-power vacuum tubes.

Most tubes contain two sources of heat when operating. The first one of these is the filament or heater. While some vacuum tubes contain directly heated cathode others employ the indirectly heated cathode. This usually consists of a nickel tube, coated on the outside with the same strontium, calcium, barium oxide mix used in directly heated filaments and fitted with a tungsten filament inside the tube to heat it. This tungsten filament is usually uncoiled and coated in a layer of alumina to insulate it from the nickel tube of the actual cathode. This form of construction allows for a much greater electron emitting area and, because the heater is insulated from the cathode, the cathode can be positioned in a circuit at up to 150 volts more positive than the heater or 50 volts more negative than the heater for most common types. It also allows all the heaters to be simply wired in series or parallel rather than some requiring special isolated power supplies such as specially insulated windings on power transformers or separate batteries.

The second source of heat is generated at the anode, when electrons, accelerated by the voltage applied to the anode, strike the anode and impart a considerable fraction of their energy to it, raising its temperature. In tubes used in power amplifier or transmitting circuits, this source of heat will exceed the power dissipated in the cathode heater.

This heat usually escapes the device by black body radiation from the anode/plate as infra red light. Some is conducted through the connecting wires going to the base but none is convected in most types of tube because of the vacuum and the absence of any gas inside the bulb to convect. It is the way tubes get rid of heat which most affects their overall appearance, next to the type of unit (triode, pentode, etc.) they contain, or whether they contain more than one of these basic units.

For devices required to radiate more than 500 mW or so, usually indirectly heated cathode types, the anode or plate is often treated to make its surface less shiny, and to make it darker, either gray or black. This helps it radiate the generated heat and maintain the anode or plate at a temperature significantly lower than the cathode, a requirement for proper operation.

Other variations

Pentagrid converters were generally used for frequency conversion in super heterodyne receivers in favour of a combination of a triode and hexode vacuum tube combination. Other methods of frequency conversion were based on octode tubes which featured eight electrodes and in which the additional grids were either control grids with different signals applied to each one, or screen grids. In many designs a special grid acted as a second anode and provided a built-in oscillator, which coupled this oscillator signal with the incoming radio signal to create a single, combined effect on the anode current. This was equivalent to the same effect as an analogue multiplier whereas the useful component of the output was the difference frequency between that of the incoming signal and that of the oscillator.

This tube above is the 12SA7 Pentagrid converter tube. Can you break down the word Pentagrid and tell me how many grids this tube has?. Penta and grid? Yes it is a tube which has five grids. Let us learn more about Pentagrid and other vacuum tube variants.

Just before vacuum tubes were phased out in favour of solid state devices precision control and screen grids, called frame grids, and offered enhanced performance. Instead of the typically elliptical fine-gauge wire supported by two larger wires, a frame grid was a metal stamping with rectangular openings that surrounded the cathode. The grid wires were in a plane defined by the stamping, and the control grid was placed much closer to the cathode surface than traditional construction would permit.

Multiple Vacuum Tube Designs

In order to reduce the cost and complexity of radio equipment, it was common practice to combine more than one tube function, or more than one set of elements in a single tube. An example is the RCA Type 55 was a double diode used as a detector, automatic gain control, rectifier and audio preamplifier in early AC powered radios. The Dual Triode is a typical example of a multi-section tube which, for most purposes, can perform the functions of two triode tubes, while taking up half as much space and costing less.

This is an RCA 12AX7 dual-triode tube produced in 1947. Can you think of five advantages and five disadvantages of multi-section vacuum tubes? I will ask you the same question later on.

The invention of the 9-pin miniature tube base allowed the accommodation of many multi section tubes like the triode pentode along with a host of similar tubes which were quite popular in television receivers while some colour Television receivers used tubes which were constructed with two anodes and beam deflection electrodes (often referred to as 'sheet beam' tubes). These tubes were specifically designed for demodulation of synchronous signals of colour signals in colour television receivers.

The desire to include many functions in one envelope resulted in the construction of very compact tubes which contained two triodes and two diodes although many in the compact tubes series had only triple triodes. An early example of multiple devices in one envelope was available in the 1920s and had 3 triodes in a single glass envelope together with all the fixed capacitors and resistors required to make a complete radio receiver. It had only one tube socket and was able to substantially reduce the cost of radio receiver construction at that time in Germany where state tax was levied by the number of sockets in a set. This unique vacuum tube construction, even though it compromised reliability was a precursor to the modern day integrated circuits.

During the vacuum tube era, this tube in the picture could be referred to as a modern miniature vacuum tube why do you think it was “modern”? Let us find out below.

3.4 Special-Purpose Tubes

Various inert gasses such as Argon, Neon or Helium will ionise at predictable voltages and this served as the basis for the construction of such special-purpose devices as voltage regulator tubes.

The Thyratron vacuum tube is a special-purpose tube filled with low-pressure gas or mercury, some of which vaporizes. It contains a hot cathode and an anode, but also includes a control electrode, which behaves like the grid of a triode. When the control electrode starts conduction, the gas ionizes, and the control electrode no longer can stop the current; the tube "latches" into conduction. Removing the anode voltage allows the gas to de-ionize, restoring it to its non-conductive state. This function is similar to that of a modern Silicon Controlled Rectifier (SCR). Thyratrons are known to carry large currents in comparison to their physical size with the hydrogen filled versions being widely applied for radar transmitters because of their very consistent time delay between turn-on pulse and full conduction

Tubes usually have glass envelopes, but metal, fused quartz silica), and ceramic are possible choices. Initial versions of some vacuum tubes used a metal envelope sealed with glass beads, while a glass disk fused to the metal was used in later versions. Metal and ceramic are used almost exclusively for power tubes above 2 kW dissipation. In some power tubes, the metal envelope is also the anode and air is blown through an array of fins attached to this anode, thus cooling it. Power tubes using this cooling scheme are available up to 150 kW dissipation. Above that level, water or water-vapour cooling are used. The highest-power tube currently available is a forced water-cooled power tetrode capable of dissipating 2.5 megawatts. (By comparison, the largest power transistor can only dissipate about 1 kilowatt.) while another very high power tube is a 1.25 megawatt tetrode used in military and commercial radio-frequency installations.

The Klystron is extremely specialized tubes which is used for extremely precise, rapid high-voltage switching and are heavily controlled at an international level due to their intended purpose; which is the initiation of the precise sequence of detonations used to set off nuclear weapons.

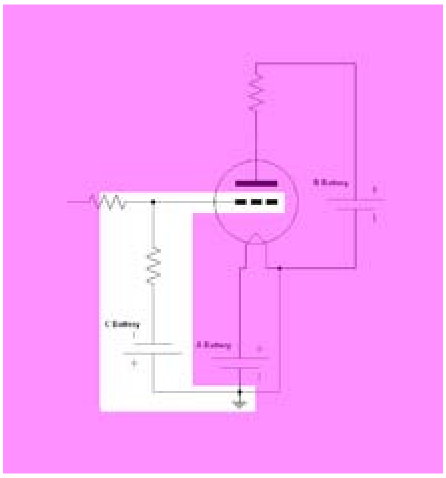
Medical radiographic and nuclear imaging, use special vacuum tubes which use a specially designed vacuum tube diode with a rotating anode to dissipate the large amounts of heat developed during operation. Radiographic, fluoroscopic, and CT X-ray imaging equipment use a focused cathode and are housed in aluminium which is filled with dielectric oil. Nuclear imaging equipment uses photomultiplier tube arrays to detect radiation.

3.4 Vacuum Tube Power Requirements

Batteries

The voltages required by vacuum tubes were provided by batteries in early radio sets and as many as three different voltages were required which meant three different batteries were used. The low voltage battery provided the filament voltage and vacuum tube heaters were designed for single, double or triple-cell lead acid batteries, giving nominal heater voltages of 2 V, 4 V or 6 V respectively. Portable radios sometimes used, dry cell batteries 1.5 or 1 Volt heaters. Reducing filament consumption improved the life span of batteries and by 1955, radio receiver tubes requiring between 50 mA and 10 mA for the heaters had been developed, This did not last long however as the advent of transistors rendered many vacuum tube applications obsolete.

Anode voltage was provided by high tension supply or batteries which were generally dry cells containing many small 1.5 volt cells in series. They typically came in ratings of 22.5, 45, 67.5, 90 or 135 volts. Some sets used a grid bias battery and although many circuits used grid leak resistors, voltage dividers or cathode bias to provide proper tube bias, they had very low battery drain.



This diagram shows batteries for a vacuum tube circuit where the C battery is highlighted.

AC power

Battery replacement represented a major cost of operation for early radio receiver users and the development of battery eliminators reduced operating costs and contributed to the growing popularity of radio. A power supply using a transformer with several windings, one or more rectifiers and large filter capacitors provided the required direct current voltages from the alternating current source.

As a cost reduction measure, especially in high-volume consumer receivers, all the tube heaters could be connected in series across the AC supply, and the plate voltage derived from a half-wave rectifier directly connected to the AC input, eliminating the need for a heavy power transformer. As an additional feature, these radios could be operated on AC or DC mains. While this arrangement limited the plate voltage and indirectly, the output power that could be obtained, the resulting supply was adequate for many purposes. A filament tap on the rectifier tube provided the 6 volt, low current supply needed for a dial light.

Vacuum tube cathodes can be directly heated as in the RS242 triodes in this picture

Direct and indirect heating

It became common to use the filament to heat a separate electrode called the cathode, and to use this cathode as the source of electron flow in the tube rather than the filament itself. This minimized the introduction of hum when the filament was energized with alternating current. In such tubes, the filament is called a heater to distinguish it as an inactive element. Development of vacuum tubes that could use alternating current for the heater supply allowed elimination of one rectifier element.

4.0 CONCLUSION

This Unit has enabled us know much more about Vacuum Tubes that we knew before reading it. We Learnt about the progression of Vacuum tubes from the discovery of the "Edison Effect" right through the evolutionary tree to the sophisticated and highly specialized Vacuum tubes we have today.

We learnt about the thermionic diode, triode, tetrode, pentode and are now comfortable with Vacuum tube nomenclature. We also learnt about additional grids such as the screen grid and the consequences resulting from these additional grids as well as the solution to these developments.

We also treated power requirements of, and heat dissipation in Vacuum tubes.

5.0 SUMMARY

- The discovery of the Edison effect led to the development of Vacuum rubes

- In Vacuum tubes, electrons travel through vacuum and not through a conducting material.

- The triode vas invented by Lee De Forest when he placed a metal grid in the middle of vacuum tube and by using a small input current to change the voltage on the grid controlled the flow of a more powerful current, through the tube.

- Vacuum tubes are constructed of a glass tube surrounding a vacuum with embedded electrical contacts at the ends.

- Some Vacuum tubes are filled with gas under low pressure.

- Vacuum tube nomenclature reflects the number of functional electrodes embedded in the tube, the minimum being two which connect the Anode and the Cathode. This tube is the diode. The next is the triode followed by the tetrode, pentode and hexode in the progression.

- Initially Vacuum tubes were made by lamp manufacturers. Later improvements led to oxide coated filaments which reduced tube temperature allowing for greater precision in construction.

- Heat generated in Vacuum tubes are mainly from the cathode and from electron bombardment of the Anode.

- Multiple Vacuum tube designs comprising more than one functional device in an envelope exist.

- Even though Vacuum tube functions have been largely taken over by Solid State devices, specialized Vacuum tube applications exist such as the Klystron, Thyratron, Cathode Ray, Photomultiplier and powerful Radio Transmission tubes of today.
- Vacuum tubes require at least two sources of electrical power. The filament requires a low voltage supply for heating while the

Anode requires a high voltage supply. The screen grid, suppressor grid and other terminals' voltages may be provided separately but are often derived from the Anode supply.

6.0 TUTOR MARKED ASSIGNMENTS

1 Name three specialized Vacuum tubes and their specialized functions

2. Draw and label a Vacuum tube tetrode. Describe the function of the screen grid.

3. What is the relationship between the gain of a vacuum tube triode and its grid to cathode spacing?

4. How many grids do each of the following Vacuum tubes possess?

   - Hexode
   - Pentode
   - Diode
   - Tetrode
   - Octode

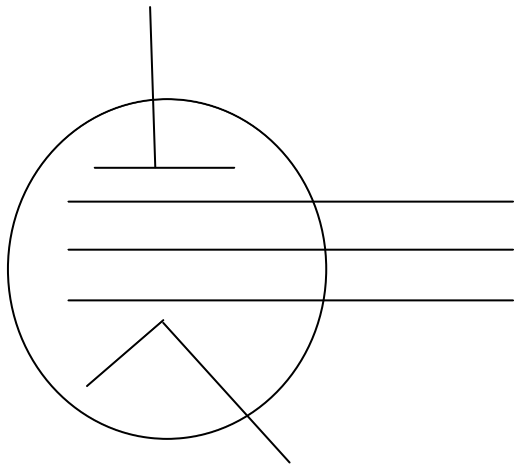
5. What is "Space Charge" and what significance does it have on the functioning of Vacuum tubes?

6. What are the benefits of combining more than one device in a Vacuum tube envelope such as in the case of the Dual Triode?

7. Vacuum tubes require multiple power supplies. Discuss this?

8. What merits do indirectly heated cathode have over direct heated cathode in Vacuum tubes?

9. Label the electrodes of the Pentode in the diagram below



10. Describe the operation of the suppressor grid.

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

Analog Filter Design
By M. E. Van Valkenburg Published by Holt, Rinehart and Winston

UNIT 2 SEMICONDUCTOR MATERIALS

CONTENTS

1.0 INTRODUCTION

Polymers and semiconductors are both highly visible engineering materials with a major impact on contemporary society and while the application of semiconductor technology has clearly revolutionized society, solid-state electronics is revolutionizing technology itself.

A relatively small group of elements and compounds have the important electrical property of semi-conduction in which they are neither good electrical conductors nor good electrical insulators but instead, their ability to conduct electricity is intermediate. Semiconductors in general do not fit into any of the four structural materials categories based on atomic bonding. Metals are inherently good electrical conductors. Ceramics and polymers (non-metals) are generally poor conductors but good insulators. The three semiconducting elements (Si, Ge, and Sn) from column IVA of the periodic table serve as a boundary between metallic and non-metallic elements.

2.0 OBJECTIVES

After reading through this unit, you should be able to

1 Describe the Electrical Properties of Semiconductor Materials
2 List Different Semiconductor Devices
3 Distinguish Between Electrons and Holes as Carriers
4 Understand the Process of Semiconductor Doping

5 Explain Charge Depletion in Semiconductor Junction
6 Describe the Flattening Of the Fermi-Dirac Distribution with Temperature
7 Explain Band Diagram of P-N Junction Operation
8 Discuss the Czochralski Process of Semiconductor Purification
9 List Different Semiconductor Materials

3.0 MAIN CONTENT

3.1 Semiconductor Materials

A semiconductor is a material with electrical conductivity due to electron flow which is intermediate in magnitude between that of a conductor and an insulator which means conductivity approximately in the range of 10^3 to 10^{-8} Siemens per centimeter. Semiconductor materials are the foundation of modern electronics, including radio, computers, telephones, and many other devices. Such devices include transistors, solar panels, various diodes such as light-emitting diode, the silicon controlled rectifier, and digital and analog integrated circuits. Similarly, semiconductor solar photovoltaic panels directly convert light energy into electrical energy.

In a metallic conduction, current is carried by the flow of electrons. In semiconductors, current is often schematized as being carried either by the flow of electrons or by the flow of positively charged "holes" in the electron structure of the material. In both cases, only electron movements are involved.

Common semiconducting materials are crystalline solids, but amorphous and liquid semiconductors are known. These include hydrogenated amorphous silicon and mixtures of arsenic, selenium and tellurium in a variety of proportions. Such compounds share with better known semiconductors intermediate conductivity and a rapid variation of conductivity with temperature, as well as occasional negative resistance. Such disordered materials lack the rigid crystalline structure of conventional semiconductors such as silicon and are generally used in thin film structures, which are less demanding for as concerns the electronic quality of the material and thus are relatively insensitive to impurities and radiation damage.

You might find it strange that there are organic semiconductors – but – it is true. Organic semiconductors are organic materials with properties resembling conventional semiconductors. Silicon is used to create most semiconductors commercially and many other materials are used, including germanium, gallium arsenide, and silicon carbide. A pure semiconductor is often called an “intrinsic” semiconductor. The electronic properties and the conductivity of a semiconductor can be changed in a controlled manner by adding very small quantities of other elements, called “dopants”, to the intrinsic material. This is achieved in crystalline silicon by adding impurities of boron or phosphorus to the molten silicon and then allowing it to solidify into the crystal. This process is called "doping"

Semiconductor materials are insulators at absolute zero temperature but conduct electricity at room temperature. The defining property of a semiconductor material is that it can be doped with impurities that alter its electronic properties in a controllable way.

Because of their application in devices like transistors and lasers, the search for new semiconductor materials and the improvement of existing materials is an important field of study in materials science. Most commonly used semiconductor materials are crystalline inorganic solids. These materials are classified according to the periodic table groups of their constituent atoms.

Different semiconductor materials differ in their properties. Thus, in comparison with silicon, compound semiconductors have both advantages and disadvantages. For example, gallium arsenide (GaAs) has six times higher electron mobility than silicon, which allows faster operation; wider band gap, which allows operation of power devices at higher temperatures, and gives lower thermal noise to low power devices at room temperature; its direct band gap gives it more favourable optoelectronic properties than the indirect band gap of silicon; it can be alloyed to ternary and quaternary compositions, with adjustable band gap width, allowing light emission at chosen wavelengths, and allowing e.g. matching to wavelengths with lowest losses in optical fibers. GaAs can be also grown in a semi insulating form, which is suitable as a lattice-matching insulating substrate for GaAs devices. Conversely, silicon is robust, cheap, and easy to process, whereas GaAs is brittle and expensive, and insulation layers cannot be created by just growing an oxide layer; GaAs is therefore used only where silicon is not sufficient.

By alloying multiple compounds, some semiconductor materials are tuneable in band gap or lattice constant. The result is ternary, quaternary, or even quinary compositions. Ternary compositions allow adjusting the band gap within the range of the involved binary compounds; however, in case of combination of direct and indirect band gap materials there is a ratio where indirect band gap prevails, limiting the range usable for optoelectronics. Lattice constants of the compounds also tend to be different, and the lattice mismatch against the substrate, dependent on the mixing ratio, causes defects in amounts dependent on the mismatch magnitude; this influences the ratio of achievable radiative/non radiative recombination and determines the luminous efficiency of the device. Quaternary and higher compositions allow adjusting simultaneously the band gap and the lattice constant, allowing increasing radiant efficiency at wider range of wavelengths; for example AlGaInP is used for LEDs. Materials transparent to the generated wavelength of light are advantageous which facilitates efficient extraction of photons from the material. In such transparent materials, light production is not limited to just the surface. Index of refraction is also composition-dependent and influences the extraction efficiency of photons from the material

3.2 Electrical Conduction

In crystalline semiconductors, electrons can have energies only within certain bands which are located between the ground state, corresponding to electrons tightly bound to the atomic nuclei of the material, and the free electron energy. The free electron energy being the energy required for an electron to escape entirely from the material. The energy bands each correspond to a large number of discrete quantum states of electrons, and most of the states with low energy which are closer to the nucleus are full up to a particular band called the valence band.

Semiconductors and insulators are distinguished from metals because their valence band is nearly filled with electrons under usual operating conditions, while very few in the case of semiconductor or virtually none in the case of insulator are available in the conduction band; the conduction band being the band immediately above the valence band.

The band gap between the bands determines the ease with which electrons in a semiconductor can be excited from the valence band to the conduction band.

With covalent bonds, an electron moves by hopping to a neighbouring bond. The Pauli Exclusion Principle requires the electron to be lifted into the higher anti-bonding state of that bond.

For delocalized states, for example in one dimension, for every energy there is a state with electrons flowing in one direction and another state with the electrons flowing in the other. For a net current to flow, more states for one direction than for the other direction must be occupied. For this to occur, energy is required, as in the semiconductor the next higher states lie above the band gap.

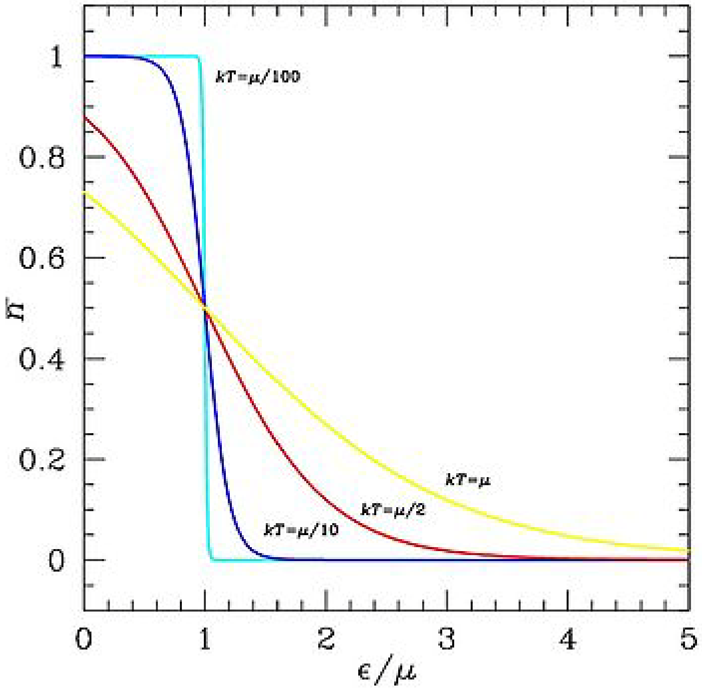
Often this is stated as: full bands do not contribute to the electrical conductivity. However, as the temperature of a semiconductor rises above absolute zero, there is more energy in the semiconductor to spend on lattice vibration and on lifting some electrons into an energy state of the conduction band.

Electrons excited to the conduction band also leave electron holes behind which are unoccupied states in the valence band. Both the conduction band electrons and the valence band holes contribute to electrical conductivity.

The holes themselves don't actually move, but a neighbouring electron can move to fill the hole, leaving a hole at the place it has just come from, and in this way the holes appear to move, and the holes behave as if they were actual positively charged particles. Because one covalent bond between neighbouring atoms in solid is ten times stronger than the binding of a single electron to the atom, freeing an electron does not imply destruction of the crystal structure.

Holes

The concept of holes can also be applied to metals too where the Fermi level lies within the conduction band. With most metals the Hall Effect indicates electrons are the charge carriers. However, some metals have a mostly filled conduction band. In these, the Hall Effect reveals positive charge carriers, which are not the ion-cores, but holes. In the case of a metal, only a small amount of energy is needed for the electrons to find other unoccupied states to move into, and hence for current to flow.



Take a loo at this table, It is the Fermi-Dirac distribution; Let us discuss its relevance to semiconductor materials.

Fermi-Dirac distribution

The States with energy below the Fermi energy have higher probability of being occupied, and those above are less likely to be occupied and the energy distribution of the electrons determines which of the states are filled and which are empty. This distribution is described by Fermi-Dirac statistics. The distribution is characterized by the temperature of the electrons, and the Fermi energy or Fermi level. Under absolute zero conditions the Fermi energy can be thought of as the energy up to which available electron states are occupied. At higher temperatures, the Fermi energy is the energy at which the probability of a state being occupied has fallen to 0.5. The dependence of the electron energy distribution on temperature also explains why the conductivity of a semiconductor has a strong temperature dependency, as a semiconductor operating at lower temperatures will have fewer available free electrons and holes.

3.3 Doping

The property of semiconductors that makes them most useful for constructing electronic devices is that their conductivity may easily be modified by introducing impurities into their crystal lattice. The process of adding controlled impurities to a semiconductor is known as doping and the amount of impurity, or dopant, added to an intrinsic (pure) semiconductor varies its level of conductivity. Doped semiconductors are often referred to as extrinsic semiconductors.

By adding impurity to pure semiconductors, the electrical conductivity may be varied not only by the number of impurity atoms but also, by the type of impurity atom and the changes may be thousand folds and million folds. For example, 1 cm^3 of a metal or semiconductor specimen has a number of atoms on the order of 10^{22}. Since every atom in metal donates at least one free electron for conduction in metal, 1 cm^3 of metal contains free electrons on the order of 10^{22}. At the temperature close to 20 °C , 1 cm^3 of pure germanium contains about 4.2×10^{22} atoms and 2.5×10^{13} free electrons and 2.5×10^{13} holes (empty spaces in crystal lattice having positive charge) The addition of 0.001% of arsenic (an impurity) donates an extra 10^{17} free electrons in the same volume and the electrical conductivity increases about 10,000 times.

The materials chosen as suitable dopants depend on the atomic properties of both the dopant and the material to be doped. In general, dopants that produce the desired controlled changes are classified as either electron acceptors or donors. A donor atom that activates (that is, becomes incorporated into the crystal lattice) donates weakly bound valence electrons to the material, creating excess negative charge carriers. These weakly bound electrons can move about in the crystal lattice relatively freely and can facilitate conduction in the presence of an electric field. (The donor atoms introduce some states under, but very close to the conduction band edge.

Electrons at these states can be easily excited to the conduction band, becoming free electrons, at room temperature.) Conversely, an activated acceptor produces a hole. Semiconductors doped with donor impurities are called n-type, while those doped with acceptor impurities are known as p-type. The n and p type designations indicate which charge carrier acts as the material's majority carrier. The opposite carrier is called the minority carrier, which exists due to thermal excitation at a much lower concentration compared to the majority carrier.

You know, for example that pure semiconductor silicon has four valence electrons. In silicon, the most common dopants are IUPAC group 13 (commonly known as group III) and group 15 (commonly known as group

V) elements. Group 13 elements all contain three valence electrons, causing them to function as acceptors when used to dope silicon. Group 15 elements have five valence electrons, which allow them to act as a donor. Therefore, a silicon crystal doped with boron creates a p-type semiconductor whereas one doped with phosphorus results in an n-type material. Remember this – you will need it later on.

Carrier concentration

The concentration of dopant introduced to an intrinsic semiconductor determines its concentration and indirectly affects many of its electrical properties. The most important factor that doping directly affects is the material's carrier concentration. In an intrinsic semiconductor under thermal equilibrium, the concentration of electrons and holes is equivalent.

$$n = p = n_i.$$

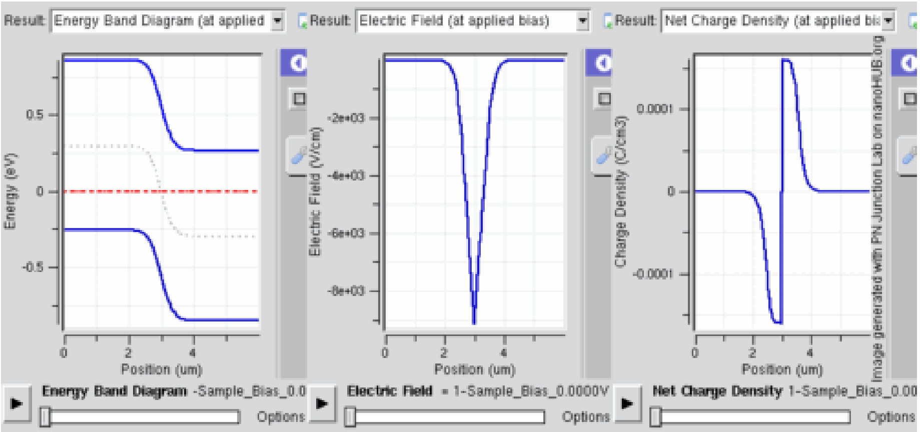
If we have a non-intrinsic semiconductor in thermal equilibrium the relation becomes:

$$n_0 \cdot p_0 = n_i^2$$

where n_0 is the concentration of conducting electrons, p_0 is the electron hole concentration, and n_i is the material's intrinsic carrier concentration. Intrinsic carrier concentration varies between materials and is dependent on temperature. Silicon's n_i, for example, is roughly 1.08×10^{10} cm^{-3} at 300 Kelvin (room temperature)

In general, an increase in doping concentration affords an increase in conductivity due to the higher concentration of carriers available for conduction. Very highly doped semiconductors have conductivity levels comparable to metals and are often used in modern integrated circuits as a replacement for metal. Often superscript plus and minus symbols are used to denote relative doping concentration in semiconductors. For example, n^+ denotes an n-type semiconductor with a high, often degenerate, doping concentration. Similarly, p^- would indicate a very lightly doped p-type material. It is useful to note that even degenerate levels of doping imply low concentrations of impurities with respect to the base semiconductor. In crystalline intrinsic silicon, there are approximately 5×10^{22} atoms/cm³. Doping concentration for silicon semiconductors may range anywhere from 10^{13} cm^{-3} to 10^{18} cm^{-3}. Doping concentration above about 10^{18} cm^{-3} is considered degenerate at room temperature. Degenerately doped silicon contains a proportion of impurity to silicon on the order of parts per thousand. This proportion may be reduced to parts per billion in very lightly doped silicon. Typical concentration values fall somewhere in this range and are tailored to produce the desired properties in the device that the semiconductor is intended for.

Effect on band structure



Band diagram of PN junction operation in forward bias mode showing reducing depletion width. Both p and n junctions are doped at a 1e15/cm3 doping level, leading to built-in potential of ~0.59V. Reducing depletion width can be inferred from the shrinking charge profile, as fewer dopants are exposed with increasing forward bias

Doping a semiconductor crystal introduces allowed energy states within the band gap but very close to the energy band that corresponds to the dopant type. In other words, donor impurities create states near the conduction band while acceptors create states near the valence band. The gap between these energy states and the nearest energy band is usually referred to as dopant-site bonding energy or E_B and is relatively small. For example, the E_B for boron in silicon bulk is 0.045 eV, compared with silicon's band gap of about 1.12 eV. Because E_B is so small, it takes little energy to ionize the dopant atoms and create free carriers in the conduction or valence bands. Usually the thermal energy available at room temperature is sufficient to ionize most of the dopant.

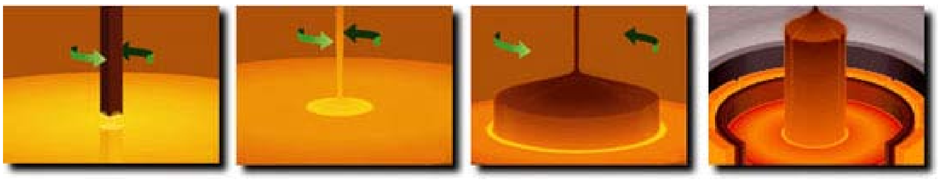
Dopants also have the important effect of shifting the material's Fermi level towards the energy band that corresponds with the dopant with the greatest concentration. Since the Fermi level must remain constant in a system in thermodynamic equilibrium, stacking layers of materials with different properties leads to many useful electrical properties. For example, the p-n junction's properties are due to the energy band bending that happens as a result of lining up the Fermi levels in contacting regions of p-type and n-type material.

This effect is shown in a band diagram. The band diagram typically indicates the variation in the valence band and conduction band edges versus some spatial dimension, often denoted x. The Fermi energy is also usually indicated in the diagram. Sometimes the intrinsic Fermi energy, E_i, which is the Fermi level in the absence of doping, is shown. These diagrams are useful in explaining the operation of many kinds of semiconductor devices.

3.4 Semiconductors Material Preparation

Semiconductors with reliably predictable electronic properties are requires for production and the level of chemical purity needed is extremely high because the presence of impurities even in very small proportions can have large effects on the properties of the material. High degree of crystalline perfection is also required in the semiconductor material since faults and imperfections in crystal structure interfere with the semiconducting properties of the material. Crystalline faults are a major cause of defective semiconductor devices. The larger the crystal, the more difficult it is to achieve the necessary perfection and conventional mass production processes use crystal ingots between 100 mm and 300 mm (4–12 inches) in diameter which are grown as cylinders and sliced into wafers.

Because of the required level of chemical purity and the perfection of the crystal structure which are needed to make semiconductor devices, special methods have been developed to produce the initial semiconductor material. A technique for achieving high purity includes growing the crystal using the Czochralski process. An additional step that can be used to further increase purity is known as zone refining. In zone refining, part of a solid crystal is melted. The impurities tend to concentrate in the melted region, while the desired material re-crystallizes leaving the solid material more pure and with fewer crystalline faults.

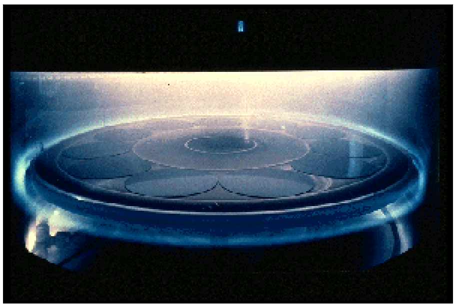


In the Czochralski process which is just one step in the production of semiconductor devices, a large (10 or more inches in diameter!) single crystal of silicon is grown. In the process, a solid seed crystal is rotated and slowly extracted from a pool of molten Silicon.

In another processing step, after wafers of silicon are sliced from the single crystal and then polished, they are thermal-annealed in an oven and in a vacuum chamber; fine, intricate patterns are etched into silicon wafers with an ion discharge which is another critical step in the production of semiconductor components. The ion plasma etching process makes possible the small geometries need for very large scale integration in silicon chips.

In manufacturing semiconductor devices involving junctions between different semiconductor materials, the length of the repeating element of the crystal structure, is very important for determining the compatibility of materials.

Intel's continued investment in state-of-the-art manufacturing facilities means the highest quality, highest-performance products for Intel customers.



4.0 CONCLUSION

I believe that having studied this Unit you have learnt that only a very small group of elements and compounds have the important electrical property of semi-conduction in which they are neither good electrical conductors nor good electrical insulators

You were exposed to many terms which include "Intrinsic semi conduction", "Dopants" and "Holes". You know that it is important for you to know what these terms and words mean in context.

The Fermi-Dirac distribution has been explained and its significance to semiconducting explained. Dopants which may either be electron acceptors or electron donors, and the effect of doping on material type the band structure and the Fermi-Dirac distribution treated.

We took a look at the processes of preparation of Semiconductor material and appreciate that preparation requires several stages of purification, and the most favoured by industry being the Czochralski process and zone refining process.

5.0 SUMMARY

- Semiconductors are materials with electrical conductivity due to electron flow which is intermediate in magnitude between that of a conductor and an insulator.

- Common semiconducting materials are crystalline solids, but amorphous and liquid semiconductors are known. Also known are organic semiconductors.

- The band gap between bands determines the ease with which electrons in a semiconductor can be excited from the valence band to the conduction band.

- In semiconductors, holes do not actually move, but neighboring electrons can move to fill holes, leaving a hole at the place it has just come from, and in this way the holes appear to move, and the holes behave as if they were actual positively charged particles.

- In the Fermi-Dirac distribution, states with energy below the Fermi energy have higher probability of being occupied, and those above are less likely to be occupied and the energy distribution of the electrons determines which of the states are filled and which are empty.

- By adding impurity to pure semiconductors, the electrical conductivity may be varied not only by the number of impurity atoms but also, by the type of impurity atom and the changes may be thousand folds and million folds. This process is called Doping.

- One technique for achieving high semiconductor purity is by growing the crystal using the Czochralski process and additional increase in purity can be achieved through another process called zone refining.

- Crystalline faults are a major cause of defective semiconductor devices and the larger the semiconductor crystal, the more difficult it is to achieve the necessary perfection.

6.0 TUTOR MARKED ASSIGNMENTS

1. What s a semiconductor

2. Name the two most common semiconductor elements and specify their group in the periodic table.

3. List and describe two structural defects often encounrtered in semiconductor manufacturing process.

4. Describe the conduction mechanism in Conductors, Insulators and Semiconductors.

5. With what group elements would you prepare a “P” type semiconductor?

6. List ten electronic components made out semiconductor materials

7. Explain the effect of doping on the Fermi-Dirac distribution for Semiconductors

8. Describe the Czochralski process and the zone refining processes for attaining extremely high yield in semiconductor preparation.

9. Consider this statement: “The most important factor that doping directly affects is the material's carrier concentration”. Discuss this.

10. Draw and label the Fermi-Dirac distribution for semiconductors.

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Semiconductor Device Fundamentals
By Robert F. Pierret Published by Prentice Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

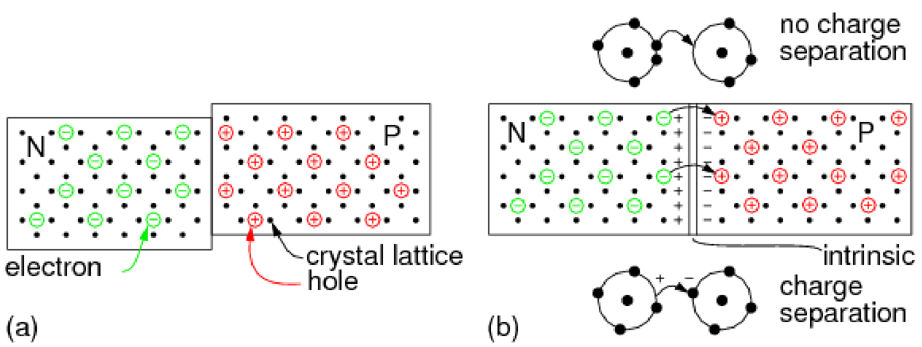
UNIT 3 P-N JUNCTION DIODES

CONTENTS

1.0 INTRODUCTION

P-N junction forms the basis not only of rectifier diodes, but of many other electronic devices such as LEDs, lasers, photodiodes and bipolar junction transistors and solar cells. A p-n junction aggregates recombination, generation, diffusion and drift effects into a single device.

If a block of P-type semiconductor is placed in contact with a block of N-type semiconductor in Figure below the result is of no value. We have two conductive blocks in contact with each other, showing no unique properties. The problem is two separate and distinct crystal bodies. The number of electrons is balanced by the number of protons in both blocks. Thus, neither block has any net charge.



However, a single semiconductor crystal manufactured with P-type material at one end and N-type material at the other has some unique properties. The P-type material has positive majority charge carriers, holes, which are free to move about the crystal lattice. The N-type material has mobile negative majority carriers, electrons. Near the junction, the N-type material electrons diffuse across the junction, combining with holes in P-type material.

P-N junctions are formed by joining n-type and p-type semiconductor materials. Since the n-type region has a high electron concentration and the p-type a high hole concentration, electrons diffuse from the n-type side to the p-type side. Similarly, holes flow by diffusion from the p-type side to the n-type side. If the electrons and holes were not charged, this diffusion process would continue until the concentration of electrons and holes on the two sides were the same, as happens if two gasses come into contact with each other. However, in a p-n junction, when the electrons and holes move to the other side of the junction, they leave behind exposed charges on dopant atom sites, which are fixed in the crystal lattice and are unable to move. On the n-type side, positive ion cores are exposed. On the p-type side, negative ion cores are exposed. An electric field forms between the positive ion cores in the n-type material and negative ion cores in the p-type material. This region is called the "depletion region" since the electric field quickly sweeps free carriers out, hence the region is depleted of free carriers. A "built in" potential due to the electric field is formed at the junction.

A p-n junction with no external inputs represents equilibrium between carrier generation, recombination, diffusion and drift in the presence of the electric field in the depletion region. Despite the presence of the electric field, which creates an impediment to the diffusion of carriers across the electric field, some carriers still cross the junction by diffusion. In the animation below, most majority carriers which enter the depletion region move back towards the region from which they originated. However, statistically some carriers will have a high velocity and travel in a sufficient net direction such that they cross the junction. Once a majority carrier crosses the junction, it becomes a minority carrier. It will continue to diffuse away from the junction and can travel a distance on average equal to the diffusion length before it recombines. The current caused by the diffusion of carriers across the junction is called diffusion current.

Minority carriers which reach the edge of the diffusion region are swept across it by the electric field in the depletion region. This current is called the drift current. In equilibrium the drift current is limited by the number of minority carriers which are thermally generated within a diffusion length of the junction.

In equilibrium, the net current from the device is zero. The electron drift current and the electron diffusion current exactly balance out (if they did not there would be a net build-up of electrons on either one side or the other of the device). Similarly, the hole drift current and the hole diffusion current also balance each other out

2.0 OBJECTIVES

After reading through this unit, you should be able to

1. Explain How a P-N Junction Functions
2. Describe the Depletion Region of A P-N Junction
3. Appreciate Forward, Reverse and Zero Voltage Effect on Depletion
4. Understand Reverse Breakdown on the P-N Junction
5. Describe the Devices Whose Operation Depend On P-N Junction
6. Know That Not All P-N Junctions Rectify Current
7. Describe the Process for Manufacturing P-N Junctions

3.0 MAIN CONTENT

3.1 P-N Junction and Junction Diodes

The P-N junction possesses some properties which have useful applications in modern electronics. A p-doped semiconductor is relatively conductive. The same is true of an n-doped semiconductor, but the junction between them can become depleted of charge carriers, and hence nonconductive, depending on the relative voltages of the two semiconductor regions. By manipulating this non-conductive layer, p–n junctions are commonly used as diodes: circuit elements that allow a flow of electricity in one direction but not in the other (opposite) direction. This property is explained in terms of forward bias and reverse bias, where the term bias refers to an application of electric voltage to the p–n junction.

The forward-bias and the reverse-bias properties of the p–n junction imply that it can be used as a diode. A p–n junction diode allows electric charges to flow in one direction, but not in the opposite direction; negative charges (electrons) can easily flow through the junction from n to p but not from p to n and the reverse is true for holes. When the p–n junction is forward biased, electric charge flows freely due to reduced resistance of the p–n junction. When the p–n junction is reverse biased, however, the junction barrier (and therefore resistance) becomes greater and charge flow is minimal.

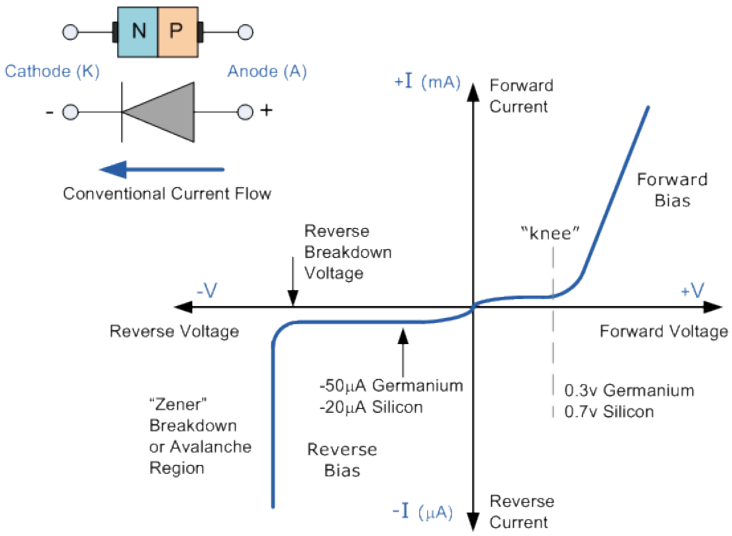
Normally, p–n junctions are manufactured from a single crystal with different dopant concentrations diffused across it. Creating a semiconductor from two separate pieces of material would introduce a grain boundary between the semiconductors which severely inhibits its utility by scattering the electrons and holes. However, in the case of solar cells, polycrystalline silicon is often used to reduce expense, despite the lower efficiency

A diode is one of the simplest semiconductor devices and it has the characteristic of passing current in one direction only. However, unlike a resistor, a diode does not behave linearly with respect to the applied voltage as the diode has an exponential voltage/current relationship and therefore cannot be described operationally by simply applying Ohm's law.

If a suitable positive voltage (forward bias) is applied between the two ends of the PN junction, it can supply free electrons and holes with the extra energy they require to cross the junction as the width of the depletion layer around the PN junction is decreased. By applying a negative voltage (reverse bias) results in the free charges being pulled away from the junction resulting in the depletion layer width being increased. This has the effect of increasing or decreasing the effective resistance of the junction itself allowing or blocking current flow through the diode.

Then the depletion layer widens with an increase in the application of a reverse voltage and narrows with an increase in the application of a forward voltage. This is due to the differences in the electrical properties on the two sides of the PN junction resulting in physical changes taking place. One of the results produces rectification as seen in the PN junction diodes static I-V (current-voltage) characteristics. Rectification is shown by an asymmetrical current flow when the polarity of bias voltage is altered.

Look below and study both the Junction Diode Symbol and its Static I-V Characteristics curve taking note of the bon linear “knee” at the top right hand side of the curve, and also noting the reverse breakdown or Zener breakdown region at the bottom left..



P-N junctions cannot be used as a rectifying device without first biasing the junction by connecting a voltage potential across it. Reverse Bias refers to an external voltage potential which increases the potential barrier whereas an external voltage which decreases the potential barrier is said to act in the Forward Bias direction.

There are two operating regions and three possible "biasing" conditions for the standard Junction Diode:

Zero Bias — Where no external voltage potential is applied to the PN-junction.

Reverse Bias — Where the voltage potential is connected negative, to the P-type material and positive, to the N-type material across the diode which has the effect of Increasing the PN-junction width.

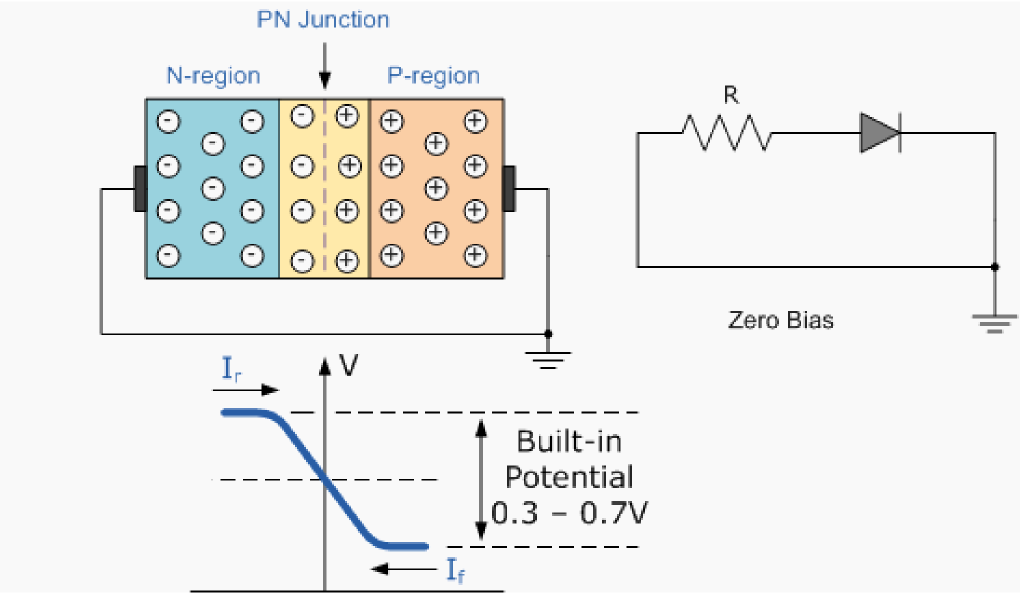
Forward Bias — Where the voltage potential is connected positive, to the P-type material and negative, to the N-type material across the diode which has the effect of Decreasing the PN-junction width.

3.2 Zero Bias

When a diode is connected in a Zero Bias condition, no external potential energy is applied to the PN junction. However if the diodes terminals are shorted together, a few holes (majority carriers) in the P-type material with enough energy to overcome the potential barrier will move across the junction against this barrier potential. This is known as the Forward Current

Likewise, holes generated in the N-type material (minority carriers), find this situation favourable and move across the junction in the opposite direction. This is known as the Reverse Current. This transfer of electrons and holes back and forth across the PN junction is known as diffusion.

Zero Biased Junction Diode



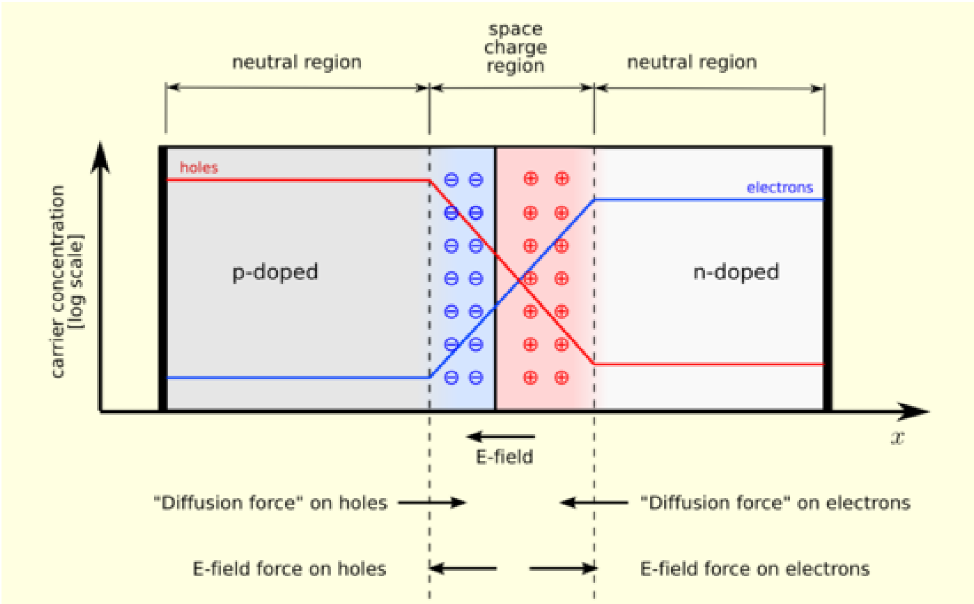
The potential barrier that now exists discourages the diffusion of any more majority carriers across the junction. However, the potential barrier helps minority carriers (few free electrons in the P-region and few holes in the N-region) to drift across the junction. Then an Equilibrium or balance will be established when the majority carriers are equal and both moving in opposite directions, so that the net result is zero current flowing in the circuit. When this occurs the junction is said to be in a state of Dynamic Equilibrium.

The minority carriers are constantly generated due to thermal energy so this state of equilibrium can be broken by raising the temperature of the

PN junction causing an increase in the generation of minority carriers, thereby resulting in an increase in leakage current but an electric current cannot flow since no circuit has been connected to the PN junction.

Equilibrium

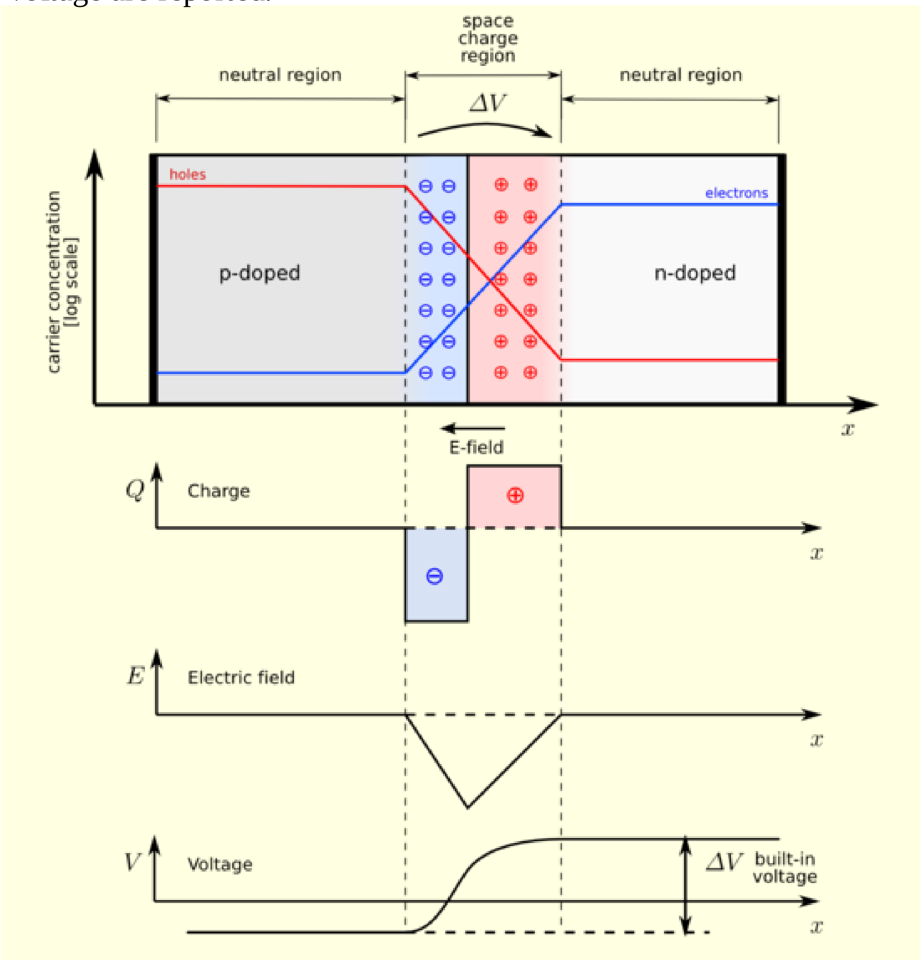
In a p–n junction, without an external applied voltage, an equilibrium condition is reached in which a potential difference is formed across the junction. This potential difference is called built-in potential. After joining p-type and n-type semiconductors, electrons near the p–n interface tend to diffuse into the p region. As electrons diffuse, they leave positively charged ions (donors) in the n region. Similarly, holes near the p–n interface begin to diffuse into the n-type region leaving fixed ions (acceptors) with negative charge. The regions nearby the p–n interfaces lose their neutrality and become charged, forming the space charge region or depletion layer.



A p–n junction in thermal equilibrium with zero bias voltage applied. Electrons and holes concentration are reported respectively with blue and red lines. Gray regions are charge neutral. Light red zone is positively charged. Light blue zone is negatively charged. The electric field is shown on the bottom, the electrostatic force on electrons and holes and the direction in which the diffusion tends to move electrons and holes.

The electric field created by the space charge region opposes the diffusion process for both electrons and holes. There are two concurrent phenomena: the diffusion process that tends to generate more space charge and the electric field generated by the space charge that tends to counteract the diffusion. The carrier concentration profile at equilibrium is shown below. Also shown are the two counterbalancing phenomena that establish equilibrium.

This p–n junction in thermal equilibrium with zero bias voltage applied. Under the junction, plots for the charge density, the electric field and the voltage are reported.

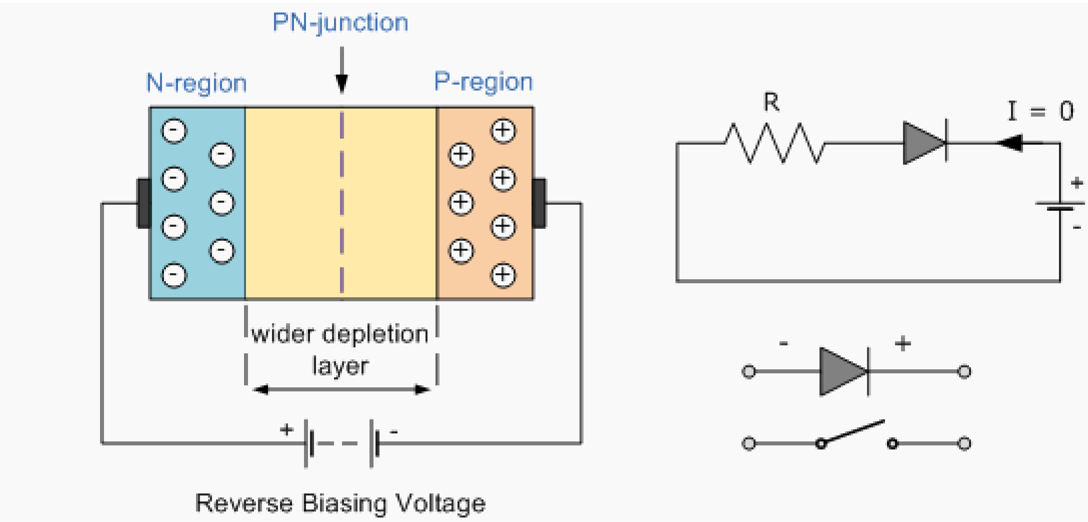


The space charge region is a zone with a net charge provided by the fixed ions (donors or acceptors) that have been left uncovered by majority carrier diffusion. When equilibrium is reached, the charge density is approximated by the displayed step function. In fact, the region is completely depleted of majority carriers (leaving a charge density equal to the net doping level), and the edge between the space charge region and the neutral region is quite The space charge region has the same charge on both sides of the p–n interfaces, thus it extends farther on the less doped side

3.3 Reverse Bias

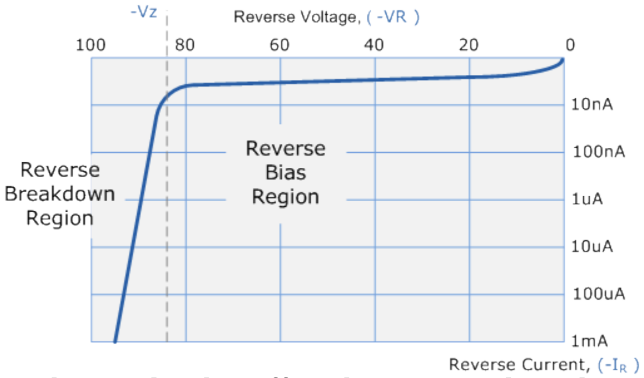
When a diode is connected in a Reverse Bias condition, a positive voltage is applied to the N-type material and a negative voltage is applied to the P-type material. The positive voltage applied to the N-type material attracts electrons towards the positive electrode and away from the junction, while the holes in the P-type end are also attracted away from the junction towards the negative electrode. The net result is that the depletion layer grows wider due to a lack of electrons and holes and presents a high impedance path, almost an insulator. The result is that a high potential barrier is created thus preventing current from flowing through the semiconductor material.

Study the figure below and observe that the reverse biased junction diode shows an increase in the depletion layer

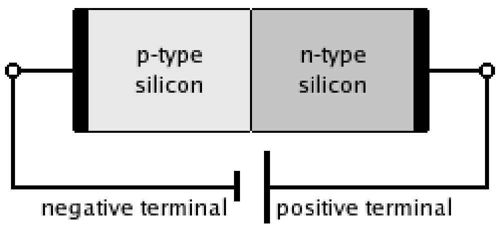


This condition represents a high resistance value to the PN junction and practically zero current flows through the junction diode with an increase in bias voltage. However, a very small leakage current does flow through the junction which can be measured in microamperes, If the reverse bias voltage applied to the diode is increased to a sufficiently high enough value, it will cause the PN junction to overheat and fail due to the avalanche effect around the junction. This may cause the diode to become shorted and will result in the flow of maximum circuit current.

Take a look at this Reverse Characteristics Curve for a Junction Diode; can you identify the reverse breakdown region?



Sometimes this avalanche effect has practical applications in voltage stabilising circuits where a series limiting resistor is used with the diode to limit this reverse breakdown current to a preset maximum value thereby producing a fixed voltage output across the diode. These types of diodes are commonly known as Zener Diodes.



You must take note of the battery polarity for the silicon p–n junction in reverse bias. Take a look at the correct reverse bias battery connection in the diagram above.

Now for emphasis:

As reverse bias usually refers to how a diode is used in a circuit and a reverse biased diode cathode voltage is higher than that at the anode. No current will flow until the diode breaks down. Connecting the P-type region to the negative terminal of the battery and the N-type region to the positive terminal, corresponds to reverse bias.

Because the p-type material is now connected to the negative terminal of the power supply, the 'holes' in the P-type material are pulled away from the junction, causing the width of the depletion zone to increase. Similarly, because the N-type region is connected to the positive terminal, the electrons will also be pulled away from the junction. Therefore the depletion region widens, and does so increasingly with increasing reverse-bias voltage. This increases the voltage barrier causing a high resistance to the flow of charge carriers thus allowing minimal electric current to cross the p–n junction. The increase in resistance of the p-n junction results in the junction to behave as an insulator. This is important for radiation detection because if current was able to flow, the charged particles would just dissipate into the material. The reverse bias ensures that charged particles are able to make it to the detector system.

The strength of the depletion zone electric field increases as the reverse-bias voltage increases. Once the electric field intensity increases beyond a critical level, the p–n junction depletion zone breaks-down and current begins to flow, by either the Zener or avalanche breakdown processes. Both of these breakdown processes are non-destructive and are reversible, so long as the amount of current flowing does not reach levels that cause the semiconductor material to overheat and cause thermal damage.

As previously mentioned the reverse leakage current of under a µA for silicon diodes was due to conduction of the intrinsic semiconductor. This is the leakage that can be explained by theory. Thermal energy produces few electron hole pairs, which conduct leakage current until recombination. In actual practice this predictable current is only part of the leakage current. Much of the leakage current is due to surface conduction, related to the lack of cleanliness of the semiconductor surface. Both leakage currents increase with increasing temperature, approaching a µA for small silicon diodes.

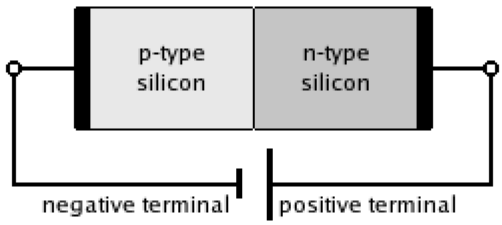
For germanium, the leakage current is orders of magnitude higher. Since germanium semiconductors are rarely used today, this is not a problem in practice.

Zener Diode

If the diode is reverse biased, only the leakage current of the intrinsic semiconductor flows. This current will only be as high as 1 µA for the most extreme conditions for silicon small signal diodes. This current does not increase appreciably with increasing reverse bias until the diode breaks down. At breakdown, the current increases greatly and the diode will be destroyed unless a high series resistance limits current. Normally a diode with a higher reverse voltage rating than any applied voltage is selected to prevent this reverse breakdown. Silicon diodes are typically available with reverse break down ratings of 50, 100, 200, 400, 800 V and higher.

It is possible to fabricate diodes with a lower rating of a few volts for use as voltage standards and this effect is used to advantage in Zener diode regulator circuits. Zener diodes have a certain - low - breakdown voltage. A typical value for breakdown voltage is for instance 5.6V. This means that the voltage at the cathode can never be more than 5.6V higher than the voltage at the anode, because the diode will break down - and therefore conduct - if the voltage gets any higher. This effectively regulates the voltage over the diode.

Varicap



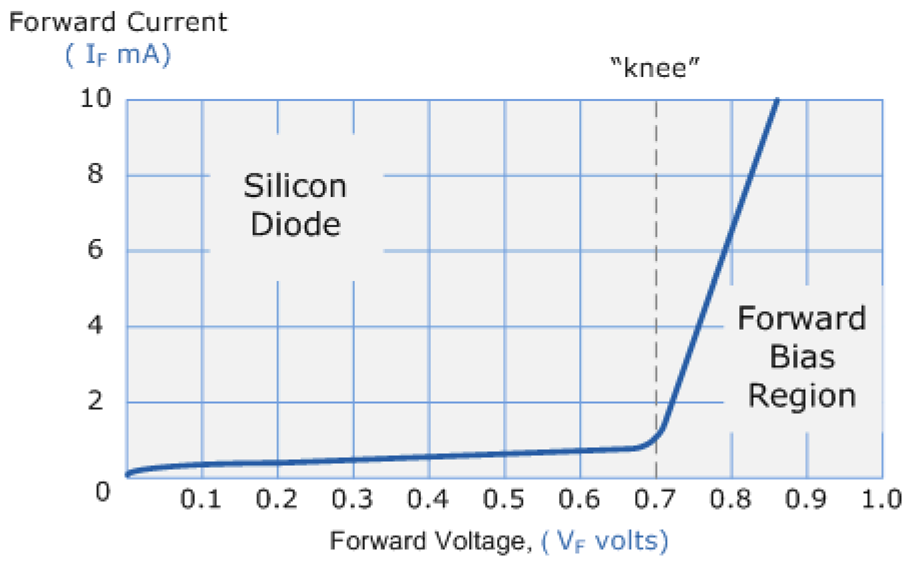
Do you know that a depletion region capacitance is the consequence of a reversed bias P-N junction? And that the value of this capacitance varies with the reverse bias voltage? Find out how below.

Another application where reverse biased diodes are used is in Varicap diodes. The width of the depletion zone of any diode changes with voltage applied. This varies the capacitance of the diode.

3.4 Forward Bias

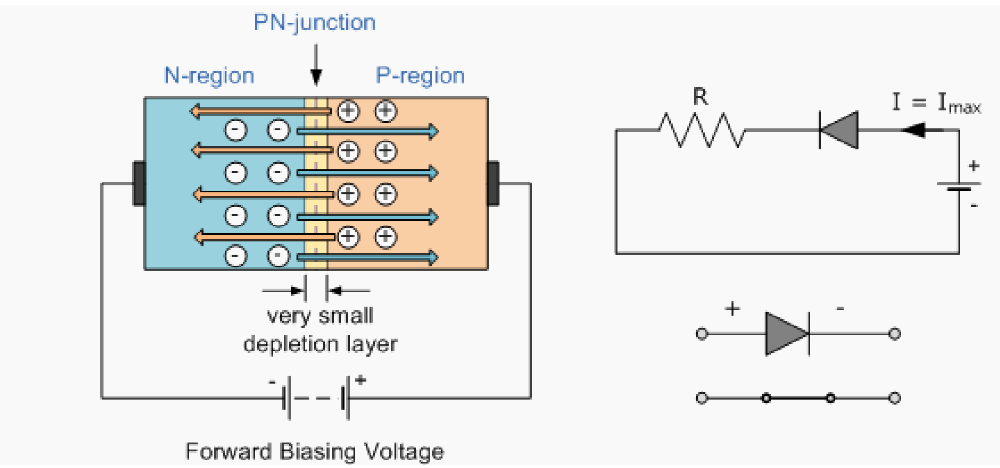
When a diode is connected in a Forward Bias condition, a negative voltage is applied to the N-type material and a positive voltage is applied to the P-type material. If this external voltage becomes greater than the value of the potential barrier, approx. 0.7 volts for silicon and 0.3 volts for germanium, the potential barriers opposition will be overcome and current will start to flow. This is because the negative voltage pushes or repels electrons towards the junction giving them the energy to cross over and combine with the holes being pushed in the opposite direction towards the junction by the positive voltage. This results in a characteristics curve of zero current flowing up to this voltage point, called the "knee" on the static curves and then a high current flow through the diode with little increase in the external voltage as shown below.

The curve below shows you the Forward Conduction Characteristics for Junction Diodes



The application of a forward biasing voltage on the junction diode results in the depletion layer becoming very thin and narrow which represents a low impedance path through the junction thereby allowing high currents to flow. The point at which this sudden increase in current takes place is represented on the static voltage/current characteristics curve above as the knee point.

When you forward bias a Junction Diode it results in a reduction in the depletion layer. Did you know that? Now you do and you can see this illustrated below.



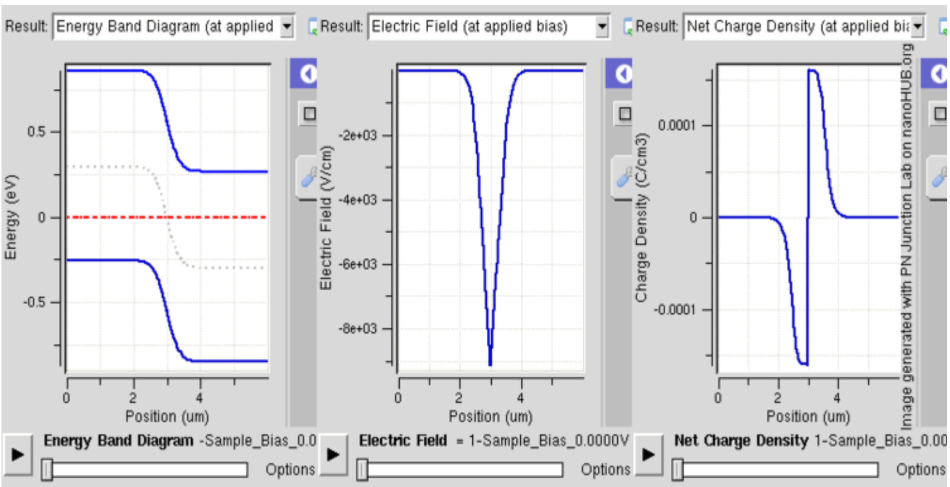
This condition represents the low resistance path through the PN junction allowing very large currents to flow through the diode with only a small increase in bias voltage. The actual potential difference across the junction or diode is kept constant by the action of the depletion layer at approximately 0.3v for germanium and approximately 0.7v for silicon junction diodes. Since the diode can conduct very high current above this knee point as it effectively becomes a short circuit and exceeding the maximum forward current specification causes the device to dissipate more power in the form of heat resulting in quick failure.

Once again for emphasis:

In forward bias, the p-type is connected with the positive terminal and the n-type is connected with the negative terminal.

PN junction operation in forward bias mode showing reducing depletion width. Both p and n junctions are doped at a 1e15/cm3 doping level, leading to built-in potential of ~0.59V. Reducing depletion width can be inferred from the shrinking charge profile, as fewer dopants are exposed with increasing forward bias.

With a battery connected this way, the holes in the P-type region and the electrons in the N-type region are pushed towards the junction. This reduces the width of the depletion zone. The positive charge applied to the P-type material repels the holes, while the negative charge applied to the N-type material repels the electrons. As electrons and holes are pushed towards the junction, the distance between them decreases.



This lowers the barrier in potential. With increasing forward-bias voltage, the depletion zone eventually becomes thin enough that the zone's electric field can't counteract charge carrier motion across the p–n junction, consequently reducing electrical resistance. The electrons which cross the p–n junction into the P-type material (or holes which cross into the N-type material) will diffuse in the near-neutral region. Therefore, the amount of minority diffusion in the near-neutral zones determines the amount of current that may flow through the diode.

Only majority carriers (electrons in N-type material or holes in P-type) can flow through a semiconductor for a macroscopic length. With this in mind, consider the flow of electrons across the junction. The forward bias causes a force on the electrons pushing them from the N side toward the P side. With forward bias, the depletion region is narrow enough that electrons can cross the junction and inject into the P-type material. However, they do not continue to flow through the P-type material indefinitely, because it is energetically favourable for them to recombine with holes. The average length an electron travels through the P-type material before recombining is called the diffusion length, and it is typically on the order of microns.

Although the electrons penetrate only a short distance into the P-type material, the electric current continues uninterrupted, because holes (the majority carriers) begin to flow in the opposite direction. The total current (the sum of the electron and hole currents) is constant in space, because any variation would cause charge build-up over time (this is Kirchhoff's current law). The flow of holes from the P-type region into the N-type region is exactly analogous to the flow of electrons from N to P (electrons and holes swap roles and the signs of all currents and voltages are reversed).

Therefore, the macroscopic picture of the current flow through the diode involves electrons flowing through the N-type region toward the junction, holes flowing through the P-type region in the opposite direction toward the junction, and the two species of carriers constantly recombining in the vicinity of the junction. The electrons and holes travel in opposite directions, but they also have opposite charges, so the overall current is in the same direction on both sides of the diode, as required.

Increasing the voltage well beyond 0.7 V in Silicon diodes may result in high enough current to destroy the diode. The forward voltage is a characteristic of the semiconductor: 0.6 to 0.7 V for silicon, 0.2 V for germanium, a few volts for Light Emitting Diodes (LED). The forward current ranges from a few mA for point contact diodes to 100 mA for small signal diodes to tens or thousands of amperes for power diodes.

3.5 Non-Rectifying Junctions

I want you to know that not all P-N junctions rectify. Schottky junction is a special case of a p-n junction, where metal serves the role of the n-type semiconductor. The Shockley diode equation models the forward-bias operational characteristics of a p–n junction outside the avalanche (reverse-biased conducting) region.

Contact between the metal wires and the semiconductor material also creates metal-semiconductor junctions called Schottky diodes. In a simplified ideal situation a semiconductor diode would never function, since it would be composed of several diodes connected back-to-front in series. But in practice, surface impurities within the part of the semiconductor which touches the metal terminals will greatly reduce the width of those depletion layers to such an extent that the metal-semiconductor junctions do not act as diodes. These "non rectifying junctions" behave as ohmic contacts regardless of applied voltage polarity.

4.0 CONCLUSION

This Unit has introduced the P-N junction to us in detail and made us aware that it forms the basis not only of rectifier diodes, but of many other electronic devices such as LEDs, lasers, photodiodes and bipolar junction transistors and solar cells. It also aggregates recombination, generation, diffusion and drift effects into a single device.

We learn also that when the junction is biased, the response of the junction depends on the polarity of the biasing voltage. We therefore have forward, reverse and zero bias.

Further, we took note that when subjected to increasing reverse bias, every P-N junction will suffer reverse breakdown which becomes catastrophic if the reverse current is not limited.

We have become familiar with the diagrams which render the details of processes, structure, layout and the connection of the P-N junction. We have also become familiar with the characteristic curves of the different bias configuration.

And finally, we learn that there are P-N junctions which do not rectify, but present Ohmic properties.

5.0 SUMMARY

- P-N junction forms the basis of rectifier diodes and many other electronic devices such as LEDs, lasers, photodiodes and bipolar junction transistors and solar cells.

- The physical contact of a P type material with a separate N type material exhibits no unique properties as the rectifying properties of a single P-N junction

- A single semiconductor crystal manufactured with P-type material at one end and N-type material at the other has the unique properties associated with the P-N junction

- A p-n junction with no external inputs represents equilibrium between carrier generation, recombination, diffusion and drift in the presence of the electric field in the depletion region,

- There are two operating regions and three possible "biasing" conditions for the standard Junction Diode**:** Zero Bias; where no external voltage potential is applied to the PN-junction, Reverse Bias; where the voltage potential is connected negative, to the P-type material and positive, to the N-type material across the diode which has the effect of Increasing the PN-junction width, and Forward Bias where the voltage potential is connected positive, to the P-type material and negative, to the N-type material across the diode which has the effect of Decreasing the PN-junction width.

- Zener diodes which are useful for Voltage calibration are a consequence of reverse breakdown.

- In the Varicap, the width of the depletion zone of a diode changes with voltage applied. This varies the capacitance of the diode and the value of this capacitance is a function of the applied voltage.

- Not all P-N junctions rectify. Schottky junction is a special case of a P-N junction, where metal serves the role of the n-type semiconductor and the junction presents Ohmic properties.

6.0 TUTOR MARKED ASSIGNMENTS

1. Describe the electrical properties of a mono-crystalline P-N junction and distinguish this from a mere P and N type semiconductor material in physical contact.

2. What do Light Emitting Diodes, Lasers, Photodiodes and Bipolar Junction Transistors and Solar Cells have in common?

3. Sketch and label the characteristic curve of a rectifying P-N junction highlighting the important parts of the curve.

4. How many ways are there of biasing a P-N junction? Describe the electrical properties of each of them.

5. Can you name a non- rectifying P-N junction? What electrical properties does it present?

6. The forward bias voltage of a Germanium P-N junction is 0.25 volts. What is the forward bias voltage of a Silicon P-N junction?

7. Can you describe the process of equilibrium in a zero biased P-N junction?

8. What do you understand by Avalanche effect of a P-N junction? Describe this phenomenon stating two practical applications.

9. Briefly explain how a Varicap works and state three practical applications of the Varicap.

10. What is a Shockley diode?

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Semiconductor Device Fundamentals
By Robert F. Pierret Published by Prentice Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

UNIT 4 TRANSISTORS

CONTENTS

1.0 INTRODUCTION

The problems with vacuum tubes fuelled the search for alternative ways to make three terminal devices instead of using electrons in vacuum and researches and scientists alike began to consider how one might control electrons in solid materials, like metals and semiconductors. In the 1920's, solid state two terminal devices were already being realised through point contact between a sharp metal tip and pieces of naturally occurring semiconductor crystal. These point-contact diodes were used to rectify signals and make simple AM radio receivers (crystal radios). It was not until 1947 however, when John Bardeen and Walter Brattain, working at Bell Telephone Laboratories, were trying to understand the nature of the electrons at the interface between a metal and a semiconductor, that they realized that by making two point contacts very close to one another, they could make a three terminal device – which was the point contact transistor.

Bell Telephone Company immediately realized the potential power of this new technology which sparked off huge research effort in solid state electronics. Bardeen and Brattain received the Nobel Prize in Physics, 1956, together with William Shockley, "for their researches on semiconductors and their discovery of the transistor effect." Shockley had developed a so-called junction transistor, which was built on thin slices of different types of semiconductor material pressed together, was easier to understand theoretically, and could be manufactured more reliably.

The transistor is a three terminal, solid state electronic device in which it is possible to control electric current or voltage between two of the terminals by applying an electric current or voltage to the third terminal. These three terminal characteristics of the transistor are what make it possible to build an amplifier for electrical signals, such as the one in radios and televisions. With the three-terminal transistor it is possible to construct an electric switch, which can be controlled by another electrical switch and by cascading these switches (switches that control switches that control switches, etc.) it is possible to build up very complicated logic circuits.

Today, these logic circuits are very compact and a silicon chip can contain up to 1,000,000 transistors per square centimetre. These switches can alternate between the on and the off states every 0.000000001 seconds and are the heart of today’s personal computer and many other electronic appliances seen every day.

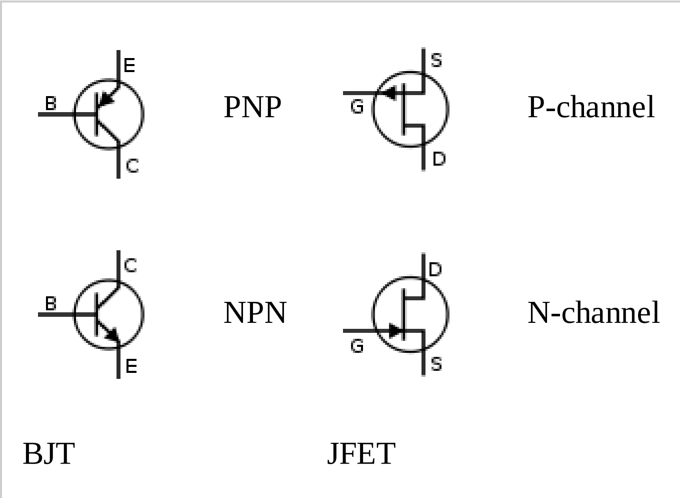
2.0 OBJECTIVES

After reading through this unit, you should be able to

1. Describe the similarities and the differences between transistor and vacuum tube triode
2. Distinguish between Bipolar Junction Transistors and Field Effect Transistors
3. Describe the operation of a transistor
4. Calculate the current gain of a Bipolar Junction Transistor
5. Draw and label the terminals of BJT and FET transistors
6. List major advantages and limitations of FET over BJT
7. Sketch simple diagrams of transistor as a switch and as an amplifier

3.0 MAIN CONTENT

3.1 Types of Transistors



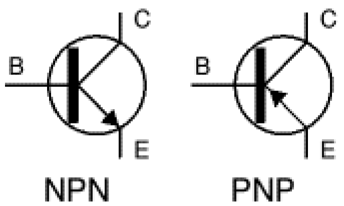
You should sketch, label and memorize these Bipolar Junction Transistor and Junction Field Effect Transistor symbols as you will come across them frequently.

Transistors can generally be classified into two. The first classification is the Bipolar Junction Transistors often referred to as BJT while the second is the Field Effect Transistors referred to as FET,

Bipolar Junction Transistors fall into two types which have slight differences in how they are used in a circuit. These are NPN PNP transistors. Do you remember that a bipolar transistor has terminals labelled base, collector and emitter, and that a small current at the base terminal (that is, flowing from the base to the emitter) can control or switch a much larger current between the collector and emitter terminals?

Now for a field-effect transistor, the terminals are labelled gate, source, and drain, and a voltage at the gate can control a current between source and the drain.

3.2 Bipolar junction transistor



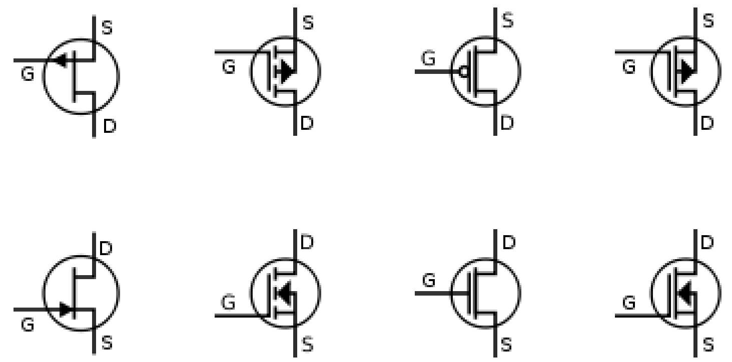
Bipolar transistors are so named because they conduct by using both majority and minority carriers. The bipolar junction transistor (BJT), the first type of transistor to be mass-produced, is a combination of two junction diodes, and is formed of either a thin layer of p-type semiconductor sandwiched between two n-type semiconductors (an n-p-n transistor), or a thin layer of n-type semiconductor sandwiched between two p-type semiconductors (a p-n-p transistor). This construction produces two p-n junctions: a base–emitter junction and a base–collector junction, separated by a thin region of semiconductor known as the base region (two junction diodes wired together without sharing an intervening semiconducting region will not make a transistor).

The BJT has three terminals, corresponding to the three layers of semiconductor - an emitter, a base, and a collector. It is useful in amplifiers because the currents at the emitter and collector are controllable by a relatively small base current." In an NPN transistor operating in the active region, the emitter-base junction is forward biased (electrons and holes recombine at the junction), and electrons are injected into the base region. Because the base is narrow, most of these electrons will diffuse into the reverse-biased (electrons and holes are formed at, and move away from the junction) base-collector junction and be swept into the collector; perhaps one-hundredth of the electrons will recombine in the base, which is the dominant mechanism in the base current. By controlling the number of electrons that can leave the base, the number of electrons entering the collector can be controlled. Collector current is approximately β (common-emitter current gain) times the base current. It is typically greater than 100 for small-signal transistors but can be smaller in transistors designed for high-power applications.

Unlike the FET, the BJT is a low–input-impedance device. Also, as the base–emitter voltage (V_{be}) is increased the base–emitter current and hence the collector–emitter current (I_{ce}) increase exponentially according to the Shockley diode model and the Ebers-Moll model. Because of this exponential relationship, the BJT has a higher transconductance than the FET.

Bipolar transistors can be made to conduct by exposure to light, since absorption of photons in the base region generates a photocurrent that acts as a base current; the collector current is approximately β times the photocurrent. Devices designed for this purpose have a transparent window in the package and are called phototransistors.

3.3 Field-effect transistor



All the transistors shown above are Field Effect Transistors. Try to sketch, label and identify each as you will also come across them in future.

The field-effect transistor (FET), sometimes called a unipolar transistor, uses either electrons (in N-channel FET) or holes (in P-channel FET) for conduction. The four terminals of the FET are named source, gate, drain, and body (substrate). On most FETs, the body is connected to the source inside the package, and this will be assumed for the following description.

In FETs, the drain-to-source current flows via a conducting channel that connects the source region to the drain region. The conductivity is varied by the electric field that is produced when a voltage is applied between the gate and source terminals; hence the current flowing between the drain and source is controlled by the voltage applied between the gate and source. As the gate–source voltage (V_{gs}) is increased, the drain–source current (I_{ds}) increases exponentially for V_{gs} below threshold, and then at a roughly quadratic rate ($I_{ds} \propto (V_{gs} - V_T)^2$) (where V_T is the threshold voltage at which drain current begins) in the "space-charge-limited" region above threshold. A quadratic behaviour is not observed in modern devices, for example, at the 65 nm technology node.

For low noise at narrow bandwidth the higher input resistance of the FET is advantageous.

FETs are divided into two families: junction FET (JFET) and insulated gate FET (IGFET). The IGFET is more commonly known as a metal–oxide–semiconductor FET (MOSFET), reflecting its original construction from layers of metal (the gate), oxide (the insulation), and semiconductor. Unlike IGFETs, the JFET gate forms a PN diode with the channel which lies between the source and drain. Functionally, this makes the N-channel JFET the solid state equivalent of the vacuum tube triode which, similarly, forms a diode between its grid and cathode. Also, both devices operate in the depletion mode, they both have a high input impedance, and they both conduct current under the control of an input voltage.

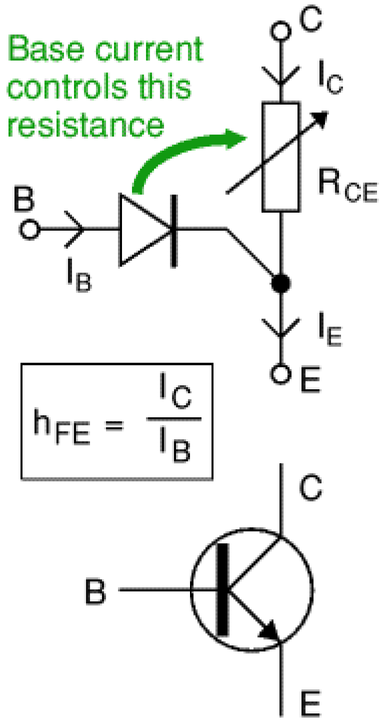
Metal–semiconductor FETs (MESFETs) are JFETs in which the reverse biased PN junction is replaced by a metal–semiconductor Schottky-junction. These, and the HEMTs (high electron mobility transistors, or HFETs), in which a two-dimensional electron gas with very high carrier mobility is used for charge transport, are especially suitable for use at very high frequencies (microwave frequencies; several GHz).

Unlike bipolar transistors, FETs do not inherently amplify a photocurrent. Nevertheless, there are ways to use them, especially JFETs, as light-sensitive devices, by exploiting the photocurrents in channel–gate or channel–body junctions.

FETs are further divided into depletion-mode and enhancement-mode types, depending on whether the channel is turned on or off with zero gate-to-source voltage. For enhancement mode, the channel is off at zero bias, and a gate potential can "enhance" the conduction. For depletion mode, the channel is on at zero bias, and a gate potential (of the opposite polarity) can "deplete" the channel, reducing conduction. For either mode, a more positive gate voltage corresponds to a higher current for N-channel devices and a lower current for P-channel devices. Nearly all JFETs are depletion-mode as the diode junctions would forward bias and conduct if they were enhancement mode devices; most IGFETs are enhancement-mode types.

3.4 Transistor Operation

A transistor is a semiconductor device used to amplify and switch electronic signals. It is made of a solid piece of semiconductor material, with at least three terminals for connection to an external circuit. A voltage or current applied to one pair of the transistor's terminals changes the current flowing through another pair of terminals. Because the controlled (output) power can be much more than the controlling (input) power, the transistor provides amplification of a signal. Today, some transistors are packaged individually, but many more are found embedded in integrated circuits.



The operation of a transistor is difficult to explain and understand in terms of its internal structure. It is more helpful to use this functional model which presupposes that the base-emitter junction behaves like a diode.

A base current I_B flows only when the voltage V_{BE} across the base-emitter junction is 0.7V or more for Silicon transistors.

The small base current I_B controls the large collector current Ic.

$Ic = h_{FE} \times I_B$ (unless the transistor is fully saturated)

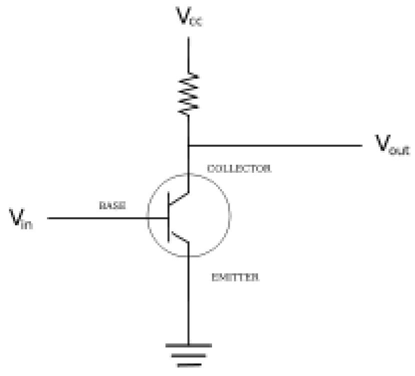
h_{FE} is the direct current gain and is typically 100

The collector-emitter resistance R_{CE} is controlled by the base current I_B:

$I_B = 0$ R_{CE} = infinity transistor off

I_B small R_{CE} reduced transistor partly on

I_B increased $R_{CE} = 0$ transistor fully on (saturated)

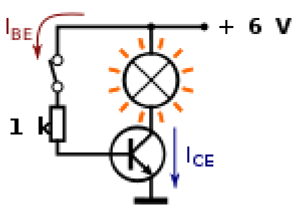


This simple circuit above shows you the Emitter, Base and Collector which are the labels of a bipolar transistor

The usefulness of a transistor is as a result of its ability to use a small signal applied between one pair of its terminals to control a much larger signal at another pair of terminals. This is called gain and it can be a voltage or a current gain. A transistor can control its output in proportion to the input signal which means it acts as an amplifier.

Alternatively, you can use the transistor to turn current on or off in a circuit whereby it acts as an electrically controlled switch.

Transistor as a switch



Bipolar Junction Transistor can be used as an electronic switch as shown above, Transistors are commonly used as electronic switches, for both high power applications including switched-mode power supplies and low power applications such as logic gates.

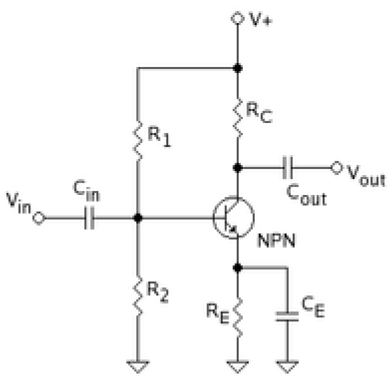
In the diagram above, as the base voltage rises the base and collector current rise, and the collector voltage drops because of the collector resistor.

Let us take a look at the relevant equations which describe this process:

> $V_{RC} = I_{CE} \times R_C$, the voltage across the resistor (which might be a lamp with resistance R_C)
>
> $V_{RC} + V_{CE} = V_{CC}$, the supply voltage (shown as 6Volts in the diagram above)

If V_{CE} could fall to 0 (as in a perfect closed switch) then Ic will have an upper limit of V_{CC} / R_C, even when base current is increased after the transistor collector current has reached its maximum. At this point the transistor is saturated. We can therefore choose values of input voltage such that the output is either completely off, or completely on. When this happens the transistor acts as a switch, a mode of operation which is common in digital circuits where only "on" and "off" values are desired.

Transistor as an amplifier



We will use this transistor amplifier circuit in the common-emitter configuration to demonstrate the transistor as an amplifier.

The common-emitter amplifier is desirable where a small change in input voltage (V_{in}) changes the small current through the base of the transistor and the transistor's current amplification combined with the properties of the circuit mean that small swings in V_{in} produce large changes in V_{out}.

Various configurations of single transistor amplifier are possible, with some providing current gain, some voltage gain, and some both and from mobile phones to televisions, vast numbers of products include amplifiers for sound reproduction, radio transmission, and signal processing.

Advantages, Limitations and Categories of Transistors

Before you proceed, you should list five advantages and five disadvantages of transistors over vacuum tube diode valves.

Have you done so?

Now let us list the key advantages that have allowed transistors to replace their vacuum tube predecessors in most applications:

- Small size and minimal weight, allowing the development of miniaturized electronic devices.
- Highly automated manufacturing processes, resulting in low per-unit cost.
- Lower possible operating voltages, making transistors suitable for small, battery-powered applications.
- No warm-up period for cathode heaters required after power application.
- Lower power dissipation and generally greater energy efficiency.
- Higher reliability and greater physical ruggedness.
- Extremely long life. Some transistorized devices have been in service for more than 50 years.
- Complementary devices available, facilitating the design of complementary-symmetry circuits, something not possible with vacuum tubes.

- Insensitivity to mechanical shock and vibration, thus avoiding the problem of micro phonics in audio applications.

Now let us consider those attributes which have set transistors behind their vacuum tube predecessors in terms of performance:

- Silicon transistors do not operate at voltages higher than about 1,000 volts (SiC devices can be operated as high as 3,000 volts). In contrast, electron tubes have been developed that can be operated at tens of thousands of volts.

- High power, high frequency operation, such as that used in over-the-air television broadcasting, is better achieved in electron tubes due to improved electron mobility in a vacuum.

- Silicon transistors are much more vulnerable than electron tubes to an electromagnetic pulse generated by a high-altitude nuclear explosion.

And finally, in summary, we will categorize transistor by the following:

- Semiconductor material: germanium, silicon, gallium arsenide, silicon carbide, etc.

- Structure: BJT, JFET, IGFET (MOSFET), IGBT, "other types"

- Polarity: NPN, PNP (BJTs); N-channel, P-channel (FETs)

- Maximum power rating: low, medium, high

- Maximum operating frequency: low, medium, high, radio frequency (RF), microwave (The maximum effective frequency of a transistor is denoted by the term f_T, an abbreviation for "frequency of transition". The frequency of transition is the frequency at which the transistor yields unity gain).

- Application: switch, general purpose, audio, high voltage, super-beta, matched pair

- Physical packaging: through hole metal, through hole plastic, surface mount, ball grid array, power modules
- Amplification factor h_{fe} (transistor beta)

- Thus, a particular transistor may be described as silicon, surface mount, BJT, NPN, low power, high frequency switch.

4.0 CONCLUSION

In this concluding Unit you have seen the reason why transistors have over the years replaced Vacuum tubes for most applications. You will appreciate the fact that for most of these applications however, the Vacuum tube's electrical and performance characteristics often exceed those of their solid state counterparts.

You can recognize the symbols of many transistors and can classify them into Field Effect Transistors and Bipolar Junction Transistors. Aside from further classifying into PNP/NPN, and JGFET/IGFET respectively, you now know how transistors operate.

Functionally the transistor is discriminated as either as a switch or as an amplifier and I am sure you can by now sketch simple transistor diagrams to illustrate these. Also you can carry oft simple calculations for bias and saturation.

You will be able to state several advantages of transistors over Vacuum tube valves as well as many corresponding disadvantages and finally, you can classify transistors by many of their common attributes including Transistor type, Transistor casing, Transistor application and Semiconductor material; all of which contribute to the classifications which you will later on become familiar with in components' specification sheets and data books.

5.0 SUMMARY

- Transistors are three terminal, solid state electronic devices in which it is possible to control electric current or voltage between two of the terminals by applying an electric current or voltage to the third terminal.

- Logic circuits are very compact and a silicon chip can contain up to 1,000,000 transistors per square centimetre which can alternate between the on and the off states every 0.000000001 seconds.

They are the heart of personal computer and many other electronic appliances.

- Transistors can be classified into Bipolar Junction Transistors (BJT) or Field Effect Transistors (FET). Bipolar Junction Transistors can be further classified into NPN and PNP while Field Effect Transistors are subdivided into P or N channel, and JFET or MOSFET.

- In operation, transistors are semiconductor device used to amplify and switch electronic signals and to understand the operation is helpful to use a functional model which presupposes that the base-emitter junction behaves like a diode.

- Transistors have many advantages over valves some of which are small size and minimal weight, allowing the development of miniaturized electronic devices. Highly automated manufacturing processes, resulting in low per-unit cost. Lower operating voltages, making transistors suitable for small, battery-powered applications. No warm-up period for cathode heaters required after power application. Low power dissipation and greater efficiency. Higher reliability and greater physical ruggedness. Extremely long life. Some transistorized devices have been in service for more than 50 years. Complementary devices are available, facilitating the design of complementary-symmetry circuits, something not possible with vacuum tubes. Insensitivity to mechanical shock and vibration, thus avoiding the problem of micro phonics in audio applications.

6.0 TUTOR MARKED ASSIGNMENTS

1. Sketch and label a PNP BJT and a P channel MOSET transistor

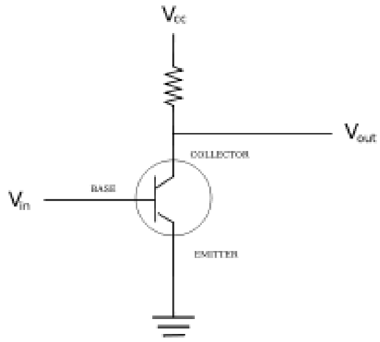
2. List two primary applications of transistors

3. What is the ratio of the change in collector current to the change in base current called?

4. What are those factors that make transistors attractive over valves?

5. Sketch a transistor amplifier circuit in the common emitter configuration and give hypothetical values to the passive components.

6. In the sketch below, the supply voltage is 12 volts while the resistor connected to the collector is 10 Kilo Ohms. We wish to use this transistor as a switch which means we must drive it to saturation. If the current gain of this transistor is 100, then what must the minimum base current be to drive the transistor collector to saturation?



7. Explain enhancement and depletion mode of operation of Field Effect Transistors

8. How does a phototransistor work?

9. List five disadvantages of transistors when compared with valves.

10. Why is it helpful to use a functional model which presupposes that the base-emitter junction behaves like a diode in understanding the operation of transistors?

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Semiconductor Device Fundamentals
By Robert F. Pierret Published by Prentice Hill

UNIT 1 RESONANT CIRCUITS AND PASSIVE FILTERS

CONTENTS

1.0 INTRODUCTION

LC circuit are resonant circuit or tuned circuit that consists of an inductor, and a capacitor,. When we connect them together, an electric current reaches a maximum at the circuit's resonant frequency. We can use LC circuits to either generate signals at a particular frequency, or for picking out a signal at a particular frequency from a more complex mix of signals. They are the key components in several applications such as oscillators, filters, tuners and frequency mixers and LC circuits are an idealization since it is assumed that there is no dissipation of energy due to resistance. The more practical assumption is that we adopt a model incorporating resistance which dissipates some energy because no practical circuit exists without losses, In order to gain a good understanding of the process, we shall study LC circuits in the pure form which assumes lossless elements.

Take a look at this picture of a Television signal splitter consisting of a high-pass filter on the left and a low-pass filter on the right. This is a practical example of filters which you encounter every day. Let us find out more about passive filters in particular as we progress in this Unit.

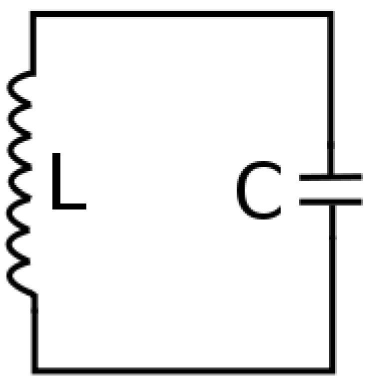
2.0 OBJECTIVES

After reading through this unit, you should be able to

1. Describe an LC Circuit
2. Explain Resonance
3. Understand the Energy Exchange Mechanism In LC Circuits
4. Calculate the Frequency Of Oscillation Of A Resonant Circuit
5. Relate the "Q" Factor To Frequency Selectivity
6. Know Why Unsustained Oscillations Die Down Asymptotically
7. Work with both Parallel and Series LC Circuits
8. Use LC Circuits As Frequency Band Pass or Rejector Filters
9. Distinguish Between Active and Passive Filters
10. Identify Filter Topology
12. Calculate Filter Response To Frequency.

3.0 MAIN CONTENT

3.1 LC Circuit Operation



An LC circuit can store electrical energy vibrating at its resonant frequency. A capacitor stores energy in the electric field between its plates, depending on the voltage across it, and an inductor stores energy in its magnetic field, depending on the current through it.

If a charged capacitor is connected across an inductor, charge will start to flow through the inductor, building up a magnetic field around it, and reducing the voltage on the capacitor. Eventually all the charge on the capacitor will be gone and the voltage across it will reach zero. However, the current will continue, because inductors resist changes in current, and energy to keep it flowing is extracted from the magnetic field, which will begin to decline. The current will begin to charge the capacitor with a voltage of opposite polarity to its original charge. When the magnetic field is completely dissipated the current will stop and the charge will again be stored in the capacitor, with the opposite polarity as before. Then the cycle will begin again, with the current flowing in the opposite direction through the inductor.

The charge flows back and forth between the plates of the capacitor, through the inductor. The energy oscillates back and forth between the capacitor and the inductor until (if not replenished by power from an external circuit) internal resistance makes the oscillations die out. Its action, known mathematically as a harmonic oscillator, is similar to a pendulum swinging back and forth, or water moving back and forth in a tank and for this reason, the LC circuit is also called a tank circuit. The oscillation frequency is determined by the capacitance and inductance values used. In typical tuned circuits in electronic equipment the oscillations are very fast, thousands to millions of times per second.

Time domain solution

By Kirchhoff's voltage law, the voltage across the capacitor, V_C, must equal the voltage across the inductor, V_L:

$$V_C = V_L.$$

Likewise, by Kirchhoff's current law, the current through the capacitor plus the current through the inductor must equal zero:

$$i_C + i_L = 0.$$

From the constitutive relations for the circuit elements, we also know that

$$V_L(t) = L\frac{di_L}{dt}$$

and

$$i_C(t) = C\frac{dV_C}{dt}.$$

Rearranging and substituting gives the second order differential equation

$$\frac{d^2i(t)}{dt^2} + \frac{1}{LC}i(t) = 0.$$

The parameter ω, the radian frequency, can be defined as: $\omega = (LC)^{-1/2}$. Using this can simplify the differential equation

$$\frac{d^2i(t)}{dt^2} + \omega^2 i(t) = 0.$$

The associated polynomial is $s^2 + \omega^2 = 0$, thus

$$s = +j\omega$$

or

$$s = -j\omega$$

where j is the imaginary unit.

Thus, the complete solution to the differential equation is

$$i(t) = Ae^{+j\omega t} + Be^{-j\omega t}$$

and can be solved for A and B by considering the initial conditions.

Since the exponential is complex, the solution represents a sinusoidal alternating current.

If the initial conditions are such that A = B, then we can use Euler's formula to obtain a real sinusoid with amplitude 2A and angular frequency $\omega = (LC)^{-1/2}$.

Thus, the resulting solution becomes:

$$i(t) = 2A\cos(\omega t).$$

The initial conditions that would satisfy this result are:

$$i(t=0) = 2A$$

3.2 Resonance effect

The resonance effect occurs when inductive and capacitive reactances are equal in absolute value. (Notice that the LC circuit does not, by itself, resonate. The word resonance refers to a class of phenomena in which a small driving perturbation gives rise to a large effect in the system. The LC circuit must be driven, for example by an AC power supply, for resonance to occur (below).) The frequency at which this equality holds for the particular circuit is called the resonant frequency. The resonant frequency of the LC circuit is

$$\omega = \sqrt{\frac{1}{LC}}$$

where L is the inductance in Henries, and C is the capacitance in Farads. The angular frequency ωhas units of radians per second.

The equivalent frequency in units of hertz is

$$f = \frac{\omega}{2\pi} = \frac{1}{2\pi\sqrt{LC}}.$$

LC circuits are often used as filters; the L/C ratio is one of the factors that determines their "Q" and so selectivity. For a series resonant circuit with a given resistance, the higher the inductance and the lower the capacitance, the narrower the filter bandwidth. For a parallel resonant circuit the opposite applies. Positive feedback around the tuned circuit ("regeneration") can also increase selectivity (see Q multiplier and Regenerative circuit).

Stagger tuning can provide an acceptably wide audio bandwidth, yet good selectivity.

Series LC circuit Resonance

Here L and C are connected in series to an AC power supply. Inductive reactance magnitude (X_L) increases as frequency increases while capacitive reactance magnitude (X_C) decreases with the increase in frequency. At a particular frequency these two reactances are equal in magnitude but opposite in sign. The frequency at which this happens is the resonant frequency (f_r) for the given circuit.

Hence, at f_r :

$$X_L = -X_C$$

$$\omega L = \frac{1}{\omega C}$$

Converting angular frequency into hertz we get

$$2\pi f L = \frac{1}{2\pi f C}$$

Here f is the resonant frequency. Then rearranging,

$$f = \frac{1}{2\pi\sqrt{LC}}$$

In a series AC circuit, X_C leads by 90 degrees while X_L lags by 90. Therefore, they cancel each other out. The only opposition to a current is coil resistance. Hence in series resonance the current is maximum at resonant frequency.

At f_r, current is maximum. Circuit impedance is minimum. In this state a circuit is called an acceptor circuit.

Below f_r, $X_L \ll (-X_C)$. Hence circuit is capacitive.

Above f_r, $X_L \gg (-X_C)$. Hence circuit is inductive.

Impedance

First consider the impedance of the series LC circuit. The total impedance is given by the sum of the inductive and capacitive impedances:

$Z = Z_L + Z_C$

By writing the inductive impedance as $Z_L = j\omega L$ and capacitive impedance as $Z_C = (j\omega C)^{-1}$ and substituting we have

$$Z = j\omega L + \frac{1}{j\omega C}.$$

Writing this expression under a common denominator gives

$$Z = \frac{(\omega^2 LC - 1)j}{\omega C}.$$

The numerator implies that if $\omega^2 LC = 1$ the total impedance Z will be zero and otherwise non-zero. Therefore the series LC circuit, when connected in series with a load, will act as a band-pass filter having zero impedance at the resonant frequency of the LC circuit.

Parallel LC circuit Resonance

Here a coil (L) and capacitor (C) are connected in parallel with an AC power supply. Let R be the internal resistance of the coil. When X_L equals X_C, the reactive branch currents are equal and opposite. Hence they cancel out each other to give minimum current in the main line. Since total current is minimum, in this state the total impedance is maximum.

Resonant frequency given by: $f = \frac{1}{2\pi\sqrt{LC}}.$

Note that any reactive branch current is not minimum at resonance, but each is given separately by dividing source voltage (V) by reactance (Z). Hence I=V/Z, as per Ohm's law.

At f_r, line current is minimum. Total impedance is maximum. In this state a circuit is called a rejector circuit.

Below f_r, circuit is inductive.

Above f_r,circuit is capacitive.

Impedance

The same analysis may be applied to the parallel LC circuit. The total impedance is then given by:

$$Z = \frac{Z_L Z_C}{Z_L + Z_C}$$

and after substitution of Z_L and Z_C and simplification, gives

$$Z = \frac{-j\omega L}{\omega^2 LC - 1}.$$

Note that

$$\lim_{\omega^2 LC \to 1} Z = \infty$$

but for all other values of ω^2LC the impedance is finite (and therefore less than infinity). Hence the parallel LC circuit connected in series with a load will act as band-stop filter having infinite impedance at the resonant frequency of the LC circuit.

Applications of Resonance Effect

Most common application is tuning. For example, when we tune a radio to a particular station, the LC circuits are set at resonance for that particular carrier frequency.

A series resonant circuit provides voltage magnification.

A parallel resonant circuit provides current magnification.

A parallel resonant circuit can be used as load impedance in output circuits of RF amplifiers. Due to high impedance, the gain of amplifier is a maximum at resonant frequency.

LC circuits behave as electronic resonators, which are a key component in such applications as Oscillators, Filters, Tuners, Mixers, Foster-Seeley discriminator, Contactless cards, Graphics tablets and Electronic Article Surveillance (Security Tags)

3.3 Electronic filters

Do you know that electronic filters are electronic circuits which perform signal processing functions, specifically when you want to remove unwanted frequency components from signals, to enhance wanted signals, or when you want to do both? You will discover that the family of electronic filters is quite a large one; comprising passive, active, analog, digital, High-pass, low-pass, band pass, band-reject, all-pass, discrete-time, continuous-time, linear, non-linear, infinite impulse response and finite impulse response. You must agree that this is quite a large family when you consider that each member comprises many variants.

Most common types of electronic filters are linear filters regardless of other aspects of their design. For the purpose of our study of filters we will limit our discussion to passive filters which incidentally represent the oldest forms of electronic filters. They are passive analog linear filters, constructed using only resistors and capacitors or resistors and inductors. These are known as RC and RL single-pole filters respectively. More complex multipole LC filters have also existed for many years, and their operation is well understood.

3.4 Passive filters

Passive implementations of linear filters are based on combinations of resistors (R), inductors (L) and capacitors (C). These types are collectively known as passive filters, because they do not depend upon an external power supply and/ they do not contain active components such as transistors.

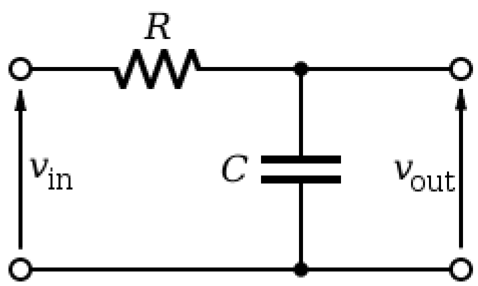
Inductors block high-frequency signals and conduct low-frequency signals, while capacitors do the reverse. A filter in which the signal passes through an inductor, or in which a capacitor provides a path to ground, presents less attenuation to low-frequency signals than high-frequency signals and is a low-pass filter. If the signal passes through a capacitor, or has a path to ground through an inductor, then the filter presents less attenuation to high-frequency signals than low-frequency signals and is a high-pass filter. Resistors on their own have no frequency-selective properties, but are added to inductors and capacitors to determine the time-constants of the circuit, and therefore the frequencies to which it responds.

The inductors and capacitors are the reactive elements of the filter. The number of elements determines the order of the filter. In this context, an

LC tuned circuit being used in a band-pass or band-stop filter is considered a single element even though it consists of two components.

At high frequencies (above about 100 megahertz), sometimes the inductors consist of single loops or strips of sheet metal, and the capacitors consist of adjacent strips of metal. These inductive or capacitive pieces of metal are called stubs.

Single element types



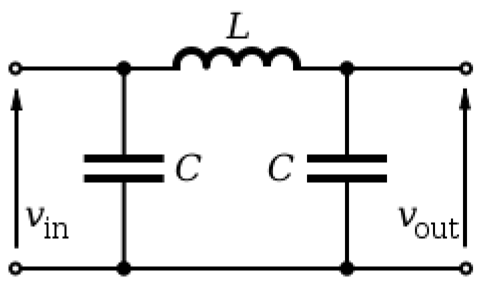
This is a low-pass electronic filter which you can realize by an RC circuit

The simplest passive filters, RC and RL filters, include only one reactive element, except hybrid LC filter which is characterized by inductance and capacitance integrated in one element.

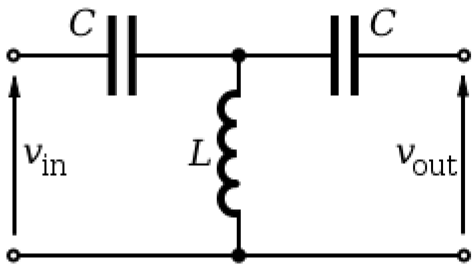
L filter

An L filter consists of two reactive elements, one in series and one in parallel.

T and π filters



This Low-pass π filter has the topology of the symbol "π"



You can see that this High-pass T filter actually has the topology of the letter "T"

Three-element filters can have a 'T' or 'π' topology and in either geometries, a low-pass, high-pass, band-pass, or band-stop characteristic is possible. The components can be chosen symmetric or not, depending on the required frequency characteristics. The high-pass T filter in the diagram above has very low impedance at high frequencies, and a very high impedance at low frequencies. That means that it can be inserted in a transmission line, resulting in the high frequencies being passed and low frequencies being reflected. Likewise, for the illustrated low-pass π filter, the circuit can be connected to a transmission line, transmitting low frequencies and reflecting high frequencies. Using m-derived filter sections with correct termination impedances, the input impedance can be reasonably constant in the pass band.

Multiple element types

Multiple element filters are usually constructed as a ladder network. These can be seen as a continuation of the L,T and π designs of filters. More elements are needed when it is desired to improve some parameter of the filter such as stop-band rejection or slope of transition from pass-band to stop-band.

4.0 CONCLUSION

We learnt about Resonant circuits and Passive filters in his unit. You now know that LC circuits can be either parallel or serial connection and that the idealization of lossless capacitance and inductance are far from what is practically attainable. You know that this is because of the inductor coil resistance and capacitive leakages.

Many of the consequences of resonance, both the desirable and the undesirable are recalled and you can now distinguish between resonance and sustained oscillation which requires a replenishment of energy lost through resistance in the LC tank circuit.

Having been exposed to the very large family of filter types, you can state what a passive filter is and determine the frequency response of low pass, high pass and band pass filters. Finally you can now sketch the important filter topologies and you know the electrical elements which determine filter response.

5.0 SUMMARY

- An LC circuit is a tuned circuit which when connected together will cause an electric current to reach its maximum value at the resonant frequency.

- LC resonant circuits are key components in several applications such as oscillators, filters, tuners and frequency mixers and LC circuits and are an idealization since we assume that there is no dissipation of energy due to resistance.

- Resonance occurs when inductive and capacitive reactances are equal in absolute value.

- Resonance frequency of LC circuit elements is given by

$$f = \frac{1}{2\pi\sqrt{LC}}$$

- Electronic filters are used to remove unwanted frequency components from signals, to enhance wanted signals, or to do both.

- Passive linear filters are based on combinations of resistors, inductors and capacitors which do not depend upon an external power supply and/do not contain active components such as transistors.

- An L filter consists of two reactive elements, one in series and one in parallel and can either be low pass or high pass filter..

- Three-element filters can have a 'T' or 'π' topology and in either geometries, a low-pass, high-pass, band-pass, or band-stop characteristic is possible. The components chosen can be symmetric or not, depending on desired frequency response.

- Multiple element filters are usually constructed as a ladder network.

6.0 TUTOR MARKED ASSIGNMENTS

1. What is resonance and how can it be achieved with circuit elements

2. How does an LC network attain oscillation

3. Is there any difference in the resonant frequency if a given inductor and capacitor initially connected in series are now connected in parallel? Prove your answer

4. What is the relationship between the reactance of the inductor and the capacitor in ac LC circuit at resonance?

5. Give 5 practical applications of LC circuits.

6. How would you describe passive filters?

7. What electrical components are passive filters made of?

8. Using the frequency domain, explain how low pass, high pass and band pass filters work?

9. Sketch a single element High Pass filter.

10. Can you explain why an “L” topology filter can never be a band pass filter?

7.0 REFERENCES/FURTHER READINGS

Analog Filter Design
By M. E. Van Valkenburg Published by Holt, Rinehart and Winston

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

UNIT 2 ATTENUATORS AND IMPEDANCE MATCHING

CONTENTS

1.0 INTRODUCTION

An attenuator is an electronic device that reduces the amplitude or power of a signal without appreciably distorting its waveform and they are usually passive devices made from simple voltage divider networks.

Impedance matching schemes were originally developed for electrical power, but it can also be applied to any other field where a form of energy (not necessarily electrical) is transferred between a source and a load.

An alternative to impedance matching is impedance bridging, where the load impedance is chosen to be much larger than the source impedance and where maximizing voltage transfer, rather than maximum power transfer, is the goal.

2.0 OBJECTIVES

After reading through this unit, you will

1 Understand what attenuators are and what they do
2 Be able to sketch the basic attenuator topologies
3 Be able to distinguish between balanced and unbalanced attenuators
4 Know how to qualify attenuators by their specifications

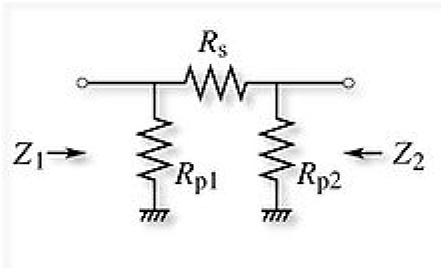
5 Easily explain the meaning of impedance matching
6 Know the significance of complex conjugate in complex load matching
7 Be able to explain reflectionless impedance matching and maximum power transfer
8 Easily identify impedance matching devices in the real world
9 Recognise the role of the Smith Chart in Transmission Line matching networks

3.0 MAIN CONTENT

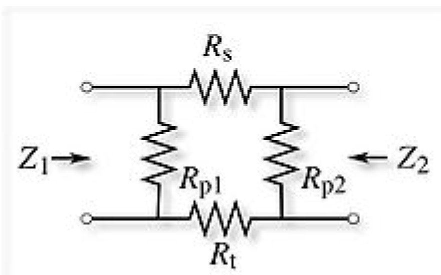
3.1 Attenuators

Fixed attenuators in circuits are used to lower voltage, dissipate power, and to improve impedance matching. In measuring signals, attenuator pads or adaptors are used to lower the amplitude of the signal a known amount to enable measurements, or to protect the measuring device from signal levels that might damage it. Attenuators are also used to 'match' impedances by lowering apparent Standing Wave Ratio.

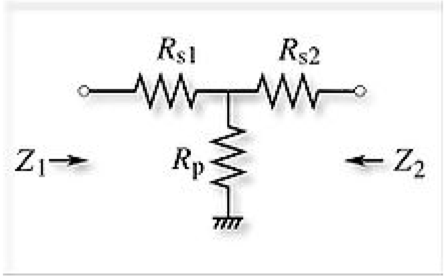
These are four typical attenuator circuits



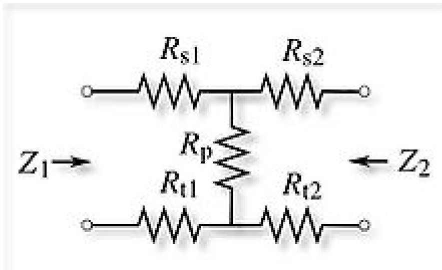
As the topology implies, you can see that this is a π-type unbalanced attenuator circuit Unbalanced because it is not totally symmetrical.



This π-type attenuator circuit is balanced. Can you say why?



This is another one. It is a T-type attenuator circuit. Is it balanced or unbalanced? Yes you guessed right. It is unbalanced because of its asymmetry.



This is a T-type balanced attenuator circuit

Basic circuits used in attenuators are pi pads (π-type) and T pads. These may be required to be balanced or unbalanced networks depending on whether the line geometry with which they are to be used is balanced or unbalanced. Attenuators used with coaxial lines are the unbalanced form while attenuators for use with twisted pair are required to be the balanced form.

Four fundamental attenuator circuit diagrams are given in the figures above. Since an attenuator circuit consists solely of passive resistor elements, it is linear and reciprocal. If the circuit is also made symmetrical which is usually the case since it is often required that the input and output impedances Z_1 and Z_2 are equal, then the input and output ports are not distinguished, but by convention the left and right sides of the circuits are referred to as input and output.

3.2 Attenuator characteristics

Do you recognise the device in this picture; it is a Radio Frequency Microwave Attenuator

Attenuators Specifications

The following parameters are used in commercial attenuator specifications

Attenuation is expressed in decibels of relative power and you may roughly say that a 3dB attenuator reduces power to one half while a 6dB attenuator to one fourth, 10dB to one tenth, 20dB to one hundredth, 30dB to one thousandth and so on.

Frequency bandwidth is expressed in Hertz

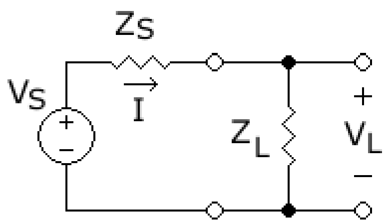
Power dissipation depends on mass and surface area of resistance material as well as possible additional cooling fins and heat sinks.

Standing Wave Ratio is the standing wave ratio for the input and the output ports

Other specifications include Accuracy and Repeatability

3.3 Impedance Matching

Impedance matching is the practice of designing the input impedance of an electrical load or the output impedance of its corresponding signal source in order to maximize the power transfer and minimize reflections from the load.



In the case of a complex source impedance Z_S and load impedance Z_L, matching is obtained when

$$Z_S = Z_L^*$$

where * indicates the complex conjugate.

Description

The term impedance is used for the resistance of a system to an energy source. For constant signals, this resistance can also be constant. For varying signals, it usually changes with frequency. The energy involved can be electrical, mechanical, magnetic or even thermal. The concept of electrical impedance is perhaps the most commonly known. Electrical impedance, like electrical resistance, is measured in ohms. In general, impedance has a complex value, which means that loads generally have a resistance to the source that is in phase with a sinusoidal source signal **and** reactance that is out of phase with a sinusoidal source signal. The total impedance (symbol: Z) is the vector sum of the resistance (symbol: R; a real number) and the reactance (symbol: X; an imaginary number).

In simple cases, such as low-frequency or direct-current power transmission, the reactance is negligible or zero and the impedance can be considered a pure resistance, expressed as a real number. In the following summary, we will consider the general case when the resistance and reactance are both significant, and also the special case in which the reactance is negligible.

Reflectionless matching

Impedance matching to minimize reflections and maximize power transfer over a (relatively) large bandwidth (also called reflectionless matching or broadband matching) is the most commonly used. To prevent all reflections of the signal back into the source, the load (which must be totally resistive) must be matched exactly to the source impedance (which again must be totally resistive). In this case, if a transmission line is used to connect the source and load together, it must also be the same impedance: $Z_{load} = Z_{line} = Z_{source}$, where Z_{line} is the characteristic impedance of the transmission line. Although source and load should each be totally resistive for this form of matching to work, the more general term 'impedance' is still used to describe the source and load characteristics. Any and all reactance actually present in the source or the load will affect the 'match'.

Complex conjugate matching

This is used in cases in which the source and load are reactive. This form of impedance matching can only maximize the power transfer between a reactive source and a reactive load at a single frequency. In this case,

$$Z_{load} = Z_{source}^{*}$$

(where * indicates the complex conjugate).

If the signals are kept within the narrow frequency range for which the matching network was designed, reflections (in this narrow frequency band only) are also minimized. For the case of purely resistive source and load impedances, all reactance terms are zero and the formula above reduces to

$$Z_{load} = Z_{source}$$

as would be expected.

3.4 Power transfer

Whenever a source of power with a fixed output impedance, such as an electric signal source, a radio transmitter, or even mechanical sound (e.g., a loudspeaker) operates into a load, the maximum possible power is delivered to the load when the impedance of the load (load impedance or input impedance) is equal to the complex conjugate of the impedance of the source (that is, its internal impedance or output impedance). For two impedances to be complex conjugates, their resistances must be equal, and their reactances must be equal in magnitude but of opposite sign.

In low-frequency or DC systems, or in systems with purely resistive sources and loads, the reactances are zero, or small enough to be ignored. In this case, maximum power transfer occurs when the resistance of the load is equal to the resistance of the source (see maximum power theorem for a proof).

Impedance matching is not always desirable. For example, if a source with a low impedance is connected to a load with a high impedance, then the power that can pass through the connection is limited by the higher impedance, but the electrical voltage transfer is higher and less prone to corruption than if the impedances had been matched. This maximum voltage connection is a common configuration called impedance bridging or voltage bridging and is widely used in signal processing. In such applications, delivering a high voltage (to minimize signal degradation during transmission and/or to consume less power by reducing currents) is often more important than maximum power transfer.

In older audio systems, reliant on transformers and passive filter networks, and based on the telephone system, the source and load resistances were matched at 600 ohms. One reason for this was to maximize power transfer, as there were no amplifiers available that could restore lost signal. Another reason was to ensure correct operation of the hybrid transformers used at central exchange equipment to separate outgoing from incoming speech so that these could be amplified or fed to a four-wire circuit. Most modern audio circuits, on the other hand, use active amplification and filtering, and they can use voltage bridging connections for best accuracy.

Strictly speaking, impedance matching only applies when both source and load devices are linear, however useful matching may be obtained between nonlinear devices with certain operating ranges.

3.5 Impedance matching devices

Adjusting the source impedance or the load impedance, in general, is called "impedance matching".

There are three possible ways to improve an impedance mismatch, all of which are called "impedance matching":

- devices intended to present an apparent load to the source of $R_{load} = R_{source}^*$ (complex conjugate matching). Given a source with a fixed voltage and fixed source impedance, the maximum power theorem says this is the only way to extract the maximum power from the source.

- devices intended to present an apparent load of $R_{load} = R_{line}$ (complex impedance matching), to avoid echoes. Given a transmission line source with a fixed source impedance, this "reflectionless impedance matching" at the end of the transmission line is the only way to avoid reflecting echoes back to the transmission line.

- devices intended to present an apparent source resistance as close to zero as possible, or presenting an apparent source voltage as high as possible. This is the only way to maximize energy efficiency, and so it is used at the beginning of electrical power lines. Such an impedance bridging connection also minimizes distortion and electromagnetic interference, and so it is also used in modern audio amplifiers and signal processing devices.

There are a variety of devices that are used between some source of energy and some load that perform "impedance matching".

To match electrical impedances, engineers use combinations of transformers, resistors, inductors, capacitors and transmission lines.

These passive and active impedance matching devices are optimized for different applications, and are called baluns, antenna tuners (sometimes called ATUs or roller coasters because of their appearance), acoustic horns, matching networks, and terminators.

Transformers

Transformers are sometimes used to match the impedances of circuits with different impedances. A transformer converts alternating current at one voltage to the same waveform at another voltage. The power input to the transformer and output from the transformer is the same (except for conversion losses). The side with the lower voltage is at low impedance, because this has the lower number of turns, and the side with the higher voltage is at a higher impedance as it has more turns in its coil.

Resistive network

Resistive impedance matches are easiest to design and can be achieved with a simple L pad consisting of only two resistors. Power loss is an unavoidable consequence of using resistive networks and they are consequently only usually used to transfer line level signals.

Stepped transmission line

Most lumped element devices can match a specific range of load impedance. For example, in order to match an inductive load into a real impedance, a capacitor needs be used. And if the load impedance becomes capacitive for some reason, the matching element must be replaced by an inductor. In many practical cases however, there is a need to use the same circuit to match a broad range of load impedance, thus simplify the circuit design. This issue was addressed by the stepped transmission line where multiple serially placed quarter wave dielectric slugs are used to vary transmission line's characteristic impedance. By controlling the position of each individual element, a broad range of load impedance can be matched without having to reconnect the circuit.

Some special situations - such as radio tuners and transmitters - use tuned filters such as stubs to match impedances at specific frequencies. These can distribute different frequencies to different places in the circuit.

In addition, there is the closely related idea of

- power factor correction devices intended to cancel out the reactive and nonlinear characteristics of a load at the end of a power line. This causes the load seen by the power line to be purely resistive. For a given true power required by a load, this minimizes the true current supplied through the power lines, and so minimizes the power wasted in the resistance of those power lines.

For example, a maximum power point tracker is used to extract the maximum power from a solar panel, and efficiently transfer it to batteries, the power grid, or other loads. The maximum power theorem applies to its "upstream" connection to the solar panel, so it emulates a load resistance equal to the solar panel source resistance. However, the maximum power theorem does not apply to its "downstream" connection, so that connection is an impedance bridging connection—it emulates a high-voltage, low-resistance source, to maximize efficiency.

L-section

One simple electrical impedance matching network requires one capacitor and one inductor. One reactance is in parallel with the source (or load) and the other is in series with the load (or source). If a reactance is in parallel with the source, the effective network matches from high impedance to low impedance. The L-section is inherently a narrowband matching network.

The analysis is as follows. Consider a real source impedance of R_1 and real load impedance of R_2. If a reactance X_1 is in parallel with the source impedance, the combined impedance can be written as:

$$\frac{jR_1X_1}{R_1 + jX_1}$$

If the imaginary part of the above impedance is completely canceled by the series reactance, the real part is

$$R_2 = \frac{R_1X_1^2}{R_1^2 + X_1^2}$$

Solving for X_1

$$X_1 = \sqrt{\frac{R_2R_1^2}{R_1 - R_2}}$$

If $R_1 \gg R_2$the above equation can be approximated as

$$X_1 \approx \sqrt{R_1R_2}$$

The inverse connection, impedance step up, is simply the reverse, e.g. reactance in series with the source. The magnitude of the impedance ratio is limited by reactance losses such as the Q of the inductor. Multiple L-sections can be wired in cascade to achieve higher impedance ratios or greater bandwidth. Transmission line matching networks can be modelled as infinitely many L-sections wired in cascade. You can attain optimal matching circuits can be designed for a particular system with the use of the Smith chart.

4.0 CONCLUSION

As you have read this Unit, I presume that you now know what an attenuator is and what attenuators do. Similarly, you now recognize the practical significance of impedance matching and its role in rower transfer; particularly in transmission lines, coaxial cable systems and waveguides. No doubt you can now identify and mention impedance matching devices you see every day; such as your neighborhood or community transformer.

5.0 SUMMARY

- Attenuators are electronic devices that reduce the amplitude or power of a signal without appreciably distorting its waveform .
- Impedance matching is used to maximize power transfer
- Fixed attenuators in circuits are used to lower voltage, dissipate power, and to improve impedance matching
- Attenuators can either be balanced or unbalanced
- Attenuators are specified by their attenuation, frequency, power dissipation, standing wave ratio, accuracy and repeatability
- Impedance matching which is used to minimize reflections and maximize power transfer over a relatively large bandwidth is called reflectionless matching or broadband matching.
- Complex conjugate impedance matching can only maximize the power transfer between a reactive source and a reactive load at a single frequency.
- Common impedance matching devices are Transmission Lines, Transformers, Resistive networks, Waveguides and Transmission Lines.
- The Smith Chart is used to design Transmission Lines for optimal impedance matching

6.0 TUTOR MARKED ASSIGNMENTS

1. What is impedance matching?

2. How is a Smith Chart useful in impedance matching

3. Sketch and label two Balanced Attenuator circuits

4. Which is balanced and which is unbalanced between a coaxial transmission line, a twisted pair transmission line and a parallel pair transmission line?

5. When do you apply complex conjugate in impedance matching?

6. By what performance parameters do you qualify attenuators?

7. What is reflectionless impedance matching? Give one physical application.

8 Describe how a transformer matches impedance between a load and a source

9. In what ways are attenuation and impedance matching related?

7.0 REFERENCES/FURTHER READINGS

Electronic Devices and Circuit Theory 7th Edition
By Robert E. Boylestad and Louis Nashesky Published by Prentice Hall

Network Analysis with Applications 4th Edition
By William D. Stanley Published by Prentice Hall

Fundamentals of Electric Circuits 4th Edition
By Alexander and Sadiku Published by Mc Graw Hill

Semiconductor Device Fundamentals
By Robert F. Pierret Published by Prentice Hill

Electrical Circuit Analysis
By C. L. Wadhwa Published by New Age International

NATIONAL OPEN UNIVERSITY OF NIGERIA

Form QST1
Questionnaire

Dear Student,

While studying the units of this course, you may have found certain portions of the text difficult to comprehend . We wish to know your difficulties and suggestions, in order to improve the course. Therefore we request you fill out and send us the following questionnaire, which pertains to this course. If you find the space provided insufficient, kindly use additional sheet.

Course Code ________________ Course Title ________________________________

1. How many hours did you need for studying each of these units?

Unit	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
No. Of Hours															
Unit	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
No. Of Hours															

2. Which of these units do you find most difficult to understand?

3. Please give specific problems you find difficult with the unit.

____
____
____
____

4. How would you like the unit improved?

____
____
____

Please mail to
The Course Coordinator ______________ THROUGH the Study Centre Manager
National Open University of Nigeria
Victoria Island,
Lagos.

NATIONAL OPEN UNIVERSITY OF NIGERIA

Form QST2
Questionnaire

In the questions below, we ask you to reflect on your experience of the course as a whole.

1	Course Code and Title ___________________________
2	Mother Tongue ___________________________
3	I am registered for a ___________________________Degree/Programme
4	Why did I choose to take this course? ___________________________
5	Which study unit did I enjoy the most and why? ___________________________
6	Which study unit did I enjoy the least and why? ___________________________
7	Was the course material easy to understand or difficult? ___________________________
8	Which particular topic do I understand better than before and how? ___________________________ ___________________________
9	Does the course have any practical applications in the real world, e.g. for the work I currently do? ...YES/NO? EXPLAIN ___________________________
10	What aspect would I like to know more about or read further? ___________________________
11	How could the course be improved? ___________________________
12	Other comments about the course (please tick) ___________________________

Items	Excellent	Very Good	Good	Poor	Give specific examples, if poor
Presentation Quality	☐	☐	☐	☐	________
Language and Style	☐	☐	☐	☐	________
Illustrations Used (diagrams, tables, Etc.)	☐	☐	☐	☐	________
Conceptual clarity	☐	☐	☐	☐	________
Self Assessment Questions	☐	☐	☐	☐	________
Facilitators response To TMA Questions	☐	☐	☐	☐	________