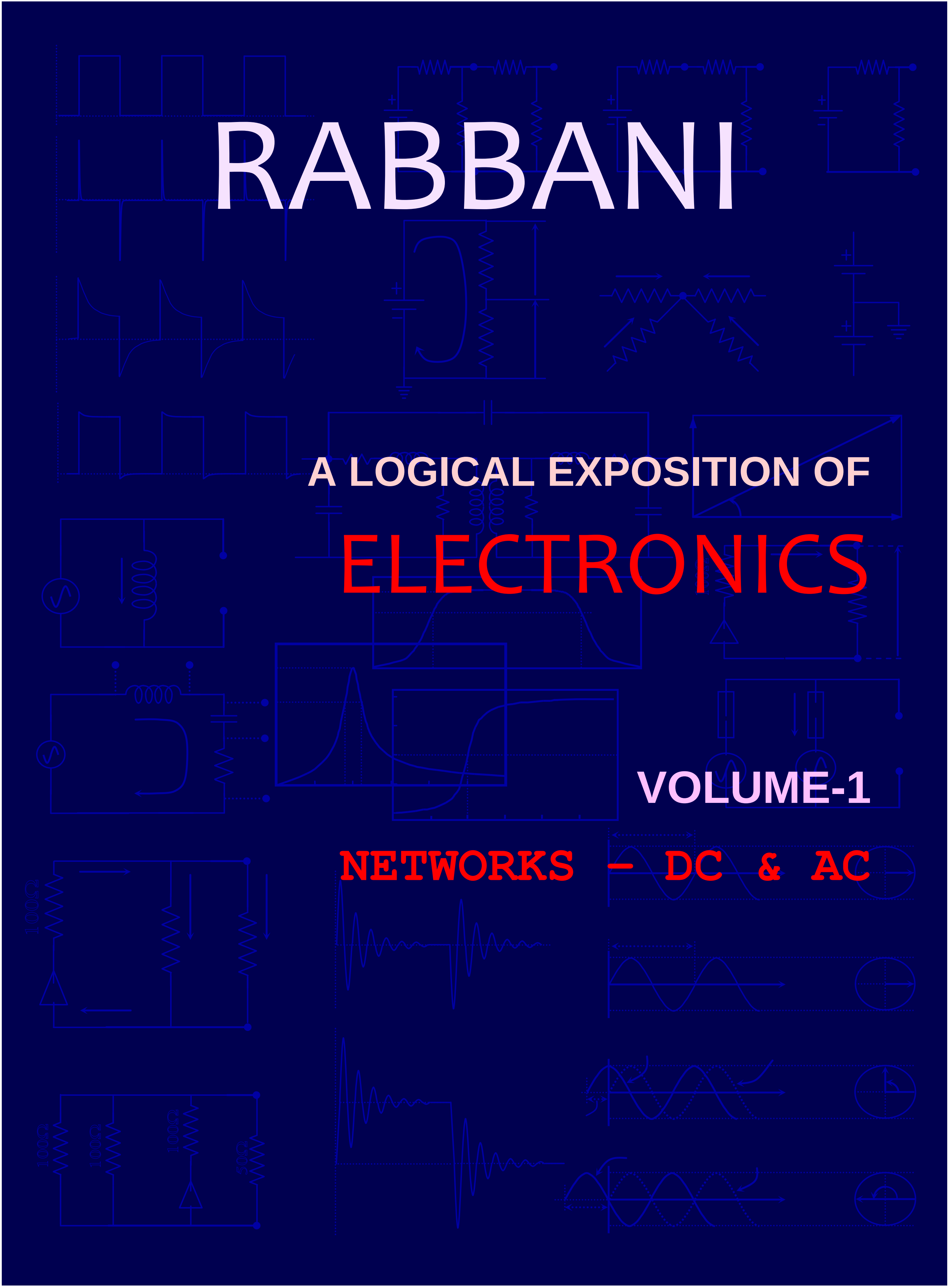


RABBANI

A LOGICAL EXPOSITION OF

ELECTRONICS

VOLUME-1

NETWORKS – DC & AC

Dr. K Siddique-e- Rabbani is a Professor and first Chairman of the recently formed Post Graduate Department of Biomedical Physics & Technology at the University of Dhaka. He has been a teacher in the Physics department since 1978, is a leading electronic designer, and a successful entrepreneur of self designed electronic products in Bangladesh. He obtained a Ph.D. in Electronics (specifically, Microelectronics) in 1978 from the University of Southampton, UK, under a Commonwealth Scholarship. Realising that Bangladesh cannot put an IC chip on the market in the next 30 years he switched to Medical Physics and Biomedical Engineering on his return to his homeland, initially inspired by Late Professor M Shamsul Islam of Physics. Dr. Rabbani is a pioneer researcher in this field in Bangladesh. Side by side he has contributed greatly in promoting indigenous technology based electronic industry in Bangladesh. Innovative electronic products designed by him in the fields of Power, Medicine and Education has been manufactured by several firms over the last two decades.

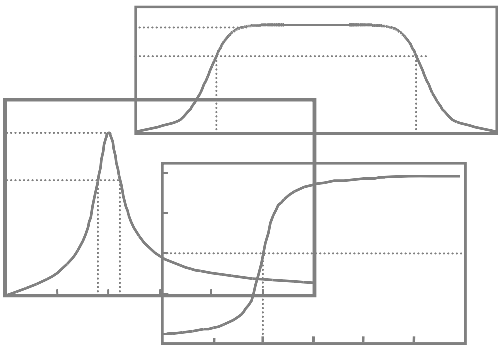
Dr. Rabbani has put into this book more than 30 years of teaching and practical design experience in Electronics. Primarily written as a text for Physics undergraduate students, this book will supplement the needs of engineering students and professionals alike. The author knows first hand what are the stumbling blocks that a student faces and has addressed these with special care. A detailed physical explanation coupled with a logical progression of mathematical treatment and practical application hints makes this book stand out among other texts. Presentation in simple English makes it suitable for self study by students of non-English speaking countries as well.

A LOGICAL EXPOSITION OF

ELECTRONICS

VOLUME-1

NETWORKS – DC & AC



K SIDDIQUE-E- RABBANI

Department of Biomedical Physics & Technology

University of Dhaka

Dhaka, Bangladesh

Published by
Dhaka University Prakashana Samstha
University of Dhaka
Dhaka-1000, Bangladesh

First published
May 2010

Cover Design, text illustrations & layout:
Author

Price: Bangladesh Tk-230 (Two hundred thirty) only
SAARC Countries US$-6 (Six dollars) only
Other Countries US$ - 10 (Ten dollars) only

Printed by
Mother Printers
8,10 Nilkhet, Babupura, Dhaka 1205
Phone: 8615959, 8626542

ISBN- 984-70283-0006-1

Dedication

To
My parents,
Wife and children

Preface

This book has been written mainly for Physics students at the undergraduate level, who have a natural craving to know the how's and why's of everything. Therefore a logical presentation of the subject has been attempted in this book with many subtleties uncovered that most of the texts keep silent about. Mathematical treatments have been developed from the beginning as far as practicable and have been presented in a step by step approach. This relieves the students from unnecessary memorising and gives them the power to develop analytical solutions on their own. There is no meaning in learning a subject as Electronics without knowing where it is applied in our lives, and how to apply the knowledge in a practical design. Therefore practical application examples have been included wherever possible to cater to these natural demands of an inquisitive student.

Through a long teaching experience I have discovered many stumbling blocks in the subject that confront a student. Concepts that are supposed to have been cleared at pre-university level have never been with many students coming from different backgrounds. That an electrical current direction is taken to be opposite to the direction of electron flow just for historical reasons, and that Ohm's law has independent and dependent variables which need careful attention, have been clearly spelled out which other texts hardly mention. Basic differences between electricity and electronics, emf and potential drop, signal and noise have also been spelled out. To some students these questions remain ever unanswered. Therefore I have started at a very basic level and have tried to clear these points as much as possible. Most questions like how a current flows though a capacitor circuit in spite of the fact that there is an insulator interposed, why the reactance is lower at higher frequencies, have been answered from a physical point of view. Some physical explanations to observed phenomena like the dc transient current through a capacitor and that in a series LCR circuit under different damping situations, have been attempted by me intuitively, which appear to be original, found in no other book.

An overview of the methods and concepts has been given before going into the analytical details of each topic which places a student on a better footing. While dealing with Thevenin's equivalent circuits, that we are following two different strategies – one for which the circuit is known and for the other, it is unknown – have been clearly spelled out which remains mostly obscured in other texts. Again that an RC filter is nothing but a voltage divider, a familiar circuit, and that this simple statement makes a lot of difference to a student in comprehending the analytical treatment is an interesting evolution of my own teaching methodology.

Most available texts on Electronics are aimed at Engineering students who have to go into the subject in great details. These texts concentrate more in the analytical aspects and numerical exercises while descriptions on the physical aspects and logics behind an issue are kept to a minimum. On the other hand there are compact volumes on Electronics for scientists which give a broad overview of the subject without providing adequate analytical details as demanded by a formal course offering in Physics. So this book will fill a gap which has been felt acutely by all of us teaching electronics to Physics students.

Volume 1 deals with electrical networks that form the foundation of electronics as a formal subject of study. Without a proper understanding and grasp of these techniques one cannot proceed to the realm of electronics. It has been pointed out clearly that electronic devices cannot be handled directly for analytical purposes. They have to be modeled to quantities that we are familiar with in electrical networks, which is the subject of this volume, and then solved analytically. Therefore, electrical networks is very much a study of electronics and one should try to grasp the matters covered in this book thoroughly.

In writing this book the syllabus of the University of Dhaka to which the author was a key contributor has been followed mostly. Bangladesh National University which administers all other colleges in the country follows the same syllabus, so this book will hopefully address the requirements of a large number of students. Engineering students and professionals will also find this book a useful addition to their existing ones, to clear out conceptual difficulties that contribute in creating confusions now and then, and to find logical ways to tackle analytical solutions on their own.

This book will be followed by at least another volume which will deal with the rest of the Electronics syllabus for the Physics undergraduates in our country.

I am grateful to the University of Dhaka for granting me a Sabbatical leave which gave me the time and opportunity for writing this book. My wife and children's persistent goading to return to the book has helped me from diverting to other things which I do very easily. Their whole hearted co-operation in relieving me of many household duties has allowed me to concentrate and complete this first volume in time. I must also acknowledge the contribution of the hosts of students who attended my classes regularly filling me with enthusiasm and posing intelligent questions that put me into deep thinking without which I could not have answered many of the matters raised in this book. I am also grateful to students Samir, Sayem and Zaid, who painstakingly made brilliant notes from my mostly impromptu lectures and have later given me copies of their notes to help me in writing the book.

I have written the texts and drawn all the graphics directly on the computer, and have taken care that errors are at a minimum. Since preparing the first manuscript in October 2004 I have given copies to many colleagues and students in order to provide me feedback on the book and to point out errors which resulted in improvements in several places. However, I would still welcome error reporting from all the readers. I am a human being, definitely my knowledge has gaps and limitations, so if there are conceptual or analytical mistakes, I would be grateful if the learned readers point these out to me together with the necessary corrections. Their contributions will be duly acknowledged in the subsequent editions.

I hope to embark upon writing the second volume next, which will go directly into semiconductor devices and their applications. I am indeed grateful to the Almighty for the ability and opportunity given to me to write this first volume and pray I am granted further opportunities.

K Siddique-e- Rabbani

Department of Biomedical Physics & Technology
University of Dhaka
Dhaka, Bangladesh

Contents

Chapter 1: Introduction

1.1 Why and how of Electronics

Why do you want to study electronics? Firstly you may want to understand how different electronic equipment work. Secondly you may want to design and develop electronic equipment yourself. It is the latter that will take you to the real world of electronics and you will enjoy electronics to the most. As an electronic designer you dive into the world of unseen, juggle with ideas developed within your trained mind, and then it is a real joy to see the magic unfolding in the seen world. At the same time you will find that you can help everybody around you with some of your acquired capabilities. Again, developing a small useful gadget for the house may end up as a successful commercial product. Besides, you will suddenly find that you understand the working of almost all appliances around you whether big or small, which others without the knowledge of electronics can hardly comprehend. Therefore think yourself as lucky to have taken up a serious study of electronics.

Electronics is a practical science leading to the world of technology. Therefore as you go through this book, gather a multimeter, a breadboard and few electronic components and try things out yourself. You may need the help of other practical oriented books to start with, but when you combine that to what you study in this book, you will get the real feeling of entering the world of electronics.

Designing in electronics is basically organising some logical functional units to achieve an overall objective, a desired overall operation or function. So the first requirement is common intelligence and a general idea of technology which you all have already. Next comes the task of building up the logical functional units using electronic circuits. This is where your knowledge and skill in electronics comes in. You will need to know the following:

A. Electrical network concepts and analysis
B. Available electronic devices, their operations and functions
C. Modelling of electronic devices in terms of known electrical parameters so that the performance of the circuit may be analysed before it is made out of real devices and components.
D. Based on the above models, ways of integrating the electronic devices into an electrical network to achieve the desired objective.

A brief note on the 3rd point: we already know simple electrical devices or components like voltage and current sources, resistors, inductors, capacitors, etc., and ways of analysing networks employing these components. *Ohms law* forms a basis for all these analyses. Since electronic devices are new to us and are varied in forms and functions, we do not know how to analyse them directly. So we replace these devices with simple equivalent models that are expressed in terms of known devices and functions. Once we have done it, it boils down to the same old simple techniques of electrical network analysis! Such modelling is done in every branch of science. Remember the picture of an atom with electrons circling around a nucleus? Has any one seen an atom? – No, the picture you see is that of a model that we have made up in terms of concepts that we can understand. That means you only need to know the simple and old *Ohm's law* to understand and analyse all electronic circuits, however complex !

Let us talk about a common application of electronics – a public address amplifier system as shown in the simplified block diagram in Fig.1-1. When a person speaks into the *microphone* it generates a minute *electrical signal* (with power in the range of microwatts), which mimics the *pattern of sound waves* created by the human speaker. This signal is first amplified to a suitable level using a *pre-amplifier*, then modified using a *signal conditioner* to eliminate noise or to enhance some features of the signal. The modified signal is then amplified sufficiently in a *power amplifier* stage in order to drive a *loud speaker* that converts the *electrical signal* to a *sound signal with the original pattern* that the human speaker produced, but with *hundreds or thousands times the power* of the original speech so that a large crowd can hear it. This is the first step in designing a public address system. Does it sound difficult? – Not at all, it is simply common logical sense.

Next comes working out the details of each functional unit where we will use electronic devices integrated into an electrical network – the electronic circuit. Just as a sculptor or a carpenter needs to know his tools well, how to use them to create a masterpiece, similarly you as an electronic designer need to have a knowledge of available tools in electronics (electronic devices such as diodes, transistors, integrated circuits, etc.), and how to combine them in a circuit to get the desired performance. You can work out the same function in different ways using a variety of devices. Since the workings of these tools or devices are mostly unseen, you need to know the Physics of these devices. As is common with most of the physical sciences and engineering practices, mathematical representation and analyses help greatly in understanding, and designing such circuits. A study of electronics involves all of the above.

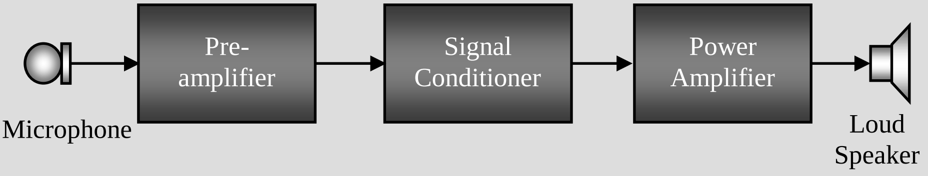


Fig.1-1: A public address amplifier system

Electronics has seen the sharpest development in the history of science and technology and is continuing to do so. A complete knowledge encompassing every small aspect of every device is not possible, as mysteries are still unfolding. Many subtle features and characteristics are revealed to an observant electronic experimenter, which may not be apparent simply by studying books. Therefore set out to designing and fabricating electronic circuits yourself, starting with a very simple one. You will find that it increases your understanding manifold, and as you go along, an inner confidence will build up which will easily distinguish yourself from any other learner in electronics who chose not to go into the designing business.

1.2 Electricity versus Electronics

Both the subjects of Electricity and Electronics deal with the movement of electrons. "Then what is the difference?" – people often ask. One answer frequently given is that in Electronics we can control the flow of electrons. Through an inductive fan regulator, or using a rheostat, we can also control a current, i.e., the flow of electrons, to some extent – but this is a domain of electricity. So the above answer is not complete. In modern days we can control the same current in a much better way using a *TRIAC* or a *Transistor*. We also did the same using a *Vacuum Tube* in the past. In these cases it is said that we have used electronic devices. So what is the fundamental issue that divides the two subjects? – If we look deeper into the working of the devices mentioned above, in an inductive fan regulator and in a rheostat electron flow occurs in metallic conductors, while in the TRIAC and in the Transistor, electron flow occurs in a semi-conducting crystal, which is not a good conductor of electricity. In a Vacuum Tube the electron flow occurs in 'vacuum' – an extreme non-conductor, or an insulator of electricity! These are against our common notions about electricity. Through special techniques, we have been able to move electricity in semi-conductors and in insulators, and these offer much greater opportunity of controlling the movement of electrons according to our desires. Electronic devices provide us with control over the movement of electrons in ways never possible with electrical devices. In a rheostat we can change the resistance by adding or subtracting lengths of the conductor material, while in a transistor we can change the resistance within the same material through the application of a very minute controlling current or voltage. Therefore we possibly can answer the question posed above as follows: ***in Electricity we deal with the flow of electrons through conducting materials, while in Electronics we deal with the flow of electrons through materials that are not good conductors, i.e., semi-conductors and insulators*.**

1.3 Brief History of Electronics

It has been a continuous development that led to electronics and many a scientist's contributions may be cited. The following list attempts at putting up a chronological picture.

1850: German scientist Geissler observed that the inside of an evacuated tube (not 100% vacuum) lits up when a voltage is applied through electrodes.

1878: British scientist Sir William Crookes identified the above as a flow of particles through the vacuum.

1895: Bengali Scientist J.C. Basu performed first radio transmission

Around 1895: Edison observed that current passed from a plate electrode to a heated filament in a vacuum tube if the plate electrode is made positive with respect to the hot filament, and not in the reverse direction. This is known as Edison effect. He could not explain it and left it there.

1904: Fleming explained Edison Effect in terms of a flow of electrons emitted from the hot filament material and developed the vacuum diode based on this effect, which could be used to convert ac power to dc power.

1907: Lee de Forest developed the vacuum triode by adding a grid electrode to the diode. This made signal amplification possible, and real development in electronics started.

1930: Julius Lilienfield, a former Professor of University of Leipzig who migrated to the US, disclosed the principle of Field Effect Transistor through a US patent.

1943: General purpose computer made of vacuum tubes (ENIAC)

1948: Shockley, Brattain and Bardeen fabricated the first semiconductor Transistor.

1951: Commercial production of transistor

1960's: IC production

1970's: Microprocessor, a whole Central Processing Unit of a computer in an Integrated circuit chip

Around 1975: Desktop microcomputer

1.4 Vacuum diode

Let us try to explain the inner working of the Vacuum Diode, where electron flow occurs through vacuum, an insulator in normal terms. Fig. 1-2a shows an evacuated glass tube with a filament at one end, and a plate electrode at the other. The filament is heated using a battery B_F as shown. When hot certain materials release electrons (called *'thermionic emission'*) and thus the filament produces an electron cloud around it. If now another battery B_P is connected with polarity as shown, the plate electrode becomes positive with respect to the filament. This creates an electric field between the filament and the plate which drives the electrons in the cloud towards the plate within the evacuated tube. The positive plate takes up the electrons and pushes them towards the positive terminal of the battery through the meter and the resistor shown, and of course, through the connecting wires made of conductors. On the other side of the battery, electrons pushed downwards by the negative terminal of the battery flow towards the filament and replenish the electron cloud. Thus a continuous flow of

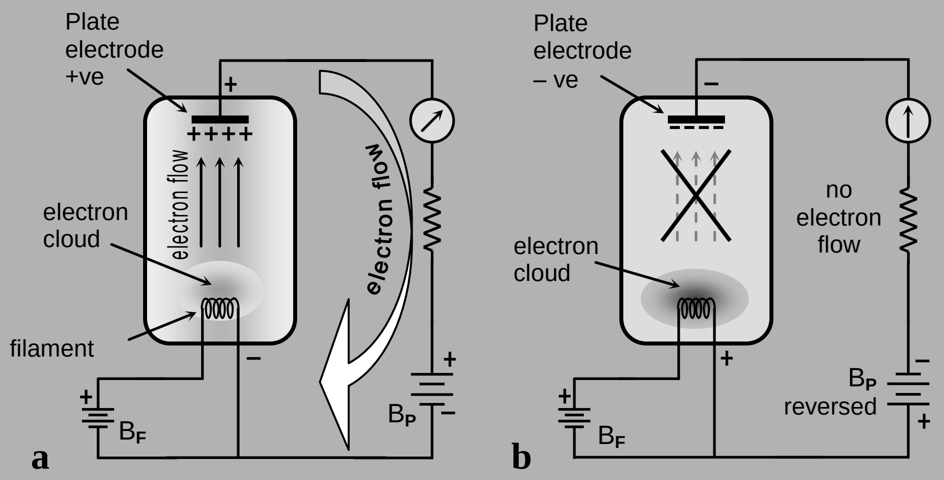


Fig.1-2a: Electron flows through vacuum due to a favourable electric field.
Fig.1-2b: No electron flow due to a reverse electric field.

electrons takes place in the loop (circuit), and an electric current is established through the vacuum tube (an insulator!) by the battery B_P.

On the other hand if the polarity of the battery B_P is reversed (Fig.1-2b), no current flows. This is because the electric field established between the plate electrode and the filament forces the negative electron cloud away from the plate. Thus we have a current flow if the battery is connected as in Fig. 1-2a and no current when it is connected in the reverse direction as in Fig.1-2b. Interestingly Edison observed this phenomenon but could not explain the mechanism that we just described above, as he did not have the necessary scientific background. He just noted the experimental observation in his diary and left it there. It was later named 'Edison Effect'.

The positive plate at the top is called the *'anode'* while the filament, connected to the negative end of B_P is called the *'cathode'*. The battery B_F and its circuit heats the filament only, it does not have any other function. The Edison effect was duly explained later and a practical device called the *Vacuum Diode* was invented by Alexander Fleming in 1904. The vacuum diode was used by Fleming to convert an ac power to dc power, the process being called *rectification* (as if ac was impure and dc was pure!).

In later devices, the filament was covered by an electrode material which is more efficient in emitting electrons, and this electrode was heated indirectly by the filament. The negative of the plate battery B_P was connected to this electrode instead of the filament and this new electrode was then called the cathode. Diode *(di-ode)* stands for *'two electrodes'* – the *cathode* and the *anode* necessary for the main current. The filament is treated as a supporting mechanism and is not counted in the naming of the device. In a practical diode the filament is usually arranged in a vertical configuration and a cylindrical cathode covers the filament closely without touching it. The anode is another co-axial (having the same axis) cylinder with a slightly larger radius.

Box 1.1: That the *direction of electron flow* is opposite to the *direction of conventional electric current* usually confuses the student. The electric current was conceived of earlier than the discovery of electrons. At that time electric *'current'* was thought to be the flow of *'positive charges'* from the positive terminal of a battery to its negative terminal through the conductors in the outer circuitry. Later it was revealed that in solid metallic conductors positive charges do not move, they are associated with the fixed nucleuses of atoms. It is the negatively charged electrons in the loosely bound outermost orbits of a conductor *(called free electrons)* that move in *the opposite direction* to cause the current. The scientific community has a tendency to preserve history unless compelled otherwise. So, it was argued that imagining a flow of a positive charges in the conductors do not affect the results of relevant mathematical analyses, and in doing so we can keep some historical 'rules' unchanged. Thus the wrong direction for current was maintained and accepted generally.

Hopefully this should clear the confusion. We all say of a *'conventional current direction'* while at the back of our mind we know that this is wrong! This is a '*lie'* that we all have agreed to perpetuate.

However, it should be borne in mind that *in a liquid conductor*, called an *electrolyte*, both *positive and negative charges* (in the form of *ions*) move to constitute the current.

Fig.1.3 shows a rectifier circuit to convert ac power to dc power employing such a vacuum Diode. The load resistor is connected between the cathode and the common terminal (common to both input ac source and the output). The ac source provides both positive and negative voltages on the anode alternately with respect to the cathode. The vacuum diode conducts only when the anode is positive as described above. Therefore the resulting current flows through the load resistor only in one direction half the time

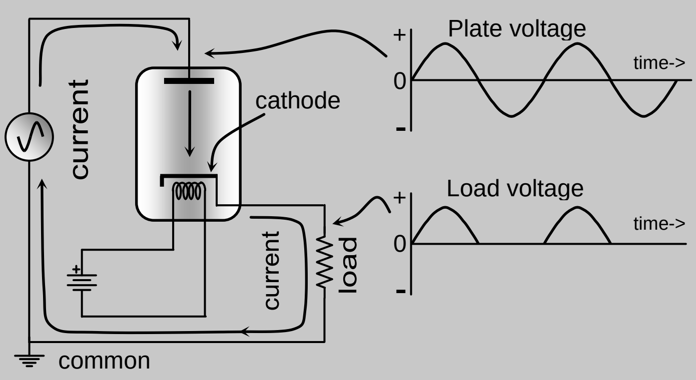


Fig.1-3: Rectification using a vacuum diode. Note direction of conventional current as against electron flow in previous diagrams.

in each cycle, while no current flows during the other half. Thus a unidirectional current or a dc is produced in the load resistor, though its magnitude varies with time. The voltage existing at the source (applied to the plate) and the voltage developed at the load resistor (both with respect to the common point, or ground) are shown graphically on the right hand side of Fig.1.3 for a sinusoidal ac source. The conventional current direction is shown in this figure (which is reverse to the direction of electron flow, see box). Since current is allowed to pass only in one direction and not the other, a vacuum diode was also called a *'valve'* – a term generally associated with water valves having a similar function.

1.5 Vacuum Triode and Amplifier

By adding a mesh of wire (called a *'grid'*) surrounding the cathode at close proximity (Fig.1.4) it became possible to control the flow of *anode-to-cathode current*

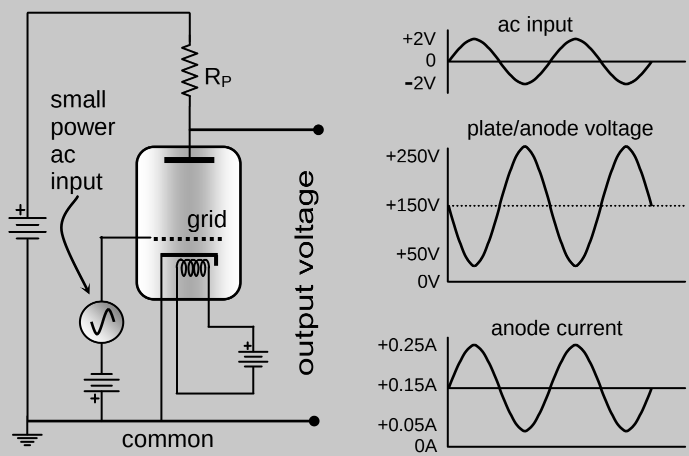


Fig.1-4: Amplification using a vacuum Triode.

significantly by applying a small negative voltage to the grid with respect to the cathode (typical anode voltage ~few hundred volts, grid voltage ~few volts). The electrons could still flow from the cathode to the anode through the holes in the grid mesh because of the large attractive force created by the high anode voltage, but because of its close proximity to the cathode, the negative grid could easily influence the magnitude of electron flow by the repulsive force it provided. By varying this grid voltage in a pattern according to our desire, it is possible to vary the large current flowing in the anode-cathode circuit in the same pattern. In Fig.1.4 a small ac input voltage is added to the negative dc voltage applied to the grid. The grid takes negligible current, so the source has to deliver only a small power. However, this grid voltage compels the anode current through load resistor R_P to vary in proportion to the ac input voltage. This will also cause the anode voltage to vary. The relevant voltages and currents are also shown through graphs in Fig.1.4. The variations in anode voltage comes in association of a dc voltage, called a dc bias, which can easily be removed later to obtain the desired ac pattern at the output. The same applies for the anode current. The anode power is very large because of the high anode voltage (obtained from battery B_P) and current. Thus a small varying power applied to the grid can control a much larger power obtained from the battery B_P.

Here the power in the output signal (the varying pattern) is much larger than the power in the input signal. One can say that the small ac input signal has been multiplied or amplified many times at the output. This is usually called *'amplification'* and this capability has allowed electronics to bring a technological revolution to the whole world. Lee-de-Forest is credited with the invention of this vacuum triode (standing for '*three electrodes*') in 1907.

Vacuum tubes with more electrodes were devised to improve performances. Different sizes of vacuum tubes were available, starting from finger-sized miniature ones to giant chest sized ones used for powerful radio transmitters. With the advent of semiconductor devices the vacuum tubes have almost become obsolete except in applications like high-powered radio transmitters. The cathode ray tube used in Television receivers and in computer monitors is also a special kind of vacuum tube.

However, it is also getting a serious challenge from various semiconductor and organic semiconductor displays and may become obsolete in not too distant a future. Interestingly vacuum tubes have recently made a comeback into our homes through microwave ovens where special vacuum tubes known as 'Klystrons' generate the high power microwave radiation which heats up food without heating the containers.

In this book we shall not go any further into vacuum tubes, rather we shall concentrate on semiconductor devices, which are still ruling modern electronics, and are expected to stay on for quite a long time. However, before we go into the heart of electronics we need to review and build up some background knowledge on electrical networks and circuits, which form the topic of the next few chapters.

Does amplification defy conservation of energy? Have we created a much higher energy out of a small input energy in a vacuum triode amplifier? - Not at all. The higher energy comes from the powerful battery B_P, and the small input energy has exercised a control over the battery to deliver power in its own pattern through the use of the vacuum triode. There lies the power of electronic devices!

Doesn't a transformer do the same? No. In a transformer the output power is equal to the input power (this is for the ideal situation, practically some percentage is lost), i.e., there is no power amplification as in an electronic amplifier. Thus in a transformer if the output voltage is greater than the input voltage, the output current has to be less than the input current. On the other hand in an electronic amplifier both output voltage and current can be greater than the corresponding input voltage and current. Thus there is power amplification.

Chapter 2: Background
Electrical Networks

In this chapter we will review some background knowledge on electricity and electrical networks that have a direct bearing on what we are going to study.

2.1 *EMF* and Potential drop

EMF stands for ***Electro-Motive Force***. *EMF is a property of an electrical source (battery cell or generator) which can create and maintain a difference of electrical potential between two points and the EMF is taken to equal this potential difference.* A common dry battery cell has the ability to create and maintain a potential difference of 1.5 volts between its two terminals (Fig.2-1) which is its *EMF* (e).

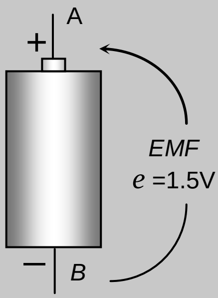


Fig.2-1: *EMF* of a dry battery cell

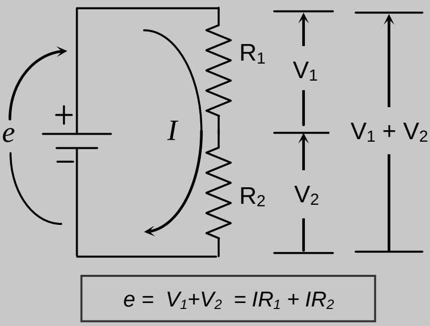


Fig.2-2: *EMF* and potential drop

Because of the *EMF*, an electrical source can drive a current through an electrical load. In Fig.2-2 the battery (represented by its symbol) drives a current I through resistors R_1 and R_2 in series, which together act as the load. As a result there are potential drops, V_1 across R_1 and V_2 across R_2, given by IR_1 and IR_2 respectively. So a potential drop (also called a voltage drop) is always associated with a load when a current flows through it. In Fig.2-2 the *EMF*,

$$e = V_1+V_2 = IR_1 + IR_2$$

assuming the battery cell to be ideal without any internal resistance. The *EMF* and the potential drop have the same unit, *volts*.

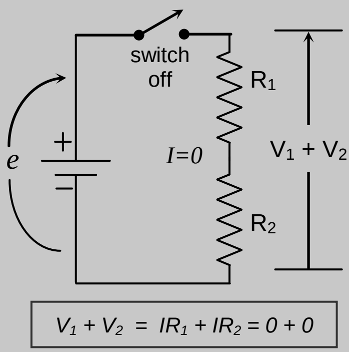


Fig.2-3: Zero potential drop if $I = 0$

If no current flows ($I = 0$) through a resistor no potential drop appears across it. In Fig.2-3 the current has been stopped by breaking the circuit at a suitable point using a switch. The *EMF* or the potential difference across the battery is still *1.5V* but the potential drop across R_1 and R_2 is zero, ie, $V_1+V_2 = 0$ since both IR_1 and IR_2 are zero. You can stop the current flow by breaking the circuit at any other point in the circuit and it will give the same result.

Box 2.1 *The electrical potential is very similar to gravitational potential energy. Compare the above example to the gravitational potential drop of a massive object when you drop it from the rooftop of a two-storied house. We can associate a potential drop for falling through each floor. If you are the agent carrying the massive object to the rooftop, you are acting as the source of* ***Gravitational motive force*** *(equivalent to the battery in the circuit) raising the gravitational potential of the object.*

Some books define *EMF* as the voltage measured across a battery cell when it is not connected to any circuit. In fact this is not a definition, rather it is a way of measuring *EMF*. All practical battery cells have some internal resistance (or source resistance, R_s) and an equivalent circuit of such a cell is shown in Fig.2-4a where the internal resistance is shown in series with an ideal cell. The *EMF* of the ideal cell is still *e*, but you cannot access the internal terminal *X* of this battery cell, you have access only to the external terminals *A* and *B*. If you do not connect the cell to any circuit (open

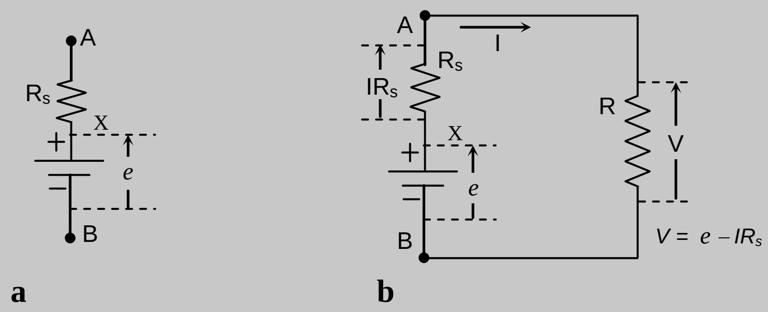


Fig.2-4: Practical battery cell with internal resistance

circuit), no current flows through R_s and therefore there is no potential drop across R_s. Therefore the voltage measured across *A* and *B* is equal to the *EMF*. However, if you connect the battery cell to a resistor circuit so that a current flows as shown in Fig.2-4b,

a voltage IR_s is dropped across the internal resistance, and the measured voltage across terminals A and B of the cell,

$$V = e - IR_s$$

which is obviously less than the *EMF*. Therefore to measure the *EMF*, no current should flow, i.e., a cell cannot be connected to any closed circuit.

A voltmeter has also an internal resistance, R_m. Therefore a current, however small, flows through the voltmeter when we attempt to measure the *EMF* of a cell as shown in Fig. 2-5. The measured voltage, V_m, is basically the voltage dropped across this voltmeter resistance R_m (ie, $V_m = IR_m = e - IR_s$). Thus the measured voltage is always less than the *EMF*. To measure the *EMF* accurately, an ideal voltmeter is needed which has infinite internal resistance so that the current equals zero, which is never possible.

Practically, if a voltmeter with a very high internal resistance is used so that $R_m >> R_S$, then the voltage drop across R_S will be negligible and the measured potential would represent the *EMF* effectively. There is a very good age-old technique using a potentiometer circuit where no current is taken from the cell by balancing against an exactly equal potential, but the procedure is not straightforward for quick measurements.

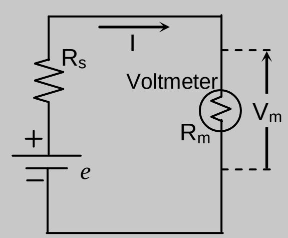


Fig.2-5: Voltmeter measurement

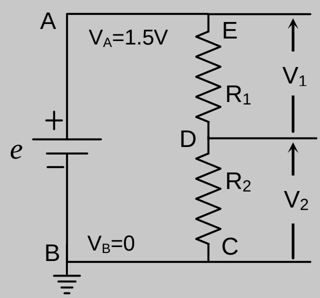


Fig.2-6: Ground reference

2.2 Ground

Potential does not have any absolute value, it is always expressed between two points. In Fig.2-6 the potential difference between the points *A* and *B* is 1.5 volts, we can say that $V_{AB} = 1.5V$. For convenience we usually choose a single point in a circuitry as the reference point where we choose the potential to be zero. Potentials at all other points are expressed with respect to this single point. This point is usually called *common point*, *ground* or *earth* in a circuit and one of the symbols frequently used is shown in Fig.2-6 where we have chosen the point *B* to be the ground (i.e., $V_B = 0$). Once we have chosen a *ground* in a circuit, we do not need to mention the potential difference between two points all the time, we can just say that 'the potential at such and such point is so many volts'. This would naturally mean potential with respect to the ground. Thus in Fig.2-6, $V_A = +1.5V$ is enough to describe the potential at point *A* with respect to the ground. We can also see that

$$\text{since } V_B = 0, \qquad V_{AB} = V_A - V_B = V_A.$$

Potentials at other points can be similarly defined. Thus,

$$V_D = V_2 \qquad \text{and} \qquad V_E = V_1 + V_2.$$

To describe potentials between any two points not including the ground, we have to express using the difference as before. Thus

$$V_E - V_D = V_{ED} = V_1.$$

Here we would like to mention that in all drawn circuits we assume the line segments as connections having zero resistance, and therefore potentials at all points along an unbroken line is the same. Thus in Fig.2-6, points *B* and *C* have the same potential, *0V* while points *A* and *E* are both at *1.5V*.

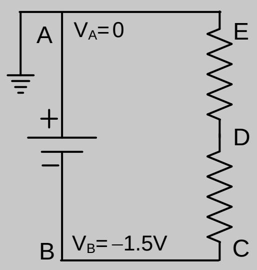


Fig.2-7: Choice of Ground

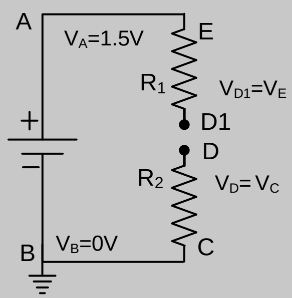


Fig.2-8: Voltage at break

As mentioned above, choosing the ground is entirely a matter of convenience. We could have chosen any other point as the ground, but once we choose one, we have to stick to it for the entire circuit. Thus in Fig.2-7 we have chosen point *A* as the ground ($V_A = 0$), so that $V_B = -1.5V$, but we cannot choose both *A* and *B* as grounds simultaneously in the same circuit.

We would like to emphasise one further point before we leave this section, if there is no current through a resistor, the potential drop across it is zero and therefore potentials at both its ends are the same. Thus in Fig.2-8, since the circuit is broken in the middle, $V_D = V_C = 0V$ while $V_{D1} = V_E = 1.5V$. The whole picture will change as soon as we connect together points *D* and *D1*.

2.3 Single or multiple cell symbol?

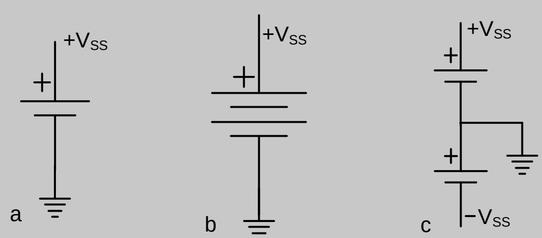


Fig.2-9:
Single (a) or multiple cell (b) symbols and dual power supplies (c)

It needs to be mentioned that though it is logical to show the symbol of a single electrical cell as in Fig.2-9a and a battery (multiple) of cells as in Fig.2-9b, this is not given much importance. We do not bother at all regarding the number of cells. We just use either of the symbols. However, to show a suggestion for high voltages where required, people often use the multiple cell symbol, having many cells.

2.4 Dual power supply

Sometimes we need both a positive and a negative *dc* supply in the same circuit. This is achieved by connecting two *dc* power supplies (or batteries) together with their common point grounded as shown in Fig.2-9c. This is called a dual power supply. In this figure, the two supply terminals are shown at $+V_{SS}$ and $-V_{SS}$ respectively with respect to the ground.

2.5 AC fundamentals

What is an *ac?* We know it stands for *alternating current*, but what happens *physically*? When we have a *dc* battery as the power source in a closed circuit, the current always flows in one direction in the circuit-loop (that is the reason for the name *direct current*). On the other hand, the polarity of an *ac* generator reverses repeatedly (for which it is sometimes called an alternator) so that the current through a closed

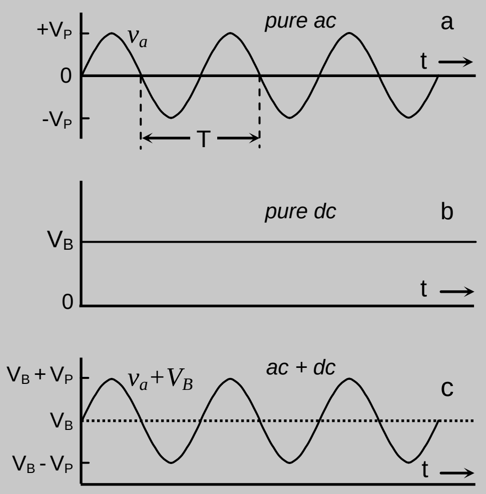


Fig.2-10: Pure Sinusoidal *ac*, pure *dc* and combined *ac* + *dc* (or, varying *dc*) represented graphically

circuit alternates in direction – once clockwise then anticlockwise. For this alternately changing direction of current we call it an *alternating current* or *ac*. However, *ac* has become more of an adjective than a noun, and we frequently use terms like *ac voltage* and *ac current*, none of which strictly make any sense if we expand the abbreviation.

Mathematically, an *ac* voltage means that the voltage changes in magnitude with time, once becoming positive then negative, then positive again. An *ac* current is similarly defined. Which direction is taken to be positive and which one negative is arbitrary, it is simply a matter of choice, but once we define a direction, that becomes the reference for all further descriptions involving that circuit. Again the pattern of change of this voltage or current may be anything, sudden or slow. A very common pattern is called a *sinusoidal* shape or waveform where the quantity (voltage or, current) changes exactly like a Sine function in mathematics. The *ac* generator mentioned above also produces sinusoidal *ac* waveforms.

A sinusoidal *ac* voltage v is given by,

$$v = V_P \, Sin \, \omega t \qquad \text{... 2.1}$$

where V_P is the peak value of voltage and ω is the angular frequency in radians per second. Again, ω equals $2\pi f$ where f is the frequency in number of repetitions per second. This is usually expressed in the unit of Hz (Hertz). We also see square, triangular or saw-tooth or other complex *ac* waveforms.

Fig.2.10a shows the time response of a sinusoidal *ac* voltage, v_a , having peak voltages of $+V_P$ and $-V_P$ on positive and negative sides respectively. It repeats with a period T such that the frequency, $f = 1/T$. Most of the analyses of *ac* circuits performed in electronics also assume sinusoidal shapes for voltages and currents, as this shape is mathematically simple to handle. Besides, all complex waveforms can be expressed as a combination of many pure sine waveforms with different amplitudes and frequencies (*Fourier's theorem*).

Just for comparison, Fig.2.10b shows the time response of a pure *dc* voltage, V_B. We can see that it has a constant value which remains unchanged with time (obviously, it does not change direction). Such a *dc* is called a *smooth dc*. On the other hand, we may have a *varying dc*, where the magnitude changes with time, but remains always on one side of the time axis. Fig. 2.10c shows a *sinusoidally varying dc* voltage. Physically, in a *varying dc* the current is also varying in magnitude, but flows in only one direction in the circuit. A varying *dc* may also be considered as a sum of an *ac* and *dc*. As in Fig.2.10c, we can write the total voltage as $V_{tot} = v_a + V_B$ where V_B is a *dc* which forms a *baseline* above which v_a, a sinusoidal *ac* is superposed. The *dc* baseline is also called a *dc bias*, or simply *bias*, in all electronic analysis.

Almost *all electronic devices* (vacuum diodes, transistors) are basically *dc devices*, they can work only in one direction. Therefore to amplify *ac signals* (where current flows in both directions) we have to add a *dc bias* to it so that the resulting voltage is a *varying dc* as shown in Fig.2.10c. When the whole job of amplification is complete, the *dc* is blocked using any suitable technique, either using a capacitor-resistance circuit or a transformer, and a pure *ac* is taken out at the output.

Any periodic *ac* will have a frequency f as defined above. Sometimes we like to describe *dc* as a special case of *ac*. We say that as f approaches 0, an *ac* becomes a *dc*. This is particularly useful in mathematical analysis. For ease of understanding, we try to follow a standard convention in designating *ac*, *dc* and mixed electrical quantities using symbols. These are given in Box 2.2. For mathematical analysis, sometimes we represent *ac* quantities using Complex numbers, and the conventions are also given in the same box. (See Chapter 4 for the basics of Complex representation)

Box 2.2 Convention for *ac* and *dc* variables

In this book we will follow the conventions for representing voltages and currents by variables with appropriate subscripts as given below.

Pure dc:

Variable: *Capitals*, Subscript: *Capital*.

Example: V_{IN}, V_{OUT}, V_{AB}, I_{AB}, V_C, I_C, etc.

Pure ac:

Variable: *Small*, Subscript: *Small*.

Example: v_{in}, v_{out}, v_{ab}, i_{ab}, v_c, i_c, etc.

Mixed dc + ac:

Variable: *Capitals*, Subscript: *Small*.

Example: V_{in}, V_{out}, V_{ab}, I_{ab}, V_c, I_c, etc.

Complex form(ac):

Variable: *Bold capital*, Subscript: *bold capital*

Example: $\mathbf{V_{IN}}$, $\mathbf{V_{OUT}}$, $\mathbf{V_{AB}}$, $\mathbf{I_{AB}}$, $\mathbf{V_C}$, $\mathbf{I_C}$, $\mathbf{X}$, $\mathbf{Z}$, etc.

While writing by hand you cannot use bold faced font. Use a bar instead, either below or above, such as $\overline{V}$, $\overline{I}$, $\overline{X}$, $\overline{Z}$, etc.

Magnitude (absolute value) of a Complex quantity:

It is simply shown without the bold face.

Example (for variables given above): V_{IN}, V_{OUT}, V_{AB}, I_{AB}, V_C, I_C, X, Z, etc.

2.6 Ohm's Law

In Fig.2-11, a source of *EMF* creates a potential difference V across a conductor (if the source is ideal, having zero internal resistance, then $V = e$, otherwise $V < e$). This potential difference V causes a current I to flow through the conductor. Ohm's law states that ***temperature remaining constant, the current through a conductor is proportional to the potential across it.***

This means,	$\boldsymbol{I \propto V}$	... 2.2a

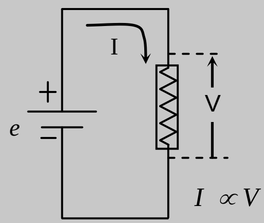


Fig.2-11: Ohm's law definition

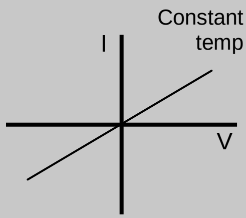


Fig.2-12: I-V curve of an Ohmic conductor

This is the basic form of Ohm's law. Note that there is no mention of resistance in this law. When we form an equation from this law we define *Conductance*, *G*, as the proportionality constant,

$$I = G\,V \qquad \text{... 2.2b}$$

where G relates to a property of the conductor which determines how much current it will allow to pass due to a certain applied voltage across it.

Resistance, *R*, is defined as the inverse of *Conductance* as,

$$R = \frac{1}{G} \qquad \text{... 2.2c}$$

so that in terms of R, 2.2b becomes

$$I = \frac{1}{R}\,V \qquad \text{... 2.2d}$$

This is the more well known equation formed from Ohm's law. R is a characteristics of the conductor which attempts to oppose a current through it (I decreases if R increases), therefore the name *Resistance* was chosen for this property of the conductor. A piece of material having this property is called a *Resistor*. Therefore the name *resistor* is a product of Ohm's law [note that we use conductor and resistor to represent the same thing]. The basic statement of Ohm's law is graphically shown in Fig.2-12 (I – V

curve). We can see that the graph is a straight line passing through the origin extending to both the 1st and 3rd quadrants. The line has a positive slope and the inverse of the slope is the resistance *R*. A material having such features is called an *Ohmic conductor*.

If the I – V curve of any material deviates from this straight line behaviour it is called a *non-Ohmic conductor*. Thus a semiconductor diode (to be described later) having non-linear I – V curves (Fig. 2-13) is said to be non-Ohmic in nature.

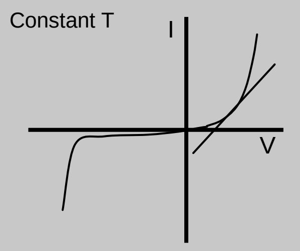


Fig.2-13: I-V curve of a non-ohmic conductor

Box 2.3 The statement *V* = *IR* is not strictly Ohm's law as is popularly used, it is an inference of Ohm's law. The same can be said about *R* = *V/I* . However, these forms are useful in finding any one of the three parameters when the other two are known. You should also note that in Eq.2.2, *V* is the independent parameter and *I* is the dependent parameter, as *I* depends on *V*, so they are drawn with the axes as shown in these graphs. We usually do not plot with the axes other-way round.

Box 2.4 Incidentally some books tend to show a non-linear behaviour for tungsten filaments of light bulbs (Fig.B-2-1) implying that it is non-ohmic. This is not right and they miss out one important point – the temperature is not constant throughout the curve, it increases with increasing voltage. Had the temperature been kept constant (through appropriate heat removal), it would have been a straight line. Of course, we would get separate straight lines with different slopes at different temperatures indicating that the resistance vary with temperature. For a metal the slope decreases, i.e., the resistance increases with temperature (Fig.B-2-2).

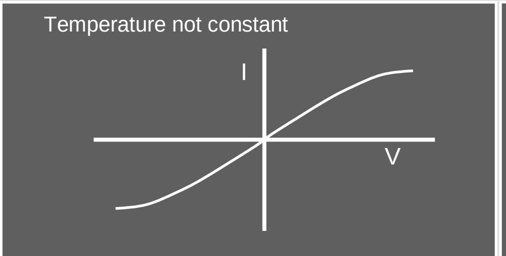


Fig.B-2-1: I-V curve of a tungsten filament bulb

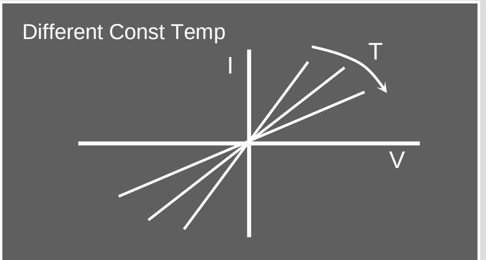


Fig.B-2-2: Ohmic conductor at different temperatures

2.7 Incremental resistance from I – V curve

Even for a non-linear I – V curve we can define an incremental resistance by taking the inverse of the slope of the tangent at any point on the curved line as,

$$r = \left(\frac{dI}{dV} \right)^{-1} \qquad \text{... 2.3}$$

which may change with voltage unlike an ideal resistor whose value is constant. This is also shown in Fig. 2-13 where the tangent represents the slope at the point of interest. For an ohmic resistor, this is the same as the normal resistance obtained from $R = V/I$.

> You may ask why we did not use dV/dI in Eq.2.3. Well, according to Ohm's law, I is dependent on V (V does not depend on I), so there can be dI/dV, and not dV/dI.

2.8 Series and parallel combination of resistors, which one dominates?

You know that a series combination of two resistors R_1 and R_2 (Fig.2-14a) is simply

$$R_{ser} = R_1 + R_2 \qquad \text{... 2.4}$$

and a parallel combination is (Fig.2-14a),

$$R_{par} = \left[\frac{1}{R_1} + \frac{1}{R_2} \right]^{-1} = \frac{R_1\, R_2}{R_1 + R_2} \qquad \text{... 2.5}$$

If they have widely different values which one dominates? In series combination it is obviously the larger one. If you have to cross equal lengths of a very difficult route and an easy route respectively to go to your destination, it is always the harder one which will bother you and will determine the difficulty of your journey. The same happens for electrons! Suppose $R_1 = 1000\Omega$ and $R_2 = 100\Omega$, the equivalent combined resistance is 1100Ω which is closer to the larger 1000Ω value.

On the other hand when you have two resistors having widely different values in parallel (Fig.2-14b), it is always the smaller resistor that dominates the final equivalent value. Suppose you have two parallel roads – a very rough one and a smooth one. If a hundred students are asked to go to a destination on the other end of these parallel roads, which one will most of them take? – Obviously, the smooth one with less resistance. Electrons do the same! Taking the previous values for R_1 and R_2, the parallel

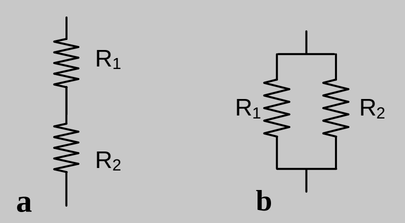


Fig.2-14: Series (a) and parallel (b) combination of resistors

combination is 90.9Ω which is closer to the smaller 100Ω, supporting the above argument. Note also that the combined result is less than even the smaller one in the group. This you should remember to verify your results. You can take note of another point, 1000Ω is 10 times greater than 100Ω, and the equivalent value is about 10% below 100Ω. This information comes handy in making quick estimates when you design a circuit.

2.9 Constant Voltage source

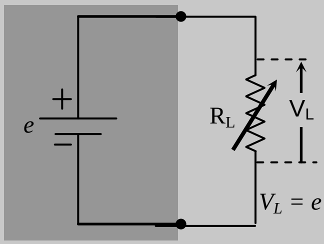


Fig.2-15: Ideal Constant Voltage Source, $R_S = 0$

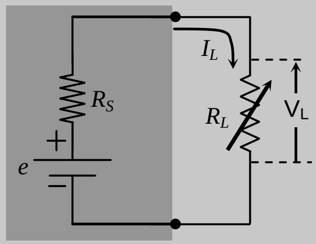


Fig.2-16: Practical Constant Voltage Source, R_S<<R_L

As shown in Fig.2-15 an ideal battery cell with zero internal resistance ($R_S = 0$) can be considered a Constant Voltage Source (shaded portion) as its output or load voltage will remain equal to its *EMF* irrespective of the load resistance. In practice such a source can never be found, therefore, a Constant Voltage Source can be defined as a source of *EMF* with an internal resistance R_S (shaded part in Fig.2-16) such that for all practical values of load resistance (R_L) in a particular application, the voltage dropped across R_S is negligible. *This is clearly possible for $R_S << R_L$, which is the requisite condition for a Constant Voltage Source*. Therefore, for a good constant voltage source R_S should be as low as possible. For a fixed e and R_S the above condition will be satisfied if R_L is above a certain minimum value R_{Lmin} as shown in Fig.2-17a. Below R_{Lmin} the load voltage will decrease and can no more be called constant.

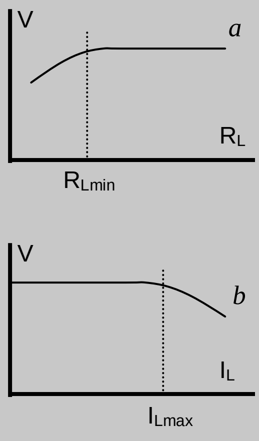


Fig.2-17: Characteristics of a practical constant voltage source, with R_L (*a*) and with I_L (*b*)

Since the load current,

$$I_L = \frac{e}{R_S + R_L} \qquad \text{... 2.6}$$

it is clear that for the above situation, there will be a corresponding maximum load current I_{Lmax} (Fig. 2-17b) above which the source will no longer act as a Constant Voltage Source. How much reduction in load voltage can be tolerated depends upon the particular application. For some applications, 10% may be satisfactory, for some 1% may be adequate, some may require 0.1% or even lower figures. It should be noted that for a constant voltage source, the load current would vary if R_L changes.

2.9.1 Constant Voltage Source for *ac*: In an *ac* source the voltage is always changing, so how do we define a constant source? If the amplitude (peak value) of an *ac* remains constant, we call it a constant voltage *ac* source. Obviously its rms or average values would also be constant. The symbol of a sinusoidal *ac* constant voltage source is shown in Fig. 2-18. Though the direction continuously alternates, we need a reference direction for algebraic analyses. Therefore you will see that we have placed a "+" sign on one side of the source. Do not confuse it with a *dc* source!

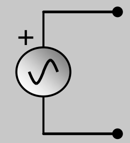


Fig.2-18: Symbol of a sinusoidal *ac* constant voltage source

2.10 Constant Current source

An ideal Constant Current Source will supply a constant current to any load irrespective of the value of R_L as shown in Fig.2-19 (shaded part); again it cannot be achieved strictly in practice. Besides, we do not know of any simple source besides a battery which can provide a current, but a battery is very nearly a constant voltage device. Therefore in terms of a constant voltage source with an internal resistance R_S (shaded part in Fig.2-20) we can visualise a practical constant current source such that $R_S >> R_L$. From equation 2.6 above we can see that under this condition

$$I_L \cong \frac{e}{R_S} \quad ... 2.7$$

which is constant if e and R_S are fixed. We can show the behaviour graphically as in

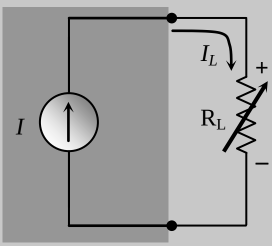


Fig.2-19: Constant Current Source, I = constant for any R_L

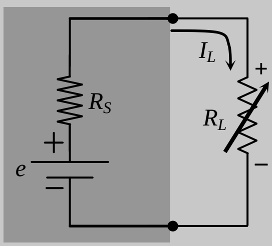


Fig.2-20: Practical Constant Current Source, $R_S >> R_L$

Figs. 2.21. With reference to Eq.2.6 we can see that a maximum load R_{Lmax} exists beyond which it can no longer be called a constant current source. As opposed to the constant voltage source, to keep the current constant for different R_L, the load voltage will vary in a Constant Current Source (since $V = IR_L$).

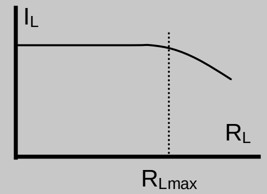


Fig.2-21: Characteristics of a practical constant current source

Here it can also be noted that for a good constant current source, R_S should be as large as possible. However, this also lowers the value of current I_L which can then only be increased by increasing *e* to a large value. We can also visualise that *an ideal Constant Current Source can be made up of a constant voltage source with infinite internal resistance* though it would be of no use as the current would be zero. You can make a good practical constant current source using large external resistors in series to the source of *EMF*. To overcome the current limitation problem, special circuits have been designed using semiconductor transistors which can provide large constant currents even with low voltage sources.

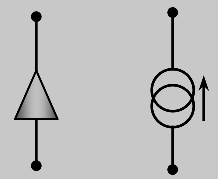


Fig.2-22: Alternative symbols of a constant current source

You should note that the sign of potential developed across the load would depend on the current direction. For the current source direction shown by the arrow in Fig.2.19, the sign of the potential developed across R_L is also shown in the figure. Alternative symbols for a current source are also used as shown in Fig.2-22 where the directions of current are implied (for the triangular one the apex gives the direction).

2.10.1 Constant Current Source for *ac*: The same symbol is used for both *ac* and *dc*. The peak, RMS, or average current values are constant, as for an *ac* voltage source.

2.11 Source Resistance Measurement

The source resistance R_S , if it is unknown, can be measured in both the above cases using a very simple technique.

We know that $V_L = I_L R_L$, and using equation 2.6 we get,

$$V_L = e\frac{R_L}{R_S + R_L} \qquad \text{... 2.8}$$

We can see that,

when $R_L = \propto$ (open circuit), $V_L \simeq e$,

and $V_L = e / 2$ when $R_L = R_S$

The following technique evolves from the above analyses and is widely used experimentally. In either Fig. 2-16, or Fig.2-20, open the load, connect a voltmeter with a very high input resistance and measure the output voltage, which can be treated as the open circuit voltage. Next, connect a variable R_L and adjust it to obtain an output voltage exactly equal to half the open circuit voltage. Take the load resistance out from the circuit and measure its value. This would give you the value of R_S. (If you measure R_L while connected to the circuit a wrong value will be obtained because of R_S connected in parallel, and the *EMF* of the source may complicate the measurement.)

Do not try the above method with batteries and voltage supply units, which will require very high currents in bringing down the output voltage to 50%, and you may damage the equipment and cause a fire hazard. In such cases, first measure the open circuit value, and then bring the output down to about 95% of the open circuit value by adjusting R_L. These measurements will give you two equations based on equation 2.8, which you can easily solve to get R_S. We will leave it to you to find an analytical solution. ***Try it!***

2.12 Kirchoff's laws

Kirchoff's laws are direct descendants of Ohm's law to make analyses simpler in certain circuit situations.

2.12.1 Kirchoff's voltage laws

Kirchoff's law for voltage says the sum of all the voltages within a closed circuit loop, considered with appropriate signs, is zero, or,

$$\sum_i V_i \Big|_{closed\ loop} = 0 \qquad \text{... 2.9}$$

This is natural to expect because when you move out from a point and come back to the same original point there cannot be any net gain or loss in potential energy. Now let us apply this law to sample circuit-1 shown in Fig. 2-23. We have to assume any reference direction for the current in the loop to start with (for complex circuits with branches, you have to consider the currents through each component separately, as appropriate). Because of the assumed current direction, the high and low potential ends of any resistor in the path will be determined as shown in the figure. Let us start our journey from *a* and travel clockwise through *bcd* to come back to the starting point. As we make our first journey through V_1, we have the freedom to choose this potential difference as either positive, or negative. However, once we make the choice, we will have to stick to the same choice for the rest of the journey. Let

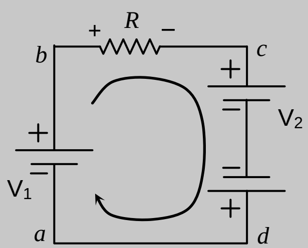


Fig.2-23: Sample circuit-1

us choose V_1 as positive as we go from a lower potential at *a* to a higher potential at *b*. Then as we go along, the potential drop across *R* makes a travel from a higher to a lower potential, so we have to give it a negative sign. Travel through V_2 is also similar, so it gets a negative sign too. Finally travel through V_3 is again from a low to high potential giving it a positive sign. Thus we get,

$$V_1 - IR - V_2 + V_3 = 0$$

wherefrom, $I = (V_1 - V_2 + V_3)/R$

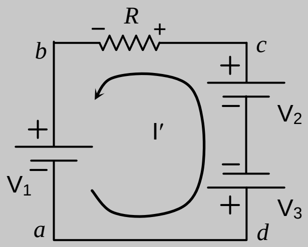


Fig.2-24: Sample circuit-2

The result is the same whatever current direction you choose and wherever you start from. Fig.2-24 shows the same circuit but here we have chosen the reference current in the opposite direction, and we would also start from a different point. Let us start form point *c* and choose an anticlockwise travel through points *bad*. Following the above convention (+ ve for going from low to high potential) we get,

$$-I'R - V_1 - V_3 + V_2 = 0$$

wherefrom, $I' = -(V_1 - V_2 + V_3)/R$

Had we chosen the opposite convention (+ ve for going from high to low potential) the result would have been the same (try it!). The magnitude of the current calculated in both the above procedures is the same; only the sign is different. This is because we chose a reference current in the opposite direction in the latter case. Following such methods you can determine the voltage loop equations in any circuit. Remember, if a circuit is not closed, you cannot apply Kirchoff's voltage law there.

2.12.2 Kirchoff's current law for nodes

This law applies to a node (junction) of a circuit where currents come in and go out through different branches as shown in Fig.2-25, and states that the total current going into a node is zero, i.e.,

$$\sum_i I_i \Big|_{node} = 0 \qquad \text{... 2.10}$$

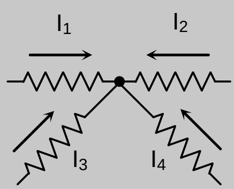


Fig.2-25: Current node

This is also common sense, if there is a net non-zero current going into a node, there will be a build up of charge there. Where would this excess charge go? Here the reference direction has been taken as the current going into a node for algebraic analysis, current going out of the node would be taken as negative then. In practice there has to be currents in both directions, charges brought in by one or more branches have to be carried away by the remaining branches. If the total algebraic sum of such currents is non-zero in a given problem, the problem itself is wrong!

2.13 Voltage Divider Network

In spite of its simplicity, this circuit needs special attention. You will need it almost in any circuit design or analysis, therefore you have to familiarise yourself with different approaches to this circuit and its variations. In Fig.2-26, the loop current,

$$I = \frac{e}{R_1 + R_2} \qquad \text{... 2.11}$$

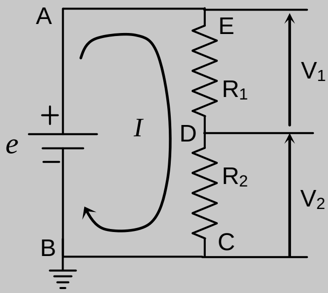


Fig.2-26: Voltage Divider

Therefore, voltage V_1 across R_1 and V_2 across R_2 can be given as:

$$V_1 = I\,R_1 \qquad \text{... 2.12a}$$

and $$V_2 = I\,R_2 \qquad \text{... 2.12b}$$

where $$e = V_1 + V_2 \qquad \text{... 2.12c}$$

You can substitute the value of current I from Eq.2-11 to see the values of V_1 and V_2 in terms of e and the resistors. Equations 2.11 and 2.12 show how a voltage e can be divided into two parts using a circuit as shown in Fig.2-26; hence the name *'voltage divider'*. Here V_2 is especially useful in many circuits as this may be an output voltage where e is the input voltage with a common ground, and is given by,

$$V_2 = e\frac{R_2}{R_1 + R_2} = \frac{e}{1 + \dfrac{R_1}{R_2}} \qquad \text{... 2.13a}$$

which can also be written as a ratio of output to input voltages, called the *voltage gain*,

$$\frac{V_2}{e} = \frac{R_2}{R_1 + R_2} = \frac{1}{1 + \dfrac{R_1}{R_2}} \qquad \text{...2.13b}$$

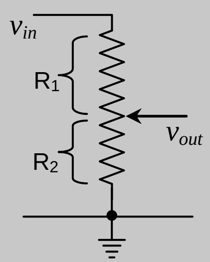


Fig.2-27: Gain control

It can be seen from equation 2.13 that if e and R_1 are kept fixed, V_2 increases if R_2 increases. If both R_1 and R_2 vary then V_2 increases if the ratio R_1/R_2 decreases.

This is the basis of all gain control circuits using a variable three terminal resistor, called a potentiometer as shown in Fig.2-27 where the central contact (called *'brush'*) can be moved to vary the ratio R_1/R_2 from $\propto$ (infinity) with the brush at the lower extreme, to 0 with the brush at the upper extreme. If the input voltage is v_{in} the corresponding output voltage then varies from 0 to v_{in}. The volume control in your radio or amplifier uses this very device.

Equations 2.13 also gives us,

$$\frac{V_1}{V_2} = \frac{R_1}{R_2} \qquad \text{... 2.14}$$

which has to be remembered very well. The ratio of potentials or voltages across the two resistors is equal to the ratio of their resistance values, which means that the higher resistance of the two drops more voltage. Besides, the sum equals the input voltage.

Example: If e = 10V, R_1 = 6kΩ and R_2 = 4kΩ , then V_1 would be 6V and V_2 would be 4V which we can evaluate by a glance without performing any calculation. The voltages will remain the same if the 6kΩ and 4kΩ resistors are replaced by 300Ω and 200Ω resistors respectively. Here the current will increase, but that is not of much concern when we are interested only in voltages (of course, we do not want to waste power, so would not like the current to be unnecessarily high).

2.14 Current Divider Network

Fig. 2-28 shows a current divider network. Here the current I obtained from a constant current source is divided between two resistors. Since the voltages across R_1 and R_2 are the same always, therefore,

$$I_1R_1 = I_2R_2$$

so that, $$\frac{I_1}{I_2} = \frac{R_2}{R_1} \qquad \text{... 2.15}$$

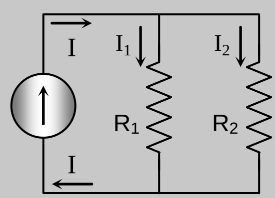


Fig.2-28: Current Divider

Note that there is inverse dependency. The smaller resistor will take the larger amount of current. This was discussed before with parallel resistor combinations. In fig.2.28, if you are asked to find the voltage across R_2 you can calculate it either from I_1R_1 or I_2R_2. The same applies for the voltage across R_1.

Example 2.1: If I = 1A, R_1 = 6Ω and R_2 = 4Ω , then I_1 = 0.4A and I_2 = 0.6A, which we can evaluate by a glance without performing any calculation. What is the voltage across the resistors in this case? Find out yourselves.

2.15 Thevenin's equivalent circuit

Equivalent circuits are needed to simplify and analyse complex circuits. Electronic devices like diodes and transistors which cannot be directly described in terms of any known electrical parameters like resistance, capacitance, inductance, etc. are not amenable to analysis unless we represent them in terms of the above parameters. Therefore we cannot do without equivalent circuits in electronics, and Thevenin's equivalent circuit is possibly the most widely used circuit. This is based on Thevenin's theorem whose statement is given below.

Thevenin's theorem states that any complex network can be replaced by a constant voltage source, V_{Th} in series with a resistor, R_{Th} as shown in Fig.2-29, where V_{Th} is given by the open circuit voltage at the output and the R_{Th} is given by the total resistance measured between the output terminals with all internal voltage sources replaced by shorts, and all internal current sources replaced by open circuits.

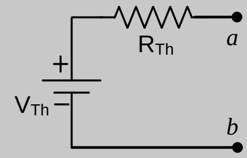


Fig.2-29: Thevenin's equivalent circuit

(Remember, an ideal voltage source has zero internal resistance, and an ideal current source has infinite internal resistance.)

We can see in Fig.2.29 that since there is no current through R_{Th} under open circuit condition, the open circuit voltage V_{OC} across *ab* is equal to V_{Th}. Again looking back through terminal *ab*, the total resistance is R_{Th} since V_{Th} has zero internal resistance. This is just a model proposed by Thevenin and we have to represent a whole complex circuit in this way. There could have been other models, but this simple model has gained wide acceptance. We will discuss another model based soon on a current source instead of the voltage source in the above model, which is also widely used.

2.15.1 Determination of Thevenin's equivalent circuit

This basically means reducing any complex circuit to the arrangement shown in Fig.2-29 and determining the values of the Thevenin parameters V_{Th} and R_{Th}. There are two common approaches – one for an unknown circuit, and the other for a known circuit.

Unknown Circuit

If the internal circuit details are not known, or if it is too complex to be analysed easily, we can determine the Thevenin parameters (V_{Th} and R_{Th}) through a knowledge of the open circuit output voltage V_{OC} and short circuit output current I_{SC}. These values may be measured directly as shown in Fig.2.30. We can measure the open circuit voltage V_{OC} using a voltmeter with very high internal resistance as shown in Fig.2.30a and the short circuit current I_{SC} using a current-meter (Ammeter) with very low internal resistance (a short) as shown in Fig.2.30b. With reference to Fig.2-29, V_{Th} itself is the open circuit voltage V_{OC}. Now R_{Th} can be easily obtained from the ratio of V_{Th} to I_{SC} .

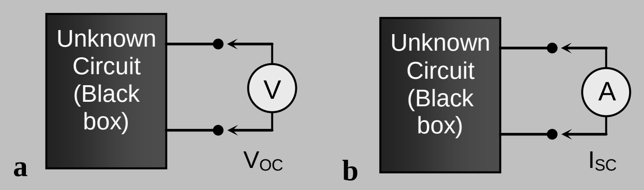


Fig.2-30: Measuring V_{OC} and I_{SC} of an unknown circuit

Thus we have, $V_{Th} = V_{OC}$ ... 2.16

and $R_{Th} = V_{Th} / I_{SC}$ ... 2.17

which are determined from the measured values completely. Thus we obtain the values of V_{Th} and R_{Th} for this unknown circuit which completes the requirement to draw the corresponding Thevenin's equivalent circuit as given in Fig.2-29.

We frequently use the alternative form of Eq.2.17,

$$I_{SC} = V_{Th} / R_{Th} \quad \text{... 2.18}$$

Known Simple Circuit
If the circuit is reasonably simple and the internal circuit details are known, we usually perform network analysis directly to reduce the circuit. This method will be clarified using a few examples below.

Example 2.1 *(for a known simple circuit)***:** Let us determine the Thevenin's equivalent circuit for the simple voltage divider circuit shown in Fig.2-31a between its two output

Fig.2-31: Voltage Divider and its output resistance

points shown as *p* and *q*.

The Thevenin voltage (open circuit voltage between *p* and *q*) is, following Eq.2.13:

$$V_{Th} = V_{OC} = e\frac{R_2}{R_1 + R_2} \quad \text{... 2.19}$$

To determine R_{Th}, we short the internal voltage source e in Fig.2.31a following Thevenin's theorem. This leads us to the parallel combination of R_1 and R_2 as shown in Fig.2.31b. According to Thevenin's theorem this combination looking back from the terminals p and q is the required series resistance R_{Th} which is given by,

$$R_{Th} = \frac{R_1 \, R_2}{R_1 + R_2} \qquad \text{... 2.20}$$

Thus we obtain both V_{Th} and R_{Th} in terms of the circuit parameters e, R_1 and R_2 and Fig.2-29 is the required equivalent circuit.

A slightly more complex network can be systematically simplified using Thevenin's equivalent network as shown in the example below.

Example 2.2 *(for a known circuit, slightly more complex)***:** Let us determine the Thevenin's equivalent circuit for the circuit shown in Fig.2-32a between its two output points r and s. Firstly we determine the Thevenin's equivalent circuit for the segment

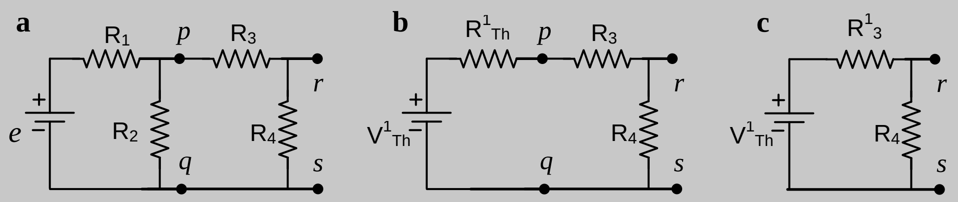


Fig.2-32: A complex circuit and its sequential simplification

to the left of points p and q, which is exactly the same as the simple voltage divider given in Ex.2.1. This first simplified equivalent circuit is shown in Fig.2-32b and we can use the previous solutions to write down the following intermediate values directly. Thus,

$$V^1_{Th} = e \frac{R_2}{R_1 + R_2} \qquad \text{... 2.21}$$

and,

$$R^1_{Th} = \frac{R_1 \, R_2}{R_1 + R_2} \qquad \text{... 2.22}$$

Now R^1_{Th} and R_3 can be combined to get the series combination in the next equivalent circuit shown in Fig.2.32c as,

$$R^1_3 = R^1_{Th} + R_3 \qquad \text{... 2.23}$$

The circuit in Fig.2.32c is again a simple voltage divider and we can use the previous results, as before, to get values for the final equivalent circuit as given by Fig.2.29,

$$V_{Th} = V^1_{Th} \frac{R_4}{R^1_3 + R_4} \qquad \text{... 2.24}$$

and,
$$R_{Th} = \frac{R^1_3 \, R_4}{R^1_3 + R_4} \qquad \text{... 2.25}$$

where the intermediate values are as given by equations 2.21 to 2.23.

Example 2.3 *(to determine the current through a resistor in a network)*: In the circuit of Fig.2.32a if it is asked to find the current through resistor R_4, this needs a slight

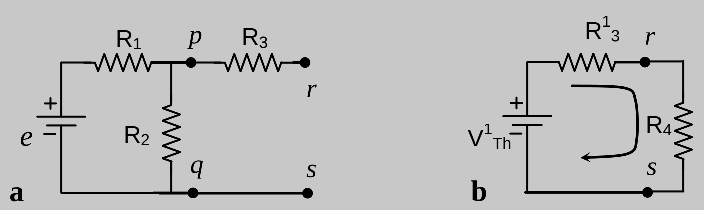


Fig.2-33: Determining current through a resistor

modification. Here we have to take R_4 out of the circuit as shown in Fig. 2.33a and determine the Thevenin's equivalent circuit between points *r* and *s* as shown in Fig.2.33b, without R_4. Here V^1_{Th} and R^1_3 are the same as in Fig.2.32c and given by equations 2.21 and 2.23. Then we connect R_4 as a load and determine the current through R_4 straightway as,

$$I_4 = \frac{V^1_{Th}}{R^1_3 + R_4} \qquad \text{... 2.26}$$

Thus we can simplify complex circuits step by step to reach its final Thevenin's equivalent circuit as given by Fig.2-29, and perform the required analysis.Next we will discuss another model as mentioned before.

2.16 Norton's equivalent circuit

This is basically a complement to Thevenin's equivalent circuit. Norton's equivalent circuit uses a Constant Current Source as against a Constant Voltage Source in a Thevenin's equivalent circuit.

Norton's theorem states that any complex network can be replaced by a constant current source, I_N in parallel to a resistor, R_N as shown in Fig.2-34a, where I_N is given by the short circuit current at the output and R_N by the total resistance measured *between the output terminals with all internal voltage sources replaced by shorts, and all internal current sources replaced by open circuits.*

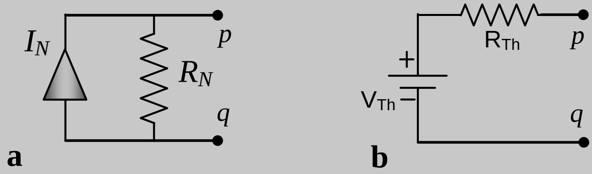


Fig.2-34: Norton's and Thevenin's equivalent circuits

(Remember, as before, an ideal voltage source has zero internal resistance, and an ideal current source has infinite internal resistance.)

To determine a Norton's equivalent circuit it may sometimes be easier to determine the Thevenin's equivalent circuit first and then convert. Therefore we need to know the correspondence between the two equivalent circuits.

Let the Norton's equivalent circuit and the Thevenin's equivalent circuit shown in Figs. 2.34a and 2.34b represent the same original circuit. Now, in Fig.2.34a, when the output is open all the current I_N flows through resistor R_N. Therefore, the open circuit output voltage is,

$$V_{OC} = I_N R_N$$

This should equal the open circuit voltage in Fig.2.-34b which is V_{Th} .

Therefore, $$I_N R_N = V_{Th} \qquad \text{... 2.27}$$

or, $$I_N = \frac{V_{Th}}{R_N} \qquad \text{... 2.28}$$

Again, when the output is shorted in Fig.2-34a, the whole of I_N flows through this shorted path (no current goes through the resistor R_N, having non-zero value). Therefore, I_N is the short circuit current in Fig.2-34a. On the other hand in Fig.2-34b, the short circuit current is given by V_{Th} / R_{Th} (Eq.2.18). Therefore equating these two we get,

$$I_N = \frac{V_{Th}}{R_{Th}} \qquad \text{... 2.29}$$

Comparing equations 2.28 and 2.29, we get,

$$R_{Th} = R_N \qquad \text{... 2.30}$$

which means that the equivalent resistances in both the circuits are the same. This also comes from the statements of both the equivalent circuits.

Thus we get the Thevenin parameters V_{Th} and R_{Th} in terms of the Norton's parameters from equations 2.27 and 2.30 respectively while the Norton's parameter I_N and R_N are obtained from the Thevenin parameters using equations 2.28 and 2.30 respectively.

Example 2.4: A Thevenin's equivalent circuit has a V_{Th} = 2V and R_{Th} = 1kΩ. What is the Norton's equivalent circuit?

Answer: From the above circuits and analyses,

$$I_N = V_{Th} / R_{Th} = 2V / 1k\Omega = 2 \text{ mA},$$

$$\text{and } R_N = R_{Th} = 1k\Omega$$

... as simple as this!

2.17 Superposition Principle:

This is mostly required where more than one source (voltage or current) is present. *Suppose you are to determine the voltage between two specific points in such a circuit. According to this principle, you break the solution into several parts, in each you obtain the voltage between the specified points considering only one source at a time while replacing each of the other voltage sources by a short circuit, and replacing each of the other current sources by an open circuit. After doing this for all the sources, you add up all the partial answers taking care of the signs (algebraic addition) to get the final result. You can determine the current through a branch instead of the voltage in the same manner.* An example will clarify the principle.

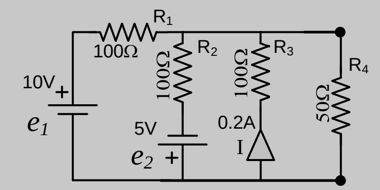


Fig.2-35: The circuit for analysis

Example 2.5: We have to find out the current through R_4 in the network shown in Fig. 2.35. Since there is more than one source, we shall use the superposition principle in three steps as follows.

Step-i:
To analyse the effect of e_1 only, we redraw the circuit first as in Fig.2-36a, where e_2 is

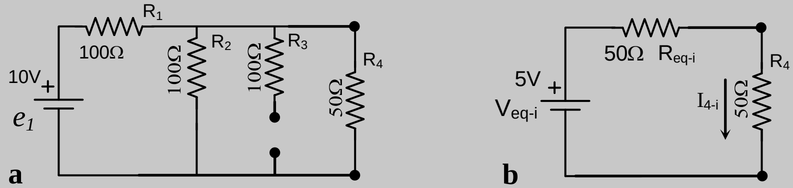


Fig.2-36: Step – i: Solution for e_1

replaced by a short and I is replaced by an open circuit.
Next we draw the Thevenin's equivalent circuit for the part to the left of R_4, which is essentially a voltage divider, in Fig.2-36b (note that R_3 has no contribution as one of its end is open).

Now we can evaluate,

$$V_{eq-i} = 10V \frac{100\Omega}{(100+100)\Omega} = 5V\ ,$$

$$\text{and,}\ R_{eq-i} = \frac{100\Omega \times 100\Omega}{(100+100)\Omega} = 50\Omega$$

Therefore,

$$I_{4-i} = \frac{5V}{(50+50)\Omega} = 0.05A = 50mA$$

Step-ii:

Next we analyse the effect of e_2 only and redraw the circuit first as in Fig.2-37a, where e_1 is replaced by a short, and I is replaced by an open circuit. Here we redraw the

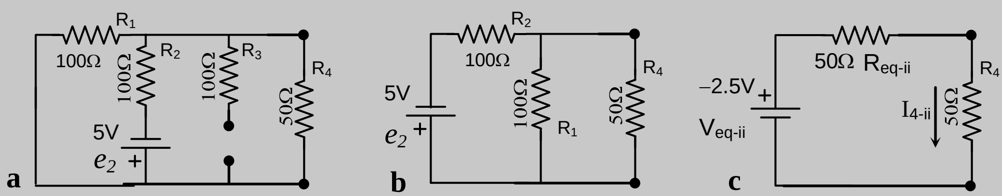


Fig.2-37: Step –ii: Solution for e_2

circuit as in Fig.2-37b where the circuit to the left of R_4 is essentially a voltage divider (note repositioning of R_1 and R_2; R_3 has been taken out as it is of no use). Note that the sign of e_2 is in opposite sense to that of e_1 and we have already represented the polarity of e_1 as positive (which we have chosen as our reference direction). Therefore, when we draw the Thevenin's equivalent circuit in Fig.2-37c, we draw the symbol with our reference polarity but place a (–) sign for the voltage value. As before, we evaluate,

$$V_{eq-ii} = -5V \frac{100\Omega}{(100+100)\Omega} = -2.5V\ ,\quad \text{and}\ R_{eq-ii} = \frac{100\Omega \times 100\Omega}{(100+100)\Omega} = 50\Omega$$

Therefore, $$I_{4-ii} = \frac{-2.5V}{(50+50)\Omega} = -0.025A = -25mA$$

Again, note that we have shown the direction of current in Fig.2-37c according to our chosen reference. The negative result above signifies that the actual current is in the opposite direction.

Step-iii:

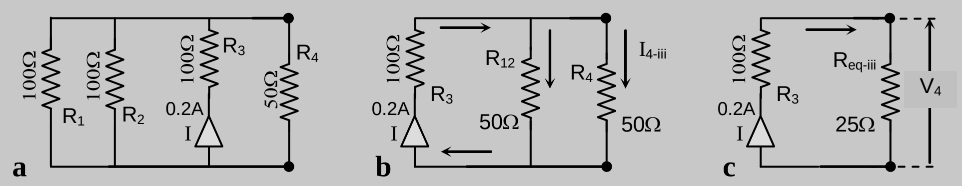


Fig.2-38: Step –iii: Solution for *I*

Next we analyse the effect of *I* only and redraw the circuit first as in Fig.2-38a, where both e_1 and e_2 are replaced by shorts. To determine the current through R_4 we draw the equivalent circuit shown in Fig.2-38b which essentially is a current divider network. Here R_{12} is the parallel combination of R_1 and R_2. Since both of these resistors equal 100Ω each, R_{12} should be exactly half of this, ie, 50Ω which we can determine even without calculation (check it for yourself through calculation). So the current of 0.2A from the source passes first through R_3 and then branches out into two paths, through R_{12} and through R_4. They then recombine together and reaches back the source. Again since these two resistances are exactly equal, they will share half of the total current each (check it yourself using Eq.2.15). Therefore the current through R_4 is,

$$I_{4\text{-}iii} = 0.2\text{A} / 2 = 100 \text{ mA}$$

Note that its direction is also the same as for $I_{4\text{-}i}$ and hence has a positive sign.

Final step, Combining all the three currents:

We shall now combine all the three part currents obtained above by adding these algebraically (ie, with appropriate signs). This gives,

$$I_4 = I_{4\text{-}i} + I_{4\text{-}ii} + I_{4\text{-}iii} = (50 - 25 + 100) \text{ mA} = 125 \text{ mA}.$$

Since this is positive, we can conclude that the resulting current direction is downwards in Fig.2-38, i.e., in the same direction as for $I_{4\text{-}i}$ and $I_{4\text{-}iii}$ and opposite to that for $I_{4\text{-}ii}$.

Thus we have solved a complex problem using Superposition Principle. Note that it has employed all the techniques for a Voltage divider, a Current divider and Thevenin's *equivalent circuit. We have not used Norton's equivalent circuit, but we could have done that as well. (Try it yourself!)*

Alternative approach for $I_{4\text{-}iii}$

In obtaining the current $I_{4\text{-}iii}$ through R_4 we could have taken another approach which we shall show now. In this approach we determine the voltage across R_4 first from which we determine the current. To do this, we redraw the circuit as in Fig.2-38c where the parallel combination of R_{12} and R_4 are combined to form the equivalent resistor,

$$R_{eq-iii} = \frac{50\Omega \times 50\Omega}{(50+50)\Omega} = 25\Omega$$

Now if V_4 is the voltage across R_4 then it is also the voltage across the parallel combination $R_{eq\text{-}iii}$. Since the whole of I (0.2A) passes through $R_{eq\text{-}iii}$ we can determine the voltage across it as,

$$V_4 = I \times R_{eq\text{-}iii} = 0.2\text{A} \times 25\ \Omega = 5\text{V}$$

Therefore, the current through R_4 would be that due to V_4 and is,

$$I_{4\text{-}iii} = V_4 / R_4 = 5\text{V} / 50\ \Omega = .1\text{A} = 100\text{mA}$$

Obviously the former current divider approach is simpler in this case where we only asked for the current through R_4. However, the latter method shows how the output voltage needs to be calculated if you have two parallel resistors driven by a current source.

2.18 Modeling devices with Single port

How to handle devices like microphones, Loudspeakers, etc. in electrical circuit analysis? Obviously, we have to model these devices either using Thevenin's or Norton's equivalent circuits. A microphone has only one signal output port as shown in Fig.2-39a, (actually there are at least two wires which are needed to convey the signal

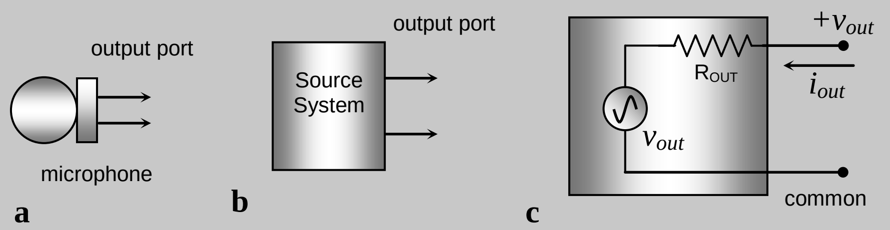


Fig.2-39: A microphone as a single port Source system and its Thevenin equivalent circuit

potential, but it is essentially a single port carrying a single signal). Why is it called a port? Well, a port is an outlet or inlet to a country through which it interacts with the outside world. Similarly any electrical device may have one or more ports to interact with the outside world. If it only sends out signals through this port, we call it an *Output Port*. Similarly if it only receives signals through this port, as happens in a Loudspeaker, we call it an *Input Port*. The microphone is a source of signal therefore we can generalise such devices as *Source Systems* as shown in Fig.2-39b which has an output port only. Now this source system can be modeled in terms of a Thevenin equivalent circuit which has been shown in Fig.2-39c. At the output we have two readily measurable parameters, output voltage v_{out} and output current i_{out}. The voltage has been shown to be + ve with respect to common for algebraic reference though the symbol v_{out} represents an ac which changes direction alternately. The reference direction of current is shown inwards for the same reason, i.e., for algebraic reference. This direction somewhat corresponds to Kirchoff's current law discussed before where all positive currents are assumed to flow inwards towards a node. It is clear that the Thevenin voltage source has a driving voltage of v_{out} while the output resistance R_{OUT} is directly related as

$$R_{OUT} = \frac{v_{out}}{i_{out}} \qquad \text{... 2.31}$$

Therefore Fig.2-39c having the above relationship for its parameters represents any *Source System* completely.

We can in the same way talk about a *Load System* such as a Loudspeaker shown in Fig.2-40a which has only an *input port*. It receives an input voltage v_{in} from a source which is not shown here and an input current i_{in} is driven into this load system. This load system can also be represented by a Thevenin's equivalent circuit consisting of a

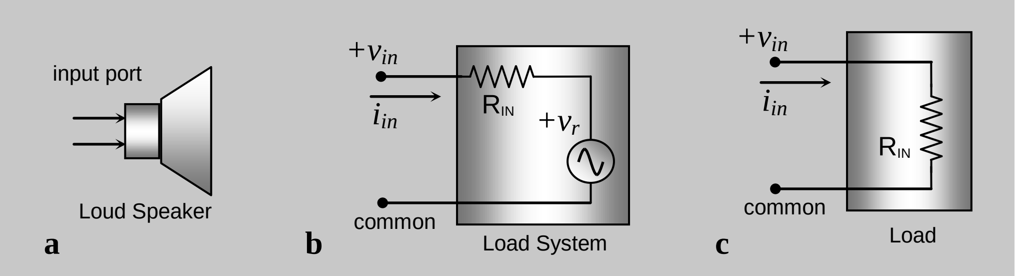


Fig.2-40: A Loudspeaker as a single port load system and its Thevenin equivalent circuits

voltage source v_r and a series resistance which we call the input resistance R_{IN} of the load system as shown in Fig.2-401b. Here v_r is a voltage source within the load system, usually called a *reverse voltage*, which is different from, and may be independent of v_{in} (the loud speaker may produce a very small voltage due to the movement of the speaker coil in a magnetic field). Therefore current i_{in} would be determined by the difference $v_{in} - v_r$ as,

$$i_{in} = \frac{v_{in} - v_r}{R_{IN}} \qquad \text{... 2.32a}$$

In most practical applications reverse voltage v_r would be negligible compared to input voltage v_{in} and we can simplify the load system as Fig.2-40c, without v_r so that the above equation becomes,

$$i_{in} = \frac{v_{in}}{R_{IN}} \qquad \text{... 2.32b}$$

The load is simply represented by a resistor, called the input resistance of the load system, or the *Load Resistance*.

Remember, we could have represented both the source and the load using the Norton's equivalent circuit as well.

For general situations with frequency dependent source and load, we shall use *Impedance* in place of *Resistance* for both the above cases. However, sometimes when we deal with a single frequency for which the impedance has a definite fixed value, we occasionally use the term *resistance* to simplify expressions, but you have to remember that this is not strictly correct.

2.19 Maximum Current, Voltage and Power transfer

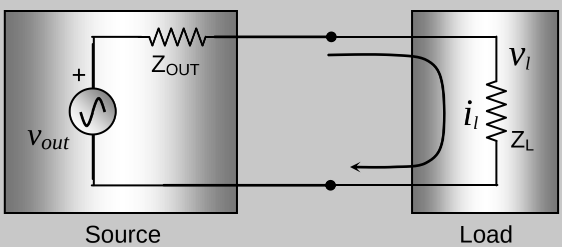


Fig.2-41: Signal Transfer system

In many occasions we have to use the electrical output of a source system to drive a load system as shown in Fig.2-41. Here we have modeled the output of the source system using a Thevenin's network, i.e., a voltage source, v_{out} with a series impedance Z_{OUT}. We have represented the loading system by simply a load impedance Z_L as discussed before. Here we have used the general *Impedance* term instead of *Resistance* used above. We have not shown any input voltage to the load here, rather we have shown the load voltage v_l which results because of the current i_l passing through Z_L. Here the load current is given by,

$$i_l = \frac{v_{out}}{Z_{OUT} + Z_L} \qquad \text{... 2.31}$$

and the voltage available to the load is given by,

$$v_l = i_l \; Z_L = v_{out} \; \frac{Z_L}{Z_{OUT} + Z_L} \qquad \text{... 2.32}$$

Therefore the power available at the load system is given by,

$$P_L = v_l \; i_l = \frac{v_{out}^2 \; Z_L}{\left(Z_{OUT} + Z_L\right)^2} \qquad \text{... 2.33}$$

The above three equations give us the resultant values obtained though such a linkage between two systems. Sometimes we are interested to find out the conditions of maximum voltage, current and power transfer from a source to a load and the above equations allow us to determine those conditions as given below.

Maximum current transfer:

From Eq.2.31, we can see that for a given Z_{OUT}, the load current i_l will be maximum (= v_{out} / Z_{OUT}) if Z_L is zero. For practical purposes we require that for maximum current transfer,

$$Z_L << Z_{OUT} \qquad \text{... 2.34}$$

Maximum voltage transfer:

From Eq.2.32, we can see that for a given Z_{OUT}, the load voltage v_l will be maximum, and equal to v_{out} if Z_L is infinity. Again for a given Z_L the load voltage v_l will be maximum if Z_{OUT} is zero. For practical purposes we require that for maximum voltage transfer,

$$Z_L >> Z_{OUT} \qquad \text{... 2.35}$$

Maximum power transfer:

From Eq.2.33 the condition for maximum power transfer is not so clear as the previous two cases. Therefore we differentiate P_L with respect to Z_L and set it to zero to find the maximum condition. Doing this exercise gives us,

$$\frac{dP_L}{dZ_L} = \frac{Z_{OUT} - Z_L}{\left(Z_{OUT} + Z_L\right)^3} = 0$$

This leaves us with, $Z_{OUT} - Z_L = 0,$

i.e.,

$$Z_{OUT} = Z_L \qquad \text{... 2.36}$$

as the required condition for maximum power transfer.

At maximum power transfer how much of the power is transferred? We can easily see from Fig.2-41 that the power will be equally divided between the output impedance and the load impedance as they are equal and are in series. Therefore the maximum power transfer can only be 50% of the input power. If Z_L is greater or lower than Z_{OUT}, the power transferred will be less.

This means then that a source circuit can transfer only 50% of its output signal power at the most to the next circuit. Note that in case of maximum current and voltage transfers, the power transfer will be much lower (analyse Eqs.2.31 and 2.32 and find out yourself).

2.20 Two Port Network

Previously we dealt with source or load systems having one port only. Frequently we encounter networks or circuits that have both an input port and an output port as shown

Fig.2-42: A generalised two-port network

in Fig.2-42. An amplifier is a good example of such a two port network. An amplifier has both an input and an output. A filter circuit, or, any signal processing circuit that takes something as an input and gives out something as an output is a two port device. Some devices may have more than two ports, i.e., they may have more than one input and more than one output. However, we will restrict our analysis to two-port devices here. The concept may easily be extended to such multi port devices.

How do we model such a complex device? Well, very simple, we separate out the input and the output of the two port network and replace them using the single port representations as discussed before. Since we have two types of equivalent circuits – Thevenin's and Norton's, and we have two ports, so we can make four combinations of these, making four types of equivalent circuits. We can also make an analytical approach which is a very elegant one, and the combination will give us a deeper understanding. Let us start with the analytical one.

2.19.1 Analytical presentation

In Fig.2.42 we can see that there are four measurable parameters – input and output voltages, input and output currents. For brevity and ease of analysis we represent the subscripts to the symbols using numbers as, v_1, v_2, i_1 and i_2 respectively. Taking any two of these as the variables, we can form six pairs of equations (= ${}^4C_2$) from which the other two parameters can be determined. In each equation-pair, the two variables are the chosen independent parameters while the other two are the dependent parameters. We are showing three of these equation-pairs below that are of interest to us from a practical viewpoint. These equation-pairs are,

$$v_1 = Z_{11}\, i_1 + Z_{12}\, i_2 \qquad \text{... 2.37a}$$

$$v_2 = Z_{21}\, i_1 + Z_{22}\, i_2 \qquad \text{... 2.37b}$$

$$i_1 = Y_{11}\, v_1 + Y_{12}\, v_2 \qquad \text{... 2.38a}$$

$$i_2 = Y_{21}\, v_1 + Y_{22}\, v_2 \qquad \text{... 2.38b}$$

$$v_1 = h_{11}\, i_i + h_{12}\, v_2 \qquad \text{... 2.39a}$$

$$i_2 = h_{21}\, i_i + h_{22}\, v_2 \qquad \text{... 2.39b}$$

If we analyse the dimensions of all the terms in these equations, we will understand why the letters representing the constant co-efficients have been chosen as shown. In the equation-pair 2.37, all the co-efficients have the dimension of impedance *(= v/i)*, therefore we used the symbol *Z*. In the next pair Eq.2-38, all the co-efficients have the dimension of admittance (= *i/v*, inverse of impedance), therefore we used the symbol *Y (=1/Z)*. In the next pair of equations, h_{11} is an impedance, h_{12} and h_{21} are plain numbers without dimensions, and h_{22} is an admittance, i.e., the co-efficients have mixed or *'hybrid'* dimensions, for which we used the first letter *h*.

Now, can we make up any circuit from the above equations? Definitely yes. Fig.2-43 gives a circuit implementation of both the equations 2.37a and 2.37b. How? First let us look at Eq.2.37a and to the left hand part of Fig.2-43. Here, $Z_{11}\, i_1$ is the potential across Z_{11} because of the current i_1. Next, $Z_{12}\, i_2$ is the magnitude of the voltage source shown

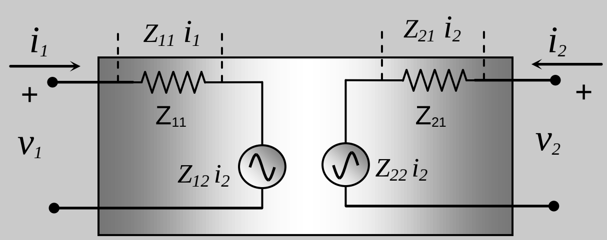


Fig.2-43: Circuit Implementation of Eq.2.37

in series (check the dimensions to satisfy yourself). Therefore the input voltage v_1 equals the sum of these two potential differences which satisfies Eq.2.37a. Similarly the right hand part of Fig.2-43 is just the representation of Eq.2.37b. Now if we look at Fig.2-43, we can see that each part is similar to a Thevenin's equivalent circuit. Therefore we can say that the equation-pair 2.37 gives a representation where both the input and the output ports are implemented by a Thevenin's equivalent circuit each.

Significance of the parameters

Let us copy Equations 2.37 a & b here for a closer look.

$$v_1 = Z_{11}\, i_1 + Z_{12}\, i_2 \qquad \text{... 2.37a}$$

$$v_2 = Z_{21}\, i_1 + Z_{22}\, i_2 \qquad \text{... 2.37b}$$

What is Z_{11} ? From Eq.2.37a, we can see that if i_2 is made 0, then $Z_{11} = v_1 / i_1$ which is the *input impedance* (= input voltage / input current). Now what condition does $i_2 = 0$ represent? It means that there is no output current, which is only possible if the output is an open circuit. From these two conditions we can say that

Z_{11} = Input Impedance with output open.

Similarly, $Z_{12} = v_1 / i_2$ for $i_1 = 0$ from Eq.2.37a. Now what is the physical significance of v_1 / i_2 ? It is the ratio of input voltage and output current, i.e., it involves the effect of an output parameter to an input parameter. We name such effects as *reverse transfer effects*. Since the ratio has the dimensions of impedance, and the input current is zero, we would say that,

Z_{12} = Reverse Transfer Impedance with input open.

Without going into further explanation we shall put the names of the other two co-efficients below. Try yourself to justify the names based on the above explanations. Just remember that any effect on output due to an input is a *forward transfer effect.*

Z_{21} = Forward Transfer Impedance with output open.

Z_{22} = Output Impedance with input open.

Therefore we can now also say which of the components in Fig.2-43 means what and how to measure them if need arises, and how the equations 2.37 a & b are represented by Thevenin's equivalent circuits.

Now, let us copy Equations 2.38 a & b here for a closer look.

$$i_1 = Y_{11}\, v_1 + Y_{12}\, v_2 \qquad \text{... 2.38a}$$

$$i_2 = Y_{21}\, v_1 + Y_{22}\, v_2 \qquad \text{... 2.38b}$$

Here, $Y_{11} = i_1 / v_1$ with $v_2 = 0$. Clearly i_1 / v_1 is the inverse of the input impedance, i.e., it is the input admittance. What condition does $v_2 = 0$ represent? A zero output voltage is only possible if the output terminals are shorted together. Therefore we would say that,

***Y_{11}* = Input Admittance with output shorted.**

Similarly we can name all the other three co-effcients as (justify yourself),

***Y_{22}* = Reverse Transfer Admittance with input shorted.**

***Y_{21}* = Forward Transfer Admittance with output shorted.**

***Y_{22}* = Output Admittance with input shorted.**

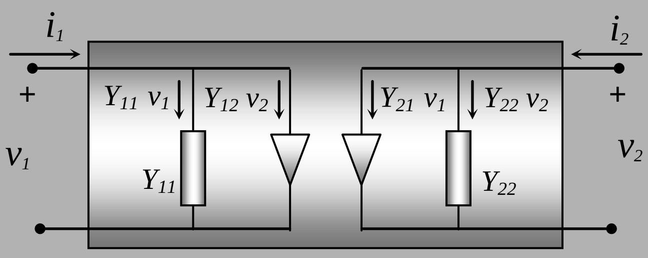


Fig.2-44: Circuit Implementation of Eq.2.38

We can see that Fig.2-44 is a good representation of equations 2.38 a & b and we can now also say which of the components in Fig.2-44 means what and how to measure them if need arises. We can also see that the circuit representations of equations 2.38a & b are nothing but Norton's equivalent circuits.

Now, let us copy Equations 2.39 a & b here for a closer look.

$$v_1 = h_{11}\ i_i\ + h_{12}\, v_2 \qquad \text{... 2.39a}$$

$$i_2 = h_{21}\ i_i\ + h_{22}\, v_2 \qquad \text{... 2.39b}$$

Following the guidance provided by the above explanations we can name these co-efficients directly as,

***h_{11}* = Input Impedance with output shorted.**

***h_{12}* = Reverse Voltage Gain with input open.**

***h_{21}* = Forward Current Gain with output shorted.**

***h_{22}* = Output Admittance with input open.**

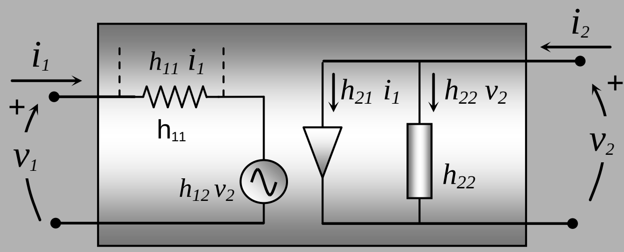


Fig.2-45: Circuit Implementation of Eq.2.39

Note that v_1/v_2 is a voltage gain term, without any dimensions, but is a reverse effect, from output to input. Similarly, i_2/i_1 is a dimensionless forward current gain term, from input to output. These also formed the basis of the above nomenclatures.

We can see that Fig.2-45 is a good representation of equations 2.39 a & b and we can now also say which of the components in Fig.2-45 means what and how to measure them if need arises. We can also see that the circuit representation of Eq.2.39a is a Thevenin's equivalent circuit while that of Eq.2.39b is nothing but a Norton's equivalent circuit. This also shows the significance of the name *'hybrid'*, which means *'mixture of various kinds'* as indicated before.

In our dealings with transistors and other devices later, we will use the two-port networks and their simplified forms extensively for analysis. Therefore one needs to have a grasp of the basics of the above representations well.

2.21 Signal and Noise

Frequently we talk of these two terms but the difference is very subtle and the terms are purely subjective. At your examination time your roommate playing a song on a cassette player may irritate you. Here the song is a signal for your roommate because (s)he is interested in it. On the other hand the same song is a noise for you as you feel disturbed by it. Therefore any pattern of electrical voltage that is of interest to a subject is an electrical signal, while all unwanted electrical voltages will be termed as electrical noise. It may happen that at a different situation, the same subject may become interested in some of voltages considered as noise before (as you may get interested in the same songs after the examination is over). The names would therefore automatically change with the interest of the subject who is describing the event.

Chapter 3
Capacitors, Inductors; dc transients

3.1 Capacitors and Capacitance

Any two conductors, placed in such a way that they do not touch each other, can store electrical charge if they are suitably connected to a source of *emf*. This arrangement is called a *Capacitor*, and this property of storing charges is described by the term *Capacitance.* The amount of charge, Q, that can be stored depend on the geometry and the nature of the arrangement, and also on the voltage, V, existing across the capacitor, and is given by,

$$Q = CV \qquad \text{... 3.1}$$

Here C is the term which is related to the geometry and the nature of the arrangement of the conductors, and is called the *Capacitance* in the technical sense. The unit for measurement of capacitance is a *Farad*, which equals a *Coulomb per Volt*, and is a very large quantity. Practical devices have capacitances of the order of microfarads, nanofarads, etc.

For a simple parallel plate capacitor, capacitance is given by,

$$C = \frac{\varepsilon A}{d} \qquad \text{... 3.2}$$

where A is the overlapping area between the plates, d is the separation between them, and ε relates to an electrical property, called the *permittivity,* of the insulating material (called a dielectric) in between the plates. The dielectric could be vacuum, air, paper, mica, plastic sheet or any other good insulating material.

3.2 How does a capacitor work?

3.2.1 Charging of a capacitor

Suppose the capacitor shown in Fig.3-1 (a & b) has vacuum as the dielectric. When the switch is just closed (Fig.3-1a), the battery (electron pump, or, source) pushes electrons out of its negative terminal to plate J of the capacitor in the direction shown, as a result of which we get an excess accumulation of negative charges there. These excess electrons repel an equal number of electrons away from plate K (Coulomb effect) which are conducted through the connecting wire and are subsequently taken in by the battery at its positive terminal. (Inside the battery, the electrons are again pushed down to its negative terminal through an chemical energy transfer process.) Thus due to a lack of electrons, plate K is charged positively and a potential difference is created across the capacitor plates. An electron flow is also initiated in the conducting parts of the circuit which can be detected by a current sensor (shown as an ammeter, A) at this moment. This is a bit awkward; the capacitor appears as a break in the circuit, there should not have been a current flow at all! Well this will happen in the long run, but for

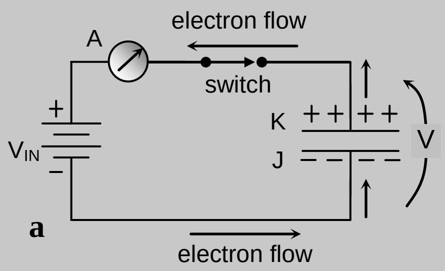


Fig.3-1: Charging (a) and discharging (b) a vacuum capacitor. Conventional current is taken to flow in the opposite direction

a brief period just after switching, we will experience a *transient (short lasting) current* in the conducting parts of the circuit due to charging of the capacitor. The current will be large at the moment of switching and will eventually die away to zero exponentially. The charge on the capacitor plates and the potential across them will also increase gradually in a corresponding manner (remember, $q=CV$). How the current dies away is explained below.

For our thinking purposes, suppose we move electrons from the battery to plate *J* in packets. Just after switching on, suppose we have moved a packet of electrons to plate *J*. This will force a similar packet to move from plate K to the battery. Since plate *J* now has an excess number of negatively charged electrons, these charges try to oppose further incoming electron-packets from the negative terminal of the battery (Coulomb repulsion again). However, if the *emf* of the battery is larger than the potential difference across the capacitor plates, it will be able to overcome the opposition and push more electron packets to plate *J* of the capacitor, but their flow rate will be somewhat reduced than before because of the opposing forces. These new electron packets accumulating on the plate *J* will in turn repel an equal number of electrons away from plate *K* at the same rate thereby increasing the potential difference across the capacitor. In this way the electron flow will continue for some time but its rate will be progressively reduced. This happens in an exponentially decaying fashion until the potential difference between the capacitor plates becomes exactly equal to the *emf* of the source. At this point the source of *emf* can no longer overcome the opposing forces, and the electron flow becomes zero. We say that the capacitor has become fully charged. The amount of excess negative charge on plate *J* is exactly equal to the amount of excess positive charge on plate *K*. During this charging transient and in equilibrium, Eq.3.1 above is valid at each point in time. We just mentioned coulomb repulsion as a cause for the decaying current. However, the resistance of the conducting wires will also have a role in controlling the actual magnitude of the current.

At this stable position, if we suddenly increase the *emf* of the source to a greater value, some more electrons will make their way to plate *J*, a transient current will flow again, and the amount of charge on the plates of the capacitor will increase to a value determined by the new *emf*, increasing the potential difference between the plates as well. Therefore we said earlier (Eq.3.1) that the quantity of charge stored on the capacitor is proportional to the voltage across it.

In the above description we referred to an electron flow as this is the physical mechanism that takes place. However, we have to remember that conventional current has the opposite direction and so we can say that when the switch is flipped 'on', a clockwise transient current flows in the circuit of Fig.3-1a, which dies away exponentially.

3.2.2 Discharging of a capacitor

Now, if we suddenly remove the battery and replace it by a short as shown in Fig.3-1b, the capacitor would not be able to hold the excess charge that it has. Therefore the excess electrons from its plate *J* will flow through the conducting wires in a reverse direction to that before in order to neutralise the lack of electrons on plate *K* of the capacitor (shown by a reverse current in the ammeter). This current will also be large in the beginning but will die away to zero exponentially when there will be no excess charge left on the capacitor plates. We can say that the capacitor with the potential *V* across it acts similar to a battery and pushes the excess electrons from its negative plate *J* to the positive plate *K* for a while, but unlike a battery the potential decreases exponentially, reducing the current too. Eq.3.1 is also valid at any point of time during discharge.

Instead of replacing the battery by a short, if we reduce the battery *emf* suddenly, the capacitor will discharge too since its potential becomes greater than the *emf* of the source; the transient will last till the two voltages become equal.

In Fig.3-1b we get a current in the circuit even when there is no battery in the circuit. Where does the energy come from? In fact the capacitance stores energy during charging, which it releases while discharging. Therefore a capacitor is basically an energy storing device and this quality has been used to obtain different desired circuit functions in electronics about which we will study more in this book.

3.2.3 Gradual change in applied potential

We only talked about sudden changes above to make a phenomenological description simple. In fact any change in the applied voltage to the capacitor, whether sudden or gradual will cause charging or discharging of the capacitor as appropriate. In fact we can define a current through a capacitor as,

$$i = \frac{dq}{dt} = C\,\frac{dv}{dt} \qquad \text{... 3.3}$$

where *C* is a constant for a particular capacitor. Note we have used small letters for the symbols as these are changing.

3.3 What is Capacitance?

In the above description, the quantity of stored charge is also proportional to the capacitance of the capacitor. The larger the area of the capacitor plates, more charge can be stored. What is *Capacitance* then? By rearranging Eq.3.1 we can see that

$$C = Q/V$$

which says that the capacitance is the *quantity of charge stored per unit potential difference.*

This can be compared with the water storage capacity of a bucket, though the analogy may not be 100% compatible. For a cylindrical bucket the capacity can be stated as the mass of water it can hold per unit height (say, 100 kg/m) similar to the *charge per unit potential difference (dimension: Coulombs/Volt) in an electrical capacitor*. In the bucket the mass stored will be proportional to the height of water, similarly in a capacitor, the charge stored is proportional to the voltage. Again by increasing the capacity of the bucket by increasing the area of cross section, we can increase the mass stored for the same height. Similarly, in a capacitor, by increasing its capacitance by changing ε, *A or d* appropriately, we can increase the quantity of charge stored for the same voltage.

3.4 Effect of ε, *A or d*

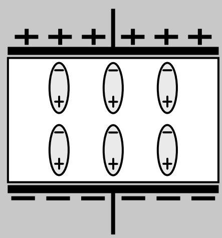


Fig.3-2: Polarisation within a dielectric in a capacitor

Increasing area *A* allows more charge to be stored at the same voltage resulting in an increased capacitance, and is simple to understand. The increase in capacitance with a reduced separation *d* (in vacuum) may be thought of as due to an increased interaction of the excess charges on the two plates. Increased Coulomb force will cause more electrons to be repelled from the positive plate, and thus both the plates will have a greater number of excess charge. For any other insulating material except vacuum as the dielectric, the electric field between the plates of the capacitor will polarise the molecules of the dielectric through displacement of the centres of positive and negative charges as shown in Fig.3.2. This causes the opposite charges between the plates and the adjacent polarised molecules to come at close proximity to each other and thereby interact more (similar to that for a reduced *d*). This in turn allows the plates to hold a greater number of excess charges, resulting in an increased capacitance. The higher the polarisability of the dielectric (which is given by the permittivity ε), higher is the number of polarised charges that come near the plates and higher is the capacitance. Another way of looking at it is that the charges of the dielectric molecules near the plates tend to neutralise the effect of some of the opposite charges on the respective plates, which in turn tends to reduce the potential between the plates. If the potential is held constant by a battery, more excess charges need to be pushed into the plates, and the capacitance is increased.

There are two types of dielectric - polar and non-polar. In a polar dielectric, individual molecules are normally polarised but are randomly oriented, so that there is no net polarisation. In the presence of an electric field these become oriented as in Fig.3-2, i.e., become polarised. On the other hand non-polar molecules are not normally polarised, i.e., the centres of positive and negative charges coincide. However, when an electric field is applied, they also become polarised, same as that for a polar dielectric as shown in Fig.3-2.

3.5 Displacement current

We can see that while a transient current flows in the outside circuit, there is no electron flow within the region between the plates. Only the electric field in this region changes during the time that the capacitor is getting charged. If there is a dielectric material inside, the amount of polarisation will change, thus changing the amount of displaced charges both on the plates and inside the dielectric. However, we would like to imagine a continuity of current in the whole circuit. We say that a *displacement current* flows within the capacitor (due to displacement of charges), which equals the conducting current elsewhere in the circuit.

3.6 How a capacitor affects a dc circuit, dc transients

The capacitor essentially looks like a break in a circuit. So there should not be any stable current through a capacitor with a dc source as mentioned before. However, at the moment of switching to a dc supply, transient currents flow as explained above. Transient currents also flow after the dc supply is switched off. The natures of these flow patterns when such step voltages are applied are of interest to us and are dealt with analytically in the following sections.

3.6.1 Switching ON, charging of a capacitor

Let us consider the circuit of Fig.3-3 with the change-over switch initially at position 2.

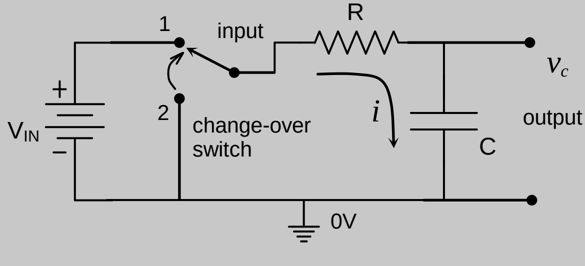


Fig.3-3: dc transients through a capacitor

The circuit is in a stable condition with the capacitor fully discharged, the voltage v_C across it being zero. As soon as the switch is flipped from 2 to 1 (say, at time t = 0), a transient current *i* flows for a certain period as explained above, charging up capacitor *C*. The conventional current direction is shown in the figure. The current gradually decreases becoming zero at infinite time.

Analysis

To analyse the circuit we can use Kirchoff's law for voltage around the loop to get,

$$V_{IN} = iR + q/C \qquad \text{... 3.4a}$$

where *q* is the instantaneous charge on the capacitor, and the corresponding voltage across capacitor $v_C = q/C$. We have used small letters for *i* and *q* as these are changing. Note that we have shown a common or ground terminal which has the reference voltage of 0V and all voltages are referred to this terminal.

Differentiating Eq.3.4 and replacing *dq/dt* by *i* we get,

$$0 = R\frac{di}{dt} + \frac{i}{C} \qquad \text{... 3.4b}$$

Reorganising, we get, $\frac{1}{i}\frac{di}{dt} = -\frac{1}{RC}$

Integrating with respect to time, $\int \left[\frac{1}{i}\frac{di}{dt}\right] dt = -\frac{1}{RC}\int dt$

Now, L.H.S. $= \int \left[\frac{1}{i}\frac{di}{dt}\right] dt = \int \frac{1}{i}\,di = \ln i$ (ignoring constants)

Therefore, $\ln i = -\frac{t}{RC} + K$

where *K* is a constant. Taking exponentials, we get,

$$i = e^K\, e^{-t/RC}$$

where we have to find the unknown constant *K* from known conditions. If the initial current at time $t = 0$ is I_0, then from the above equation, we get,

$$e^K = I_0$$

Therefore we can write,

$$\boxed{i = I_0 \; e^{-t/\tau_C}} \qquad \text{... 3.5}$$

where, $\tau_C = RC$ ... 3.6

is called the *time-constant* of the *RC* circuit.

Checking the fundamental dimensions of *R* (= *V/I*) and *C*(= *Q/V*) you will find that the dimension of the product *RC* reduces to that of time as shown below,

$$\frac{voltage}{current} \times \frac{charge}{voltage} = \frac{charge}{charge/time} = time$$

At time $t = 0$, there is no excess charge on the capacitor plates so the initial current sees no obstruction or resistance in the capacitor, the capacitor appears as a short circuit. The only resistance in the circuit is *R* that limits this current. So the initial current at time $t = 0$ is given by,

$$I_0 = V_{IN}/R \qquad \text{... 3.7}$$

The resulting temporal behaviour of the current is shown in Fig. 3-4 (white line, note scale on the left).

To obtain the voltage across the capacitor, v_C (= q/C), we rewrite Eq. 3.4 as,

$$v_C = V_{IN} - iR.$$

Now using Eq.3.5 and Eq.3.7, we get,

$$v_C = V_{IN} - \frac{V_{IN}}{R} R \, e^{-t/\tau_C}$$

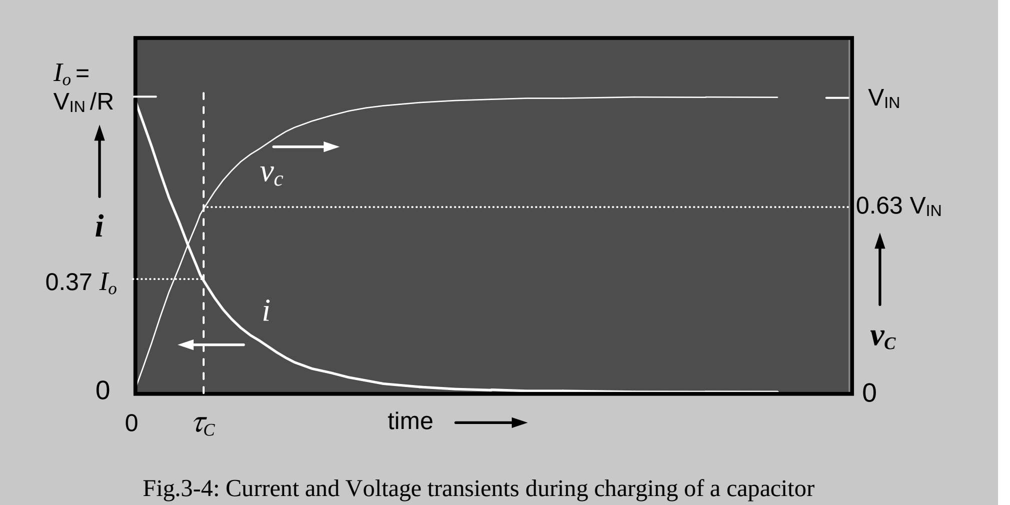


Fig.3-4: Current and Voltage transients during charging of a capacitor

or, $$\boxed{v_C = V_{IN}\ (1 - e^{-t/\tau_C}\ )} \qquad \text{... 3.8}$$

This shows that the capacitor voltage, v_C, increases asymptotically from an initial value of zero, reaching V_{IN} at time $t = \propto$ (infinity). The resulting temporal behaviour of the capacitor voltage is shown in Fig. 3-4 (black line, note scale on the right).

3.6.2 Significance of time constant

What is special about the time constant? To find it let us determine the values of the voltages and currents after a delay of one time constant.

From Eq.3.5 and Eq.3.8, we get, when $t = \tau_C$,

$$i = I_0 / e$$

and $$v_C = V_{IN}\ (1 - 1/e).$$

This signifies that after a time equal to the time constant, the current becomes *1/e* th of the initial maximum value while the voltage becomes *(1-1/e)* th of the final maximum value (remember, e ≅ 2.718, so that these values are about 37% and 63% respectively). These points are also indicated in Fig.3-4. After a period of *n* time constants, you will find that the current has become $1/e^n$ th and the voltage has become $(1 - 1/e^n\ )$ th of the respective maximum values.

Since $e^5 \cong 148$ and $e^6 \cong 403$ which are far greater than 1, we can say that after a period of about 5 or 6 time constants, both the current and the voltage may be considered to have reached almost the final long term values for practical purposes, though ideally it ought to take an infinite time.

Therefore the time constant gives us an idea about the time it takes for the capacitor to charge and discharge to any specific percentage of the maximum value. Besides, time constants allow us to compare the temporal behaviours of different circuits.

Example Question: Let V_{IN} = 10V, R=1kΩ and C=100μF in the circuit of Fig.3-3. Let the capacitor be in a fully discharged condition initially with switch at position 2. After switching to position 1, at approximately what time would the voltage across the capacitor be 6.3V? What would be the current at that instant?

Answer: 6.3V is 63% of the maximum value (=10V), and this is attained after one time constant, which equals, *RC* = 1kΩ x 1μF = <u>100mSec</u>. The current at that instant is about 37% of the maximum, which should be, *i* = 0.37 x 10V/1kΩ = <u>3.7mA</u>

Note: For any other value of voltage or time or current, use Eq.3.5 to Eq.3.7 as appropriate, and solve.

3.6.3 Application: Timer

One useful application of such an *RC* circuit is in producing a time delay circuit. Fig.3.5a shows the basic scheme of such a circuit. Here the capacitor voltage drives a voltage sensitive (VS) circuit. Suppose the output of this VS circuit is at 0V if its input is below a certain percentage of the input battery voltage (called a threshold level, V_{Th}).

If the input voltage crosses the threshold level the output of the VS circuit becomes high, say equal to the battery voltage. The basic working of the timer circuit is described below with the help of Fig.3-5 b & c which shows the time variations of the voltages at the input and output of the VS circuit respectively.

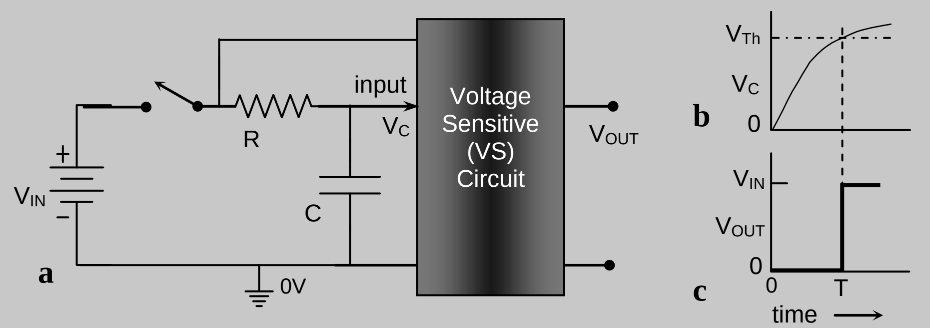


Fig.3-5: Timer circuit basics (a) and the relevant waveforms (b & c)

Suppose at time t = 0 the *RC* circuit is switched to the dc source resulting in a step voltage on the left side of *R*. The voltage of the capacitor will rise asymptotically according to Eq.3.8 and as shown in Fig.3.5b. As long as the capacitor voltage is below the threshold level V_{Th}, the output of the VS circuit is low, at 0V, as shown in Fig.3.5c. After a time *T* the capacitor voltage v_C crosses V_{Th}, and the output of the VS circuit suddenly goes high. This output can be used to drive any alarm circuit or to switch any other device according to requirement. Thus we have made a timer device which is actuated after a certain time interval of switching it to a dc supply. The threshold level may be set at any suitable level depending on the design of the VS circuit. The time delay may be changed using either *R* or *C*, though changing *R* is practically more convenient and cost effective. To use the circuit repetitively an arrangement to reset the device by discharging the capacitor needs to be made. Such practical circuits will be discussed later in the 2nd volume of the book where it will be shown that such automatic switching can be performed using semiconductor devices like transistors and integrated circuits. Timings from nanoseconds to tens of seconds can be achieved in practice using appropriate choice of component values.

3.6.4 Energy storage

Note that during discharge in Fig.3-5, there is no battery in the circuit. What is the source of *emf* needed to drive a current, and where does the energy come from? Well, the previous charging process has left excess charges stored on the capacitor having a potent energy that can be calculated as follows.

The rate of energy storage, or, the energy E stored per unit time (which is equal to power) is given by (= voltage × current),

$$\boxed{\frac{dE}{dt} = \frac{q}{C}\frac{dq}{dt}} \quad \text{... 3.8a}$$

Now the total energy stored till time T starting from 0 (during which an excess charge Q is transferred starting from an initial excess charge of 0) is,

$$E = \frac{1}{C}\int_0^T q\frac{dq}{dt}\,dt = \frac{1}{C}\int_0^Q q\,dq$$

i.e.,

$$\boxed{E = \frac{1}{2}\frac{Q^2}{C} = \frac{1}{2}QV = \frac{1}{2}CV^2} \quad \text{... 3.8b}$$

This energy is stored in the electric field that is created by the excess charges on the capacitor plates, which the capacitor can release on discharge.

3.6.5 Switching OFF, discharging the capacitor

Suppose initially the capacitor is in fully charged state with the switch at position 1. Let us see what happens when the switch is flipped to position 2 suddenly as shown in Fig.3-5. Here we can see that the electrons on the lower plate of the capacitor will flow through the closed circuit through R and will gradually neutralise the positive charges on the upper plate of the capacitor as mentioned before. This process is called *discharging* of capacitor. It is also apparent that there will be an initial current in the circuit, but it will be in the reverse direction to the charging current discussed previously. The capacitor initially has a potential difference exactly equal to that of the battery, i.e., V_{IN}, with the polarity shown (upper plate at a higher potential than the lower one). So the capacitor now acts as the new source of *emf* and produces an initial current of the magnitude (= V_{IN}/R, same as that obtained initially during charging) but in the opposite direction. The current gradually dies away as the capacitor gets discharged. Since we denoted the previous charging current as positive, we have to call this discharging current negative.

Analysis

Kirchoff's law for voltage around the loop gives us, initial voltage being zero,

$$0 = i\,R + q/C \qquad \text{... 3.9a}$$

Differentiating Eq.3.8 and replacing $i = dq/dt$ as before,

$$0 = R\,\frac{di}{dt} + \frac{i}{C} \qquad \text{... 3.9b}$$

This is exactly the same as equation as Eq.3.4b. Therefore the same solution can be applied except for the initial condition. In this case, the capacitor has an initial voltage of $-V_{IN}$ (minus sign to show reverse direction compared to that of the battery in the original loop.) as shown in Fig.3-5 and therefore the initial current will be $-I_0$ $(= -V_{IN}/R)$. The current will also be opposite to that during charging. Therefore we can write,

$$\boxed{i = -I_0\, e^{-t/\tau_C}} \qquad \text{... 3.10}$$

where τ_C is the time constant RC as before. To obtain the voltage across the capacitor, v_C $(= q/C)$, we rewrite Eq. 3.9a as,

$$v_C = -\,i\,R$$

which becomes, using Eq.3.10,

$$\boxed{v_C = V_{IN}\; e^{-t/\tau_C}} \qquad \text{... 3.11}$$

The transients are shown in Fig.3-6 a & b.

Note that for current i the polarity is reversed (– ve) but its magnitude is decreasing similar to that during charging from an initial magnitude of I_0. Capacitor voltage v_C has the same polarity as before, except that it is decreasing from a maximum initial value towards zero exponentially. (We did not draw these on the same graph as before to point out that the current goes negative from an initial zero value)

3.6.6 Voltage across Resistor

How would be the behaviour of the voltage across the resistor R ? Since this voltage equals the product $i\,R$ where R is a constant in a particular circuit, its behaviour would be exactly the same as that for the current i in both the above cases for charge and discharge (Fig.3-4 & 3-6a). In these cases the initial values will be $I_o\,R$ and $-\,I_o\,R$ respectively.

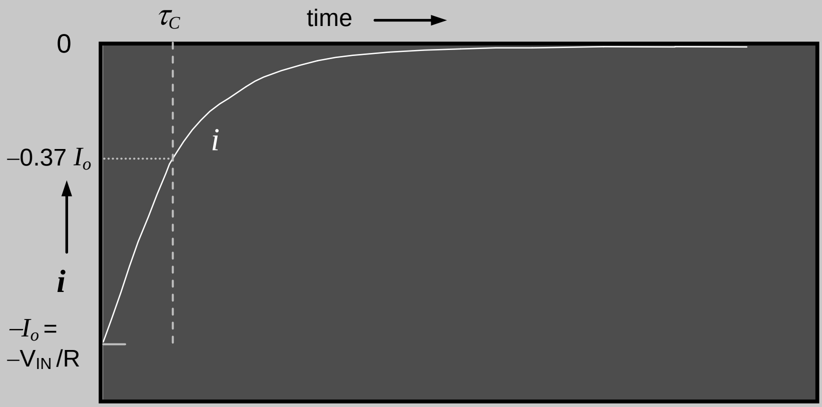


Fig.3-6a: Current transient during discharging of a capacitor

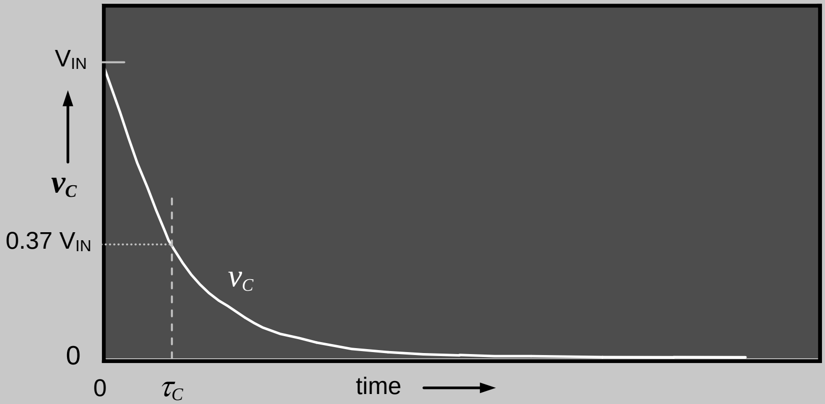


Fig.3-6b: Voltage transient during discharging of a capacitor

3.6.7 Repetitive switching

If we continue to flip the switch between positions 1 and 2 in Fig.3-3, we will get interesting patterns at the output, either across the capacitor or across the resistor. Such switching can be done electronically (using a signal generator) when we call it a square wave signal. This switching can be very fast (less than nanoseconds), or very slow. This arrangement is shown in Fig.3-7 where an electronic square wave generator drives an *RC* circuit and the output is taken once across the capacitor (a) and once across the resistor (b). Several graphs of these two types of outputs are shown in Figs. 3-8 and 3-9 for varying relationships between the period *T* of the square wave and the *RC* time constant of the circuit. Such waveforms can be the basis of many wave-shaping circuits.

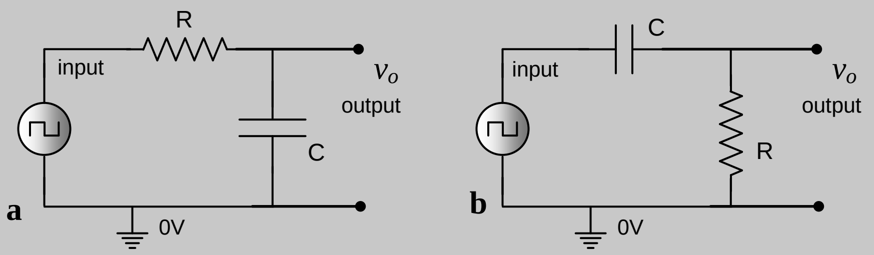


Fig.3-7: Two forms of an RC circuit driven by a square wave generator

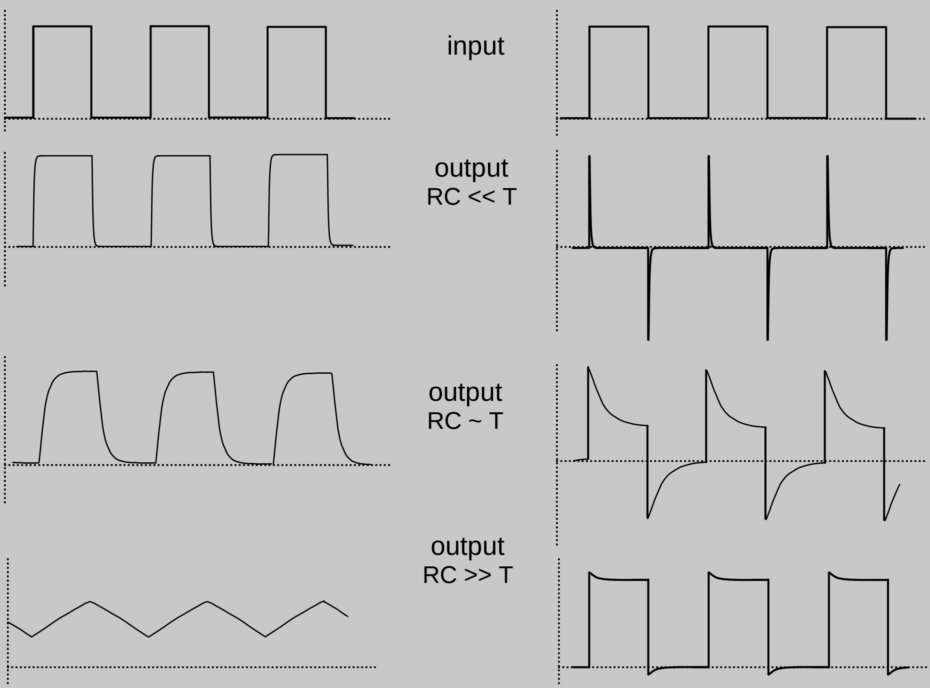


Fig.3-8: Output across capacitor

Fig.3-9: Output across resistor

3.6.8 Capacitor behaviour

We have seen in the above discussions that whenever the voltage at the input goes suddenly from zero to some nonzero value, the capacitor in parallel to the output takes some time to reach that voltage; it does not change its voltage instantly. Similarly when the voltage goes suddenly from a non-zero value to zero, the capacitor voltage does not follow the change instantly, it takes a while. Therefore, we can say that a capacitor is conservative, it tries to hold on to its previous value of voltage and does not allow it to change suddenly. This quality of the capacitor is used in many applications to introduce a delay in the circuit as we have seen in the timer example before, or to filter out high frequency signals, or to smooth out sharp variations in voltage. Here the capacitor usually appears in parallel to the output as in Fig.3-7a.

On the other hand for a step voltage change at the input the capacitor initially allows all the current the circuit can pass, as if the capacitor was a short circuit. This quality may be used to pass sudden changes in current but to block slowly varying, or dc currents in a circuit. Here the capacitor usually appears in series to the output as in Fig.3.7b.

Both the above behaviours of a capacitor will have important bearings in the case of *ac* voltages and currents. The working of a capacitor with *ac* is described in detail the next chapter.

Box3.1: Practical Capacitor model

A practical capacitor may have leakage resistance associated with it. Its model is shown in Fig. B.3-1 which is basically an ideal capacitor in parallel with its leakage resistance. In most cases this leakage resistance is very high and can be taken to be an open circuit, and therefore, ignored. One might also consider the inductance of the lead wires, which might become significant only at very high frequencies. These inductances may be modeled in series with the above assembly.

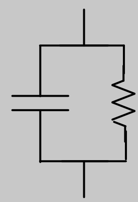


Fig.B3-1: The equivalent model of a practical capacitor

Box 3.2: Series and parallel capacitance, which one dominates?

We discussed series and parallel resistance combination in section 2.8. For a capacitor it is the reverse, i.e., for a series combination of two capacitors C_1 and C_2 (Fig.B3.2.1) the equivalent capacitance C_{ser} is,

$$C_{ser} = \left[\frac{1}{C_1} + \frac{1}{C_2}\right]^{-1} \quad = \frac{C_1 C_2}{C_1 + C_2} \qquad \text{... B3.2.1}$$

while for a parallel combination (Fig.B3.2.2), the equivalent capacitance C_{par} is

$$C_{par} = C_1 + C_2 \qquad \text{... B3.2.2}$$

If they have widely different values which one dominates? In series combination it is the smaller one since the plates of all the capacitors should hold equal amount

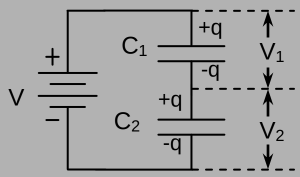


Fig.B3-1: Series combination of capacitors

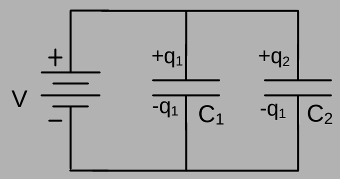


Fig.B3-2: parallel combination of capacitors

of charge, and it is the smaller one which determines what is the maximum amount of charge that can be held under a certain situation. This obviously requires that the voltages across the two capacitors are unequal, being smaller for the larger capacitance (since $q = CV$). On the other hand for the parallel combination, the voltage across both the capacitors are equal, therefore the charge on each would be unequal. Here it is the larger one that will dominate since the total charge that can be held by both the capacitors determines the combined capacitance which is the algebraic sum of the two charge amounts.

3.7 Mutual Inductance

Fig.3-10 shows two conducting wires *ab* and *cd* placed very close together. The left wire is connected to a source e_{in} of time varying *emf* while the right wire is left open for the time being. Due to a varying current i_l through the left wire, a varying electric field will be produced around it (shown by white lines with arrow). Now under the influence of this varying electric field, an *emf* e_r will be induced in the right wire (remember, from basic electromagnetic theory that if the current does not vary with time, the magnetic flux will not vary and no *emf* will be induced). If a load is connected across the right wire, it will be able to drive a current through the load because of the induced *emf*.

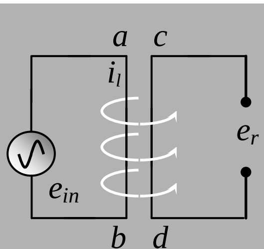


Fig.3-10: Mutual induction between straight wires

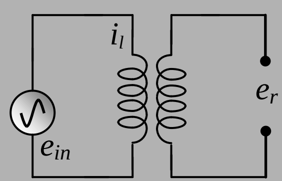


Fig.3-11: Mutual induction between coils of wires

Now, if we wind the straight wires into coils as shown in Fig.3-11, the induction effect will increase manifold as each turn can affect all other turns of wire. If we have a magnetic core passing through both the coils, the induced *emf* will increase many times more. The magnitude of the induced *emf* in the right coil e_r will be proportional to the rate of change of current (di_l/dt) in the left coil. This we can define analytically as,

$$e_r \propto -\frac{di_l}{dt}$$

or, introducing a proportionality constant *M*, we have,

$$\boxed{e_r = -\ M\ \frac{di_l}{dt}} \qquad \text{... 3.12}$$

Here a negative sign is conventionally introduced to indicate that the induced *emf* opposes the changing current direction.

Again if we have a varying current driven through the right coil, an *emf* will be induced on the left coil that can be described using a similar expression as Eq.3.12, except that the subscripts for *left* and *right* will be interchanged. The constant M will remain the same if nothing else is changed in the two-coil arrangement. Because each coil affects the other this is a case of *mutual induction*.

Here the constant *M* depends on the physical parameters of both the coils, their geometry and magnetic characteristics of materials in the neighbourhood or within the coils, which determine how much of the electric field of the left coil will affect the induced *emf* in the right coil, and *vice-versa*. Therefore this constant *M* is called the *Mutual Inductance* of the two-coil arrangement. Its unit is called *Henry*. From Eq.3.12 it can be seen that *Henry* (or *H* in short) is equivalent to *volt-sec per amp*.

3.8 Self Inductance, Inductor

Suppose in Fig.3-10 we bring the right wire closer and closer to the left wire, till at the extreme situation they become the same as shown in Fig.3-12. Now the wire is in a magnetic field due to a current going through itself. Now if the driving current varies an induced *emf* will be generated within the wire itself. In this case also if we wind the straight wire into a coil as shown in Fig.3-13a, the induction effect will increase manifold as each turn can affect all other turns of wire. If we have a magnetic core passing through the coil, the induced *emf* will increase many times more. This induced *emf* e_l is different from the source e_{in} that supplies the driving current, but it will appear in parallel with the input voltage as shown in the equivalent circuit in Fig.3-13b and it will always try to oppose the change in the driving current. The components Z_1 and Z_2 represent the respective internal impedances of the source and the coil respectively.

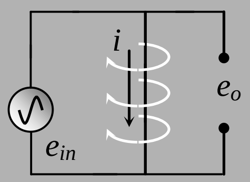


Fig.3-12: Self induction in a straight wire

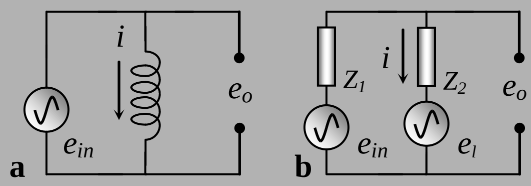


Fig.3-13: Self induction in a coil of wire (a) and its equivalent circuit (b)

We should realise that when current is increasing in the direction shown through the coil in Fig.3-13, the polarity of the induced *emf* will be positive at the upper terminal of the coil to oppose the increase. On the other hand when the current is decreasing, but still flowing in the same direction, the inductor will pump in extra current in the same direction as the input current in order to keep it from falling, so the polarity of the induced *emf* will be reversed.

Where does the inductor get energy to pump in this extra current? In fact when the current is increasing, the inductor, by opposing it, stores some energy in its magnetic field. This stored energy is released later, during the fall of input current. Just as a capacitor stores energy in its electric field, similarly an inductor stores energy in its magnetic field.

The induced *emf*, e_l, in this case is proportional to the rate of change of current *(di /dt)* that flows through the wire (or the coil) itself and is given by,

$$e_l \infty -\frac{di}{dt}$$

where the minus sign indicates the opposition discussed above.

Introducing a proportionality constant *L*, we have,

$$\boxed{e_l = -L\frac{di}{dt}} \qquad \text{... 3.13}$$

Here the constant *L* depends on the physical parameters of the coil, its geometry, and magnetic characteristics of materials in the neighbourhood or within itself, which will determine the magnitude of the induced *emf*. Since this is a case of self induction, this constant *L* is called the *Self Inductance* of the coil or of the straight wire, as the case may be. From the above discussion, it is clear that even a single straight piece of wire will have a self inductance, but obviously its magnitude is very small. Usually we use coils where self inductance is needed, with or without a magnetic core depending on requirements. In a coil, magnetic field produced by each turn is coupled not only to itself, but to many other adjacent coils thus increasing the inductance many times. Such a coil is known as an *Inductor*. The unit of self inductance is also *Henry* as for mutual inductance.

From the above discussion it is also clear that when the input current is increasing, the induced *emf* will have a polarity to oppose the input. For this reason the induced *emf* of an inductor is sometimes called a *back-emf*. However, when the input current is decreasing, the induced *emf* has the same polarity as the input *emf*. Therefore the term *back-emf* is not always valid.

3.9 Lenz's law and Conservation of energy

In the above discussion we have seen that the induced *emf* tries to oppose the change in the input current. This is stated through *'Lenz's Law'* which states that, *'The induced current will appear in such a direction that it opposes the change that produced it'*. The essence of this law can also be applied to the case of the capacitor circuit described before. When the initial electrons flow from the negative terminal of the battery to become excess stored charges on the lower plate of the capacitor (Fig.3.1), these excess electrons oppose further transfer of electrons from the battery, opposing the very cause which produced them. Had it been the other way round, i.e., if the storage of excess electrons on the lower capacitor plate aided further transfer, then the current would continuously grow to infinity producing energy from nowhere, which is an impossibility, and would violate *conservation of energy* principles. Thus we can generalise Lenz's law to all activities in nature. ***"A product of a change opposes the very cause that produced it"***.

3.10 LR circuit, dc transients

Fig.3-14 shows an LR circuit where the output voltage is taken across the inductor. Let us consider the circuit with the switch initially at position 2. The circuit is in a stable condition with no current in the circuit. As soon as the switch is flipped to 1 (say, at time t = 0), the current does not become equal to the final value instantaneously, rather it grows gradually over a certain period as explained above, because of the opposition provided by the inductor, storing magnetic energy in the process.

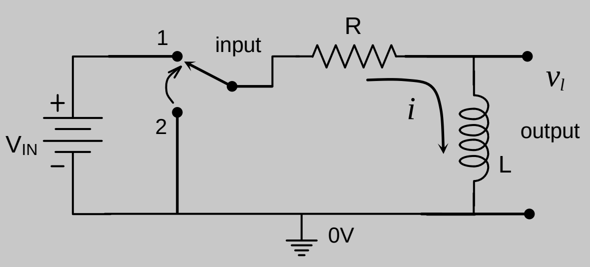


Fig.3-14: dc transients through an inductor

Analysis

To analyse the circuit we can use Kirchoff's law for voltage around the loop to get (voltage across inductor = $L\, di/dt$, with the appropriate sign),

$$V_{IN} = iR + L\, di/dt \qquad \text{... 3.14a}$$

Reorganising, we can write,

$$\frac{di}{dt} = \frac{1}{L}\left(V_{IN} - iR\right) = \frac{R}{L}\left(\frac{V_{IN}}{R} - i\right) \qquad \text{... 3.14b}$$

We need a trick, called a change of variable, to integrate this equation. Let us define a new variable,

$$z = \left(\frac{V_{IN}}{R} - i\right) \qquad \text{... 3.15}$$

then $\dfrac{dz}{dt} = -\dfrac{di}{dt}$, and Eq.3-14b becomes

$$\frac{dz}{dt} = -\frac{R}{L}\, z \,, \quad \text{or,} \quad \frac{1}{z}\frac{dz}{dt} = -\frac{R}{L}$$

Integrating with respect to time,

$$\int \left[\frac{1}{z}\frac{dz}{dt}\right] dt = -\frac{R}{L}\int dt$$

Now, LHS = $\displaystyle\int \frac{1}{z}\, dz = \ln z$ (ignoring constants),

Therefore we have,

$$\ln z = -t\,\frac{R}{L} + K \,, \qquad \text{where K is a constant.}$$

Taking exponentials, we get, $z = e^{K}\, e^{-tR/L}$

Replacing z in Eq.3.15, $\displaystyle i = \frac{V_{IN}}{R} - e^{K}\, e^{-tR/L}$

Now, we will find the constants using known conditions. At time $t = 0$, the current i is zero (as the inductor does not allow the current to change). Therefore,

$$e^{K} = \frac{V_{IN}}{R}$$

Again, at $t = \propto$ *(infinity)*, there is no obstruction or resistance from the inductor so the final current I_o sees only the resistor R in the circuit as,

$$I_0 = \frac{V_{IN}}{R} \qquad \text{... 3.16}$$

Therefore we can write,

$$\boxed{i = I_0\ (1 - e^{-t/\tau_L})} \qquad \text{... 3.17}$$

$$\text{where,} \quad \tau_L = L / R \qquad \text{... 3.18}$$

is called the *time-constant* of the *LR* circuit.

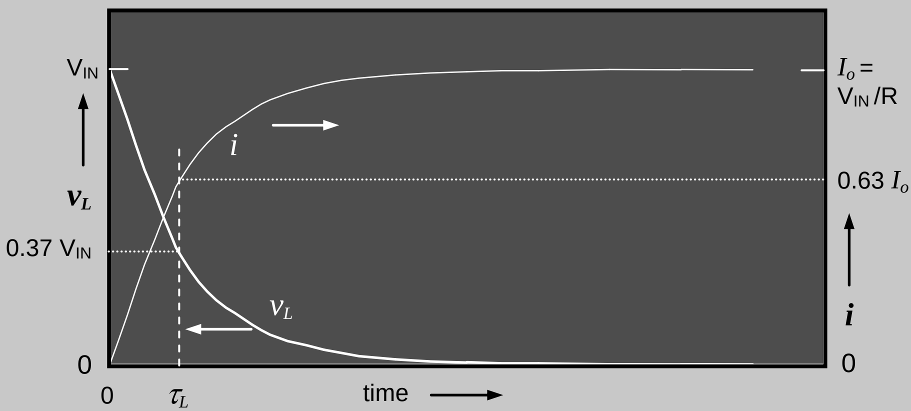


Fig.3-15: Current and Voltage transients in an inductor while switching on

The resulting temporal behaviour of the inductor current is shown in Fig. 3-15 (white line) which shows that the current increases asymptotically from zero to the maximum value with a time constant τ_L.

The voltage across the inductor, v_L, is obtained from Eqs. 3.13 and 3.17 as,

$$v_L = L\frac{di}{dt} = \frac{L\,I_o}{\tau_L}\,e^{-t/\tau_L}$$

Evaluating the constant term using Eq.3.16 and Eq.3.18,

$$\boxed{v_L = V_{IN}\ e^{-t/\tau_L}} \qquad \text{... 3.19}$$

This shows that the inductor voltage, v_L, decreases exponentially from an initial value of V_{IN}, reaching zero at time $t = \propto$ (infinity). The resulting temporal behaviour of the inductor voltage is shown in Fig. 3-15(black line).

Let us now consider the circuit when the switch is flipped back to point 2 from position 1 long after the transients have subsided. A similar analysis will show (do it yourself!) that

$$\boxed{i = I_0 \ e^{-t/\tau_L}} \qquad \text{... 3.20}$$

and

$$\boxed{v_L = -V_{IN} \ e^{-t/\tau_L}} \qquad \text{... 3.21}$$

The time responses of these quantities are shown in Fig.3.16 a & b.

The above equations show that an exponentially decreasing current is maintained for some time even though there is no battery in the circuit. Where does the energy come from? It is the magnetic energy stored in the inductor that supplies this current. It can also be seen that the direction of voltage across the inductor is now reversed as *di/dt* has the opposite sign.

In all the above treatment we have considered an ideal Inductor having zero resistance. In practice this is not possible since any coil of wire will have some resistance. However, in most cases this resistance may be ignored compared to the series resistances involved in the circuit. If it is not the case we can model the inductor as an ideal inductor with the coil resistance in series. While evaluating the voltage across the inductor, we have to add the contribution of the voltage dropped across its coil resistance, which makes it somewhat complex, but it can be solved.

3.10.1 Voltage across resistor

The voltage across resistor in the circuit of Fig.3.14 is simply *iR* whose behaviour is the same as that for *i* since *R* is a constant for a particular circuit.

3.10.2 Repetitive switching

If we carry on repetitive switching in between points 1 and 2 in Fig.3-14, we shall get waveforms similar to those shown in Fig.3.8 and Fig.3.9 for the RC circuit, but for appropriate situations and for appropriate points in this circuit. (We will leave it to you to figure these out yourself.)

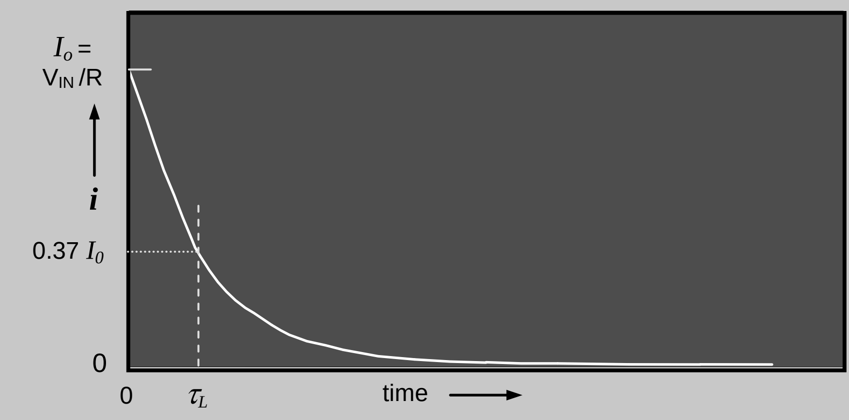


Fig.3-16a: Current transient during switching over of an inductor to ground

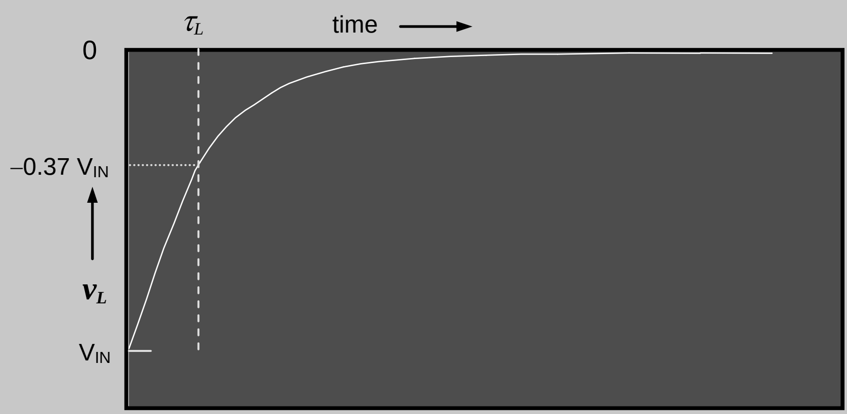


Fig.3-16b: Voltage transient during switching over of an inductor to ground

3.10.3 Energy stored by an Inductor

We have seen above that a current exists even after the battery is switched off. Where does the energy come from? In a similar way to the capacitor, the Inductor also stores energy, which it can release on demand. The energy stored per unit time, or, the power is given by (= voltage × current),

$$\frac{dE}{dt} = L\frac{di}{dt} \times i \qquad \text{... 3.22a}$$

and at any final current I, starting from an initial zero current, the total energy transferred is given by,

$$\boxed{E = \frac{1}{2}\, L\, I^2} \qquad \text{... 3.22b}$$

For the capacitor it was easier to visualise energy storage in terms of charges stored on plates. In the case of an inductor it is a bit difficult to visualise energy storage. Here it can be imagined that the energy is stored in the magnetic field created around the inductor, similar to the electric field created between the plates of a capacitor. Just as a dielectric increased the capacitance in the previous case, similarly the introduction of a magnetic material within the inductor coil increases the inductance manifold. For a capacitor we had polarisation of atoms and molecules in the dielectric, here for the inductor we have magnetisation of atoms and molecules in the magnetic material, and orientation of magnetic domains if the material is ferromagnetic.

3.10.4 Application: Voltage and Current smoothing in Power Supply units

It can be seen that the transient *current* and *voltage* in an *LR* circuit have the same behaviour as the transient *voltage* and *current* respectively in an *RC* circuit (note the sequencing of the terms *voltage* and *current*). In the *RC* circuit *the capacitor voltage did not want to change,* while in the *LR* circuit, *the inductor current does not want to change.* Therefore ***capacitors can be used to smooth out voltages in a circuit while inductors can be used to smooth out currents***.

Laboratory dc power supply units usually obtain their power from ac mains through stepping down the voltage using a transformer first, and then, rectification. This gives a varying dc voltage which needs smoothing circuitry. A combination of both inductor and capacitor gives the best of both worlds and a typical power supply unit with such a smoothing circuit is shown in Fig.3.17. Note that the two capacitors with the inductor in the middle make a graphical form which looks like the Greek letter ‘π’. Therefore such a smoothing circuit is called a ‘π-filter’. Such power supplies were used with older vacuum diode operated power supplies extensively. Inductors operating at the mains line frequency of 50Hz tend to be bulky and expensive, therefore, with the

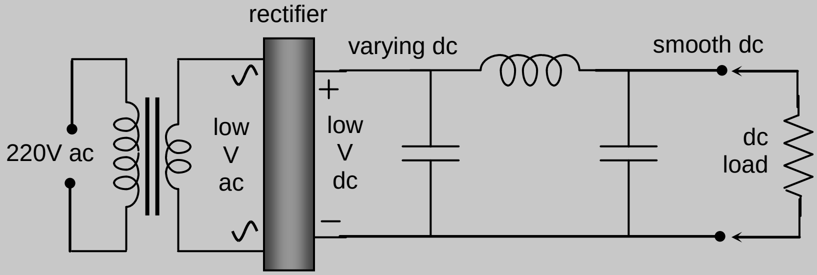


Fig.3-17: A low voltage dc power supply employing LC smoothing circuit

advent of small and cheap semiconductor devices, an alternative method of electronic voltage stabilisation became popular which gives a very smooth dc without inductors. However, inductors have recently made a come back again because of the popularity of highly efficient *switch mode power supplies* which operate at tens or hundreds of kilohertz by generating square waves of such frequencies within the power supply unit. Some of these supplies will be discussed in later volumes of this book.

3.11 Series LCR circuit, Switching to a dc supply

An LCR circuit connected to a dc voltage source is shown in Fig.3-18. When the switch is flipped from point 2 to point 1 a step voltage v_{IN} is applied to the LCR circuit as shown in Fig.3-19a and consequently a transient current i is initiated. We may remember that on the application of a step voltage a capacitor allows a sudden high current which then decreases exponentially with time, but an inductor has an opposite behaviour. It does not allow the current to change sharply. So depending on the relative values of L and C, we expect to get a combination of the two effects which are shown in the lower curves of Fig.3-19 and are discussed below.

3.11.1 Physical visualisation

To visualise the effects let us first imagine the inductor to have zero inductance. So the circuit is essentially a *CR* circuit and the current would be as shown in Fig.3-19b, rising suddenly at first, and then decreasing exponentially.

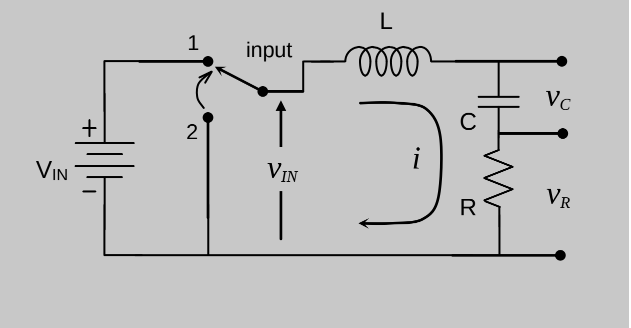


Fig.3-18: dc transients through a series LCR circuit

Now as we increase the value of L it opposes the sudden rise in capacitor current due to which we will see a gradual rise in current initially as shown in Fig.3-19c. After a while, the inductor's effect will become negligible and the current will be dominated by the capacitor charging current which is decreasing exponentially in this phase as was indicated by Fig.3-19b. Here, the inductor would again oppose this fall in current modifying the exponentially falling pattern. Since the rate of change of current is low in this phase, which is again falling with time, the opposition of the induced *emf* will be low. Therefore, the capacitor behaviour will ultimately dominate the decreasing current pattern and it will eventually reduce to zero when the capacitor will be fully charged. The combined effect would result in an overall current pattern similar to that shown in Fig.3-19c, a rounded pulse with an almost exponentially trailing end.

If we increase the value of L further beyond a threshold, the induced *emf* during the fall of the current will no more be insignificant (which works in the same direction as the driving current now, trying not to let the current drop). This will contribute significantly to the charging of the capacitor, hastening the process, and the resulting current pattern will be almost that of a half sinusoid as shown in Fig.3-19d between points p and q. When the current reaches zero at point q, the capacitor would be charged to a maximum (top plate +ve), to a voltage V_{C1} (Fig.3-19e) which is higher than V_{IN}. Where does the extra voltage come from? In the falling phase of the current the induced *emf* and V_{IN} are both in the same direction, and therefore, they will add up to make the total voltage higher than V_{IN}.

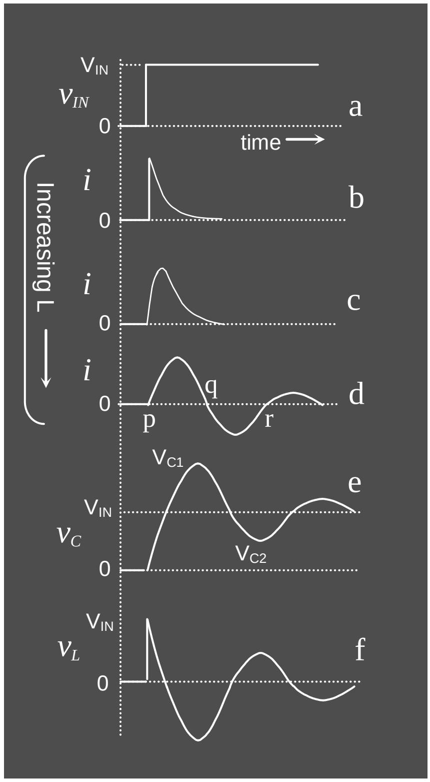


Fig.3-19: dc transients in an LCR circuit

After the current becomes zero at q, the capacitor will start discharging in the opposite direction because of the positive voltage difference $V_{C1} - V_{IN}$ and a similar situation as above will be created except that the current now would be in the reverse direction (anti-clockwise in Fig.3-18). This will again result in a rounded half sinusoidal current pattern in the reverse direction as shown between points q and r in Fig.3-19d. During this process when $v_C - V_{IN}$ becomes zero the current should have stopped, but the stored energy from the inductor carries the current further discharging the capacitor to a voltage value lower than V_{IN}, being the lowest, V_{C2} at r (Fig.3-19d,e). The positive voltage difference $V_{IN} - V_{C2}$ will start to charge the capacitor, creating a current in the clockwise direction again, and this will carry on repeating. We will get a damped sinusoidal alternating current (amplitude decreasing gradually, exponentially in this case) resulting from the sequential reversible energy storage and release by the inductor and the capacitor (also see Physics, Vol-II, by Halliday & Resnick). Note the relative phases of the current and the capacitor voltage waveforms, the latter being delayed by 90^o.

The damping occurs because of irreversible energy dissipated by the series resistor whenever a current is flowing (= i^2R, always positive irrespective of the direction of the current) which progressively reduces the energy stored by the inductor and the capacitor. Had there been no resistance in the circuit (ideal case), there would not have been any damping and the current would be purely sinusoidal in nature with a constant amplitude up to infinite time.

It would be interesting to know what the voltage patterns are across each of the *LCR* components in the circuit. Here we only discuss them for the case corresponding to that for Fig.3-19d where we get damped sinusoidal oscillations (try to find out those corresponding to the other situations yourself). The voltage v_R (= iR) across the resistor will have exactly the same pattern as for the current. The behaviour of the voltage v_C across the capacitor is shown in Fig.3-19e. Rising from zero this will eventually have a damped sinusoidal pattern but it will have a bias of V_{IN} and it will be delayed with respect to the current by a phase angle of 90^0. The maximum voltage on the capacitor will occur when the current becomes zero and vice versa. Eventually at infinite time the capacitor will be charged to its stable value of V_{IN}, its final destiny. The voltage v_L across the inductor is shown in Fig.3-19f. Initially this poses an infinite obstacle and drops all of the input voltage V_{IN} at $t = 0$. Therefore it follows the step input at this point. Then it follows a pattern completely 180^0 opposite in phase to that of the capacitor voltage except for the dc bias, which is zero in this case. After an infinite time no voltage is dropped across the inductor.

The above two oscillatory patterns are shown again with the scales suitably changed in Fig.3-20a and Fig.3-20b to visualise the patterns over a longer period. This behaviour is called *ringing* because it is similar to hitting a bell, where the bell produces an exponentially decaying ringing sound at its natural frequency of vibration. Therefore we can say that an *LCR* circuit has a *natural frequency of oscillation,* and it can be set into *ringing* by driving with a step voltage. Note that if we had an ideal situation with $R = 0$ in the circuit, there would be no power dissipation and the oscillation would go on indefinitely without any damping. However, this is not possible in practice since there will be some resistance in the inductor and in the wiring of the circuit.

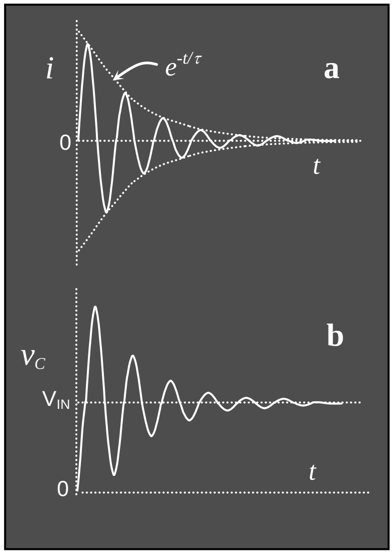


Fig.3-20: Damped oscillatory behaviour of an LCR circuit: a) current and b) voltage across capacitor

3.11.2 Analysis

To analyse we have to apply Kirchoff's law as before around the loop to get,

$$V_{IN} = L\frac{di}{dt} + \frac{q}{C} + iR \qquad \text{... 3.23}$$

Differentiating, and taking current $i = dq/dt$, we get a 2nd order differential equation,

$$L\frac{d^2i}{dt^2}+R\frac{di}{dt}+\frac{1}{C}i=0 \qquad \text{... 3.24}$$

The solution to this equation is a bit complex since it depends on the relative values of *L, C* and *R*. Therefore we follow an alternative method of trial solution based on some mathematical pre-judgment and solve it only under certain conditions. Let a trial solution be,

$$i = D\,e^{-bt} \qquad \text{... 3.25}$$

where b is a constant. *D* may be a constant or a function of *t*.

From Eq.3.25, $\frac{di}{dt}=-D\,b\,e^{-bt}$ and $\frac{d^2i}{dt^2}=D\,b^2\,e^{-bt}$

Replacing these values in Eq.3.24 we get, $b^2-\frac{R}{L}b+\frac{1}{LC}=0$

(since, *D* and e^{-bt} cannot be zero at all values of *t*).

The above is a quadratic equation in *b*, and we have as its solution,

$$b = \frac{R}{2L} \pm \sqrt{\left(\frac{R}{2L}\right)^2-\frac{1}{LC}} \qquad \text{... 3.26}$$

Let the above equation be represented as, $b = p \pm q$ where

$$p = \frac{R}{2L} \quad \text{and} \quad q = \sqrt{\left(\frac{R}{2L}\right)^2-\frac{1}{LC}} \qquad \text{... 3.27}$$

The general solution for the current is then given by,

$$i = A\,e^{-(p-q)t} + B\,e^{-(p+q)t},$$

where *A* and *B* may be constants or functions of *t.* The solution can be rewritten as,

$$\boxed{i=e^{-pt}\left(A\,e^{\,qt}+B\,e^{-\,qt}\right)} \qquad \text{... 3.28}$$

We will now try to solve it under different conditions.

Case I: Overdamped condition,

$$\text{for} \quad \frac{R}{2L}>\frac{1}{\sqrt{LC}} \qquad \text{... 3.29}$$

both p and q are positive in Eq.3.27 and the above solution (Eq.3-28) applies which has both rising and falling functions of *t* giving rise to a rounded pulse with a long trailing end similar to that shown in Fig.3.19c. The time taken for the rise and fall of the current, both are very large in this case. This situation is called *overdamped* , meaning the damping to the transients due to resistive losses (i^2R) are very high because of which the above timing behaviours are observed. The transient takes a long time to stabilise in this overdamped condition.

Case II: Critically damped condition,

$$\text{for} \quad \frac{R}{2L} = \frac{1}{\sqrt{LC}} \qquad \text{... 3.30}$$

q in Eq.3.27 is zero and we are left with only one value of b in Eq.3.26 and only one time dependent term in the solution in Eq.3.28 [i.e., $i = (A + B)\, e^{-pt}$] which is not mathematically acceptable for such a 2nd order differential equation. Therefore in such cases one of the parameters is considered to be a linear function of t (see any advanced mathematics book). The solution is then given by

$$\boxed{i = \left(A' + B't\right)e^{-pt}} \qquad \text{... 3.31}$$

We can see that this has a linearly rising component ($A' + B't$) and an exponentially decaying component (e^{-pt}) which compete with each other to give a rounded and trailing pulse, somewhat similar to the one shown in Fig.3.19c. However, the important distinction between this and the previous overdamped case is that, in the critically damped case the current rises and then decays back to zero in the shortest possible time, while for the overdamped case, the time needed is more.

Case III: Underdamped condition,

$$\text{for} \quad \frac{R}{2L} < \frac{1}{\sqrt{LC}} \qquad \text{... 3.32}$$

we have, $p \neq 0$ and q is imaginary. The solution of Eq.3.28 becomes (with $j = \sqrt{-1}$),

$$i = A'\, e^{-t/\tau}\, e^{j\omega' t} + B'\, e^{-t/\tau}\, e^{-j\omega' t} \qquad \text{... 3.33}$$

$$\text{where,} \quad \tau = \frac{2L}{R} \qquad \text{... 3.34}$$

$$\text{and,} \quad \omega' = \sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2} \qquad \text{... 3.35}$$

If $A' = B'$ in Eq.3.33, the solution becomes,

$$\boxed{i = (2A'\, e^{-t/\tau})\ Cos\,\omega' t} \qquad \text{... 3.36}$$

This solution (Eq.3.36) can be divided into two parts, the one within the brackets, and the $Cos\,\omega' t$ term. The latter term shows that there will be a sinusoidal oscillation with angular frequency ω', but the amplitude, given by the term within the bracket, is not constant, rather it decays exponentially with a time constant given by Eq.3.34. This decaying pattern is shown in dotted lines in Fig.3-20, which forms the envelope of the amplitude of oscillatory waveform given by the *Cos*$\omega' t$ term in Eq.3.36. Eventually the oscillation dies away and we get zero current at infinite time (for practical purposes we can take the current to be essentially zero after 5 or 6 time constants). Thus Eq.3.36 clearly describes the *ringing* pattern of Fig.3-19d and Fig.3-20a that we inferred earlier using physical arguments. This situation where we get an oscillation is called an *underdamped case*. The damping depends on the time constant, *2L/R*, with larger *R* making a shorter time constant, i.e., a quicker decay.

The frequency of oscillation is given by, using Eq.3.35,

$$f' = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \left(\frac{R}{2L}\right)^2} \qquad \text{... 3.37}$$

which is dependent on all the three parameters.

Case IV: Undamped condition,

$$\frac{R}{2L} << \frac{1}{\sqrt{LC}} \qquad \text{... 3.38}$$

(*R = 0* in ideal case)

We have, $p = 0$ so that $e^{-pt} = 1$, and q is imaginary. The solution becomes,

$$i = \left(A\,e^{\,j\omega t} + B\,e^{-\,j\omega t}\right) \qquad \text{... 3.39}$$

or, $$i = (A + B)\,Cos\,\omega t + j\,(A - B)\,Sin\,\omega t \qquad \text{... 3.40}$$

where $$\omega = \frac{1}{\sqrt{LC}} \qquad \text{... 3.41}$$

Taking the real part in Eq.3.40 and replacing $(A + B)$ by E, we get,

$$i = E\,Cos\,\omega t \qquad \text{... 3.42}$$

which is a continuous sinusoidal function with constant amplitude and angular frequence ω. This is a totally *undamped case* (ideal) and agrees with what we had described earlier from conceptual arguments. The frequency of oscillation, from Eq.3.41, will be,

$$\boxed{f = \frac{1}{2\pi}\sqrt{\frac{1}{LC}}} \qquad \text{... 3.43}$$

which is independent of *R*, and decreases with increasing value of the product *LC*.

We can see that the frequency in the underdamped case (Eq.3.37) reduces to that of the undamped case given in Eq.3.43 for the extreme condition $\frac{R}{2L} << \frac{1}{\sqrt{LC}}$. In practice, small values of R may be considered to contribute to a deviation from the ideal frequency given by Eq.3.37.

3.11.3 Conditions in terms of time constants

We can square and rearrange the condition for critical damping, Eq.3.30, as,

$$RC = 4\frac{L}{R} \qquad \text{... 3.44}$$

We can see that the left hand side is the time constant of the circuit without the inductor (*RC* circuit) while the right hand side is four times the time constant of the circuit without the capacitor (*LR* circuit). If we denote these time constants as τ_c and τ_L respectively, we can see that the above four cases of damping occur for:

$$\left.\begin{array}{ll} \text{Case I: Overdamped} & \tau_c > 4\tau_L \\ \text{Case II: Critically damped} & \tau_c = 4\tau_L \\ \text{Case III: Underdamped} & \tau_c < 4\tau_L \\ \text{Case IV: Undamped} & \tau_c << \tau_L \end{array}\right\} \qquad \text{... 3.45}$$

These can help visualise the relative values of L, C and R in terms of the above time constants.

3.11.3 Switching back to ground

We have to appreciate that after switching from 2 to 1 in Fig.3-18 when the current becomes zero after a long time, the capacitor remains charged to a voltage V_{IN} with the top plate positive. This is also shown in the first half of Fig.3-21c. Now if the switch is flipped from position 1 to 2 in Fig.3-18 (i.e., switched to ground), only the V_{IN} term in Eq.3.23 would be zero, and the differentiated form would be the same as Eq.3.24. Therefore the behaviour of the current as shown in Fig.3-19 b-d and Fig.3-20a would remain the same for appropriate values of the circuit components, and all of the above considerations would apply, except for the direction of the initial current, which would be reversed (negative first). This is shown by the second pattern in Fig.3-21b on downward excursion of the input voltage. The capacitor voltage will go through a baseline shift as well, as shown in Fig.3.21c, from V_{IN} to zero. After a long interval the current becomes zero again, and the capacitor now is fully discharged ($v_c = 0$). Therefore, after a long interval,in either of the switching positions, the current is zero and the capacitor is either fully charged or fully discharged depending on the switching position, and the inductor has no stored energy. For one switch position (2 to 1), energy initially comes from the battery, while for the other (1 to 2), the charged capacitor supplies the initial energy.

3.11.4 Repetitive switching & Resonance

If the switching between point 1 and 2 is repeated periodically at large enough intervals, we will see the ringing patterns shown in Fig.3-21 repeating. If the switching period is reduced gradually (you have to use an electronic switch based on transistors, manual switching will not do), we will see changing complex patterns which will depend on the point in time with respect to the damped oscillation when the *'switch'* is made. A particular sample is shown in Fig.3-22 a&b. Eventually when the switching period is made exactly equal to the period of oscillation, we will see a continuous sinusoidal oscillation having a maximum amplitude as shown in Fig.3-22 c&d. This is known as ***Resonance***, a very important and interesting natural phenomenon about which you must have studied before in mechanics. Any object with a natural frequency of oscillation will demonstrate the phenomenon of resonance when a periodic external force is applied having the same frequency. If the switching period is reduced further, there will be incomplete storing of charge in either the inductor or the capacitor. The waveform will be still be sinusoidal but will follow the period of the switching, and the amplitude will gradually decrease.

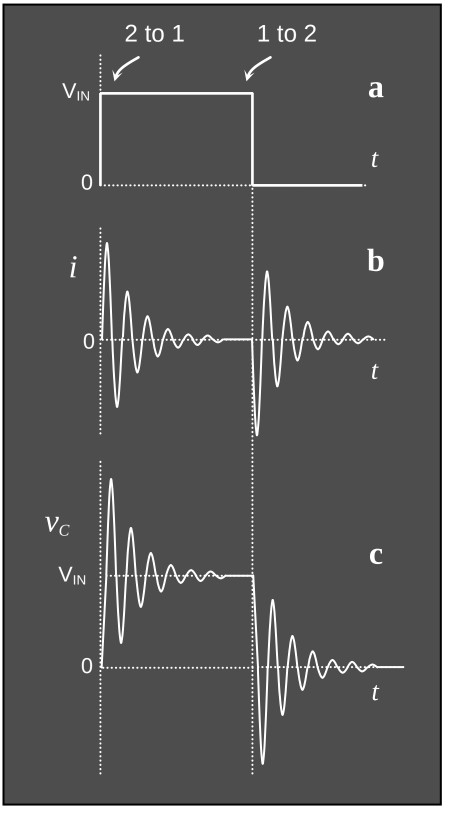


Fig.3-21: Damped oscillatory behaviour of an LCR circuit on switching on and off;
a) input, b) current and c) voltage across capacitor

We can see that even though we have square waveforms at the input, the current, and hence the output voltage taken across the resistor (= iR) is always sinusoidal at resonance. In this way we can convert the energy in a square waveform into a sinusoidal one and this has important applications in high frequency circuits, particularly in Radio transmitters. The resonance is also observed if the impressed (driving) waveform is a sinusoidal ac, and this we will study in the next chapter.

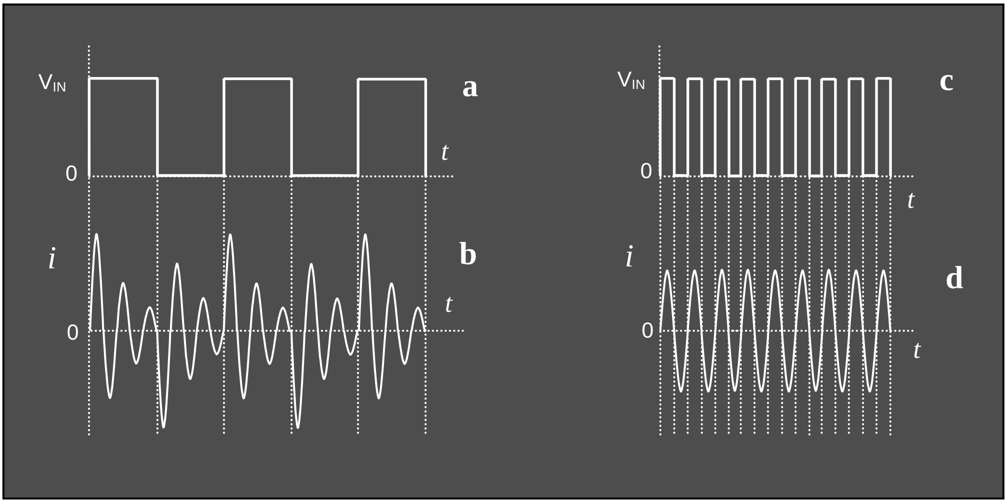


Fig.3-22: Variation of current waveform of an LCR circuit on repetitive switching frequency. Resonance occurs when switching period equals natural oscillation period (c&d).

Chapter 4: AC circuits

In this chapter we shall discuss some important ac circuits that are frequently needed in the analysis of electronic circuits. First let us recapitulate some ac fundamentals.

AC FUNDAMENTALS

4.1 Sine waveforms and phase angles

Mathematically, a sinusoidal alternating voltage (which we will call *ac voltage* from now on) with time period T and frequency f (= no. of full cycles per unit time = $1/T$) can be expressed as

$$v = V\, Sin\ \omega t \qquad \text{... 4.1}$$

where V is the amplitude, and ω is the angular frequency (= angle covered per unit time = $2\pi/T = 2\pi f$). A sinusoidal function can be generated from the linear projection of a radius vector rotating in a circle at a constant angular velocity as shown in Fig.4-1. A radius vector covers a full cycle covering an angle 2π in a time period T. This gives the above relationship between ω and T. In general, if the radius vector covers angle θ in time t, then $\omega = \theta / t$. Note that ω can also be termed as the *angular velocity* having the same unit. The unit is given in radians/sec and has the mathematical dimension of sec^{-1} since the angle is a dimensionless number. Note that frequency f has the same mathematical dimension of sec^{-1} and is measured in Hz (previously in cycles/sec).

Eq.4.1 is plotted in Fig.4-1a as a function of angle $\theta\ (= \omega t)$ where a full cycle repeats at an angular interval of 2π. For this waveform, $v = 0$ at $\theta = 0$, which corresponds to time $t = 0$. The corresponding generating circle for this sinusoidal waveform is shown on the right hand side of the figure where the projection of the radius vector A_0 on the vertical axis gives the value of v (here, vertical projection = 0 at $\theta = 0$). The radius vector can also be called a phasor and is taken to rotate counterclockwise for a positive angular displacement. At any other point in time, the phasor may have a non-zero projection on the vertical axis giving an instantaneous voltage value v; the angle subtended with the original starting position $(\theta = 0)$ is called the *phase angle* of the projected voltage value.

Since ω is constant for a sinusoidal waveform, the nature of the waveform would remain the same if we plot it as a function of time $t\ (= \theta/\omega)$ as shown in Fig.4-1b. This is more advantageous as it allows us to visualise the progression of the waveform in time. Note that the angle of full cycle 2π in Fig.4-1a corresponds to the time period T in Fig.4-1b.

Now let us consider the waveform given by

$$v = V\, Cos\ \omega t \qquad \text{... 4.2}$$

which we can also write as

$$v = V\, Sin\ (\omega t + \pi/2). \qquad \text{... 4.3}$$

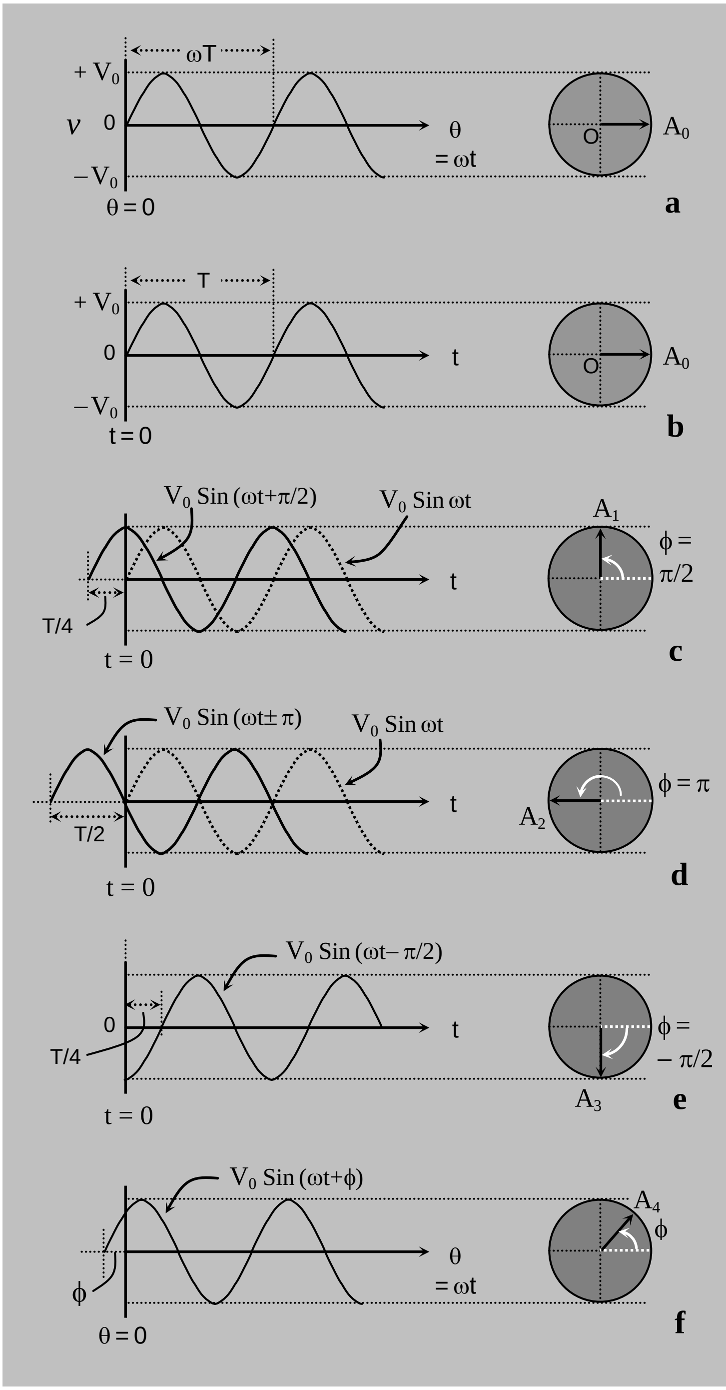


Fig.4-1: Sine waves as a function of angle and time, relationship to generating circles, and development of the concept of phasors

The wave represented by Eq.4.3 is the same as that given by Eq.4.1 except that its phase is leading (i.e., it started before) by an angle of $\pi/2$ which is equivalent to a time of

T/4. The graphical form and the corresponding generating circle is shown in Fig.4-1c where $v = V$ (maximum +ve value) at $t = 0$ and the phasor A_1 is at a phase angle of $\pi/2$ with respect to $\theta = 0$.

It easy to conceive that phasor A_1 is at an advanced position, i.e., it is leading A_0 at time $t = 0$. However, looking at the time graph how do we know that it is leading? Fig.4-1b started with $v = 0$ at $t = 0$. Therefore we have to look at the timing of the new waveform to find when $v = 0$ occurs, nearest to $t = 0$. In Fig.4-1c, by extending the graph to the left we can see that $v = 0$ occurs at time $t = -T/4$, i.e., before that in Fig.4-1b, which answers our question. The waveform given by Eq.4.1 is also plotted in Fig.4-1c in dotted lines to show the comparison. Try to appreciate that time increases towards the right in this diagram and the waveform that leads is positioned on the left side in this time graph.

Now let us consider the expression

$$v = -V\,Sin\;\omega t \qquad \text{... 4.4}$$

which can be rewritten as $v = V\,Sin\;(\omega t + \pi)$

representing an waveform which is leading that of Eq.4-1 by a phase angle of π and is shown in Fig.4-1d together with the phasor A_2 at $t = 0$. However, this can also be rewritten as

$$v = V\,Sin(\omega t - \pi)$$

representing an waveform which is lagging behind (i.e., starts later) that of Eq.4-1 by a phase angle of π. This is special since the phasor is at equal angles from both directions with respect to the starting point.

Note that $v = 0$ at $t = 0$ for this function also, but it does not have the same phase as that of Fig.4-1a which is clear from the position of the phasor A_2. In the time graph we can see that although v is the same as that for Fig.4-1b, the incremental behaviour is not the same. At time $t = 0$, v is increasing in Fig.4-1b while v is decreasing in Fig.4-1d.

Next let us consider the expression

$$v = V\,Sin\;(\omega t - \pi/2) \qquad \text{... 4.5}$$

which represents an waveform that is lagging behind that of Eq.4-1 by a phase angle of $\pi/2$ (given by the – ve sign) and is shown in Fig.4-1e together with the phasor A_3 at $t=0$. In the time graph we can see that this waveform assumes the value of zero at time $T/4$ later than that in Fig.4-1b.

When we refer to a general sinusoidal waveform, it is usual to include an arbitrary phase angle ϕ as represented graphically in Fig.4-1f and mathematically as

$$v = V\,Sin\;(\omega t + \phi). \qquad \text{... 4.6}$$

We usually refer to the phase angle as positive meaning that the waveform is leading the reference waveform of Fig.4-1b. The advancement through an angle ϕ is shown in the corresponding generating circle and in a graph against angle $\theta\,(= \omega t)$ in Fig.4-1f. If the waveform lagged behind instead of leading, then the phase angle ϕ would be negative.

Note that if we are free to choose the starting time of a sinusoidal function, we like to keep it simple and choose the phase angle as *zero.* However, if we have a prior decision about the time reference then we have to use the general expression given in Eq.4.6. Again, if we are dealing with a number of waveforms of the same frequency but having constant phase differences between them, then we can choose zero phase for only one of these. Once we have done it, the phase angles for the rest are automatically determined.

4.2 Combining ac voltages, Phasor representation

We have seen above that a radius vector, or, a phasor in the generating circle show the phase angle of a sinusoidal waveform clearly, while the length of the phasor represent the amplitude (maximum vertical projection). If we want to compare and analyse two or more sinusoidal waveforms of *the same frequency* differing in amplitude and phase only, but which maintain a *constant phase relationship with each other*, the phasor diagram makes a very useful analytical tool. Note the two conditions mentioned above in italics, you cannot compare and analyse phasors representing waveforms having different frequencies easily (as they will vary with time differently), nor those whose phase differences change with time.

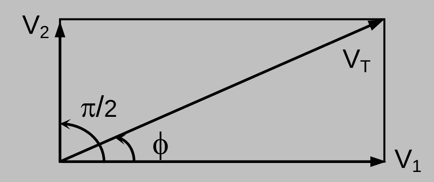


Fig.4-2: Phasor method of combining ac voltages.

Using phasor diagrams it is possible to add two waveforms to get the resultant phasor. However, remember that you can add or subtract the same physical quantities only – *you cannot add voltages with currents or with other parameters*. Furthermore, you can *add phasors representing different waveforms at a particular point in time only*. Usually we do it at $t = 0$.

So for simplified analysis of multiple waveforms using phasor representation, the waveforms should

i) have the same frequency,
ii) maintain a constant phase relationship with each other,
iii) have the same parameters for addition,
iv) this addition is to be done at a particular time, say, at $t = 0$.

The following two waveforms,

$$v_1 = V_1 \, Sin \, \omega t \qquad ... 4.6$$

and

$$v_2 = V_2 \, Sin \, (\omega t + \pi/2) \qquad ... 4.7$$

can be represented in a phasor diagram as in Fig.4-2 where the individual phasors represent the respective amplitudes V_1 and V_2 in both magnitude and phase at $t = 0$. Now the voltage sum or the total,

$$(v_1 + v_2) = V_T \, Sin \, (\omega t + \phi) \qquad ... 4.8$$

is represented simply by the diagonal of the rectangle formed by V_1 and V_2, both in amplitude (V_T) and phase (ϕ). Thus we can easily determine these resulting parameters,

$$V_T = \sqrt{V_1^{\,2} + V_2^{\,2}} \qquad ... 4.9$$

and

$$\phi = \tan^{-1} \frac{V_2}{V_1} \qquad ... 4.10$$

as the complete solution. Compare the above technique with the other possibilities – a graphical one where you have to add the two waveforms point to point in time and plot the resultant waveform – or a mathematical solution which, you can imagine, would not be simple. Therefore the above technique of adding phasors provides a simple method of adding sinusoidal waveforms.

If the phase angle between the two voltage waveforms is different from $\pi/2$ we would get a parallelogram instead of a rectangle, but still can use the same technique, i.e., the diagonal will represent the resultant. The above procedures are exactly the same as that for *vectors* that you have done in Mechanics.

Remember that there is no absolute value for phase, it all depends on the choice of the starting time. When we compare two or more waveforms, we usually choose any suitable one as the reference and refer the phase difference of the other waveforms with respect to that single reference.

4.3 Capacitors and Inductors

4.3.1 Capacitor on ac, Capacitive reactance

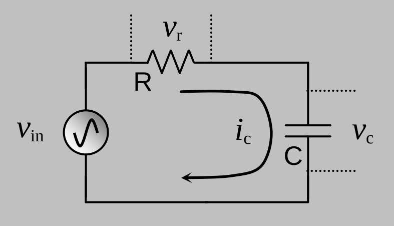


Fig.4-3: Capacitor in an ac circuit

Fig.4-3 shows a circuit comprising of an ac source v_{in}, a resistor R and a capacitor C in series. The loop current i_c, and voltages v_r and v_c across the resistor and the capacitor respectively are also shown. The current i_c is related to voltage v_c as,

$$i_c = \frac{dq}{dt} = C\frac{dv_c}{dt} \qquad \text{... 4.11}$$

where q is the instantaneous charge on the capacitor. Integrating the above, ignoring constants of integration, we get,

$$v_c = \frac{1}{C}\int i_c \; dt \qquad \text{... 4.11a}$$

Now if we choose $i_c = I_C \, Sin\, \omega t$ ... 4.12a

(This is our choice, for convenience. The reason will be apparent later.)

then $$v_c = -\frac{1}{\omega C} \; I_C \, Cos \omega t = -V_C \, Cos \omega t \qquad \text{... 4.12b}$$

where V_C is the peak capacitor voltage and related to the peak current I_C as,

$$I_C = \frac{V_C}{1/\omega C} \qquad \text{... 4.13}$$

Eq.4.13 has the form of Ohm's law if the denominator $1/\omega C$ on the right hand side is considered as a quantity *equivalent to a resistance*. This is the opposition to a current posed by the capacitor. We can also see that due to the presence of ω, this opposition is *frequency dependent*. An ideal resistor offers a ***Resistance*** to a current that is *independent of frequency*. To distinguish between these two, the frequency dependent opposition is called ***Reactance*** and is usually denoted by the symbol X. The reactance offered by a capacitor, or the ***capacitive reactance*** may be denoted by the symbol X_C. Thus

$$X_C = \frac{1}{\omega C} \qquad \text{...4.14a}$$

and $$I_C = \frac{V_C}{X_C} \qquad \text{... 4.14b}$$

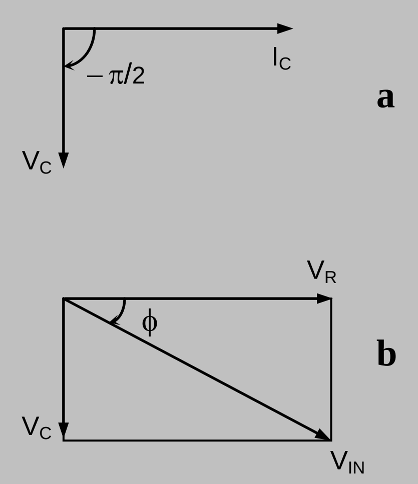


Fig.4-4: Phasor representation of currents and voltages in a capacitor-resistor circuit

Clearly, X_C decreases with increasing frequency. This also says that an ac current can pass through a capacitor and that the current increases with increasing frequency (remember dc current cannot pass through a capacitor in the stable state).

We can rewrite Eq.4.12b as,

$$v_c = -V_C\, Cos\omega t = V_C\, [-Sin(\pi/2-\omega t)] = V_C\, Sin(\omega t - \pi/2) \qquad \text{... 4.15a}$$

This indicates that the capacitor voltage lags behind the capacitor current by a phase angle of $\pi/2$ which is shown in the phasor diagram Fig.4-4a by drawing the current amplitude I_C along the + ve X-axis and the voltage amplitude V_C along the – ve Y-axis. This is a very significant result. It effectively says that the voltage across a capacitor goes slow compared to the current through it. Remember, for a dc transient we found that the voltage across the capacitor changes slowly while the current at the instant of switching changes sharply (see Chapter 3). There is similarity between these two behaviours.

> With dc (Chapter 3, Fig.3-3), at switching on, there is no charge on the capacitor, so there is no opposition to incoming electrons. As it gets charged, the electrons already accumulated on the negative plate oppose new incoming charges. The more the capacitor is charged, the more opposition it offers to current, and an infinite opposition occurs when the capacitor is fully charged. We can extend this concept to ac. At high frequency, the capacitor gets little time to get charged, therefore, it offers little opposition. At lower frequency the capacitor gets more time to get more charged, therefore it offers more opposition.

In Fig.4-4a we have shown two different quantities (current and voltage) in the same diagram to indicate their phase differences only, but they cannot be added. However, if we consider the voltage across the resistor v_r $(= i_c R)$ and draw the phasor diagrams of the two voltage-amplitudes as in Fig.4-4b, then we can add the two to get the sum-total amplitude V_{IN}, which should represent the input voltage v_{in} here. Note that i_c and v_r have the same phase since R is a constant in the product $i_c R$. Since we have considered v_r to have a phase angle of zero, we plot it along the X-axis. The voltage across the capacitor v_c is plotted along the –ve Y axis and the input voltage v_{in} is the sum of these two and is given by the sum-total or the resultant phasor. The phase angle ϕ of the input voltage has been shown with respect to the reference v_r.

> Here lies the reason for choosing the phase of i_c as zero in Eq.4.12a above. This choice has allowed us to take the voltage across the resistance v_r to have a zero phase also, and to consider it as a reference. This is a choice we usually go for. Remember, resistance is a frequency independent parameter, therefore, there is some advantage in choosing the phase of the voltage across this as the reference. Voltage across a series capacitor would be delayed by $\pi/2$ while that across a series inductor will be advanced by $\pi/2$, as we will find out soon.

From Fig.4-4b we can see that v_{in} is delayed by a phase angle ϕ with respect to v_r. So the expression for the input voltage should include this phase angle as,

$$v_{in} = V_{IN}\, Sin\, (\omega t + \phi) \qquad ...\ 4.15b$$

where the phase angle ϕ is *inherently negative* in this case, as is evident from the figure. From the phasor diagram of Fig.4-4b we can see that the peak amplitude of the input voltage is given, in terms of the respective peak amplitudes of the other voltages, by,

$$V_{IN} = \sqrt{V_R^2 + V_C^2} \qquad ...\ 4.16$$

and the phase angle by,

$$\phi = \tan^{-1} \frac{-V_C}{V_R} = -\tan^{-1} \frac{V_C}{V_R} \qquad ...\ 4.17$$

The negative sign appears since V_C is along the – ve Y- axis (phase = – $\pi/2$). This essentially says that v_{in} is lagging behind v_r in phase.

We can focus on some important results from the above discussion. If a capacitor and a resistor are in a single ac current loop (no other currents are involved) then the voltage across the capacitor will be lagging behind the voltage across the resistor by $\pi/2$. This is the same as saying that the voltage across the resistor leads the voltage across the capacitor by $\pi/2$. The sum of these two voltages will have a magnitude given by Eq.4.16 and will have a phase lying in between those of the two voltages, depending on their magnitudes. Eq.4.17 gives the phase difference between the input voltage and the voltage across the resistor taking the latter as the reference.

4.3.2 Inductor on ac and Inductive reactance

Fig.4-5 shows a circuit comprising of an ac source, v_{in}, a resistor and an inductor in series. The loop current i_l, and voltages v_r and v_l across the resistor and the inductor respectively are also shown. The voltage v_l is related to current i_l as,

$$v_l = L\frac{di_l}{dt} \qquad ...\ 4.18$$

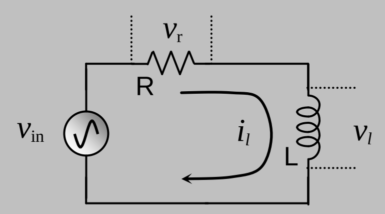


Fig.4-5: Inductor in an ac circuit

Now if we choose

$$i_l = I_L\ Sin\omega t \qquad ...\ 4.19a$$

(this is again our choice, for convenience, to have zero phase for v_r)

then $$v_l = \omega L\, I_L\, Cos\omega t = V_L\, Sin(\omega t + \pi/2) \qquad ...\ 4.19b$$

where the peak voltage V_L can be related to peak current I_L as,

$$I_L = \frac{V_L}{\omega L} \qquad \text{... 4.20}$$

Similar to the treatment with the capacitor above, Eq.4.20 has the form of Ohm's law if the denominator ωL on the right hand side is considered as a quantity equivalent to a resistance. This is the opposition to a current posed by the inductor, and is called the ***Inductive Reactance***, denoted by the symbol X_L. Thus

$$X_L = \omega L \qquad \text{... 4.21a}$$

and

$$I_L = \frac{V_L}{X_L} \qquad \text{... 4.21b}$$

Clearly, X_L increases with frequency and current decreases with frequency. Eq.4.19 above also indicates that the inductor voltage leads the inductor current by a phase angle of $\pi/2$ which is shown in Fig.4-6a. As before, here we have chosen the phase of i_l to be zero and drawn the phasor along the X-axis. Considering the voltage across the resistor v_r $(= i_l R)$ and drawing the phasors of the two voltages as in Fig.4-6b, we can add them up which should equal the input voltage, v_{in}. Note that i_l and v_r have the same phase since R is a constant in the product $i_l R$.

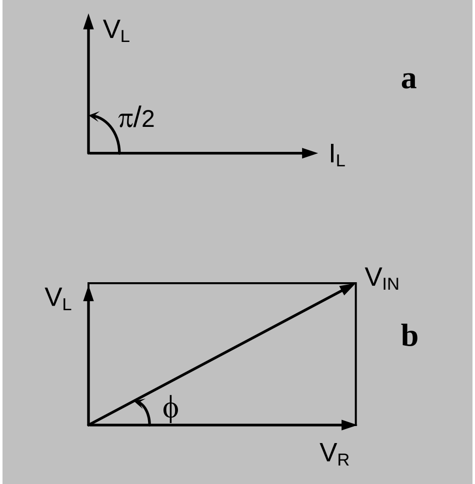


Fig.4-6: Phasor representation of currents and voltages in an inductor -resistor circuit

From the phasor diagram we can see that the peak amplitude of the combined voltage, which is the input voltage v_{in} here, is given, in terms of the respective peak amplitudes of the other voltages, by,

$$V_{IN} = \sqrt{V_R^2 + V_L^2} \qquad \text{... 4.22}$$

and the phase angle by,

$$\phi = \tan^{-1}\frac{V_L}{V_R} \qquad \text{... 4.23}$$

which is inherently positive. Here the phase angle ϕ is basically the phase difference between v_{in} and v_r with the latter as the reference. Note the essential differences with the capacitor circuit above. ***For the capacitor, the current leads the voltage, while for the inductor, the current lags behind the voltage***. Here the current grows slowly which also supports the dc transient behaviour discussed in the previous chapter.

Similar to the capacitor circuit before, we can focus on some important results from the above discussion. If an inductor and a resistor are in a single ac current loop (no other currents are involved) then the voltage across the inductor will be leading the voltage across the resistor by $\pi/2$. This is the same as saying that the voltage across the resistor lags behind the voltage across the inductor by $\pi/2$. The sum of these two voltages will have a magnitude given by Eq.4.22 and will have a phase lying in between those of the two voltages, depending on their magnitudes. Eq.4.23 gives the phase difference between the voltages across the inductor and the resistor respectively taking the latter as the reference.

4.3.3 Resistor, Inductor and Capacitor behaviour on ac

The frequency responses of resistance R, capacitive and inductive reactances X_C and X_L are represented graphically in Fig. 4-7. Note the shapes, which directly follow the definitions of the respective parameters. R is a straight line parallel to the frequency axis, meaning that its value is the same at all frequencies. Inductive reactance ($X_L = \omega L$) is a straight line going through the origin whose value increases with frequency ($\omega L = 2\pi f L \propto f$). Capacitive reactance ($X_C = 1/\omega C \propto 1/f$) is a rectangular hyperbola. Sometime when a complex circuit behaves in a certain way with frequency, we try to relate to one of these above behaviours.

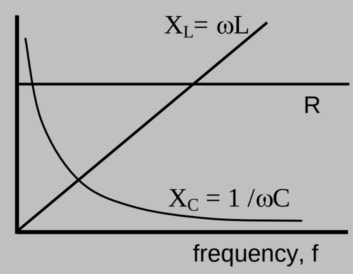


Fig.4-7: Behaviours of resistance, inductance and capacitance with frequency.

Let us summarise the important features of an ideal capacitor:

i) An ac current can pass through a capacitor.

ii) With increasing frequency, ac current through a capacitor increases.

iii) The Capacitive Reactance decreases with increasing frequency (Fig.4-7).

iv) At infinite frequency the Capacitive Reactance becomes zero, i.e., it essentially forms a short circuit.

v) At zero frequency (i.e., at dc) the Capacitive Reactance becomes infinity, i.e., it essentially becomes an open circuit.

Let us summarise the important features of an ideal inductor:

i) With increasing frequency, ac current through an inductor decreases.

ii) The Inductive Reactance increases with increasing frequency.

iii) At infinite frequency the Inductive Reactance becomes infinity, i.e., it essentially becomes an open circuit.

iv) At zero frequency (i.e., at dc) the Inductive Reactance becomes zero, i.e., it essentially forms a short circuit.

Let us summarise the important features of an ideal resistor:

i) With changing frequency, ac current through a resistor remains unchanged.

ii) A resistor has the same Resistance value at zero frequency (dc), or at infinite frequency, or at any other frequency.

We consider dc to have zero frequency, and X_C becomes infinity there. This relates to the previous finding that when the dc transients are over, the capacitor effectively works as a break (open) in the circuit allowing no current.

At the other extreme, i.e., at infinite frequency, the capacitor acts as a short. The above observations can also be related to the dc transient behaviour. At the moment of switching to a dc source, a sudden step rise in voltage can be said to contain a very high frequency content *(Fourier theorem)* and a capacitor allows all the current that can flow in a circuit.

A similar consideration applies for the inductor as well. At dc or zero frequency, an ideal inductor offers no opposition ($X_L = 0$) as *di/dt* is zero, while the opposition is infinitely high ($X_L = \infty$) at infinite frequency since *di/dt* is infinity.

4.3.4 Both Inductor and Capacitor in series, phase considerations

Fig.4-8 shows a circuit comprising of an ac source, v_{in}, a resistor, an inductor and a capacitor in series. The loop current *i*, and voltages v_r, v_l and v_c across the resistor the inductor and the capacitor respectively are also shown.

From the above treatments on a capacitor and an inductor individually in series with a resistor, we may deduce that v_l would lead the current *i* by $\pi/2$ while v_c would lag behind *i* by $\pi/2$. This combined phase relationship is shown in Fig.4-9a while Fig.4-9b shows the phasors for v_r, v_L and v_c. We can see the advantage of choosing the phase of *i* (or of v_r) as zero in this example, as it is intermediate to both the other voltages. We can add the three voltages in Fig.4-9b in terms of the respective peak voltage values as,

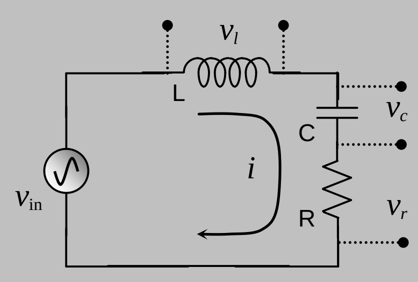


Fig.4-8: Series LCR circuit on ac

$$V_{IN} = \sqrt{V_R^2 + (V_L - V_C)^2} \qquad \text{... 4.24}$$

and the phase angle by, (with reference to that of v_r)

$$\phi = \tan^{-1}\frac{V_L - V_C}{V_R} \qquad \text{... 4.25}$$

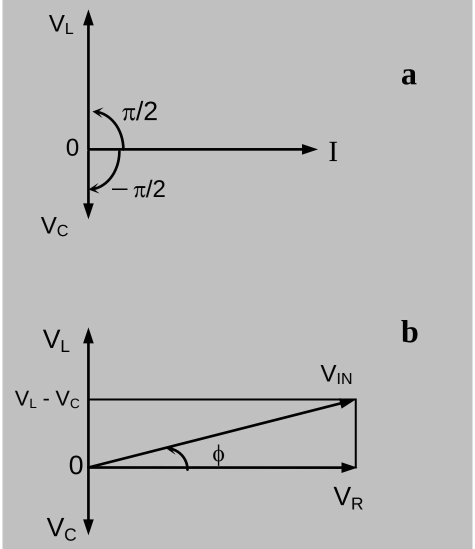


Fig.4-9: Phasor representation of currents and voltages in a series LCR circuit

From above, we can see that the magnitude of the peak voltage would be real and positive irrespective of the relative magnitudes of V_L and V_C (since they are squared) but, the sign of phase angle, ϕ, will very much depend on their relative magnitudes. When $V_L > V_C$, ϕ is positive, while ϕ is negative for $V_L < V_C$. When $V_L > V_C$, we say that the inductor dominates and the circuit has an overall inductive behaviour. On the other hand when $V_L < V_C$, we say that the capacitor dominates and the circuit has an overall capacitive behaviour.

From the above equations a special case looks interesting when $V_L = V_C$ (magnitudes of the respective voltages v_L and v_c are equal but they are opposite in phase). Then we have, $V_{IN} = V_R$ and $\phi = 0$. This means that the reactive components cancel each other and the circuit behaves as a purely resistive one. When does this happen? From Eq.4.14 and Eq.4.21 above we can see that this will happen when $X_C = X_L$ since the current through all of the three devices are the same (current in a loop has to be the same everywhere). This will have important implications to be described later.

We will deal with all the above circuits more thoroughly after we introduce a powerful mathematical technique – use of Complex numbers.

4.4 Use of Complex-number functions

[We assume you know complex number mathematics. In electricity we choose the letter j to represent the imaginary operator $\sqrt{-1}$ since i is used to represent current.]

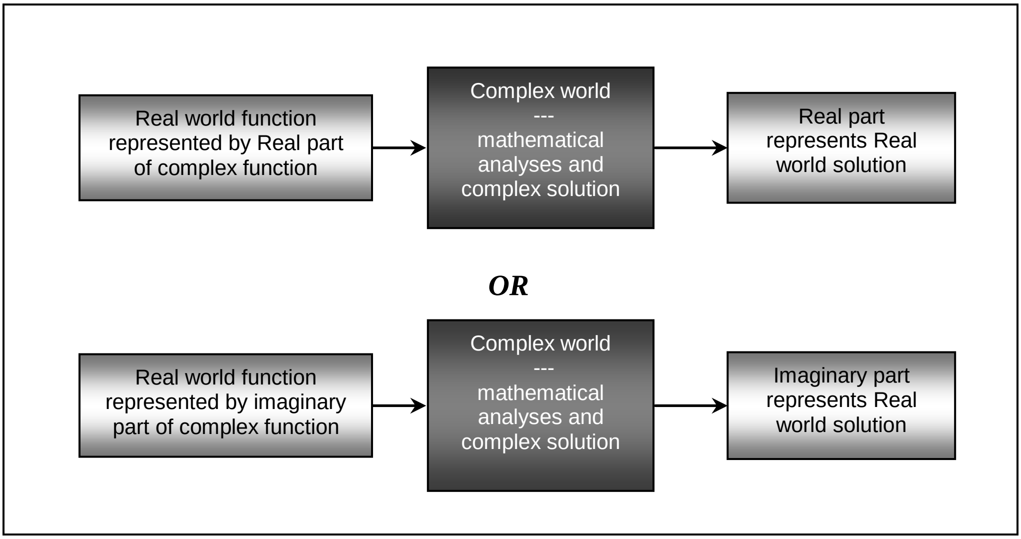
Seeking to use an exponential function

We know that differentiation or integration of a sine function gives a cosine function and vice versa. On the other hand if we subject an exponential function to the same treatments, the basic function remains unchanged. This is advantageous in performing mathematical analyses and therefore, we would like to use suitable exponential functions instead of the sinusoidal functions described above for representing electrical voltages or currents.

With the above objective let us examine the Complex exponential function

$$e^{j\omega t} = Cos\ \omega t + j\,Sin\ \omega t \qquad \text{... 4.26}$$

How can we use this to represent the sinusoidal voltages discussed before? We can see that both the real and imaginary parts of Eq.4.26 have terms that can be used to represent sinusoidal voltages or currents. So what we do is that, we choose either the real part or the imaginary part of the above exponential function, but not both at the same time, to represent our real world functions. We do all the complicated mathematical analyses involving complex functions, and finally when we arrive at a mathematical solution we simply single out our originally chosen part (real or imaginary, as appropriate) to get the real world solution. The procedure is schematically represented below.



Obviously we have to choose either one of the above two, not both in the same work. The workings will be clear when we analyse circuits of interest in the later sections. When we use a complex representation for a voltage, current, impedance or reactance, we usually use a bold faced capital (e.g. ***V, I, Z, X)*** in a book. However, for handwriting, we cannot do that, so we may use a bar above the symbols with normal face (e.g., $\overline{V}, \overline{I}, \overline{Z}, \overline{X}$ *)*. Resistance is always real as it does not have a frequency dependent term, so it is represented by a normal *R* symbol always. All magnitudes are represented by non-bold capital characters (without bars in handwriting). Remember that the magnitude of a complex number is given by the square root of its product with its complex conjugate.

As an example of representing a sinusoidal function by a complex number, we see that Eq.4.19a ($i_l = I_L\ Sin\omega t$) can be represented by the imaginary part of Eq.4.26. That gives us, the complex representation with an exponential function as,

$$\boldsymbol{I_L} = I_L\ Cos\ \omega t + j\,I_L\ Sin\ \omega t = I_L\,(Cos\ \omega t + j\ Sin\ \omega t) = I_L\ e^{\,j\omega t} \qquad \text{... 4.27}$$

remembering that it is only the imaginary term that represents our interests.

Now, from Eq.4.18 and Eq.4.12, the complex Inductor voltage,

$$\boldsymbol{V_L} = L\ d\boldsymbol{I_L}/dt = j\omega L\ I_L\ e^{\,j\omega t} \qquad \text{... 4.28}$$

Suppose this is the result we want. Then what is the real world mathematical expression for the result? Expanding the above we get,

$$\boldsymbol{V_L} = j\omega L\ I_L(Cos\ \omega t + j\ Sin\ \omega t) = -\ \omega L\ I_L\ Sin\ \omega t + j\omega L\ I_L Cos\ \omega t \qquad \text{... 4.29}$$

Since we started with the imaginary part, we have to stick to it. Therefore, taking the imaginary part from the above expression, we get, for our real world solution,

$$v_l = \omega L\ I_L Cos\ \omega t \qquad \text{... 4.30}$$

which is the same as Eq.4.19b obtained before. The advantage may not be apparent in such a simple analysis. However, when you perform more complex analysis you will appreciate the power of the Complex-number technique.

In a similar way if we want to represent a *cosine* function as in Eq.4.2 using a complex number, we could use the real part of Eq.4.26. After all analyses the result will have to be separated into its real and imaginary parts, and the real part here would represent our world solution. However, as mentioned above, we cannot mix the two representations, we use either the real part and stick to it, or we use the imaginary part and stick to it.

4.5 Representation of complex number in diagrams, Real and Imaginary axes

There is a standard form for representing complex numbers in a diagram, which is usually called an 'Argand Diagram'. Here the real part of a number is represented along the X-axis and the imaginary part is represented along the Y-axis. Suppose a complex number is given by

$$\boldsymbol{P} = a + jb \qquad \text{... 4.31a}$$

Then a is plotted along the X-axis and jb along the Y-axis as shown in Fig.4-10a, and the complex number itself is represented by the diagonal. We can see here that b is also a real number. Only when it is multiplied by the imaginary operator j it is plotted along the Y-axis. So, we can say that when the imaginary operator j operates on any real number it rotates the direction by 90^0.

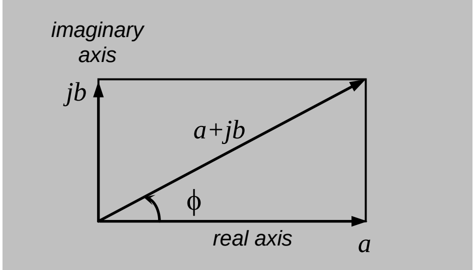


Fig.4-10a: Complex number representation

This can be extended further as shown in Fig.4-10b. The Real axis is along the X-axis, multiplying a real number by j rotates it by 90^0 to Y-axis, which we have just discussed. Multiplying again by j rotates it by a further 90^0 to negative X-axis (since $j^2 = -1$, this is also real but negative). Finally, another multiplication by j rotates the number by a further 90^0 to – ve Y-axis (since $j^3 = -j$, this is negative of the imaginary axis). Another multiplication by j will result in $j^4 = 1$ which is the original real axis again. We can also see from Fig.4-10,

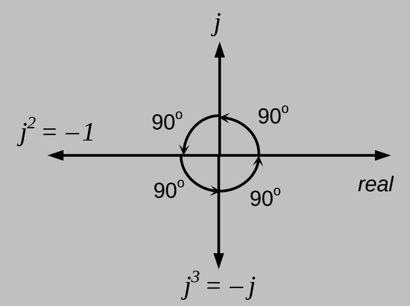


Fig.4-10b: Operation by j corresponds to a rotation by 90^0

Magnitude of $a + jb = \sqrt{(a^2 + b^2)}$ ... 4.31b

and the angle that the complex number makes with respect to the real axis is,

$$\phi = tan^{-1}\,(b/a) \qquad \text{... 4.31c}$$

4.6 Keeping the form of Ohm's law intact for ac, complex impedance

Suppose in a circuit with both resistive and reactive elements (i.e., having R, and either or both of L and C) we have

$$i = I\,Sin\,\omega t \qquad \text{... 4.32}$$

and $$v_{in} = V_{IN}\,Sin(\omega t + \phi) = I\,Z\,Sin(\omega t + \phi) \qquad \text{... 4.33}$$

where we have taken $V_{IN} = I\,Z$ for the magnitudes following Ohm's law.

According to the complex representation discussed above, we can represent the above mentioned ac current and voltage by the corresponding imaginary parts of the following complex numbers,

$$\boldsymbol{I} = I\,e^{j\omega t} \qquad \text{... 4.34}$$

and $$\boldsymbol{V_{IN}} = V_{IN}\,e^{j(\omega t + \phi)} = I\,Z\,e^{j(\omega t + \phi)} \qquad \text{... 4.35}$$

Now Eq.4.35 can be expanded and rearranged as,

$$\boldsymbol{V_{IN}} = I\,Z\,e^{j\omega t}\,e^{j\phi} = (I\,e^{j\omega t})\,(Z\,e^{j\phi}) \qquad \text{... 4.35a}$$

Here the first part of the right hand expression is simply the complex current $\boldsymbol{I}$. We can make Eq.4.35a to have the form of Ohm's Law if we represent the second part as the complex impedance $\boldsymbol{Z}$. Then we have, dropping the subscripts for a general expression,

$$\boldsymbol{Z} = Z\,e^{j\phi} \qquad \text{... 4.36}$$

and $$\boldsymbol{V} = \boldsymbol{I}\boldsymbol{Z} \qquad \text{... 4.37}$$

which is the complex form of Ohm's law for ac.

Here we can see that the complex current retains the sinusoidal frequency information while the phase information of the voltage has been transferred to the complex impedance.

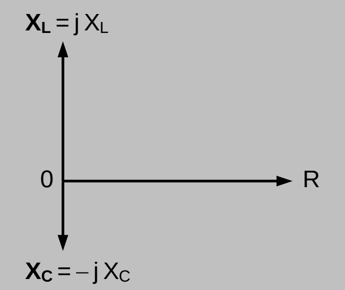


Fig.4-11: Phasor representation of resistance and reactance in a series LCR circuit

We remember from Fig.4-9 that the phase angle of the voltages across the inductor and the capacitor are 90^0 and -90^0 respectively with respect to the voltage across the resistor, which we took to have a phase angle of '0'. Following the above mentioned complex representation, we want to transfer these phase information to the respective reactances. We have seen before (complex Argand diagram) that multiplying by *j* rotates a quantity anticlockwise by 90^0 while multiplying by – *j* rotates a quantity clockwise by 90^0 (i.e., -90^0), so that the respective complex reactances can be represented as,

$$\boldsymbol{X_L} = +\ jX_L \qquad \text{... 4.38}$$

$$\boldsymbol{X_C} = -\ jX_C \qquad \text{... 4.39}$$

where X_L and X_C are real numbers, and the subscripts indicate whether the reactance is inductive or capacitive. These are shown in Fig.4-11 where *R*, the real part of the impedance is shown along the X-axis (reference 0^0).

We can represent the general form of reactance as,

$$\boldsymbol{X} = \pm\ jX \qquad \text{... 4.40}$$

where a + ve sign is appropriate for an inductive reactance and a – ve sign is appropriate for a capacitive reactance.

Then we can express complex ***Z***, following the form of Eq.4.31, as,

$$\boldsymbol{Z} = R \pm jX \quad = R + \boldsymbol{X} \qquad \text{... 4.41}$$

Here, *R* has been shown in normal font since it is always a real number.

APPLICATION TO SOME CIRCUITS OF INTEREST

We apply the above knowledge to analyse some circuits of interest – particularly RC, RL and LRC circuits, and later, transformers.

4.7 RC high pass filter circuit

Fig.4-12 shows an RC high pass filter circuit (shaded) which has an input side and an output side. It also shows an input voltage source, which we shall consider to be pure

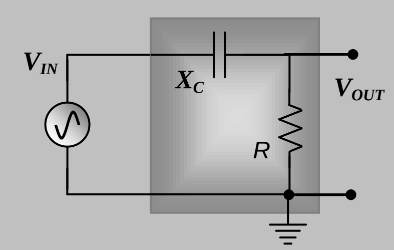


Fig.4-12: RC high pass filter

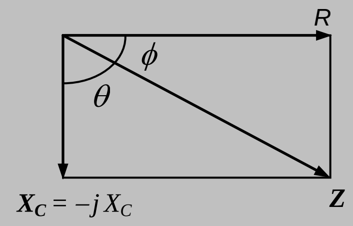


Fig.4-13: Impedance phasor diagram for the RC high pass filter

sinusoidal. Let us represent all the parameters in the circuit using relevant Complex-numbers where applicable, viz., $\boldsymbol{V_{IN}}$ and $\boldsymbol{X_C}$. Note that R is real always.

We can see that this circuit is basically a voltage divider, $\boldsymbol{V_{IN}}$ being divided between the capacitor and the resistor in series. The total impedance here is

$$\boldsymbol{Z} = R + \boldsymbol{X_C} \qquad \text{... 4.42}$$

We can use the complex phasor diagram of Fig.4-13 to represent R, $\boldsymbol{X_C}$ and the total impedance $\boldsymbol{Z}$ at an angle ϕ (which is inherently – ve) with respect to R. Here, $\boldsymbol{X_C} = -jX_C$ and $\boldsymbol{Z} = Z\,e^{\,j\phi}$

Using complex form Ohm's law, current is given by,

$$\boldsymbol{I} = \frac{\boldsymbol{V}_{IN}}{\boldsymbol{Z}} = \frac{\boldsymbol{V}_{IN}}{R + \boldsymbol{X}_C} \qquad \text{... 4.43}$$

However, we are interested only in the voltage across the resistor here and we name this as the output $\boldsymbol{V_{OUT}}$ $(= \boldsymbol{I}\,R)$ with respect to the ground. Usually we would like to see what fraction of $\boldsymbol{V_{IN}}$ is available at the output, which we call the *voltage gain*, $\boldsymbol{A_v}$. The word *gain* actually comes from an amplifier where the output is greater than input, and *gain* means how many multiples of input is available at the output. We also extend the same usage of the word *gain* to situations even where the output is less than the input, as in this case. Thus the Complex gain,

$$\boldsymbol{A_V} = \frac{\boldsymbol{V}_{OUT}}{\boldsymbol{V}_{IN}} = \frac{\boldsymbol{I}R}{\boldsymbol{I}\boldsymbol{Z}} = \frac{R}{\boldsymbol{Z}} = \frac{R}{R + \boldsymbol{X}_C} = \frac{R}{R - jX_C} \qquad \text{... 4.44}$$

whose magnitude is given by (square root of the product with its complex conjugate),

$$A_V = \frac{R}{\sqrt{R^2 + X_C^2}} = \frac{1}{\sqrt{1 + X_C^2 / R^2}} \qquad \text{... 4.45}$$

Now replacing X_C using Eq.4.14 ($X_C = 1/\omega C$) we get,

$$A_V = \frac{1}{\sqrt{1+\dfrac{1}{\omega^2 C^2 R^2}}} = \frac{1}{\sqrt{1+\dfrac{\omega_0^2}{\omega^2}}} \qquad \text{... 4.46}$$

where we have chosen

$$\omega_0 = \frac{1}{CR} \qquad \text{... 4.47}$$

Taking $\omega_0 = 2\pi f_0$, we have,

$$\boxed{f_0 = \frac{1}{2\pi CR}} \qquad \text{... 4.48}$$

and Equation 4.46 can be rewritten as,

$$\boxed{A_V = \frac{1}{\sqrt{1+\dfrac{f_0^2}{f^2}}}} \qquad \text{... 4.49}$$

Eq. 4.46 and Eq.4.49 gives two forms of the desired solution for the magnitude of the voltage gain and Eq. 4.47 and Eq. 4.48 gives the value of the constants ω_0 and f_0 that we have chosen to give us the nice simple forms for the solution. We can see that ω_0 has to have the dimension of an angular frequency (it has to cancel ω), and is usually called the *characteristic angular frequency* of the circuit as it depends on the circuit parameters R and C (remember, RC is called the *time constant* of this circuit, which is appropriate when a dc step voltage is applied). For circuits with different values of R and C, ω_0 will be different and the individual values of ω_0 will allow us to compare the behaviours or charcteristics of these different circuits. Therefore we add the adjective, *'characteristic'* in the above naming. The corresponding *characteristic frequency is given by* f_0. The significance of ω_0 or f_0 will be made clear soon.

4.7.1 Visualisation of Frequency Response

To have a preliminary idea about the frequency response, we look at the voltage gain obtained from Eq.4.49 at two extreme frequency values as follows,

when	$f = \propto$ (infinity),	$A_V = 1$
and when	$f = 0$ (dc),	$A_V = 0$

The above results indicate that at very high frequencies A_V has a value close to 1 (unity), ie, the circuit allows the input to pass through to output without much attenuation, while at low frequencies the gain decreases, becoming almost zero at very low frequencies. If we plot A_V obtained from Eq. 4.49 as a function of f, the plot would have

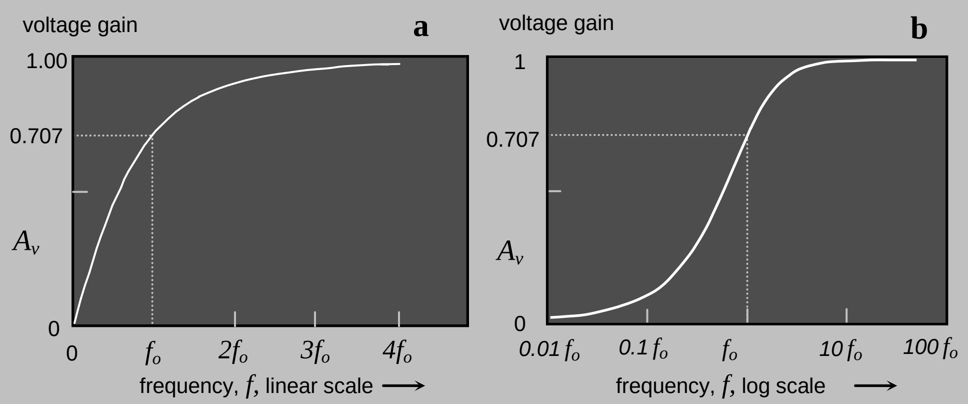


Fig.4-14: Gain Frequency response of an RC high pass filter, linear-linear scale (a) and linear-log scale (b).

a behaviour as shown in Fig.4-14a. This is called the *frequency response of voltage gain*. This figure shows that the circuit allows only high frequencies to pass, but not low frequencies. Therefore it is called a *High-Pass Filter*. Note that in Fig. 4-14a both the axes are in linear scale. Fig. 4-14b shows the gain in linear scale, but with the frequency axis plotted in log scale. A log scale accommodates a very large range of frequencies by compressing them. Note the difference in the frequency range and the essential change of shape in the two graphs. From the second graph, we can see that for practical purposes, the gain may be considered unity above $10f_o$, i.e., ten times the cut-off frequency, and negligible below $0.1f_o$, one tenth of the cut-off frequency. Eq.4.49 also tells that the maximum voltage gain of 1 is achieved only at infinity, meaning that nowhere within the finite range the gain is strictly unity, it is always less than unity.

Why is it called a filter? You know a filter paper (or, a sieve) allows liquids and very small particles to pass through but not larger particles. So a filter paper or a sieve works as a particle size selective device. Similarly the above circuit acts as an electrical frequency selective device in that it allows electrical signals of high frequencies to pass, and does not allow signals of low frequencies. Therefore it is also called a filter (more strictly, a frequency filter), but obviously it is a filter for electrical signals.

In case of the sieve, or a particle filter if you collect the smaller particles from below, and leave out the larger ones, you can call it a 'small pass filter'. On the other hand if you choose to collect larger particles from the top of the filter leaving the smaller ones out, you can call it a 'large pass filter'. Similarly in case of the electrical filters if it allows only high frequencies to pass as in the above case, then we call it a *'high pass filter'*. If we modify the above circuit (to be described shortly) to allow low frequencies to pass only, we call it a *'low pass filter'*.

4.8 Significance of ω_0 or f_0

To see the significance of ω_0 (or, of f_0), let us see what happens when the angular frequency ω becomes the same as ω_0 (or, when f equals f_0),

From Eq. 4.43 and Eq.4.46 we can see that,

when $\omega = \omega_0$ or, $f = f_0$, $$A_V = \frac{1}{\sqrt{2}} \cong 0.707 \qquad \text{... 4.50}$$

We indicate this point on the plot in Fig.4-14 corresponding to f_0 where the gain is approximately 0.707 (you now understand how this magic number has come about). Its value gives us an idea that the frequencies above it are allowed, or, passed on well, while those below are gradually *attenuated* (reduced). This *Characteristic frequency* is also called the ***lower cut-off frequency*** of this high-pass filter circuit. Different high pass filter circuits will have different lower cut-off frequencies (depending on the values of R and C) and to compare their behaviours we just quote this figure. Such circuits are almost invariably used in all amplifier circuits in order to allow a chosen range of frequencies, and remove unwanted ones.

From Eq.4.47 we can write, $$R = \frac{1}{\omega_0 C} \qquad \text{... 4.51}$$

Here the Right Hand Side is simply the Reactance of the capacitor at f_0. Therefore we may say that,

at cut-off frequency,   ***Reactance = Resistance***

Example: Suppose for a circuit, R = 100 kΩ and C = 0.1μF, then $f_0 \cong$ 16Hz.
For another circuit, let R=10 kΩ and C = 0.1μF, then $f_0 \cong$ 160Hz.

Thus the first circuit having a cut-off frequency of about 16Hz will allow frequencies above 16 Hz to pass through which is very good for a High-Fidelity (Hi-Fi) audio amplifier used for high quality music (remember, our audio range is 20Hz to 20 kHz). Since there is nothing of interest below 20Hz, we tend to cut-off these frequencies. Otherwise noise in these frequencies will cause undue power consumption, and might saturate the amplifier unnecessarily. On the other hand if there are noise signals overlapping the low frequency signals we tend to push the cut-off frequency higher to reduce noise, even at the cost of losing some signals. Noise originating from mains electricity at 50Hz poses a big problem. Pushing the cut-off to about 160 Hz simply by changing the resistor in the above example reduces the 50 Hz noise significantly. Looking at Eq.4.49 and Fig. 4-14 we can see that A_V can never be made zero ideally, but can be made very low, to get an acceptable sound quality. Later we will see how we can reduce such noise further using higher order active filters having sharper cut-off.

In choosing such higher cut-off (~ 200Hz) in audio amplifiers, we lose some quality of the resulting sound, but this is better than having a continuous noise in the background. In fact it has been found that we understand speech better if signals with frequency below 200Hz are cut out. Therefore you will find that most public address amplifiers (i.e., those that are used for speech or 'address') have cut-off frequencies of this order. However, amplifiers for music would not do this. They would try to reproduce down to 20Hz as much as possible.

Small radio and cassette recorders have cut-off frequencies at a hundred Hz or more as the loud speakers they use are small and cannot reproduce sound of frequencies below 200 Hz well. Therefore there is no point in taking all the trouble of making an amplifier working down to 20Hz!

The significance of f_0 would be more appreciated if we plot Equation 4.49 in a log-log scale instead of the linear-linear or linear-log scales shown in Fig.4-14. In a graph for a filter we often want to get information over a wide range of frequencies and voltage gains (called dynamic ranges), say, over 10^4 times each (sometimes we look for dynamic ranges of the order of 10^{12}!). This whole range cannot be satisfactorily represented in a linear scale graph. On the other hand logarithm compresses such vast ranges into smaller numerical ranges and a log scale graph can give a better representation of such quantities. A 10 times increase in a number correspond to only an increase by unity in $\log_{10}$ scale, for example: $\log_{10} 100 = 2$, $\log_{10} 10 = 1$, $\log_{10} 1 = 0$, $\log_{10} 0.1 = -1$, and so on (remember, $\log_{10} 0 = -\infty$). We have already seen its effect on the frequency range in Fig.4-14b. However, it would be useful to know about a special log scale for the voltage gain, called the '*decibel'* or dB scale for voltages, which is used widely in electronics and the basic concepts are given below.

4.9 Decibel scale

This scale was originally introduced to express signal power ratios over a large range by compressing them using a logarithmic scale.

If two signal powers are denoted by P_1 and P_2, then a Bel scale is defined in the following way. We say that *P_2 is greater than P_1 by,*

$$\log_{10} \frac{P_2}{P_1} \text{ Bel } \textit{units.} \qquad \dots 4.52$$

Now this unit appeared to be very large (the numbers smaller) for power ratios encountered practically in signals. For example for a power difference of 10^4, we would get only 4 Bels. Besides, to express smaller intervals, people seem to prefer integers, not numbers with decimal points. Therefore, people preferred to use deci-Bel unit, or the dB unit (remember, Greek *deci* means 1/10, so that 4 Bel would become 40 dB). We would now say that signal 2 is greater than signal 1 in

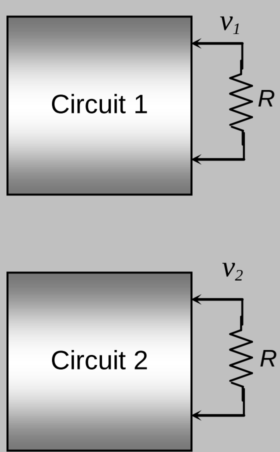


Fig.4-15: Basics of definition for decibel.

power by,

$$\boxed{10\log_{10}\frac{P_2}{P_1}\ dB} \qquad \text{... 4.53}$$

This is the basis of the dB scale.

Now suppose we would like to compare two electrical signals having voltages v_1 and v_2 respectively, each terminating into equal resistances R as shown in Fig. 4-15. So the corresponding powers are,

$$P_1 = \frac{v_1^2}{R} \quad and \quad P_2 = \frac{v_2^2}{R} \qquad \text{... 4.54}$$

$$\text{then,}\quad 10\log_{10}\frac{P_2}{P_1} = 10\log_{10}\frac{v_2^2}{v_1^2} = 20\log_{10}\frac{v_2}{v_1} \qquad \text{... 4.55}$$

where the R term cancels out on division. Therefore we can also say that signal 2 is greater than signal 1 in power by

$$\boxed{20\ \log_{10}\frac{v_2}{v_1}\ dB} \qquad \text{... 4.56}$$

This is only a different expression in terms of voltages rather than powers as given by expression 4.53, but refers to the same quantity.

4.9.1 Voltage dB scale

However, people have a tendency to carry things further on, so that some people used the above expression for dB scale using voltage ratios even when the terminating resistances in Fig.4-15 are not equal. Strictly speaking, we cannot do this, as the R terms then do not cancel in the above deductions. However, people became very fond of this expression and wanted to use it in all sorts of comparisons. They argued that let us define a new scale to express signal voltage ratios using the above expression (Exp. 4.56) knowing fully well that this does not necessarily represent a power ratio. So they called it a *Voltage dB scale*. Obviously *this equals a power ratio only if both the terminating resistances are equal, otherwise not*. This scale is widely used.

To make yourselves familiar with this scale a table for linear voltage ratios and the corresponding dB values are given below. Try to put the numbers into your brain. Note, for ratios less than one, negative dB numbers are expected.

Voltage ratio	dB difference
0.1	- 20
1	0
10	20

Voltage ratio	dB difference
100	40
1000	60
10000	80

You can see that for each ten times increase in voltage gain, the dB value increases by 20. That is, multiples are replaced by addition (this is expected, since log AB = log A + log B). This gives us an advantage. If we have two amplifiers connected sequentially, and each with a gain of 100, then the total gain is 100 x 100 =10000, which is given by a product of the individual gains. On the other hand in the dB scale we have to simply add them up algebraically. For the above example, the total gain in dB scale is 40 + 40 = 80 dB, which is just equivalent to 10000 as we can see from the above table. In the above example we have used easy numbers for voltage ratios with all 0's after 1. With not so simple numbers, which is easier – multiplication or addition?

4.10 Log-Log plot using dB scale for gain, Cut-off frequency

A log-log graph of the frequency response of voltage gain of a high pass filter, using dB scale for the voltage gain, is shown in Fig.4-16a (note: dB scale itself is a log scale).

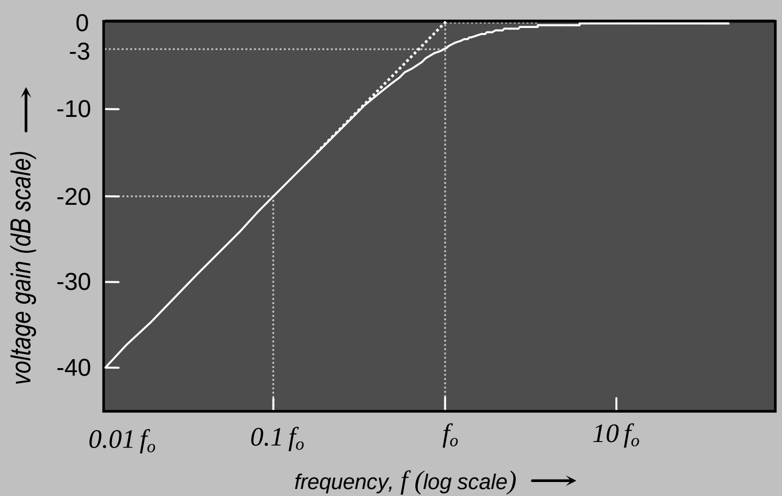


Fig.4-16a: Frequency response of Gain of RC high pass filter, log–log plot (gain in dB scale)

According to the table given above, the maximum linear voltage gain of 1 becomes 0 in the dB scale. Since the other ratios at lower frequencies are less than 1, they appear as negative dB values. What is the value of gain in dB at cut-off frequency? Since the gain is 1/√2 at cut-off,

$$20 \log_{10} \frac{1}{\sqrt{2}} = \frac{1}{2} 20 \log_{10} \frac{1}{2} = -10 \log_{10} 2 = -10 \times 0.303 \cong -3$$

That is, the *cut-off frequency f_0 occurs at – 3dB gain,* which is worth remembering. This also tells us that if the high frequency gain is anything other than 0 dB, the value at cut-off would be simply 3 dB less than that value.

There is an interesting aspect to this plot. We can see in Fig.4-16a that the curve has two approximate linear segments, one going down at low frequencies, and the other is the horizontal section (= 0 dB) at high frequencies. If we extend these straight lines they intersect at f_0 which is really significant.

4.11 Bode plot, Rolling-off slope

Some people prefer to approximate the filter using the above mentioned two straight-line segments only, forgetting the real curved plot totally. This simplified plot in the log-log scale (Fig.4-16b) is called a **'Bode plot'** and here f_0 assumes a significant role. In the Bode plot, all frequencies above f_0 are assumed to have a constant gain (here, 0 dB) while frequencies below f_0 are attenuated with a constant slope, called the *Rolling-off slope*.

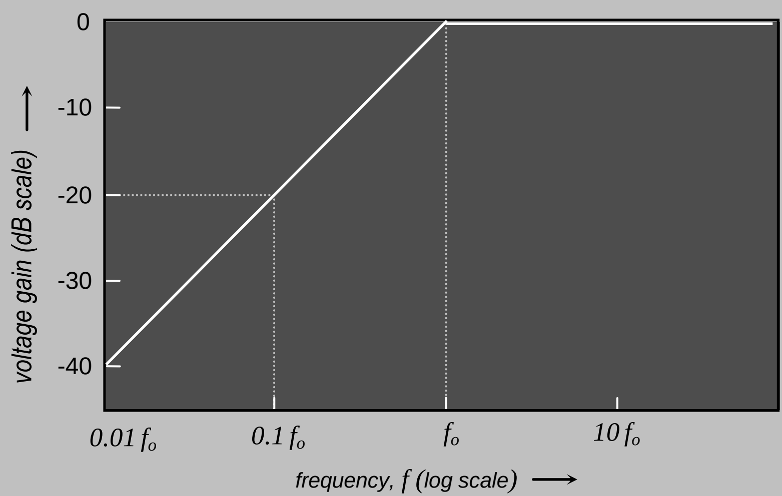


Fig.4-16b: Bode plot for an RC high pass filter

We can find the slope of the rolling-off segment of the curve using Eq.4.49. Let us choose frequencies $0.1\ f_0$, $0.01\ f_0$, etc. each 10 times less than the previous one and calculate the corresponding A_v, and see the pattern.

At $f = 0.1\ f_0$, $f_0 = 10\ f$, so that, $A_V = 1/\sqrt{101} \cong 1 / 10 = 0.1$

At $f = 0.01\ f_0$, $f_0 = 100\ f$, so that, $A_V = 1/\sqrt{10001} \cong 1 / 100 = 0.01$

These two points are also shown in the figure. The pattern will follow down to any lower gain. We can see that the gain is reduced by 10 times for each 10 times reduction in frequency. A ten times range of frequency is called a *decade of frequency* and from the above table we can see that a 10 times gain change is equivalent to a 20 dB change in the dB scale (which is additive or subtractive). Therefore we say that the ***gain reduces***

by 20dB per decade of frequency, which is the *Rolling off slope* we are looking for. Since this slope is constant and known, we can calculate the gain at any frequency for such a filter if we know f_0, or if we know the gain at any frequency lower than f_0.

There is an alternative description of the rolling-off slope. In music if the frequency is doubled or halved, this range is called an Octave. The middle *C* note has a frequency of 256 Hz, while 512 Hz is the frequency for the *C* note an octave higher. Thus the range between f_0 and $o.5f_0$ is an *Octave*. From Eq.4.49 we can calculate in a similar way that the *Rolling-off slope* is ***6 dB per Octave*** of frequencies (find out yourself!).

So memorising the above two figures for slopes help in quick mental estimation of gains in any practical design situation.

Ex.1 Let f_0 =200 Hz for a high pass filter as above. What is the gain at i) 2Hz and at ii) 25Hz?

Ans. i) 2 Hz is 2 decades (steps: 20, 2) lower than f_0, therefore the gain at 2Hz will be – 40dB (20 + 20 below 0).

ii) 25Hz is 3 octaves below 200Hz (steps: 100, 50, 25), therefore the gain at 25Hz would be –18dB (3×6 below 0).

4.12 Order of filter, ideal filter, Passive and Active filter

The above rolling off slope of *20dB per Decade* is obtained through the simple circuit shown in Fig. 4-12, and any filter with this slope is called a *First Order Filter*. If two of such filters are arranged in tandem (two in series) as shown in Fig.4-17, the slope would be doubled as the gain at each frequency is squared. Here we have to add an extra unity gain buffer circuitry in between to stop the impedance of one affecting the other. The doubling of slope is due to addition of logarithms of the two individual gains (see for yourself by analysing the total gain by squaring Eq.4.49).

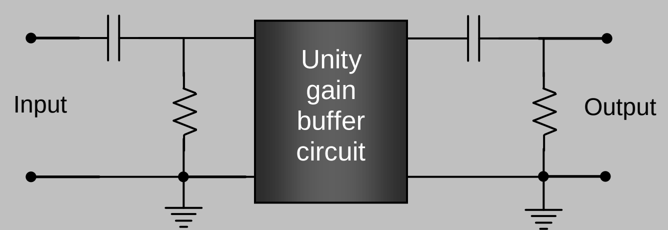


Fig.4-17: 2nd order RC high pass filter

Such a filter is called a *Second Order Filter* and has a slope of *40 dB per Decade,* or, *12 dB per Octave.* We can similarly conceive of higher order filters with greater rolling off slopes, giving sharper and sharper cut-off.

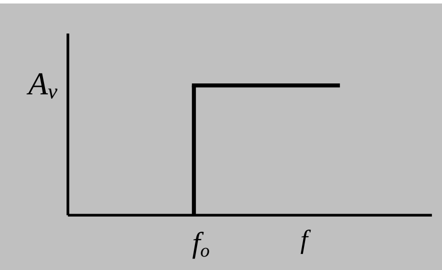


Fig.4-18: Ideal high-pass filter

What is an ideal filter? From the above, if we imagine the rolling off slope to go on increasing we would reach an ideal filter as shown in Fig. 4-18 where nothing is passed below f_0 while 100% is passed above f_0. This is an ideal characteristics which we would like to have but can never attain in practice.

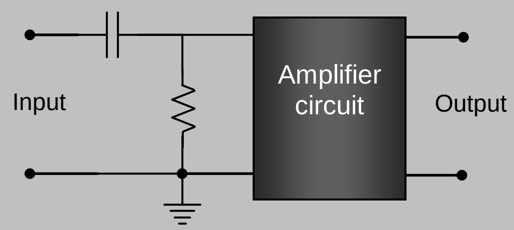


Fig.4-19: Active 1st order RC high pass filter

As mentioned above, cascading (placing in series) simple filter circuits directly is not practical as the output impedance of the first would be affected by the input impedance of the next and so on, and the gain would keep on decreasing too as it is never '1' at any finite frequency. Therefore for practical realisation we usually incorporate amplifier circuits employing transistors after each filter circuit when the whole is called an Active Filter circuit. These can be easily cascaded and the gain can also be adjusted to compensate for any reduction. In fact the gain is usually more than unity. Fig. 4-19 shows the scheme of an active 1st order High pass filter circuit. We shall study more of it later.

Circuits as in Fig.4-19 are called *Active Filter Circuits* as these take power from a source (usually a dc battery) to operate. We define *an active circuit as the one that takes power from an external source while a passive circuit takes none*. Therefore the circuit shown in Fig. 4-12 is a 1st order passive filter since it does not take power from any external source. Active filters are versatile in that they can be easily cascaded as mentioned above. However, unexpected behaviours creep in as one does such cascading with the intention of increasing the order of the filter, and therefore, active filter design has become a special sub-branch of study.

4.13 Phase response

Fig. 4-20a shows a *voltage phasor diagram* for the *RC* filter circuit of Fig.4-12. Here the voltage amplitude V_R across resistor is plotted along the X-axis and the voltage amplitude V_C across capacitor is plotted along the negative Y-axis since V_C lags behind V_R by 90°. The vector sum of the voltages amplitudes, across both the resistor and the capacitor together, is basically V_{IN} here. The phase angle between V_R and V_{IN} is ϕ with V_{IN} lagging behind V_R as shown.

The corresponding *impedance phasor diagram* for the *RC* filter circuit was shown in Fig. 4.13 and is reproduced in Fig.4-20b. The resultant impedance, ***Z***, is along the direction of the total voltage V_{IN}. Thus the phase of V_{IN} with reference to that of V_R is given by, with the help of the impedance phasor diagram (Fig.4-20b), as,

$$\phi = -\tan^{-1}\frac{X_C}{R} \qquad \text{... 4.57}$$

where a – ve sign appears as X_C lies along the – ve Y-axis.

Using $X_C = 1/\omega C$ (Eq. 4.14), we get,

$$\boxed{\phi = -\tan^{-1}\frac{1}{\omega CR}} \qquad \text{... 4.58a}$$

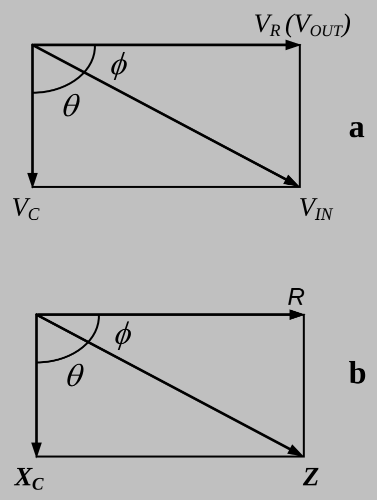


Fig.4-20: Voltage (a) and Impedance (b) phasor diagrams for RC high pass filter

The above relation gives the phase angle of v_{in} with respect to that of v_r (or v_{out}).

However, in a filter circuit, it is usual to talk about the phase of the output voltage with respect to that of the input voltage as opposed to the above choice. Therefore we will obtain the phase angle using Eq.4.58a, but will say that the output voltage v_{out} leads the input voltage v_{in} by a phase angle of ϕ. Since this description makes the phase angle positive, let us name this positive phase angle as ϕ' so that we have,

$$\boxed{\phi' = \tan^{-1}\frac{1}{\omega CR}} \qquad \text{... 4.58b}$$

Plotting Eq. 4.58b as a function of frequency gives us a phase response curve as shown in Fig.4.21. We have plotted the phase response against frequency plotted in both linear scale (Fig.4-21a) and log scale (Fig.4-21b). Note the differences in the features of these two plots.

From the above equation we find that (remember, $\omega = 2\pi f$ and $\omega_0 = 2\pi f_0$),

$$\phi' \to 0 \text{ as } f \to \propto, \qquad \text{and} \qquad \phi' \to 90^\circ \text{ as } f \to 0$$

Besides, when $\omega CR = 1$, i.e., when, $\omega = \dfrac{1}{CR} = \omega_0$,

i.e., when	$\boldsymbol{f = f_0}$,	we get,	$\boldsymbol{\phi' = 45^o}$

A small point is worth noting here. At high frequencies where the voltage gain is nearly 1, i.e., the input is almost fully allowed to pass, the phase difference is almost 0°, while at low frequencies where voltage gain is nearly zero, the phase difference is about 90°. Besides, at cut-off frequency the phase difference is 45° which indicates that although the gain is very near to unity around the cut-off frequency, there is considerable phase difference and this distorts compound waveforms which contain many frequency components. The components at higher frequencies may suffer a negligible phase difference while those near the cut-off will suffer about 45°. After filtration the resulting waveform will be distorted, having a different shape than the original one. Therefore to avoid such distortion, we usually set the lower cut-off frequency at 5 to ten times below the lowest frequency content in a signal.

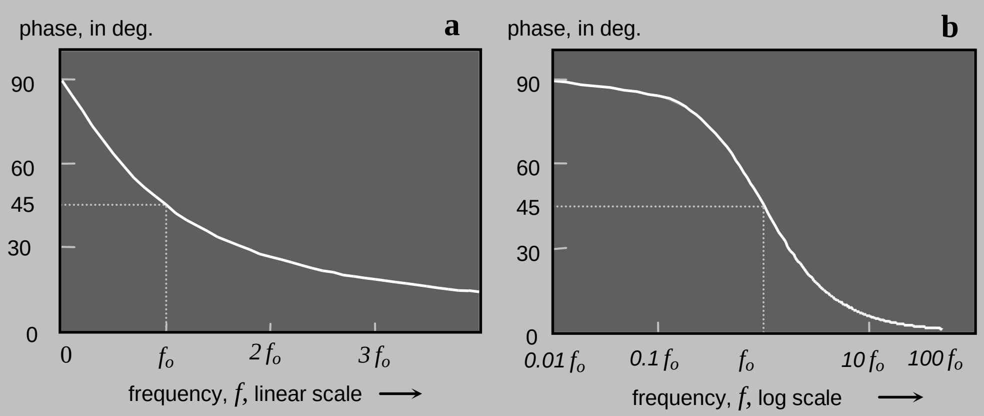


Fig.4-21: Frequency response of phase difference between input and output voltages with frequency in linear scale (a), and in log scale (b).

4.14 R-C Low pass filter

A simple rearrangement of the high pass filter circuit in Fig.4-12 will give us an *RC Low pass filter* shown in Fig.4-22 (dark shaded). As before, it has an input ac sinusoidal voltage source v_{in}, and an output v_{out}. However, here the output is taken across the capacitor instead of the resistor. Therefore we have rearranged the circuit such that one side of the capacitor is at ground potential, common to both input and output. The corresponding voltage phasor diagram is shown in Fig.4-23a where the output amplitude V_{OUT} is shown along the –ve Y-axis since it is essentially V_C in this case, different from that for the RC high pass filter shown in Figs. 4-12 and 4-20 earlier. The corresponding impedance phasor diagram is shown in Fig.4-23b. Here phasor $\boldsymbol{X_C}$ has the same direction as V_{OUT} and phasor $\boldsymbol{Z}$ has the same direction as V_{IN}. The input v_{in} is the voltage dropped across the whole impedance $\boldsymbol{Z}$, and the output v_{out} is the voltage dropped across reactance $\boldsymbol{X_C}$. As before, we would like to see what fraction of v_{in} is

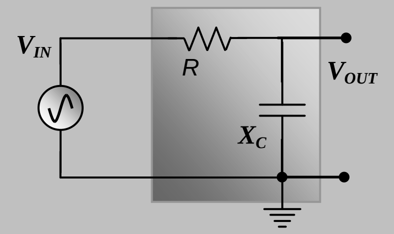


Fig.4-22: RC low pass filter

available at the output, which we call the voltage gain, A_V, as a function of frequency. Besides, we would also like to see the phase difference θ between v_{out} and v_{in} as a function of frequency.

4.14.1 Voltage gain response

To analyse, we can use the appropriate ac form of Eq. 2.13 for this voltage divider. Thus if the reactance (complex form) of the capacitor be $\boldsymbol{X_C}$, then the Complex gain,

$$\boldsymbol{A}_V = \frac{\boldsymbol{V}_{OUT}}{\boldsymbol{V}_{IN}} = \frac{\boldsymbol{I}\,\boldsymbol{X}_C}{\boldsymbol{I}\,\boldsymbol{Z}} = \frac{\boldsymbol{X}_C}{R+\boldsymbol{X}_C} \qquad \text{... 4.59}$$

Compare this with Eq.4-43 for a high-pass filter. The denominator is the same in both. However, in the previous case R was there on the numerator, while we have X_C in the present case, which relates to the circuit element across which the output voltage is measured.

We can write eq.4.59 as, (using $\boldsymbol{X_C} = -j/\omega C$)

$$\boldsymbol{A}_V = \frac{\dfrac{-j}{\omega C}}{R-\dfrac{j}{\omega C}} = \frac{1}{1+j\omega CR} \qquad \text{... 4.60a}$$

where we have used the relation, $1/-j = +j$. The magnitude of gain, A_V, is given by, using the definition of ω_o given earlier ($=1/CR$),

$$A_V = \frac{1}{\sqrt{1+\omega^2C^2R^2}} = \frac{1}{\sqrt{1+\dfrac{\omega^2}{\omega_0^2}}} \qquad \text{... 4.60b}$$

We can then rewrite Equation 4.60b in terms of the frequencies as,

$$\boxed{A_V = \frac{1}{\sqrt{1+\dfrac{f^2}{f_0^2}}}} \qquad \text{... 4.60c}$$

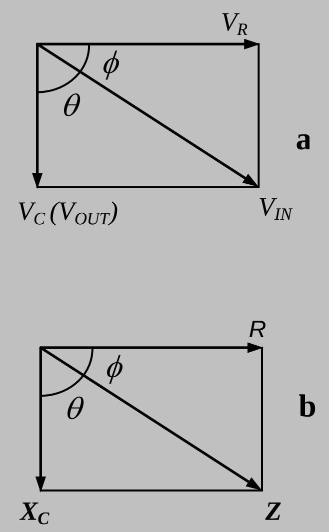


Fig.4-23: Voltage (a) and Impedance (b) phasor diagrams for RC low pass filter

Note the difference from Eq.4.49. Here we have f/f_0 while it was f_0/f in the previous case. Plotting Eq.4.60c gives us Fig.4-24a which shows that it allows only low frequencies below f_0 to pass but increasingly blocks those above f_0. Therefore it is called a *Low-pass filter* circuit. Again, the voltage gain is 0.707 at f_0, the ***upper cut-off frequency*** of this low pass filter. The corresponding log-log plot with the voltage gain expressed in dB scale is shown in Fig.4-24b where the gain is –3dB at f_0. The corresponding simplified Bode-plot is also indicated in the figure by the dotted lines. The rolling off slope is the same as that before, i.e., 20dB per decade, or 6dB per Octave.

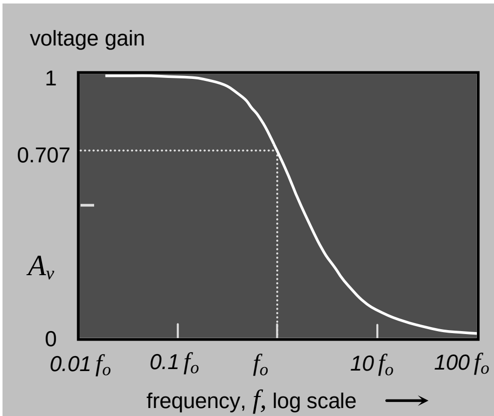


Fig.4-24a: Frequency response of voltage gain for an RC low pass filter, linear gain

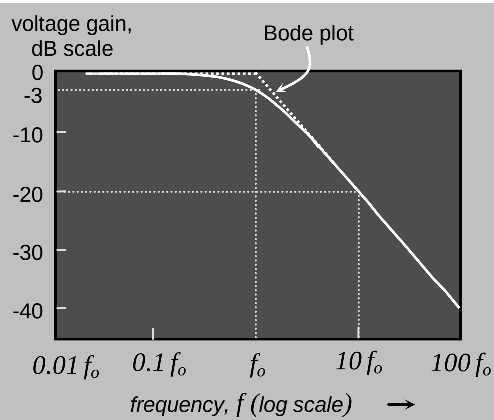


Fig.4-24b: Frequency response of voltage gain for an RC low pass filter, gain in dB scale

4.14.2 Phase response

The complex phasor diagram of impedance is shown in Fig.4-23b, and here the phasors of interest are $\boldsymbol{X_C}$ and $\boldsymbol{Z}$ with a phase angle θ between them. We can see that the phase angle would be given by, using Eq.4-14 ($X_C = 1/\omega C$)

$$\theta = -\tan^{-1}\frac{R}{X_C} = -\tan^{-1}\omega CR \qquad \text{... 4.61}$$

where –ve sign appears again since X_C is along the –ve Y-axis.

In relation to Fig.4-23a, we can see that this is the same angle as between v_c, which is the output voltage v_{out} in this case, and the input voltage v_{in}. We can also see that v_{out} is lagging behind v_{in}. Therefore, according to our usual way of expressing output with reference to input, this angle θ would be inherently negative, so we can use Eq.4.61 directly for the desired phase angle of v_{out} with respect to that of v_{in}.

This is plotted in Fig.4-25. Here $\theta = 0^o$ at $f = 0$ (dc), -90^o at $f = \infty$ (infinity), and -45^o at $f = f_0$. Comparing with Fig.4-21 for the high-pass filter we can see that it is just the

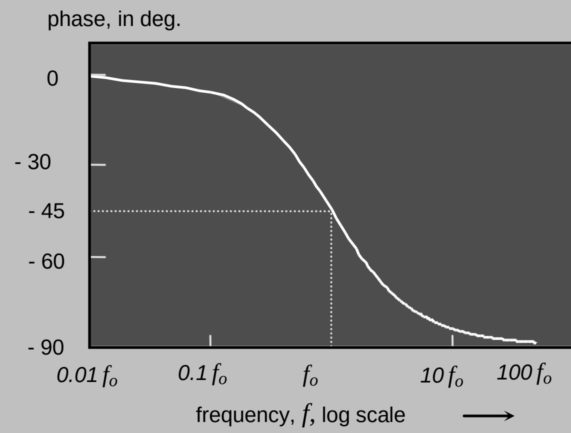


Fig.4-25: Frequency response of phase difference between input and output voltages for an RC low-pass filter

reverse. However, we can note a similarity between voltage gain and the phase difference, and can say that,

at low frequencies where the voltage gain is nearly 1, i.e., the input is almost fully allowed to pass, the phase difference is almost 0^o*, while at high frequencies where voltage gain is nearly zero, the phase difference is about* -90°.

4.15 Combination of high and low pass filters

Bandpass, narrowband and notch filters

We can get different types of composite frequency responses when we combine a high pass and a low pass filter with different cut-off frequencies in cascade (with appropriate buffers in between). Some of these responses are shown in Fig.4-26. If the cut-off frequency for the low pass filter is much greater than that for the high pass filter, we get a wide bandpass filter as in Fig.4-26a. Here we rename the cut-off frequencies as upper cut-off f_u and lower cut-off f_l as indicated in the figure. Therefore for a wide bandpass filter, $f_u >> f_l$. A new term, the bandwidth (BW) is defined as,

$$\text{BW} = f_u - f_l \qquad \text{... 4.62}$$

which signifies the range of frequencies allowed by a bandpass filter. Such filters are inherently built into typical ac amplifiers, or are intentionally introduced to achieve a particular characteristics. So the bandwidth gives an idea of the quality of an amplifier, what range of frequencies it can amplify. A good High Fidelity (Hi-Fi) audio amplifier should have a bandwidth of about 20kHz (20kHz – 20Hz ≅ 20kHz). In low cost audio amplifiers (as those used in radios and cassette recorders), the bandwidth may be less than this. In such low cost audio equipment a Tone Control is usually provided. This basically shifts the upper cut-off frequency f_u up and down. Normally we keep f_u at a high value. If there is a hissing noise (which has a high frequency content) we simply decrease f_u to eliminate the hissing noise. We lose some signal too, but a noise free

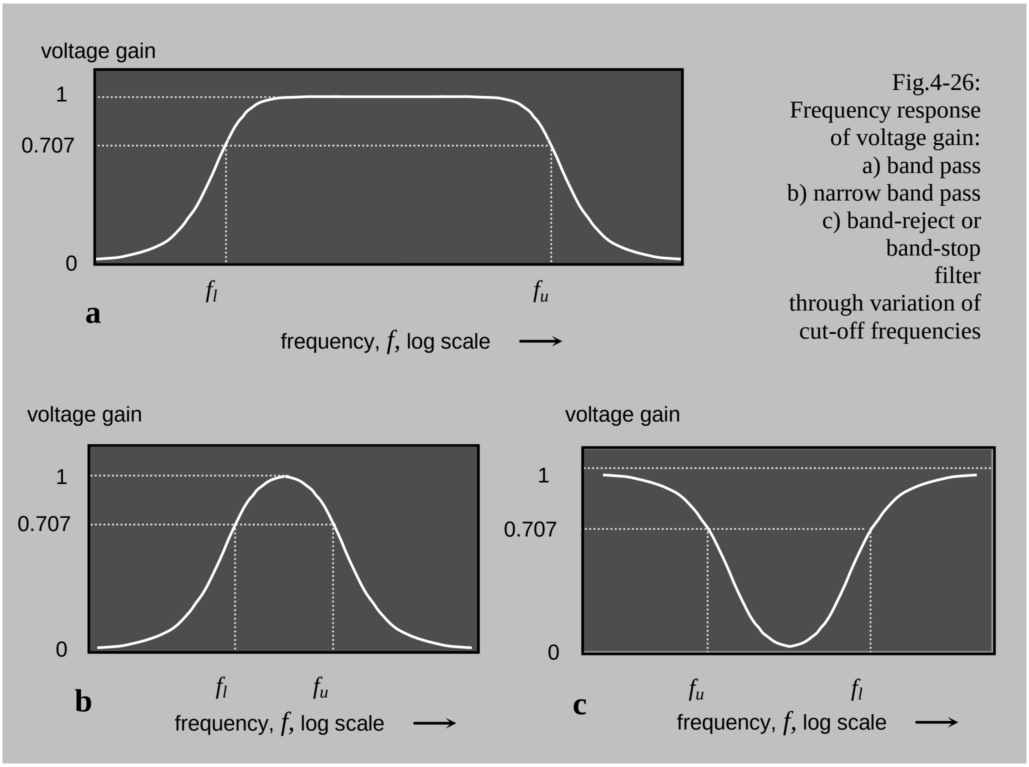


Fig.4-26: Frequency response of voltage gain: a) band pass b) narrow band pass c) band-reject or band-stop filter through variation of cut-off frequencies

signal of slightly lower quality is better than a noisy signal. This control is achieved by manipulating a variable resistor which forms the resistance element of an *RC* low pass filter, Some amplifiers have separate BASS and TREBLE tone controls. BASS refers to low frequency sound signals and one adjusts f_l of a high pass filter for BASS control, while TREBLE refers to high frequency sound signals and one adjusts f_u for TREBLE control. Bandwidth is an important parameter for electronic amplifiers. For television, the video amplifiers need a bandwidth of about 6MHz, much larger than that of an audio amplifier.

If we decrease the bandwidth, we can achieve a *narrow band pass filter* as shown in Fig.4-26b. This amplifies a narrow band of frequencies only and rejects all others. However, to make this bandwidth very narrow, when we call it a *tuned filter*, we usually go for a resonant LCR circuit to be described later in this chapter. Again if we move the cut-off frequencies such that f_u falls below f_l as shown in Fig.4-26c, it is called a *band-reject* or *band stop* filter. If the rejection is very sharp then we call it a *notch filter*. Such filters are necessary to reject noise of specific frequencies that fall within a signal range. A typical example is the 50Hz noise from the mains electricity which causes severe interference in medical signals like ECG of the heart, EMG of muscles, and EEG of the brain where the signal frequency contents spread to both sides of 50Hz. Such notch filters are often used to get rid of the interfering 50Hz.

4.16 Series LCR circuit

Let us consider the *LCR* circuit shown in Fig.4-27. It is called a *series LCR* circuit as all the elements are in a series loop with an input voltage source v_{in} having the same current i passing through them.

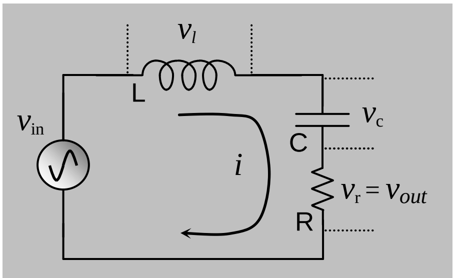


Fig.4-27: A series LCR circuit

The total reactance in the circuit is given by,

$$\boldsymbol{X} = \boldsymbol{X_L} + \boldsymbol{X_C} = j(X_L - X_C) \qquad \text{... 4.63}$$

where the –ve sign for X_C appears as it is 180^0 opposite to X_L in phase, as discussed before.

Therefore the total impedance in this circuit is,

$$\boldsymbol{Z} = R + \boldsymbol{X} = R + \boldsymbol{X_L} + \boldsymbol{X_C}$$

$$= R + j(X_L - X_C) \qquad \text{... 4.64a}$$

Note that the imaginary term can be either +ve or –ve, depending on the relative magnitudes of X_L and X_C. The magnitude of impedance is,

$$Z = \sqrt{R^2 + (X_L - X_C)^2} \qquad \text{or,}$$

$$Z = \sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} \qquad \text{... 4.64b}$$

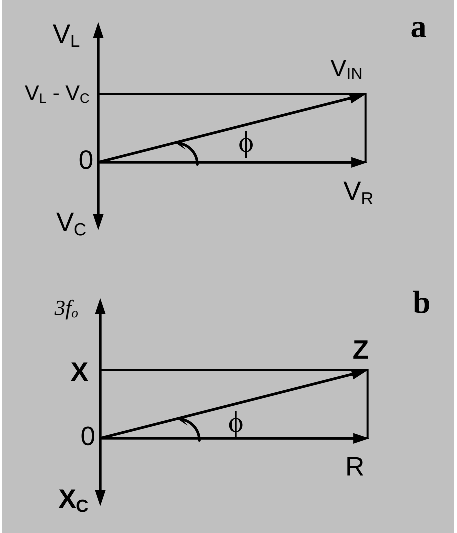


Fig.4-28: Voltage (a) and Impedance (b) Phasor diagrams of a series LCR circuit

We can use the voltage phasor diagram of Fig.4-28a and the complex impedance phasor diagram of Fig.4-28b to represent the LCR circuit. From Fig.4-28b, the total impedance, $\boldsymbol{Z}$ can be determined in terms of R, $\boldsymbol{X_L}$ and $\boldsymbol{X_C}$ which is the same as that obtained from Eq.4.64 a & b above. Let us draw the net reactance phasor $\boldsymbol{X}$ (Eq.4.63) on the vertical axis. For a general treatment, to keep the expressions outwardly positive, let us assume $X_L > X_C$ and draw the phasors accordingly. The total impedance $\boldsymbol{Z}$ (= $Z\, e^{\,j\phi}$) will be given, both in magnitude and phase, by the diagonal of the rectangle formed by the phasors R and $\boldsymbol{X}$. If $X_L < X_C$, then –ve signs will appear at appropriate places during analysis and the phasor of the total reactance will point downwards in Fig.4-28b. The phase angles of the corresponding voltages are shown in Fig.4-28a.

Applying Ohm's law for complex numbers, we get,

$$\boldsymbol{I} = \frac{\boldsymbol{V}_{IN}}{\boldsymbol{Z}} = \frac{\boldsymbol{V}_{IN}}{R + \boldsymbol{X}} \qquad \text{... 4.65}$$

However, as before we are interested in the voltage across the resistor and name this as the output ($\boldsymbol{V}_{\boldsymbol{OUT}} = \boldsymbol{I}\, R$) with respect to the ground. Therefore, the *voltage gain*, $\boldsymbol{A}_{\boldsymbol{V}}$ is given by,

$$\boldsymbol{A}_V = \frac{\boldsymbol{V}_{OUT}}{\boldsymbol{V}_{IN}} = \frac{\boldsymbol{I}R}{\boldsymbol{I}(R + \boldsymbol{X})} = \frac{R}{R + j(X_L - X_C)} \qquad \text{... 4.66}$$

Using $X_L = j\omega L$ and $X_C = -j / \omega C$, we can write Eq.4.66 as,

$$\boldsymbol{A}_V = \frac{R}{R + j(\omega L - \frac{1}{\omega C})} \qquad \text{... 4.67}$$

The magnitude of gain A_V is given by,

$$\boxed{A_V = \frac{R}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}} \qquad \text{... 4.68}$$

and the phase angle by,

$$\boxed{\phi = \tan^{-1} \frac{(\omega L - \frac{1}{\omega C})}{R}} \qquad \text{... 4.69}$$

4.16.1 Frequency Response, Resonance

Looking at Eqs.4-64, 4.68 and 4.69, we can see that if $\omega L > 1/\omega C$ the inductor dominates the impedance, and the phase ϕ is positive. On the other hand if $\omega L < 1/\omega C$ the capacitor dominates the impedance, and the phase ϕ is negative. An interesting case occurs for the frequency at which $\omega L = 1/\omega C$. Let us call this angular frequency as ω_o and the corresponding frequency as f_o. Here the value of impedance Z is minimum and is simply R, while the gain A_V is maximum at unity (since the term within bracket in the denominator is zero), and the phase angle ϕ is zero. For any other frequency, the denominator of Eq.4.68 will be greater, reducing the voltage gain. The LCR circuit is said to go through a *'Resonance'* at this special frequency. The meaning of resonance will be elaborated later. From the above discussion we can see that this resonance occurs when,

$$\omega_0 L = \frac{1}{\omega_0 C},$$

i.e., at an angular frequency,

$$\boxed{\omega_0 = \frac{1}{\sqrt{LC}}} \qquad \text{... 4.70}$$

The corresponding Resonance frequency being,

$$\boxed{f_0 = \frac{1}{2\pi\sqrt{LC}}} \qquad \text{... 4.71}$$

If we plot Eqs.4.64, 4.68 and 4.69 against frequency, giving the abcissa values as multiples of the resonance frequency f_o, we shall get the respective frequency responses as shown in Figs.4-29 a, b &c respectively .

The above mentioned features can be clearly seen in these diagrams. We can see in Fig.4-29a that below the resonant frequency the impedance decreases almost exponentially with increasing frequency – a capacitive behaviour – that we saw in Fig.4-7 before. On the other hand, above the resonant frequency, the impedance increases almost linearly with increasing frequency – an inductive behaviour. The impedance is a minimum at resonance. From Eq.4.64 it can be seen that the impedance is purely resistive (equal to R) at resonance. What happens is that the voltages across the

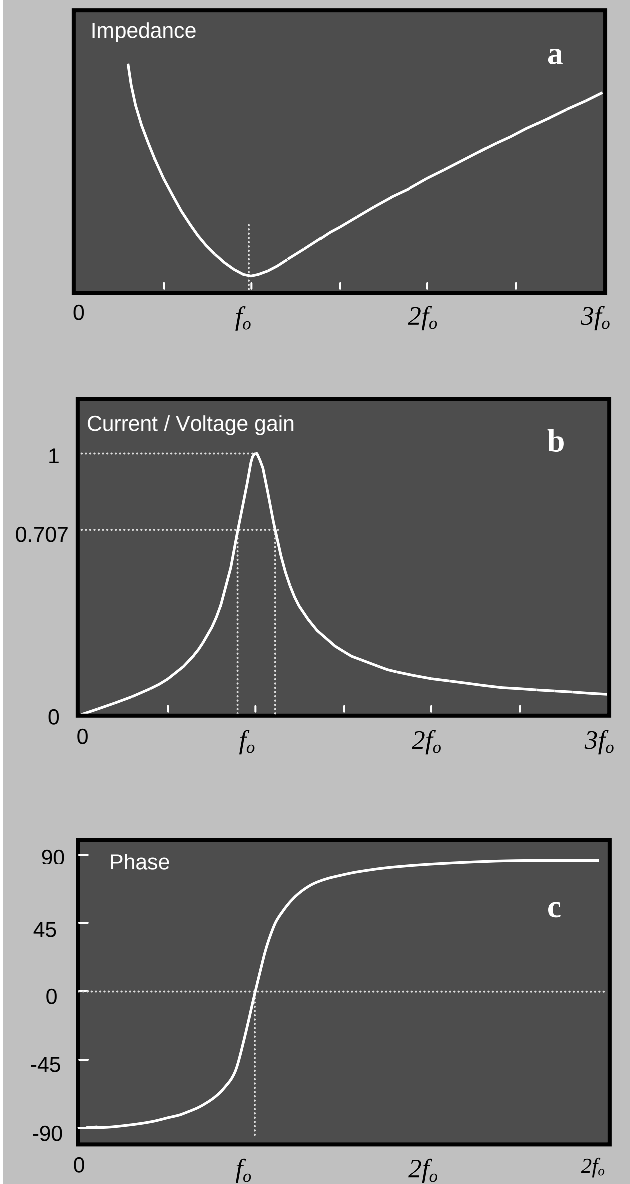


Fig.4-29: Frequency response of series LCR circuit parameters. a) Total impedance, b) Current or Voltage gain for output taken across resistor, c) Phase difference between output and input voltages.

capacitance and the inductance are exactly equal in magnitude but are in opposite phase, therefore they cancel each other completely. This can also be appreciated from the impedance phasor diagram Fig.4-28b where $X_L - X_C$ is zero.

Therefore we get the maximum current, and consequently the maximum voltage across the series resistor at resonance which is also evident in Fig.4-29b. The gain falls steeply away from f_o on both sides. This is a typical resonance pattern. Note that the gain curve is not symmetric around f_o. This is because of the different behaviour of the capacitive and inductive reactances which dominate the two sides as discussed above (on one side e^f behaviour, while $1/f$ behaviour on the other).

The phase response shown in Fig.4-29c is also interesting. The phase is positive for $f > f_o$, an inductive behaviour, while the phase is negative for $f < f_o$, a capacitive behaviour, and at resonance, the phase is zero. The circuit behaviour is simply resistive at resonance, as also observed before. At the two extremes and at resonance, we have,

at $f=+\propto$ *(infinity),* $\phi = +90^0$ *(totally inductive)*

at $f=f_0$, $\phi = 0$ *(totally Resistive)*

at $f=-\propto$ *(infinity),* $\phi = -90^0$ *(totally capacitive)*

4.16.2 What is Resonance?

What is Resonance? Resonance involves two objects one of which attempts to transfer a periodically oscillating energy to the other. Here, these two objects are the signal generator, and the LCR circuit respectively. The signal generator applies a periodic oscillatory potential on the LCR circuit. The LCR circuit has its own natural frequency of oscillation given by Eq.4.71 which was also obtained in the previous chapter while discussing dc transients. When the frequency of the periodic driving potential exactly equals the natural frequency of the LCR circuit, there is resonance, and we get a maximum voltage gain.

(see the previous chapter on dc transients and basic books on Physics, e.g., by Halliday and Resnick).

4.16.3 Cut-off frequencies, Bandwidth

(To keep the expressions simple, in the following treatment we will often refer to angular frequency as simply frequency, and bandwidth will refer to this angular frequency as well. Simple frequency and angular frequency terms will be used interchangeably and you have to understand the expression as relevant.)

To get the upper and lower cut-off frequencies of the resonance curve shown in Fig.4-29b, we use Eq.4.68. Following previous arguments, we can see that

$$A_V = \frac{1}{\sqrt{2}} = 0.707 \qquad \text{when} \qquad \left(\omega L - \frac{1}{\omega C}\right) = \pm R \qquad \text{... 4.72}$$

i.e., the voltage gain will become 0.707 times the maximum gain when the total reactance equals total resistance in the circuit. So the cut-off frequencies will occur at frequencies given by the solutions of Eq.4.72 and as shown by f_u and f_l in Fig.4-29b.

We have to note that $\omega L > (1/\omega C)$ for the upper cut-off frequency which is above the resonance frequency. So we have to use the + ve sign in Eq.4.72. Thus for the upper cut-off angular frequency ω_u, we get,

$$\omega_u^2 LC - \omega_u CR - 1 = 0 \qquad \text{... 4.73}$$

wherefrom,

$$\omega_u = \frac{CR \pm \sqrt{C^2R^2 + 4LC}}{2LC} \qquad \text{... 4.74a}$$

Since the square root term on the numerator is always greater than CR, using the –ve sign will result in a negative frequency which is not physically valid. Therefore we use the + ve sign only in Eq.4.74a and the upper cut-off frequency is given by,

$$\omega_u = \frac{CR + \sqrt{C^2R^2 + 4LC}}{2LC} \qquad \text{... 4.74b}$$

The lower cut-off frequency is below the resonance frequency, therefore, $\omega L < (1/\omega C)$ and we have to use the – ve sign in Eq.4.72. Thus for the lower cut-off frequency ω_l, we get,

$$\omega_l^2 LC + \omega_l CR - 1 = 0 \qquad \text{... 4.75}$$

wherefrom,

$$\omega_l = \frac{-CR \pm \sqrt{C^2R^2 + 4LC}}{2LC} \qquad \text{... 4.76a}$$

Since the square root term on the numerator is always greater than CR, using the –ve sign will result in a negative frequency which is not physically valid. Therefore we use the + ve sign only in Eq.4.76a and the lower cut-off frequency is given by,

$$\omega_l = \frac{-CR + \sqrt{C^2R^2 + 4LC}}{2LC} \qquad \text{... 4.76b}$$

Now the bandwidth (BW) is given by the subtraction of Eq.4.76b from Eq.4.74b as,

$$\text{BW} = (\omega_u - \omega_l) = \frac{CR}{LC} = \omega_0^2 CR \qquad \text{... 4.77a}$$

where Eq.4.70 has been used to get the last form of the expression. The above is in terms of the angular frequency. In terms of simple frequency,

$$\text{BW} = \omega_0^2 CR/2\pi \qquad \text{... 4.77b}$$

4.16.3 Q-factor

The *Quality factor*, or *Q-factor* of a resonance curve describes how steep the curve is at resonance and is defined as the ratio of its resonance frequency to bandwidth,

$$\boxed{Q = \frac{\omega_0}{\omega_u - \omega_l}} \quad \text{... 4.78a}$$

Taking the common factor 2π out of the above, the Q factor is also expressed as,

$$\boxed{Q = \frac{f_o}{f_u - f_l}} \quad \text{... 4.78b}$$

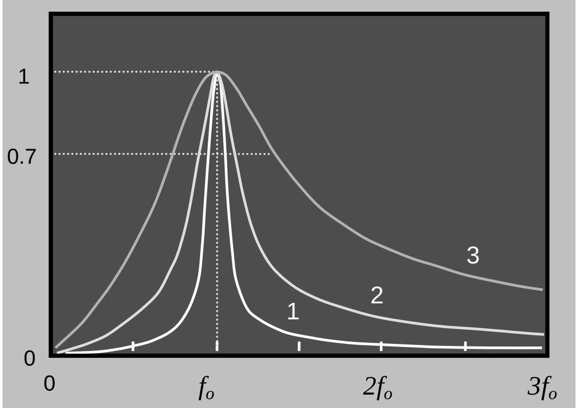


Fig.4-30: Resonance curves (current) of varying Q-factor for a series LCR circuit. $Q_1 > Q_2 > Q_3$

Looking at Fig.4-30, note that, steeper the curve, smaller is the bandwidth. This means that *Q* is large if the bandwidth is small and *vice versa*. In Fig.4-30, $Q_1 > Q_2 > Q_3$, and obviously curve 1 is the steepest. We use a resonant circuit mostly to select an individual frequency and to eliminate others. Therefore we look for as large a *Q*-factor as possible. Practically, Q >10 gives a reasonably acceptable value but in special applications much higher Q-values are required.

For the series resonant LCR circuit, we can express the Q-factor in one of the several ways given below, which use Eqs.4.70, 4.77 and 4.78,

$$Q = \frac{1}{\omega_0 CR} = \frac{1/\omega_0 C}{R} = \frac{\omega_0 L}{R} \quad \text{... 4.79}$$

This basically says that the

$$\boxed{Q - factor = \frac{Reactance\ of\ either\ L\ or\ C\ at\ Resonance}{Resistance}}$$

which is a good thing to remember.

4.17 Series-Parallel LCR circuit

Let us consider the series-parallel *LCR* circuit shown in Fig.4-31. In this arrangement the *L* and *C* are in parallel and *R* is in series. *R* could be the internal resistance of the input source. The input voltage source v_{in}, the main loop current *i*, and the voltages v_x across the parallel *LC* section and v_r across the resistor *R* are shown. The potential and the complex impedance phasor diagrams are shown in Fig.4-32 a&b.

The reactance of the parallel *LC* section is given by,

$$\boldsymbol{X} = \left(\frac{1}{\boldsymbol{X}_L} + \frac{1}{\boldsymbol{X}_C} \right)^{-1} = \left(\frac{1}{jX_L} - \frac{1}{jX_C} \right)^{-1} = j\left(\frac{1}{\omega L} - \omega C \right)^{-1}$$

$$\text{or, } \mathbf{X} = j\frac{\omega L}{1-\omega^2 LC} \qquad \text{... 4.80}$$

We can see that when, $1-\omega^2 LC = 0$, the reactance becomes infinity. This condition occurs at a frequency given by,

$$\omega_0 = \frac{1}{\sqrt{LC}} \qquad \text{... 4.81a}$$

or,

$$f_0 = \frac{1}{2\pi\sqrt{LC}} \qquad \text{... 4.81b}$$

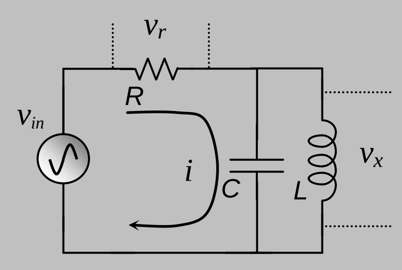


Fig.4-31: A series- parallel LCR circuit

We can also see that the reactance can be either + ve or – ve depending on the frequency. From 4.80 we can specifically see that the reactance will be +ve and inductive in nature for $\omega < \omega_o$ (when $\omega^2 LC < 1$). The reactance will be –ve and capacitive in nature for $\omega > \omega_o$ (when $\omega^2 LC > 1$), and zero for $\omega = \omega_o$ (when $\omega^2 LC = 1$). These are just the reverse of the qualitative conditions of the series LCR circuit discussed before. The reactance phasor ***X*** is shown in Fig.4-32b.

The total impedance is,

$$\boldsymbol{Z} = R + \boldsymbol{X} = R + j\frac{\omega L}{1-\omega^2 LC} \qquad \text{... 4.82}$$

The complex loop current will be given by,

$$\boldsymbol{I} = \frac{\boldsymbol{V}_{IN}}{\boldsymbol{Z}} = \frac{\boldsymbol{V}_{IN}(1-\omega^2 LC)}{R(1-\omega^2 LC) + j\omega L} \qquad \text{... 4.83}$$

If we take the voltage v_x across the parallel *LC* section, the voltage gain would be,

$$\boldsymbol{A}_V = \frac{\boldsymbol{X}}{\boldsymbol{Z}} = \frac{j\omega L}{R(1-\omega^2 LC) + j\omega L} \qquad \text{... 4.84}$$

and if we take the voltage v_r across the resistor *R*, the voltage gain would be,

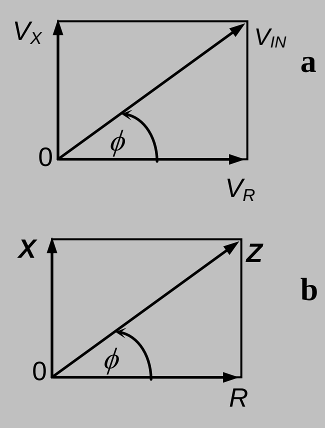


Fig.4-32: Voltage (a) and Impedance (b) Phasor diagrams for a series LCR circuit

$$A'_V = \frac{R}{\mathbf{Z}} = \frac{R(1-\omega^2 LC)}{R(1-\omega^2 LC) + j\omega L}$$

... 4.85

The phase angle ϕ between v_r and v_{in} is given by,

$$\phi = tan^{-1}\frac{X}{R} = tan^{-1}\frac{\omega L}{R(1-\omega^2 LC)}$$

... 4.86

From the above equations we can determine the respective magnitudes of the impedance and the two voltage gains as,

$$Z = \sqrt{R^2 + \frac{\omega^2 L^2}{(1-\omega^2 LC)^2}} \qquad \text{... 4.87}$$

$$A_V = \frac{\omega L}{\sqrt{R^2(1-\omega^2 LC)^2 + \omega^2 L^2}}$$

(across LC) ... 4.88

$$A'_V = \frac{R\sqrt{(1-\omega^2 LC)^2}}{\sqrt{R^2(1-\omega^2 LC)^2 + \omega^2 L^2}}$$

(across R) ... 4.89

Plots of the above equations will demonstrate the essential features of these solutions as given in Fig.4-33 a, b & c.

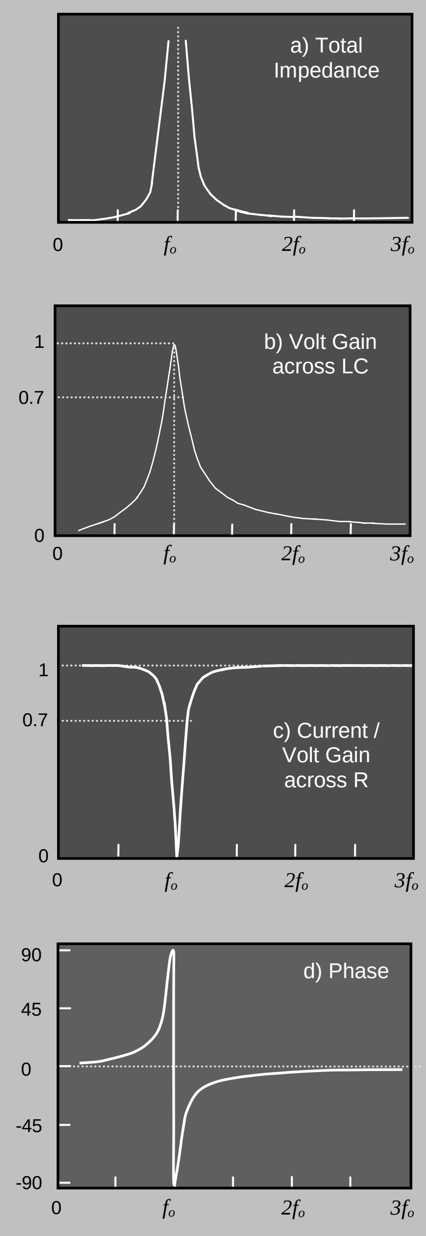


Fig.4-33: Series-Parallel LCR circuit:
a) Total impedance, b) Voltage gain across LC, c) Current / Voltage gain across R, d) Phase diff. between output across R and input.

We can see from Eqs.4.81 and 4.87 that the magnitude of the impedance would be infinity at ω_o. This behaviour is shown in Fig.4-33a where the top of the curve closing in from the two sides is left open as it tends to infinity at f_o corresponding to ω_o. At frequencies away from f_o, the impedance magnitude will be finite and decreasing bothways.

Looking at the voltage gain curve with

output taken across the *LC* section (A_v, Fig.4-33b) we can see that the voltage gain is a maximum of unity at this frequency and decreases on both sides of this frequency. This is evident from Eq.4-88 where the term $(1-\omega^2 LC)$ in the denominator becomes zero at f_o and the resulting value becomes unity. Physically, at f_o the impedance of the *LC* section is infinity, so all the voltage of the input source is dropped across this *LC* section, and the gain therefore, is unity. So here we have a phenomenon of *Resonance* as well with f_0 as the resonance frequency.

If we look at the voltage dropped across the resistor *R* instead of the *LC* section we will see a completely reversed picture. This voltage gain A_v' given by Eq.4.89 and demonstrated in Fig.4-33c is zero at resonance f_o and it increases away from f_o on both sides. We can appreciate that this voltage should have the same behaviour as the main loop current. Since the impedance of the *LC* section is infinity at f_o , the current is zero and the gain A_v' is zero too. Because of this inverse behaviour, when the output is taken across the resistor, it is sometimes called an *Anti-resonant* circuit.

The phase ϕ between the voltage v_r dropped across the resistor and the input v_{in} and is plotted in Fig.4-33d. The phase angle is positive below f_o and rises from zero gradually to +90°. It then goes through an abrupt change from +90° to –90° at resonance. Beyond f_o the phase angle again rises gradually to zero.

The cut-off frequency, bandwidth and the Q-factor can be obtained following methods carried out before for the series LCR circuit. However, as these will be somewhat more complex we do not attempt this here, but you should give it a try.

4.17.1 What happens at resonance?

What happens at resonance is that once the inductance and the capacitance get the necessary tick (pulsed energy) to get going, they sequentially store and release the total energy between themselves periodically in a sinusoidal manner. Ideally no extra energy is needed from outside to continue this activity. Therefore to the outside world (i.e., to source v_{in}) the impedance of this parallel *LC* section is infinity and no current flows in the main loop ($i = 0$). Because of the sequential storage and release of energy between *L* and *C*, there will be a sinusoidal current within the closed loop formed by these two elements (i' in Fig. 4-34), but there will be no current in the outer circuit. In practice, the inductor and the capacitance both will have some internal resistance which will dissipate energy, and therefore, there will be a finite but very high value of impedance. Correspondingly, the main loop current will be somewhat greater than zero at resonance.

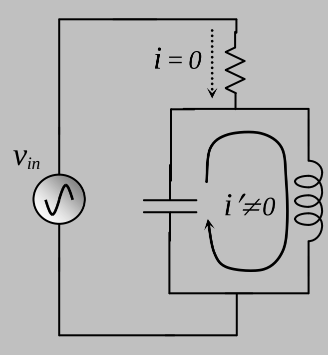


Fig.4-34: A series-parallel LCR circuit at resonance

4.17.2 Application in radio station tuning

The parallel *LC* circuit is used widely in frequency selective networks, tuned amplifiers and oscillators, particularly at high frequencies (radio frequencies). Fig.4-35 shows the basic scheme of a simple radio receiver employing a parallel *LC* circuit as its Tuning circuit. The capacitor of the *LC* tuning circuit is variable and can be manipulated by the user to vary its resonance frequency. In parallel to the tuning circuit is connected a radio receiver circuit which has the necessary capability to extract desired audio signals from the radio waves and to generate sound using a loudspeaker.

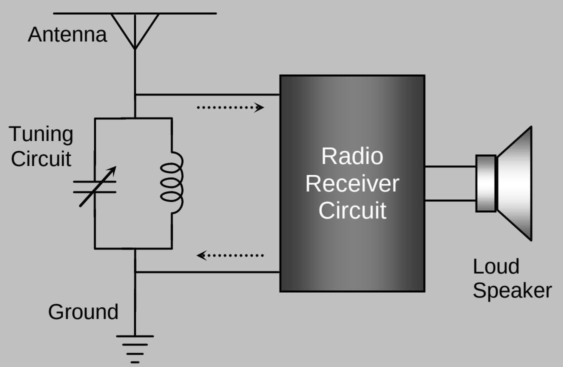


Fig.4-35: Use of a parallel LC circuit as a tuner in radio receivers

A radio wave creates a potential difference between the antenna and the ground due to which an alternating current tries to flow between the antenna and the ground. If the resonance frequency of the *LC* circuit is not the same as the incoming radio waves it allows the radio signal to go directly to the ground as its impedance is low to such frequencies. On the other hand if the resonance frequency of the *LC* circuit is exactly the same as the incoming radio signal then it blocks the signal from going directly to the ground because the *LC* circuit has almost infinite impedance at this frequency. Rather it diverts the radio signal through the parallel radio receiver circuit where the desired audio signal is extracted and we hear the audio signal. Here the radio receiver circuit offers lower impedance than the *LC* tuning circuit, and therefore, the radio signal goes from the antenna to the ground through the radio receiver circuit.

At any time we have radio signals from many stations producing emf's of different frequencies in the antenna-ground combination. To choose a particular radio station having a particular broadcasting frequency, the capacitor is adjusted so that the *LC* circuit resonates at the desired frequency. Signals of other frequencies from other radio stations are allowed to flow down to earth directly since the *LC* circuit offers negligible resistance to these signals. On the other hand the *LC* circuit because of its high impedance at the chosen frequency blocks this signal, and the signal is made to pass through the parallel radio receiver circuitry. In this way we isolate a single radio signal from many others. Just by changing the capacitance we can change the resonant frequency of the tuning circuit and choose another radio station.

4.18 Transformer, transferring ac power

The phenomenon of mutual inductance involving two coils discussed before is utilised in a popular device called a transformer whose symbol, together with a source of an alternating *emf* e_p and a load Z_L are shown in Fig.4-36. Here the two coils *P* and *S* are intimately linked through their magnetic flux. The ac source e_p applies power to the *P* side, which is transferred to a load impedance Z_L on the *S* side simply through the intermediary of the magnetic flux linkage between the coils. There is no direct electrical connection between the source of *emf* and the load. The coils may be magnetically coupled just by their physical proximity, and there may or may not be a core of magnetic material linking them. The symbol shown is for a transformer with a magnetic core, indicated by the two thick bars. An air core transformer, as shown in Fig.4.37 has no bars in the symbol.

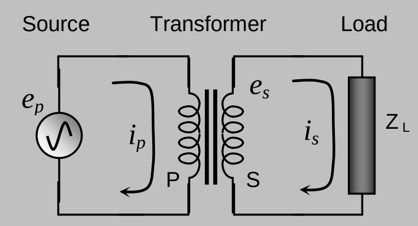


Fig.4-36: A transformer connected to a source and a load

Following Eq.3.12 in the previous chapter, the induced *emf* e_s in coil *S* would depend on the rate of change of current in coil *P*. Therefore, if an *emf* source with a varying voltage drives the coil *P*, an induced *emf* will be generated in coil *S* which can drive a current through a load Z_L in turn. It is evident from the above discussion that unless the current through the coil *P* changes with time there cannot be any *emf* in the coil *S*. Therefore the driving voltage cannot be a smooth dc. It should be a varying dc, or ac.

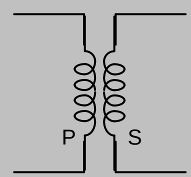


Fig.4-37: Symbol of an air-core transformer

4.18.1 Primary and secondary coils

Though the two transformer coils can easily be reversed in function, it is usual to drive a specified coil from a source, and to drive the load using the other coil. This has given rise to the names, *Primary* and *Secondary* coils, primary being the one driven by the source, while the one driving the load is the secondary. According to this scheme, the *P* and *S* coils in Figs.4-36 and 4-37 stand for the Primary and the Secondary coils respectively.

4.18.2 Analysis, ideal transformer

Here we analyse an ideal transformer. This assumes that the coils have no resistance, self inductance or capacitance; they only have mutual inductance, and the primary coil can take in any amount of current as demanded by a combination of source *emf*, the transformer and the load. That is, the transformer can supply any amount of power to the load as demanded. All the magnetic flux produced by the primary are linked to the secondary, meaning that there is no leakage flux from the primary that does not link the secondary coil. The core material, if any, also does not consume any power and is never saturated. The driving *emf* is an ac signal.

For analysis let us choose the following symbols where the subscripts p and s correspond to the respective sides:

N_p, N_s : *number of turns in respective coil*
e_p, e_s : *emf, driven or induced*
ϕ_p, ϕ_s : *magnetic flux in the two coils, should be equal for a transformer*
i_p, i_s : *current in the respective circuits*
P_p, P_s : *power in the respective circuits*

Now, from Faraday's laws of electromagnetism,

$$e_p = N_p \frac{d\phi_p}{dt}, \quad or, \quad \frac{d\phi_p}{dt} = \frac{e_p}{N_p}$$

and

$$e_s = N_s \frac{d\phi_s}{dt}, \quad or, \quad \frac{d\phi_s}{dt} = \frac{e_s}{N_s}$$

Since the two coils are intimately connected through their magnetic flux (assuming no outside leakage), therefore,

$$\frac{d\phi_p}{dt} = \frac{d\phi_s}{dt}, \text{ wherefrom, } \quad \frac{e_p}{N_p} = \frac{e_s}{N_s}.$$

This gives us the famous relationship between the voltages and the number of turns,

$$\boxed{\frac{e_p}{e_s} = \frac{N_p}{N_s}} \qquad \text{... 4.90}$$

Again, an ideal transformer does not consume any power, nor it provides any extra power. Therefore power delivered to the primary coil is equal to the power in the secondary circuit (power conservation law). This gives,

$$e_p i_p = e_s i_s \qquad \text{... 4.91}$$

from which we get,

$$\boxed{\frac{e_p}{e_s} = \frac{i_s}{i_p}} \qquad \text{... 4.92}$$

Again by combining the above two equations, we get,

$$\boxed{\frac{N_p}{N_s} = \frac{i_s}{i_p}} \qquad \text{... 4.93}$$

The above equations (Eqs.4.90 to 4.93) describe an ideal transformer function adequately. From Eq.4.90 we can see that the voltages at the two sides depend on their number of turns. By making N_s greater or less than N_p we can respectively increase or decrease an ac voltage. If $N_s > N_p$ we call it a *step-up transformer* as the voltage is increased in the secondary. Similarly if $N_s < N_p$ we call it a *step-down transformer*. Again, Eqs.4.91 and 4.92 indicate that for a fixed input power, if the voltage at the secondary is increased, the secondary current will decrease in the same proportion. The reverse is also true, i.e., a higher current in secondary is possible by decreasing the voltage.

4.18.3 Who determines the current?

From the above discussions there is a tendency to think that whatever i_p is applied is transformed according to Eq.4.92 to the secondary for a given transformer. However, this is not the right approach to think. For a given transformer and a given e_p we cannot control the current from the primary side. It is e_s and Z_L which determine i_s $(= e_s /Z_L)$ first. Then i_p is determined using Eq.4.92. Therefore, for a given e_p, i_p depends on i_s, not the reverse. On the other hand, for a given transformer if we have a freedom to choose e_p, then we can increase or decrease i_s for a given Z_L, subject to the maximum power that the transformer can handle (a practical transformer has such a power limitation, which is not there in the ideal case)

4.18.4 Reflected impedance

From the induced *emf* e_s in the secondary, the load impedance Z_L takes a current i_s such that,

$$i_s = e_s /Z_L, \qquad \text{or,} \qquad Z_L = e_s /i_s \qquad \text{... 4.94}$$

Looking from the point of view of the applied *emf* e_p at the primary in Fig.4-36, the combination of the transformer and the load impedance Z_L at the secondary produce an effective impedance Z'_p, which we call the reflected impedance of Z_L to the primary side. This is determined by the current i_p that is demanded from e_p. Thus Z'_p, seen by the source e_p is given by, using Eqs.4.90 and 4.93,

$$Z'_p = \frac{e_p}{i_p} = \left(\frac{N_p}{N_s} e_s\right)\left(\frac{N_p}{N_s}\frac{1}{i_s}\right) = \left(\frac{N_p}{N_s}\right)^2 \frac{e_s}{i_s} = \left(\frac{N_p}{N_s}\right)^2 Z_L \qquad \text{... 4.95}$$

which can also be expressed as,
$$Z'_p = \left(\frac{e_p}{e_s}\right)^2 Z_L \qquad \text{... 4.96}$$

The above equations show that the load impedance at the secondary is multiplied by the square of the ratio of the number of turns in the coils, or of the voltages, as shown to become the reflected impedance at the primary side. This reflected impedance can be used to determine the current in t he primary due to a given source *emf*.

Viewed from the secondary side, impedances in the primary side will also be reflected in a similar way.

4.18.5 Practical transformers

An ideal transformer does not consume any power itself; it just conveys the input power to the load. That means it has a 100% efficiency and can handle any amount of current and power. In practice, it is not so because of a host of reasons and some of the important ones are described below. The power consumed by the transformer itself heats it up and this limits the power that it can practically handle. This maximum depends on the materials and sizes of the magnetic core, materials and sizes of the wire making the coils, and the frequency of the ac being transferred. The maximum current that any side can handle also depends on the area of cross section of the wires making the respective coils.

The resistance of each of the coils will cause some voltage drop across it when a current flows, and the product of this voltage and current will heat the coil up.

The core material is usually made up of iron which is an electrical conductor. Therefore it can also act as a secondary coil of the transformer which form closed current loops within the core itself, called eddy currents. These also consume power and heat up the core. To reduce such eddy currents transformers never use solid blocks of iron, rather they use thin sheets of iron each insulated from the next using thin insulating varnish coatings, and arranged suitably with respect to the magnetic flux directions. Such arrangement improves the transformer performance significantly. The iron core has another magnetic property called hysteresis which results in consuming some power while the direction of the current is alternating. To reduce such hysteresis loss, special magnetic materials are used for the cores in transformers. Silicon steel is such a magnetic material which is used extensively in power line transformers.

There will be some leakage flux from the primary which complete their loops outside the magnetic core and do not link the secondary coil. There will be similar leakage fluxes in the secondary too. These will be of no use in the power transfer.

Therefore the output power is expected to be a little less than the input power. In terms of efficiency (output power/input power, described in percentage), an ideal transformer will have an efficiency of 100% while a practical one will have an efficiency less than that. Typical values are around 90%.

Again the transformer core can have various designs. Two popular ones are shown in Fig.4-38. In the top design, the two coils are wound on opposite arms of a rectangular shaped core. Such cores are used in isolating transformers where we want to reduce electrical shock hazards to a minimum. This is achieved because of the physical separation between the two coils; chances of one coil touching the other, even if there is a leak in the insulation, are minimum.

On the other hand this transformer is not very efficient. This is because some magnetic fluxes of the primary form closed loops outside the magnetic core and these do not link the secondary. Therefore there is efficiency loss due to leakage flux.

Leakage flux is minimised in the second design. Here the core has three arms as shown and both the coils are wound on the central arm, with the secondary over the primary. The secondary coil links to most of the magnetic fluxes created by the primary in this design and so it has greater efficiency.

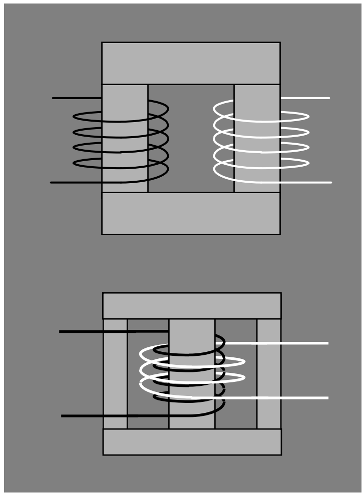


Fig.4-38: Practical transformer designs

4.18.6 Advantages in power distribution

Transformers are extensively used with the ac main line systems to increase or decrease the voltage at will, which has many advantages. One of the main advantages is that the power can be transmitted and distributed over long distances with little energy loss if the voltage is kept high. This is because for the same power, current is low, therefore i^2R loss in wire is low. Typical RMS voltage standards for transmission over long distances are: 230,000V and 132,000V, and for distribution over shorter distances within cities and towns are: 33,000V and 11,000V. At the terminal user point, typically 220V RMS for single phase connection is derived from 380V phase to phase voltage in a 3-phase system (see any book on electrical engineering for details of a 3-phase power distribution system). Some countries choose 110V RMS instead of 220V at the user end. The choice is based on a compromise between current carrying capacity of wires of practicable dimensions and human safety considerations (higher the voltage, thinner the wire, and higher the hazard). The frequency chosen is 50Hz in some countries while it is 60Hz in some others. If the frequency were higher, losses due to inter-wire capacitance and electromagnetic radiation would increase. On the other hand at lower frequencies, the size of transformers would increase. However, the small difference between 50Hz and 60 Hz is purely historical; human ego and political differences between different countries developing new technologies often result in such variations in standards.

That an ac allows voltage levels to be changed easily using transformers which makes longer transmission distances possible with little heating losses in wires is the primary reason that dc main line systems became outdated long time back.

4.18.7 Transformer cores at different frequencies

At low frequencies a ferromagnetic core (e.g., Silicon steel, in the form of sheets) is used universally to increase the mutual inductance of a transformer. At high frequencies (> few tens of kHz) the magnetic domains of iron cannot rotate so fast, therefore, special composite ferrite materials are used. At still higher frequencies(> few MHz) ferrites also cannot follow the changes so no extra core material is used, and these are called air-core transformers. You will see that a *Medium wave radio* having a frequency range of around 1MHz uses a *ferrite rod* within its coils (which is essentially a transformer converting the voltage generated by an antenna) while in a *Short wave radio*, nothing is there within the coils meaning that it is an air-core transformer.

4.18.8 Stray capacitances and inductances

Any two conductors placed close together will have a capacitance, providing a path for ac. The windings of a coil will have interwinding capacitance between each pair of its turns, and such capacitance is present in both the primary and the secondary. Since these appear in parallel to the *emf* sources, some current will be lost through these capacitances. There will also be some current directly linking the primary and secondary through the inter-coil capacitance (between the two coils). The leakage flux of the two coils that do not link the other contribute to self inductances of both the coils which will also offer resistance to current (expressed in terms of their reactance) and the combination of *L* and *C* may give rise to oscillations if the right conditions exist. The resistance of the material of the wire used to make the coils also provides an opposition to current as mentioned before.

4.18.9 Transformer Equivalent circuit

An equivalent model of a practical transformer is necessary to include the effects discussed above which will help in analysing circuits involving transformers. This is important particularly at high frequencies when stray capacitances and inductances become important. Such an equivalent circuit is shown in Fig.4-39 where the transformer symbol shown at the middle represents an ideal transformer. All the non-ideal elements have been represented by their effective equivalents outside the transformer so that these can be considered easily in an analysis. Here R_p and R_s in series to the respective coils represent the dc resistances of the windings. L_p and L_s represent the respective lumped (taken together) inductances because of leakage fluxes.

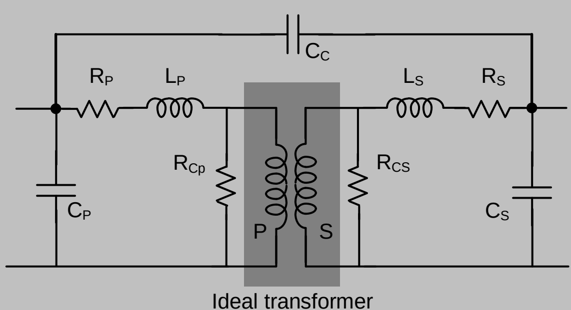


Fig.4-39: Equivalent circuit (model) of a practical transformer

C_p and C_s represent the respective lumped parallel interwinding capacitances while C_c represent the stray coupling capacitance between the two coils. R_c represents the core losses (due to hysteresis and eddy currents) which has been shown as providing an alternative path to input current. There may be some more elements contributing to non-ideal behaviour, however, the above is enough for the scope of this book.

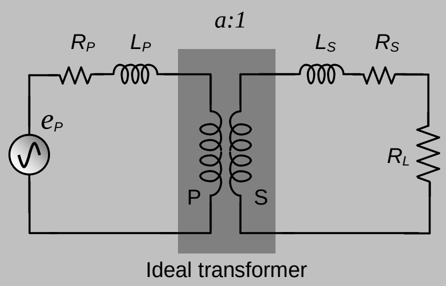


Fig.4-40: Simple Equivalent circuit of a practical transformer at low frequency

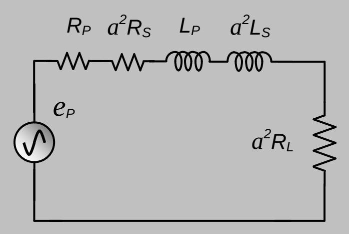


Fig.4-41: Equivalent circuit with all values reflected to the primary side

To analyse currents on any side of the transformer, the elements on the other side can be brought up through reflection as suggested in Eqs.4.95 and 4.96. For simplicity let us assume C_P, C_S and C_C to be very low (which is reasonable at low frequencies), and R_{CP} and R_{CS} to be very high (for a good core material this is also reasonable) so that they may be ignored (assumed open circuit). Then we are left with only the equivalent circuit shown in Fig.4-40, where we have added a source of *emf* at the primary and a load R_L at the secondary. If we replace the ratio N_P /N_s by *a:1*, then the equivalent circuit may be represented as in Fig.4-41 by reflecting L_S and R_S to the primary side following Eqs.4.95 and 4.96. Here the reflected values are a^2R_S, a^2L_S and a^2R_L respectively. We can now combine the two resistances R_P and a^2R_S as R_T, and the two inductances L_P and a^2L_S as L_T to get the final form as shown in Fig.4-42. This circuit can be used to calculate the current through the load for a given source potential, and to calculate the voltage dropped across R_L following techniques developed for an LR circuit earlier.

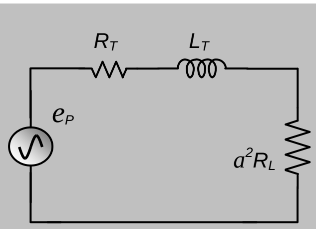


Fig.4-42: Final equivalent circuit

4.18.10 Other applications

Impedance matching and Electrical isolation

Transferring ac power at a different voltage is the most extensive use of a transformer which is also used extensively in providing low voltage dc in electronic circuits from the ac power mains. However, a transformer also provides electrical isolation between two or more circuits, which may be of the major interest in certain applications. Besides impedance-matching of two electronic circuits is also an important application of transformers in electronics which is described first below.

Impedance matching of two circuits.

In the previous chapter (section 2.17) we have seen that when two devices are connected together, a maximum power transfer occurs if the output impedance of the source equals the load impedance. If this condition is not met in a system, we can interpose an appropriately designed transformer to bring about the desired situation, and this procedure is called impedance matching. To see how it works, let us consider Eq.4.96 above which show how the load impedance is reflected to the primary side. This means that for a step down transformer the reflected load impedance at the primary is many times higher, and for a step-up transformer, it is many times lower. This property may be used through suitable design of the transformer. This is further illustrated in the example below.

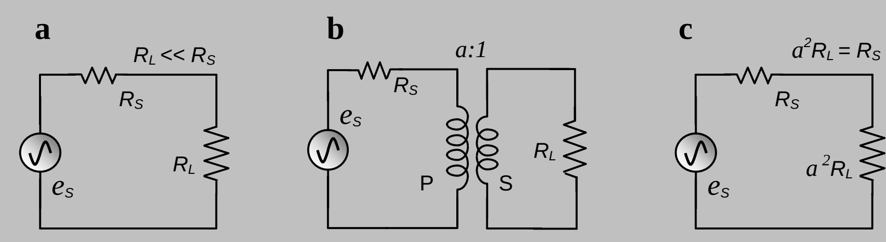


Fig.4-43: a) mismatched source and load impedance, b) Impedance matching using a transformer, and c) equivalent circuit at primary side showing impedance matching

Example 4.1: Suppose we have a source having a high output impedance (Fig.4-43b) R_S = 10,000Ω, which has to transfer power to a load with a much smaller impedance, R_L = 100Ω. We know from section 2.17 that this will not provide a good power transfer. To achieve a good power transfer we have to match the impedances by interposing a suitably designed transformer as shown in Fig.4-43b. This will give rise to the equivalent circuit in Fig.4-43c from which we can see that for maximum power transfer from the source to the load, the reflected load impedance at the primary, a^2R_L needs to be equal to R_S (=10,000Ω). This requires the turn ratio to be

$$a = \sqrt{\frac{10{,}000}{100}} = 10$$

which represents a step down transformer with a turn ratio of 10:1. Thus with the transformer, the effective load resistance is equal to the source resistance and we have maximum power transfer.

We can see the whole thing from a slightly different viewpoint. Considering the transformer itself as a two-port network, its input impedance (primary side) is high while the output impedance is low since it is a step-down transformer. So there is a good power transfer from the source to the transformer. Again at the output there is good power transfer from the low impedance of the transformer to the low load resistance.

4.18.11 Electrical isolation between two or more circuits

In a transformer electrical power is transferred through the intermediary of magnetic flux. There is no direct electrical contact between the primary and the secondary coils of a transformer. Therefore it provides an electrical isolation between the two circuits which is necessary in many situations.

In some electronic circuits we may have two or more separate parts which should not have any common electrical connections for various reasons. A transformer may provide the necessary electrical isolation.

We can obtain a low voltage dc from 220V ac without using a transformer, but that will expose us to life threatening hazards. One of the main uses of a transformer is in saving us from the hazards of getting an electrical shock from the mains 220V line when we touch the low voltage sides of an electronic appliance by isolating the exposed side electrically from the high voltage side. Therefore transformer coils should be adequately insulated from the core and from each other. In normal transformers this is achieved firstly through the use of a chemically insulated wire (called enamelled wire) for the coils, by winding the coil on a plastic bobbin which isolates the coil from the core, and through the use of varnished paper or plastic sheets between the primary and the secondary coils if they are wound one above the other. Finally, the whole assembly is dipped in liquid varnish and dried to produce a safe transformer. Normal plastic covered wires are not used in transformers as the insulations are thick and will allow only a few turns to be made in a particular volume. Instead enamelled copper wire is used to wind the coils in a transformer. Enamel is a very thin insulating coating applied on the wire chemically. This allows more turns to be wound. However, one has to be careful while winding the coils since any sharp bends in the wire will tend to break the insulation.

Even with all the precautions, there will be a capacitance between the primary and the secondary coils, and between the coils and the core. Since a capacitor allows ac to pass there will be some ac passing directly from the primary to the core. Since the core is usually in contact with the metallic cabinet of an appliance, one might feel a slight electric shock when touching exposed metallic parts of the equipment with a bare hand if the metallic body of an appliance is not connected to the *Earth* terminal of our household electrical outlets. (see any engineering book for a description of the mains power line arrangements and the necessity of the *Earth* line in providing safety)

In medical equipment, particularly the ones used in Cardiac (related to heart) clinics and hospitals, even the isolation provided by the normal transformers described above is not adequate. Therefore equipment built with ordinary transformers are not allowed in hospital equipment. For such equipment the transformers are specially designed to reduce the resistive and capacitive leakage current between the primary and the secondary so that no more than 10μA at the line frequency (50 or 60Hz) can pass through the body of the patient in the worst case of the primary touching the mains live wire by accident. This arose from the requirement that in such clinics, a direct electrical connection to the heart muscles may be brought out in some patients, and only 50μA through these links direct to the heart may kill a patient. In normal domestic appliances a higher leakage current (~100μA, as provided by normal transformers) may be allowed since current entering through a limb gets dispersed throughout the body so that very little of it actually passes through the heart.

4.18.12 Transformers in Switch Mode Power Supply

In electronic appliances, dc power is needed to drive the electronic devices, and that mostly at low voltages below 30V. Transformers are extensively used within each appliance to decrease the ac voltage first, and then to convert it to smooth dc electronically. To cater for ups and downs in line voltages, such power supplies are designed with larger voltages than needed and then the additional voltage dropped using series pass elements (usually transistors). This gives rise to lots of wasted power and low efficiency contributing to heating up of the appliances. Besides, at 50 or 60Hz the size of the transformer is large and heavy. Nowadays these are being replaced by much more efficient devices called *switch mode power supplies* that have high efficiencies (more than 90%), and the sizes of transformers are also considerably reduced. Such power supplies convert the mains line ac directly to high voltage dc first. This dc is then used to generate a high frequency (>100KHz) square waves which are then converted to low voltage using transformers. Regulation to cater for ups and downs in line voltages is carried out by modulating the width of the square wave pulses and this results in very high efficiencies. Again transformers at such frequencies can have very reduced sizes for the same power delivery. This we can guess from the reactance ωL of an inductor. Because of large ω, a very small inductor will have a high reactance and a small number of turns are required to make a transformer. The core size is also low at such frequencies.

Such novel techniques and many others depend on the wonders of electronic devices like diodes, transistors, Integrated Circuits, etc., and all these will be our subject of study in the next volumes.

APPENDICES

Appendix-1

Effective and RMS value of an ac, Power Factor

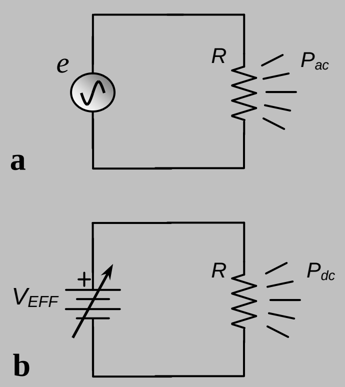


Fig.A1-1: Basis for calculation of the Effective value of an ac

We can easily describe the magnitude of a smooth dc potential by its voltage directly. On the other hand an ac potential is continuously changing in magnitude and in direction. How do we describe such a changing potential? This is done using an effective equivalent dc value that produces the same amount of average power as the ac in a resistance, as explained below. Fig.A1-1a shows an ac source with potential *e* driving a certain average power P_{AC} into a resistor *R*. Now we replace the ac source by a variable smooth dc source which is allowed to drive power into the same resistor *R*. We vary the dc voltage till the power P_{DC} dissipated by the resistor equals the average power P_{AC} dissipated when the ac source was there. The dc voltage that produced this power has the same power producing effect as the ac voltage source. Therefore we term this dc voltage value as the effective voltage value V_{EFF} of the ac. Then V_{EFF} can be used to describe the corresponding ac potential *e*.

Now let us derive a value for V_{EFF} analytically for a sinusoidal ac.

Let the ac potential be given by,

$$e = V_o \, Sin \, \omega t \qquad \text{... A1.1}$$

where V_o is the amplitude or the peak voltage and ω is the angular frequency. Then the instantaneous power p_{ac} delivered into the resistor *R* is,

$$p_{ac} = \frac{e^2}{R} = \frac{V_o^2 \, Sin^2 \, \omega t}{R} \qquad \text{... A1.2}$$

The average power P_{AC} can be obtained by integrating p_{ac} over a full cycle and dividing by the period T as follows.

$$P_{AC} = \frac{1}{T} \int_o^T \frac{V_o^2 \, Sin^2 \, \omega t}{R} \, dt \qquad \text{... A1.3}$$

Evaluating the above we get,

$$P_{AC} = \frac{V_o^2}{T} \int_o^T \frac{Sin^2 \, \omega t}{R} \, dt = \frac{V_o^2}{T} \int_o^T \frac{(1 - Cos 2\omega t)}{2R} \, dt = \frac{V_o^2}{T} \frac{T}{2R} = \frac{V_o^2}{2R}$$

i.e., $$P_{AC} = \frac{V_o^2 / 2}{R} \qquad \text{... A1.4}$$

which is the average power of the ac potential driving the resistor *R*. Note, in the above integration, the integral of the Cosine term is zero.

Now the power P_{DC} delivered to the resistor *R* by the dc voltage V_{EFF} is given by,

$$P_{DC} = \frac{V_{EFF}^2}{R} \qquad \text{... A1.5}$$

which should equal the average ac power given by Eq.A1.4. Doing this we get,

$$V_{EFF}^2 = \frac{V_o^2}{2}, \quad \text{or,} \quad \boxed{V_{EFF} = \frac{V_o}{\sqrt{2}}} \qquad \text{... A1.6}$$

which gives the value of the effective dc voltage in terms of the peak ac voltage V_o.

Evaluating √2 we get, $V_{EFF} = 0.707\, V_o$ ... A.1.6a

Looking at Eq.A1.3 and Eq.A1.5 above we can say that

$$\boxed{V_{EFF} = \sqrt{\frac{1}{T}\int_o^T V_o^2\, Sin^2 \omega t \;\; dt}} \qquad \text{... A1.7}$$

Now, how can we describe the term on the right hand side in Eq.A1.7? It is the square root of the average (or *mean*) of the square of the original ac voltage. Therefore we can call it the *Root Mean Squared* ac voltage which abbreviates to the well known term *RMS*. Therefore, *RMS voltage* of an ac is also its *Effective Voltage*, and the former name has become more popular than the latter. It would have been more meaningful if it were the other way round.

We can also appreciate from the above that the RMS voltage is the effective voltage for any ac or any varying voltage. However, the value $V_o / \sqrt{2}$ is only applicable to a sinusoidal ac. It would not hold for other waveforms.

From the above we can also deduce that the effective value of an ac current would also be given by its RMS value. So we have, for a sinusoidal ac,

$$\boxed{I_{EFF} = \frac{I_o}{\sqrt{2}}} \qquad \text{... A1.8}$$

We again appreciate that Eq.A1.8 holds only for a sinusoidal ac. Evaluating $\sqrt{2}$, we get,

$I_{EFF} = 0.707\, V_o$ ... A.1.8a

In an ac circuit, the voltage and current may have a phase difference ϕ. Corresponding to the sinusoidal voltage of Eq.A1.1 if the current is given by,

$$i = I_o Sin\,(\omega t + \phi) \qquad \text{... A1.9}$$

then we can calculate the instantaneous power as,

$$p_{ac} = e\,i = V_o I_o (Sin\,\omega t)\,[Sin\,(\omega t + \phi)] = V_o I_o (Sin\,\omega t)\,(Sin\,\omega t\,Cos\,\phi + Cos\,\omega t\,Sin\,\phi)$$

or, $\quad p_{ac} = V_o I_o (Sin^2\,\omega t\,Cos\,\phi + Sin\,\omega t\,Cos\,\omega t\,Sin\,\phi)$

$$= V_o I_o [(1 - Cos\,2\omega t)\,Cos\,\phi + Sin\,2\omega t\,Sin\,\phi]\,/\,2$$

$$= V_o I_o [Cos\,\phi - Cos\,2\omega t\,Cos\,\phi + Sin\,2\omega t\,Sin\,\phi]\,/\,2$$

If we calculate the average power by integrating the above equation over a complete period and then divide by T, the terms having $Sin\,2\omega\,t$ and $Cos\,2\omega\,t$ will give zero values and we will be left with (do it yourself),

$$\boxed{P_{AC} = \frac{V_o\,I_o}{2}\,Cos\phi = \frac{V_o}{\sqrt{2}}\,\frac{I_o}{\sqrt{2}}\,Cos\phi = V_{EFF}\;I_{EFF}\;Cos\phi} \qquad \text{... A1.10}$$

This is a very important result which says that the average power is not simply the product of effective voltage and current, there is another factor $Cos\ \phi$ which depends on the phase difference between the current and the voltage. The $Cos\ \phi$ term is called the power factor. Understandably, its maximum value is unity when $\phi = 0$. This happens only when the load is fully resistive such as for a filament lamp or a heater. For any load which has a inductive or capacitive component, the power factor is less than unity. For loads like electric fan, refrigerator where a motor is the main load, it is a mixture of inductive and resistive load and the power factor may be of the order of 0.6. What is the implication of such low values of power factor?

Suppose for a motor operating on an RMS voltage of 220V the current is 5A. This gives a product of 1100 Volt-Amp (*VA*, note: we have not used Watt). If the power factor is 0.6 then the motor will consume a power of 1100x0.6 watts (W), or, 660W. You can possibly appreciate now why we did not use *Watt* before. Now a 660W heater with a power factor of unity will take only 3A (=660W/220V) while a motor with the same power consumption takes 5A, much more than that consumed by the heater. This puts a greater demand on the current produced by a generator. Therefore power generation authorities usually require that consumers use power factor correcting devices to increase the power factor to close to unity. Since most of typical loads are inductive, power factor improvement is usually done by adding capacitors in parallel, which basically helps in decreasing the phase difference between the voltage and current.

Another point worth noting is that ac sources or generators are usually rated in *VA*, not in Watts as hinted above. This is because a generator has no way of knowing beforehand how much power a load will take as the power factor will vary from load to load. So it can only say about the maximum current that it can deliver at the rated voltage. Since the voltage is fixed, the *VA* value rather than the current is quoted because that will be closer to load power in watts with which the users are more familiar. Of course, maximum current rating is also given typically. Therefore any source like Voltage Stabilisers, UPS, or transformers which delivers or transfers ac power are rated in *VA* value.

On the other hand an ac load knows what should be its power factor. So loads are usually rated in watts; the power factor being quoted as well.

Appendix-2

Average ac voltage, Form Factor

What is the average of a sinusoidal ac voltage? Looking at Fig.A2.1 we can see that this will very well depend on the time range that is considered for this average. The average will vary with time. If we take average over a full cycle, clearly it will be zero as the +ve half cycle will cancel the –ve half cycle. If we take the average over one of these half cycles (shown shaded) we will get a non-zero average value. This is usually the average value quoted for a sinusoidal ac voltage (which is not very useful for ac, but is useful for its rectified form; see next appendix). Let us deduce a value for this average.

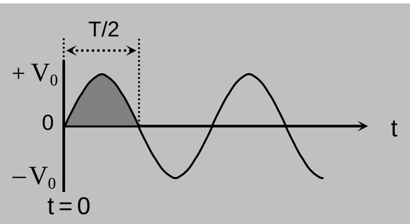


Fig.A2.1: Averaging over half a cycle

We have, $e = V_o \, Sin \, \omega t$ ... A2.1

Therefore average over the positive half wave is given by,

$$V_{AV} = \frac{1}{T/2}\int_o^{T/2} V_o \, Sin\,\omega t \, dt = \frac{V_o}{T/2}\left(-\frac{Cos\,\omega t}{\omega}\right)\Bigg|_0^{T/2} = \frac{V_o}{\pi}\,Cos\,\omega t\,\Bigg|_{T/2}^{0} = \frac{2V_o}{\pi}$$

i.e., $\boxed{V_{AV} = \frac{2V_o}{\pi}}$ ... A2.2

Evaluating π, we get $V_{AV} = 0.637\,V_o$ ... A2.3

which is the commonly accepted average of an ac sinusoidal voltage.

The average current will be similarly described.

The effective or RMS voltage (given by Eq.A1.6a) and the Average voltage both are important parameters for an ac.

The ratio of these two parameters is also an important parameter called the *Form Factor*. Therefore, we have,

$$\boxed{\text{FORM FACTOR} = V_{RMS}/V_{AV} = 1.11}$$

The above result of 1.11 holds only for a sinusoidal waveform. For any other waveform the form factor will have a different value.

Appendix-3

AC volt meters – average and RMS measurement

A galvanometer is basically a dc current measuring device. It has a certain internal resistance. Therefore at a certain applied voltage a specified current will pass through it. Therefore the galvanometer can also be calibrated in terms of a dc voltage. By adding a suitable resistance in series its dc voltage measuring capability can be increased while the ac current capability can be increased by connecting a low resistance in parallel (a shunt). Now, how to measure an ac voltage using this galvanometer which is always changing? Usually the ac voltage is firstly rectified using semiconductor diodes and the rectified dc voltage is applied to the galvanometer. The rectification can be half wave or full wave as shown in Fig. A3.1. These are varying voltage as well. So how would a galvanometer needle respond to such a changing voltage?

Well, at very low frequency (say, 1Hz) the needle will follow the voltage change which is very slow. As one increases the frequency, the needle is pulled back before it can reach the peak deflection. It will still vary but will not be able to deflect to the peak amplitude. At the line frequency of 50Hz, it is almost impossible for the needle to follow the changes, and it usually deflects to a fixed value. This fixed value is the average voltage of the rectified dc waveform. (You can appreciate that such a needle galvanometer will show zero if a 50 Hz ac is applied to it directly, since the average of the ac over a complete cycle is zero. Taking the average over a long period of time will also be almost zero)

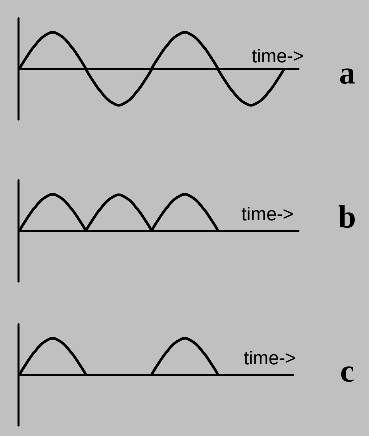


Fig.A3.1: Voltage waveforms:
a) sinusoidal ac, b) full wave rectified and c) half wave rectified

The average value of a full wave rectified waveform as shown in Fig.A3.1b will be the same as calculated in Eq.A2.2 for an ac since the full wave rectified waveform is just the repetition of a half wave (Fig.A2.1) which was considered for calculating the average. Therefore if the ac voltmeter uses full wave rectification, it will deflect to show the average value which is $0.637V_o$. Now the thing is, we usually like to describe an ac voltage in terms of its RMS value, not its average. Therefore the marking on the galvanometer dial is made to indicate the corresponding RMS voltage. For example, if the peak amplitude V_o is 10V, average of the full wave rectified voltage would be 6.37V. If we rectify this ac ourselves and feed the output to a dc voltmeter, we would get this reading. However, to read the ac directly off the meter we calibrate the dial such that at this position the needle would show 7.07V, 1.11 times the actual dc voltage reading. In this way we can read the RMS value of a sinusoidal ac voltage directly from such a meter. Even most hand held digital multimeters are also of the average reading type internally, but in the display they show the corresponding RMS value.

> Semiconductor devices used to rectify an ac waveform usually drop some voltage themselves, of the order of a fraction of a volt, therefore, voltage readings will have some errors due to this, which will be more significant at low voltages. Therefore you will see that the actual voltage markings in a needle type ac voltmeter are non linear, particularly at low voltages. However, in digital voltmeters, these errors are eliminated through the use of clever electronic circuit design.

What happens if the ac waveform is not sinusoidal? The needle would still point to the average of the full wave rectified value, but since the form factor is no more 1.11, the RMS value indicated by the dial will no longer be valid. Remember this fact while measuring a non-sinusoidal waveform using an ordinary digital meter too.

Nowadays many ac power circuits are controlled by a semiconductor device called TRIAC which control the output power by cutting off part of the sinusoidal ac. The resulting waveform is no longer a sinusoidal one and therefore any ac meter employing an average reading galvanometer would give a wrong indication for RMS value. Again at different setting for power, the waveform will be different, so the meter calibration would change too. Therefore if you use such meters, take note of the waveform. Of course there are special meters which use sophisticated techniques to obtain the RMS value, both in needle galvanometer type and in digital type. These are called RMS measuring ac meters. One would be better off using one of these meters, but as can be expected, these are a bit expensive and are not available in ordinary shops.

Some meters may use a half wave rectifier, the waveform is shown in Fig.A3.1c. During the negative half cycle, the output voltage is zero, the negative half is not passed at all. Here the average voltage will clearly be half of the value given by Eq.A2.3 for a full wave rectified waveform. Therefore the average voltage of a half wave rectified sinusoidal ac would be $0.318V_o$ (= $0.637V_o/2$) and the dial calibration will be made accordingly.

The above considerations would also apply to AC current meters.

Index

Index