M.Sc. Previous Year
Physics, MP-04

SOLID STATE ELECTRONICS

मध्यप्रदेश भोज (मुक्त) विश्वविद्यालय – भोपाल

MADHYA PRADESH BHOJ (OPEN) UNIVERSITY - BHOPAL

COURSE WRITERS

Neha Gupta, Assistant Professor, Department of Electronics and Communication Engineering, Dev Bhoomi Group of Institutions, Dehradun, Uttarakhand
(Units: 1.0-1.1, 1.2-1.4, 1.5-1.9, 2.5-2.9, 2.10-2.14, 3.2.4, 3.2.6, 3.3-3.3.6, 3.4-3.8, 4.4.3-4.4.5)

Dr. Jyoti Prasad Bandyopadhyay, Professor, Institute of Radio Physics and Electronics
(Units: 2.0-2.1, 2.2-2.4, 3.0-3.1, 3.2-3.2.3, 3.2.5, 5.0-5.1, 5.2, 5.2.1)

Dr Samiksha Sirakwar, Assistant Professor, Integrated Basic Science School for Physical and Decision Science, Babasaheb Bhimrao Ambedkar University (A Central University), Lucknow
Dr Bhupendra Pratap Singh, Former Faculty, Deptt of Physics, University of Lucknowzx
(Units: 1.4.1-1.4.4, 4.0-4.1, 4.2-4.2.8, 4.3-4.4.2, 4.6-4.13, 5.3-5.10)

Published by Registrar, MP Bhoj (Open) University, Bhopal in 2020

VIKAS®

Vikas® is the registered trademark of Vikas® Publishing House Pvt. Ltd.

VIKAS® PUBLISHING HOUSE PVT. LTD.
E-28, Sector-8, Noida - 201301 (UP)
Phone: 0120-4078900 • Fax: 0120-4078999
Regd. Office: A-27, 2nd Floor, Mohan Co-operative Industrial Estate, New Delhi 1100 44
• Website: www.vikaspublishing.com • Email: helpline@vikaspublishing.com

SYLLABI-BOOK MAPPING TABLE

Solid State Electronics

Syllabi	Mapping in Book
UNIT - I **Transistor Amplifier, Operating Point, Bias and Thermal Stability:** Operating point, factors contributing to thermal stability, Biasing technique collector to base, self bias and voltage divider bias, stabilization against variation in V_{BE} and B Bias compensation. Transistor equivalent circuits - V (Admittance) and hybrid parameters. Conversion of CB to CE hybrid parameters and CB to CC hybrid parameters. Analysis of transistor amplifier using hybrid model. R-C coupled CE amplifier and its frequency response, low and high frequency compensation, cascade stages.	**Unit-1:** Transistor Amplifier, Operating Point, Bias and Thermal Stability **(Pages 3-40)**
UNIT - II **Feedback Circuits:** Feedback in amplifiers, negative feedback and gain stability effect of feedback on input and output impedances and distortions, current and voltage feedback circuits, Emitter follower, circuits and working of Hartley, Colpitt and Phase shift oscillators. UJT and its characteristics, UJT as relaxation oscillators. Transistor as a switch. Astable, monostable and bistable multivibrators.	**Unit-2:** Feedback Circuits **(Pages 41-83)**
UNIT - III **Operational Amplifier:** Differential amplifier circuits and working of operational amplifier, Op-amp parameters, inverting and non-inverting Op-amp amplifiers. Use of 741 IC as adder, subtractor, differentiator and integrator, Op-amp as constant current source, comparator, square and triangular wave generator. Voltage multipiers circuits, wave shaping circuits, dipping, clamping, differentiating and intergrating circuits, Voltage regulated power supply, regulation sensitivity and stability factors, Over voltage and short circuit performance (Transistorised). Three terminaI IC regulated power supply circuits for positive and negative voltages.	**Unit-3:** Operational Amplifier **(Pages 85-130)f**
UNIT - IV **Communication Electronics:** Types of modulation analysis and production of AM and FM wave Generation DSB-SC modulation, of AM waves, demodulation of AM waves Generation of DSB-SC waves, coherent detection of DBS-SC waves, SSB modulation, Generation and detection of SSB waves, vestigial sideband modulation, frequencies division multiplexing.	**Unit-4:** Communication Electronics **(Pages 131-182)**

UNIT - V:

DEVICES Electronic Devices: JFET, MOSFET and MESFET, structure & working, IV characteristics under different condition, microwave devices tunnel diode and Gunn diode, impatt diodes and parametric devices.

Photonic Devices: Reditives and non radiative transmitter, LDR, photodiode detectors, solar cells (open circuit voltage, short circuit element and fill factor), LED (high frequency limit) diode lasers condition for population inversion light confinement factor, optical gain and threshold current for lasing.

Unit-5: Electronic Devices and Photonic Devices
(Pages 183-273)

CONTENTS

INTRODUCTION

NOTES

Solid-state electronics means semiconductor electronics: electronic equipment using semiconductor devices such as transistors, diodes and Integrated Circuits (ICs). The term is also used for devices in which semiconductor electronics which have no moving parts replace devices with moving parts, such as the solid-state relay in which transistor switches are used in place of a moving-arm electromechanical relay, or the Solid-State Drive (SSD) a type of semiconductor memory used in computers to replace hard disk drives, which store data on a rotating disk. The term 'Solid State' became popular in the beginning of the semiconductor era in the 1960s to distinguish this new technology based on the transistor, in which the electronic action of devices occurred in a solid state, from previous electronic equipment that used vacuum tubes, in which the electronic action occurred in a gaseous state. A semiconductor device works by controlling an electric current consisting of electrons or holes moving within a solid crystalline piece of semiconducting material such as silicon, while the thermionic vacuum tubes it replaced worked by controlling current conducted by a gas of particles, electrons or ions, moving in a vacuum within a sealed tube.

Although the first solid state electronic device was the cat's whisker detector, a crude semiconductor diode invented around 1904, solid state electronics really started with the invention of the transistor in 1947. The transistor was invented by John Bardeen and Walter Houser Brattain while working under William Shockley at Bell Laboratories in 1947. Before that, all electronic equipment used vacuum tubes, because vacuum tubes were the only electronic components that could amplify-an essential capability in all electronics. The replacement of bulky, fragile, energy-hungry vacuum tubes by transistors in the 1960s and 1970s created a revolution not just in technology but in people's habits, making possible the first truly portable consumer electronics such as the transistor radio, cassette tape player, walkie-talkie and quartz watch, as well as the first practical computers and mobile phones. Other examples of solid state electronic devices are the microprocessor chip, LED lamp, solar cell, Charge Coupled Device (CCD) image sensor used in cameras, and semiconductor laser.

NOTES

UNIT 1 TRANSISTOR AMPLIFIER, OPERATING POINT, BIAS AND THERMAL STABILITY

Structure

1.0 INTRODUCTION

Operating point is a specific point within the operation characteristic of a technical device. This point will be engaged because of the properties of the system and the outside influences and parameters. In electronic engineering establishing an operating point is called biasing.

In electronics, biasing is the setting of initial operating conditions (current and voltage) of an active device in an amplifier. Many electronic devices, such as diodes, transistors and vacuum tubes, whose function is processing time-varying (AC) signals, also require a steady (DC) current or voltage at their terminals to operate correctly. This current or voltage is a bias. The AC signal applied to them is superpositioned on this DC bias current or voltage. The operating point of a device, also known as bias point, quiescent point, or Q-point, is the DC voltage or current at a specified terminal of an active device (a transistor or vacuum tube) with no input signal applied. A bias circuit is a portion of the device's circuit which supplies this steady current or voltage.

In electrical engineering and science, an equivalent circuit refers to a theoretical circuit that retains all of the electrical characteristics of a given circuit. Often, an equivalent circuit is sought that simplifies calculation, and more broadly, that is a simplest form of a more complex circuit in order to aid analysis. In its most

NOTES

common form, an equivalent circuit is made up of linear, passive elements. However, more complex equivalent circuits are used that approximate the nonlinear behavior of the original circuit as well. These more complex circuits often are called macromodels of the original circuit. An example of a macromodel is the Boyle circuit for the 741 operational amplifier.

A transistor is a semiconductor device used to amplify or switch electrical signals and power. The transistor is one of the basic building blocks of modern electronics. It is composed of semiconductor material, usually with at least three terminals for connection to an electronic circuit. A voltage or current applied to one pair of the transistor's terminals controls the current through another pair of terminals. Because the controlled (output) power can be higher than the controlling (input) power, a transistor can amplify a signal. Some transistors are packaged individually, but many more are found embedded in integrated circuits. Austro-Hungarian physicist Julius Edgar Lilienfeld proposed the concept of a field-effect transistor in 1926, but it was not possible to actually construct a working device at that time. The first working device to be built was a point-contact transistor invented in 1947 by American physicists John Bardeen and Walter Brattain while working under William Shockley at Bell Labs. The three shared the 1956 Nobel Prize in Physics for their achievement. The most widely used type of transistor is the Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET), which was invented by Mohamed Atalla and Dawon Kahng at Bell Labs in 1959.

In this unit, you will learn about the operating point, factors contributing to thermal stability, biasing techniques collector to base, self bias and voltage divider bias, transistor equivalent circuits, V (Admittance) and hybrid parameters, analysis of transistor amplifier using hybrid model, low and high frequency compensation and cascade stages.

1.1 OBJECTIVES

After going through this unit, you will be able to:

- Understand the basic concept of operating point
- Discuss the factors contributing to thermal stability
- Analyse the biasing techniques collector to base
- Discuss about the self bias and voltage divider bias
- Explain about the stabilization against variation in VBE and B-bias compensation
- Interpret the transistor equivalent circuits
- Elaborate on the V (Admittance) and hybrid parameters
- Analysis the transistor amplifier using hybrid model
- Explain about the low and high frequency compensation
- Discuss cascade stages

1.2 OPERATING POINT

Operating point is a specific point within the operation characteristic of a technical device. This point will be engaged because of the properties of the system and the outside influences and parameters.

The working point or Q-point in a semiconductor is controlled by the upsides of the I_C (gatherer current) or V_{CE} (authority producer voltage) when no sign is applied to the information. At the point when no sign is given to the info, changes in I_C (gatherer current) and V_{CE} (authority producer voltage) happen around this point, consequently the name working point.

Since it is a point on the $I_C – V_{CE}$ trademark when the semiconductor is quiet or no info signal is given to the circuit, the working point is otherwise called the peaceful (quiet) point or essentially Q-point. The DC load line approach makes getting the working point basic. The following is a clarification of the DC load line.

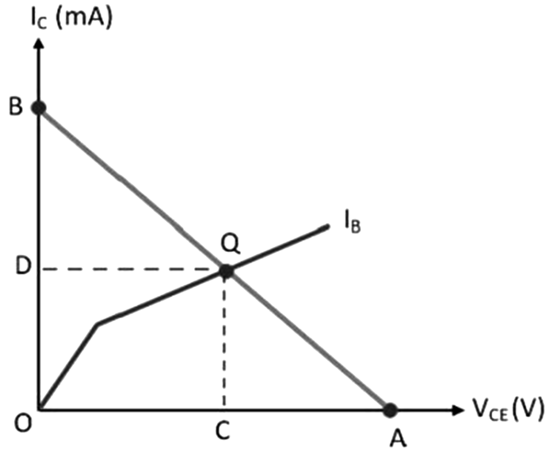


Fig. 1.1 *DC Load Line*

Let, decides the working place of specific base circuit current IB. As indicated by the heap line condition, the OA = $V_{CE} = V_{CC}$ and OB = $I_C = V_{CC}/R_C$ is displayed on the result trademark bend above. The point Q is the working point where the DC load line converges the base current I_B at the result trademark bends without even a trace of information signal.

Where I_C= OD mA

V_{CE} = OC volts.

The situation of the Q-point relies upon the utilizations of the semiconductor. Assuming that the semiconductor is utilized as a switch, then, at that point, for open switch the Q-point is in the removed locale, and for the nearby switch, the Q-point is in the immersion area. The

Q-point lies in the line for the semiconductor which works as an amplifier.

Note: In immersion locale, both the authority base district and the producer base area are in forward one-sided and substantial current course through the intersection. Furthermore, the area wherein both the intersections of the semiconductor are in switched one-sided is known as the remove locale.

1.2.1 Factors Contributing to Thermal Stability

Thermal Runaway: The authority current for the C_E circuit is given by $I_C = \beta I_B + (1+\beta) I_{CO}$. The three factors in the situation, β, I_B, and I_{CO} increment with ascend in temperature. Specifically, the opposite immersion current or spillage current I_{CO} changes extraordinarily with temperature. In particular, it copies for each 10°C ascent in temperature. The authority current I_C causes the gatherer base intersection temperature to rise which thus, increment I_{CO}, subsequently I_C will build even further, which will additionally expand the temperature at the gatherer base intersection. This cycle will become total at the authority base intersection prompting 'Warm out of Control'. Thus, the appraisals of the semiconductor are surpassed which might obliterate the actual semiconductor.

The gatherer is caused bigger in size than the producer to work with the hotness scattering at the authority intersection. Nonetheless, assuming the circuit is planned to such an extent that the base current I_B is made to diminish consequently with ascend in temperature, then, at that point, the decline in βI_B will make up for expansion in the $(1+\beta) I_{CO}$, keeping I_C practically consistent.

There are a few ways to deal with alleviate bipolar semiconductor warm out of control. For instance,

- Negative criticism can be incorporated into the biasing circuit so that expanded authority current prompts diminished base current. Thus, the expanding authority current chokes its source.
- Heat sinks can be utilized that divert additional hotness and keep the base producer temperature from rising.
- The semiconductor can be one-sided so its authority is typically not exactly 50% of the power supply voltage, which infers that gatherer producer power dissemination is at its greatest worth. Runaway is then outlandish in light of the fact that expanding gatherer current prompts a lessening in scattered power.

Thermal Stability: At the point when the intersection temperature of a semiconductor expands, the authority current increments because of expansions in 1) immersion flows and 2) DC conductance. Warm flimsiness happens when intersection temperature and gatherer current expansion in regenerative and wild style. The breaking point relies upon factors both inside and outside to the semiconductor. The inner elements are the warm obstruction, the current enhancement factor, and the base lead opposition. The outside factors are encompassing temperature, authority voltage, circuit protections, and the warm coupling between the semiconductor and temperature pay components, assuming any. Circuits with more prominent temperature steadiness grant a higher greatest power dispersal. Warm soundness is important in semiconductors, due to their negative temperature coefficient, when they begin getting hot, warm rampant is unavoidable except if there is some electrical remedy applied to the biasing of the semiconductor to reduce the flow to ensure it doesn't proceed down the way of implosion.

Further, the warmer the semiconductor creates inside, the more warm clamor in produces and in certain applications, similar to sound enhancers and low commotion R_F intensifiers, the warm commotion is an issue. Limiting that commotion is truly alluring for those applications so designs make a special effort to guarantee that intersection temperatures in the semiconductor stay cool. Some low clamor intensifiers are even cooled with fluid nitrogen to keep the intersections truly cool and the commotion level as low as could be expected.

Additionally, the primary calculate influencing warm security bipolar semiconductors is the impact of temperature on the immersion current (Is) of the base-producer diode.

The current-voltage normal for this intersection can be addressed by the Schottky condition

$I_{be} = I_s\{\exp(qV/kT) - 1\}$

where q is the electronic charge, k is Boltzmann's constant, V is the base-producer voltage and T is the outright temperature.

Since it is firmly subject to temperature, and increments with temperature, Ibe increments when the temperature increments even with steady V. Assuming the semiconductor is one-sided from a consistent voltage supply Ibe and along these lines Ic will increment with temperature. This is the essential component of warm insecurity.

1.2.2 Biasing Techniques: Collector to Base

Biasing: For appropriate working of a semiconductor, it is fundamental to apply outer voltages of right extremity across its producer base and gatherer base intersections. In this way, Biasing is a course of setting DC working voltage or current to an ideal level across a semiconductor with the goal that AC information can be amplified effectively. It is needed to actuate transistor and keep it from entering saturation or cut-off region.

Need for Biasing

To deliver mutilation free result in amplifier circuits, the inventory voltages and protections build up a bunch of DC voltage V_{CEQ} and I_{CQ} to work the semiconductor in the dynamic area. These voltages and flows are called peaceful qualities which decide the working point or Q point for the semiconductor. The method involved with giving appropriate stock voltages and protections for acquiring the ideal Q Point is called Biasing. The circuits utilized for getting the ideal and appropriate working point are known as biasing circuits. To set up the working point in the dynamic locale biasing is needed for semiconductors to be utilized as an amplifier. For simple circuit activity, the Q point is set so the semiconductor stays in dynamic mode (doesn't move to activity in the immersion locale or remove area) when information is applied. For computerized activity, the Q point is put so the semiconductor does the opposite changes from 'on' to 'off' state. Frequently, Q point is set up close to the focal point of dynamic area of semiconductor trademark to permit comparative sign swings in sure and negative ways. Q point ought to be steady. Specifically, it ought to be harsh toward varieties in semiconductor boundaries (for instance, ought not move in case semiconductor is supplanted by one more of a similar kind), varieties in temperature, varieties in power supply voltage, etc. The circuit should be useful: effortlessly executed and practical.

1.2.3 Self Bias and Voltage Divider Bias

There are four common biasing methods or techniques to bias the transistor:

1. Fixed bias (or Self bias)
2. Collector-to-base bias
3. Fixed bias with emitter resistor
4. Voltage divider bias or potential divider bias.

Fixed Bias:

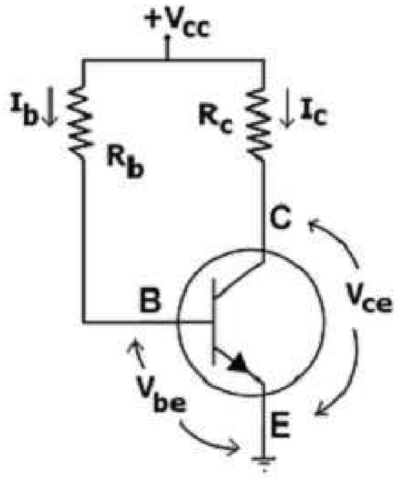


Fig. 1.2 *Circuit for Fixed Bias Arrangement*

Fixed Bias (base bias)

The single power hotspot (for instance, a battery) is utilized for both gatherer and base of a semiconductor, albeit separate batteries can likewise be utilized.

The equation of the above diagram,

$$V_{cc} = I_b R_b + V_{be}$$

Therefore,

$$I_B = (V_{cc} - V_{be})/R_b$$

For a given semiconductor, V_{be} doesn't shift altogether during use. As V_{cc} is of fixed worth, on determination of R_B, the base current I_B is fixed. Thusly, this sort is called fixed predisposition kind of circuit.

Also for given circuit,

$$V_{cc} = I_c R_c + Vce$$

Therefore,

$$V_{ce} = V_{cc} - I_C R_C$$

The normal producer current addition of a semiconductor is a significant boundary in circuit design, and is indicated on the information sheet for a specific semiconductor. It is signified as β on this page.

Since, $I_C = \beta I_B$ we can get I_C too. As such, working point given as (V_{ce}, I_C) can be set for given semiconductor.

Usage: Due to the above intrinsic disadvantages, fixed predisposition is seldom utilized in straight circuits (i.e., those circuits which utilize the semiconductor as a current source). All things considered; it is frequently utilized in circuits where semiconductor is utilized as a switch. In any case, one utilization of fixed inclination is to accomplish rough programmed gain control in the semiconductor by taking care of the base resistor from a DC signal got from the AC result of a later stage.

Self Bias: The proper predisposition plan examined in the past segment is thermally unsound. In case the temperature of the semiconductor ascends under any circumstance (because of an ascent in encompassing temperature or because of current move through it), the gatherer current will increment. This expansion in current likewise causes the DC tranquil highlight create some distance from its ideal position (level). This response to temperature is unwanted on the grounds that it influences intensifier gain (the hours of intensification) and could bring about twisting, as we will see later in this part. A superior technique for biasing, known as self-predisposition is acquired by embedding the inclination resistor straightforwardly between the base and gatherer, as shown in Figure 1.3.

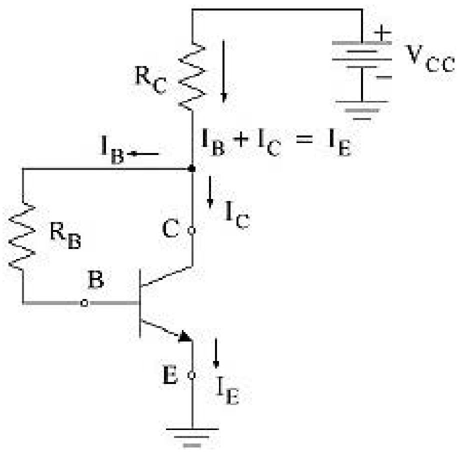


Fig. 1.3 *NPN Transistor Amplifier with Self-Bias*

By binds the gatherer to the base thusly, criticism voltage can be taken care of from the authority to the base to create forward inclination. Presently, in case an increment of temperature causes an expansion in gatherer current, the authority voltage V_C will fall on account of the increment of voltage delivered across the authority resistor R_L. This drop in V C will be taken care of back to the base and will bring about a lessening in the base current. The abatement in base current will go against the first expansion in gatherer current and will generally balance out it. The specific inverse impact is created when the authority current reductions.

From Figure 1.3,

$$R_C I_E + R_B I_B + V_{BE} = V_{CC}$$

$$R_C I_E + R_B \frac{1}{\beta + 1} I_E + V_{BE} = V_{CC}$$

$$I_E = \frac{V_{CC} - V_{BE}}{R_C + R_B/(\beta + 1)}$$

From above equation we see that to maintain the emitter current.

Collector to Base Bias:

The authority to base predisposition circuit is same as base inclination circuit with the exception of that the base resistor R_B is gotten back to gatherer, rather than to V_{CC} supply as displayed in the figure beneath.

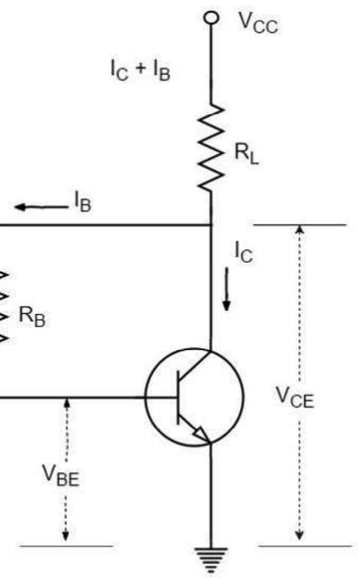


Fig. 1.4 *Collector to Base Bias Circuit*

This circuit helps in further developing the dependability significantly. Assuming the worth of I_C builds, the voltage across R_L increments and subsequently the V_{CE} additionally increments. This thus decreases the base current I_B. This activity to some degree remunerates the first increment.

The necessary worth of R_B expected to give the zero-sign gatherer current I_C can be determined as follows.

Voltage drop across R_L will be

$$R_L = (I_C + I_B)\, R_L \cong I_C R_L$$

From the figure,

$$I_C R_L + I_B R_B + V_{BE} = V_{CC}$$

Or

$$I_B R_B = V_{CC} - V_{BE} - I_C R_L$$

Therefore

$$R_B = (V_{CC} - V_{BE} - I_C R_l) / I_B$$

Or

$$R_B = (V_{CC} - V_{BE} - I_C R_L) / I_B$$

Applying KVL we have

$$(I_B + I_C)\, R_L + I_B R_B + V_{BE} = V_{CC}$$

NOTES

Or

$$I_B(R_L + R_B) + I_CR_L + V_{BE} = V_{CC}$$

Therefore

$$I_B = V_{CC} - V_{BE} - I_CR_L / R_L + R_B$$

Since V_{BE} is almost independent of collector current, we get

$$DI_B/dI_C = R_L / R_L + R_B$$

We know that

$$S=1 + \beta/1 - \beta(dI_B/dI_C)$$

Therefore

$$S=1 + \beta_1 + \beta(R_L + R_B)$$

This worth is more modest than (1 + β) which is acquired for fixed predisposition circuit. Accordingly, there is an improvement in the strength.

This circuit gives negative input which lessens the addition of the intensifier. Along these lines, the expanded soundness of the gatherer to base predisposition circuit is gotten at the expense of AC voltage gain.

Biasing with Collector Feedback Resistor

In this strategy, the base resistor R_B has its one end associated with base and the other to the gatherer as its name infers. In this circuit, the zero-signal base current is controlled by V_{CB} yet not by V_{CC}.

Unmistakably V_{CB} forward predispositions the base-producer intersection and henceforth base current I_B moves through R_B. This makes the zero-signal authority current stream in the circuit. The beneath figure shows the biasing with authority criticism resistor circuit.

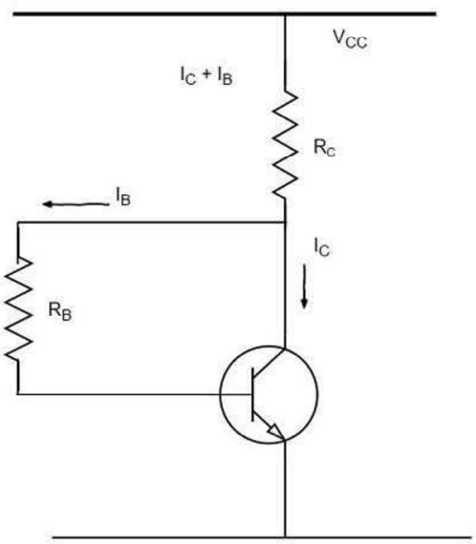


Fig. 1.5 *Biasing with collector feedback resistor circuit*

The required value of R_B needed to give the zero-signal current I_C can be determined as follows.

$$V_{CC} = I_C R_C + I_B R_B + V_{BE}$$

Or

$$R_B = V_{CC} - V_{BE} - I_C R_C / I_B = V_{CC} - V_{BE} - \beta I_B R_C / I_B$$

Since $I_C = \beta I_B$

Alternatively,

$$V_{CE} = V_{BE} + V_{CB}$$

Or

$$V_{CB} = V_{CE} - V_{BE}$$

Since

$$R_B = V_{CB}/I_B = V_{CE} - V_{BE}/I_B$$

Where

$$I_B = I_C / \beta$$

Mathematically,

Stability factor, $S < (\beta + 1)$

Therefore, this method provides better thermal stability than the fixed bias.

The Q-point values for the circuit are shown as

$$I_C = V_{CC} - V_{BE} R_B/\beta + R_C$$
$$V_{CE} = V_{CC} - I_C R_C$$

Advantages

- The circuit is straightforward as it needs just a single resistor.
- This circuit gives some adjustment, for lesser changes.

Disadvantages

- The circuit doesn't give great adjustment.
- The circuit gives negative criticism.

Voltage Divider Biasing or Emitter Bias

Among all the techniques for giving biasing and adjustment, the voltage divider predisposition strategy is the most noticeable one. Here, two resistors R_1 and R_2 are utilized, which are associated with V_{CC} and give biasing. The resistor R_E utilized in the producer gives adjustment.

The name voltage divider comes from the voltage divider shaped by R_1 and R_2. The voltage drop across R_2 forward inclinations the base-producer intersection. This causes the base current and subsequently gatherer current stream in the zero sign conditions. Shown in Figure 1.6 beneath the circuit of voltage divider inclination technique.

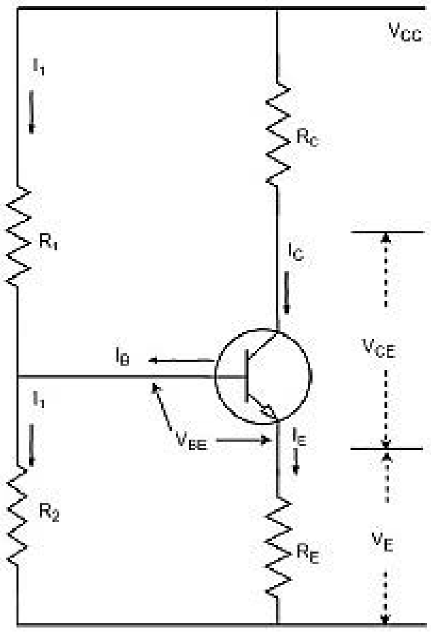


Fig. 1.6 *Voltage Divider Bias*

Assume that the current moving through opposition R_1 is I_1. As base current I_B is tiny, subsequently, it very well may be expected with sensible precision that current moving through R_2 is additionally I_1.

Presently let us attempt to determine the articulations for gatherer current and authority voltage.

Collector Current, I_C

From the circuit, it is evident that,

$$I_1 = V_{CC}/R_1 + R_2$$

Therefore, the voltage across resistance R_2 is

$$V_2 = (V_{CC}/R_1 + R_2)R_2$$

Applying Kirchhoff's voltage law to the base circuit,

$$V_2 = V_{BE} + V_E$$

$$V_2 = V_{BE} + I_E R_E$$

$$I_E = V_2 - V_{BE}/R_E$$

Since $I_E \approx I_C$,

$$I_C = V_2 - V_{BE}/R_E$$

From the above articulation, it is clear that I_C doesn't rely on β. V_{BE} is tiny that I_C doesn't get impacted by V_{BE} by any stretch of the imagination. Accordingly I_C in this circuit is practically autonomous of semiconductor boundaries and consequently great adjustment is accomplished.

Collector-Emitter Voltage, V_{CE}

Applying Kirchhoff's voltage law to the collector side,

$$V_{CC} = I_C R_C + V_{CE} + I_E R_E$$

NOTES

Since $I_E \cong I_C$

$$= I_C R_C + V_{CE} + I_C R_E$$
$$= I_C(R_C + R_E) + V_{CE}$$

Subsequently,

$$V_{CE} = V_{CC} - I_C(R_C + R_E)$$

R_E provides excellent stabilization in this circuit.

$$V_2 = V_{BE} + I_C R_E$$

Assume there is an ascent in temperature, then, at that point, the authority current I_C abatements, which makes the voltage drop across R_E increment. As the voltage drop across R_2 is V_2, which is free of I_C, the worth of V_{BE} diminishes. The decreased worth of I_B will in general re-establish I_C to the first worth.

Stability Factor

The condition for Stability component of this circuit is gotten as

Stability Factor = S = $(\beta + 1)\ (R_0 + R_3) / R_0 + R_E + \beta R_E = (\beta + 1) \times 1 + R_0 R_E\ \beta + 1 + R_0 R_E$

Where

$$R_0 = R_1 R_2 R_1 + R_2$$

Assuming the proportion R_0/R_E is tiny, R_0/R_E can be ignored when contrasted with 1 and the dependability factor becomes

Stability Factor = S = $(\beta+1) \times 1\beta+1=1$

This is the smallest possible value of S and leads to the maximum possible thermal stability.

1.2.4 Stabilization Against Variation in V_{BE} and B-Bias Compensation

These are the circuits that execute pay methods utilizing diodes to manage biasing flimsiness. The adjustment procedures allude to the utilization of resistive biasing circuits which grant I_B to differ to keep I_C generally steady.

There are two kinds of diode remuneration techniques. They are –

- Diode pay for shakiness because of V_{BE} variety
- Diode remuneration for flimsiness because of I_{CO} variety

Diode Compensation for Instability due to V_{BE} Variation

In a Silicon semiconductor, the progressions in the worth of V_{BE} results in the progressions in I_C. A diode can be utilized in the producer circuit to remunerate the varieties in V_{BE} or I_{CO}. As the diode and semiconductor utilized are of same material, the voltage V_D across the diode has same temperature coefficient as V_{BE} of the semiconductor.

The accompanying shown in Figure 1.7 self-inclination with adjustment and remuneration.

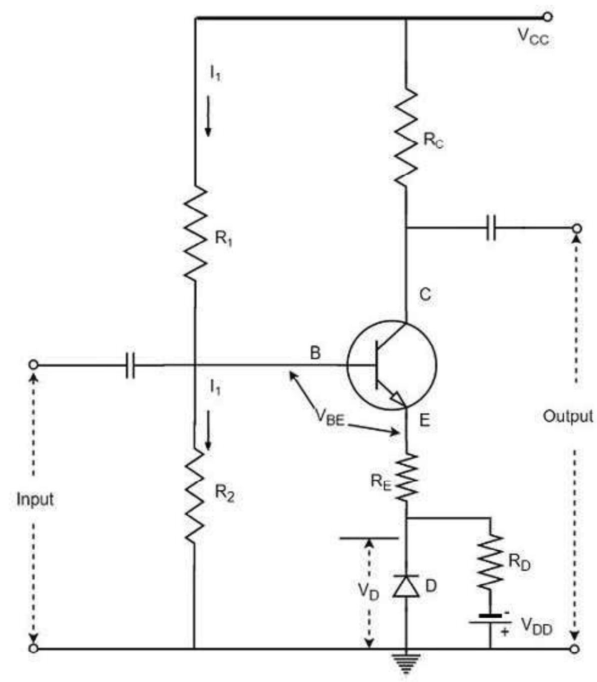


Fig. 1.7 *Diode Compensation for Instability due to V_{BE} Variation*

The diode D is forward one-sided by the source V_{DD} and the resistor R_D. The variety in V_{BE} with temperature is same as the variety in V_D with temperature, subsequently the amount $(V_{BE} - V_D)$ stays consistent. So the current I_C remaining parts steady regardless of the variety in V_{BE}.

Diode Compensation for Instability due to I_{CO} Variation

The accompanying shown in Figure 1.8 the circuit outline of a semiconductor amplifier with diode D utilized for remuneration of variety in I_{CO}.

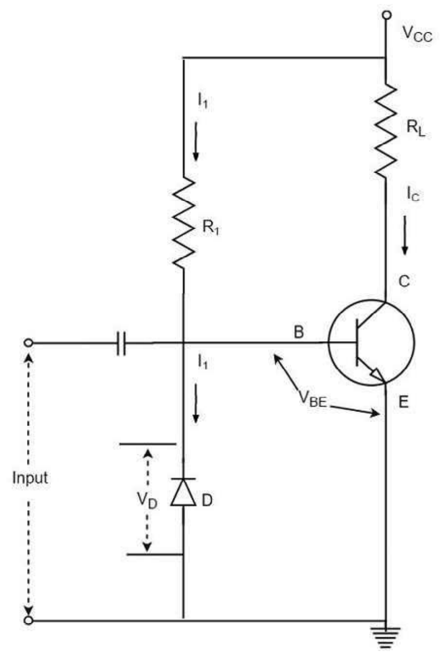


Fig. 1.8 *Diode Compensation for Instability due to I_{CO} Variation*

NOTES

Along these lines, the opposite immersion current I_O of the diode will increment with temperature at a similar rate as the semiconductor gatherer immersion current I_{CO}.

$I = V_{CC} – V_{BE}R \cong V_{CC}R$ = Constant $I = V_{CC} – V_{BE}R \cong V_{CC}R$ = Constant

The diode D is converse one-sided by V_{BE} and the current through it is the opposite immersion current I_O.

Now the base current is,

$I_B = I–I_O I_B = I – I_O$

Substituting the above value in the expression for collector current.

$I_C = \beta(I–I_O) + (1+\beta)I_{CO}$

If $\beta >> 1$,

$I_C = \beta I – \beta I_O + \beta I_{CO}$

I is practically consistent and in the event that I_O of diode and I_{CO} of semiconductor track each other over the working temperature range, then, at that point, I_C remaining parts steady.

Check Your Progress

1. What do you mean by remove locale?
2. Define thermal runaway.
3. What is biasing?
4. How will you define the working point or Q-point for the semiconductor?
5. Define fixed bias.
6. Give the statement for usage of fixed bias system.
7. What is the collector to base bias?
8. What is the main function of biasing with collector feedback resistor?
9. Write the advantages and disadvantages of biasing with collector feedback resistor.
10. How will you define the voltage divider biasing?
11. What is stability factor?
12. Name some of the diode remuneration techniques.

1.3 TRANSISTOR EQUIVALENT CIRCUITS

An analogous circuit is a theoretical circuit that retains all of the electrical characteristics of a particular circuit in electrical engineering and research. In order to help analysis, an analogous circuit is frequently sought that simplifies calculation and, more broadly, that is the simplest form of a more complex circuit. An analogous circuit is made up of linear, passive parts in their most basic form. More sophisticated equivalent circuits, on the other hand, are utilized to approximate the nonlinear behavior of the original circuit. Macromodels of the original circuit are often used to describe these more complex circuits. The Boyle circuit for the 741 operational amplifier is an example of a macromodel.

NOTES

Examples of Equivalent Circuit

Thévenin and Norton equivalents

One of the most striking aspects of linear circuit theory is the ability to treat any two-terminal circuit, no matter how complicated, as if it were merely a source and an impedance with one of two basic equivalent circuit forms:

- **Thévenin Equivalent:** A single voltage source and a series impedance can be used to substitute any linear two-terminal circuit.
- **Norton Equivalent:** A current source and a parallel impedance can be used to substitute any linear two-terminal circuit.

The single impedance, on the other hand, can be of any complexity (as a function of frequency) and irreducible to a simpler form.

DC and AC Equivalent Circuits

The output of a linear circuit is equal to the total of the output from its DC sources alone and the output from its AC sources alone, thanks to the superposition principle. As a result, the DC and AC responses of a circuit are frequently examined separately, employing separate DC and AC equivalent circuits that have the same response to DC and AC currents as the original circuit. By combining the DC and AC responses, the composite response is calculated:

- Replace all capacitances with open circuits, inductances with short circuits, and reduce AC sources to zero to create a DC version of a circuit (replacing AC voltage sources by short circuits and AC current sources by open circuits).
- By lowering all DC sources to zero, an AC equivalent circuit can be built (replacing DC voltage sources with short circuits and DC current sources with open circuits).

By linearizing the circuit around the DC bias point Q-point, using an AC equivalent circuit made by calculating the equivalent small signal AC resistance of the nonlinear components at the bias point, this technique is often extended to small-signal nonlinear circuits like tube and transistor circuits.

Two-Port Networks

Two-port networks are commonly used to describe linear four-terminal circuits in which a signal is applied to one pair of terminals and an output is taken from another. Simple equivalent circuits of impedances and dependent sources can be used to describe these. The currents applied to the circuit must satisfy the port condition to be studied as a two-port network: the current entering one terminal of a port must be equal to the current exiting the other terminal of the port. A two-port representation for transistors can be created by linearizing a nonlinear circuit around its operating point: see hybrid pi and h-parameter circuits.

Delta and Wye Circuits

Three phase sources and loads can be connected in three phase power circuits in two ways: a 'Delta' connection and a 'Wye' connection. When evaluating circuits, converting between equivalent wye and delta circuits might help simplify the analysis. The wye-delta transform can be used for this.

1.3.1 V (Admittance) and Hybrid Parameters

Admittance Parameters: Y boundaries (otherwise called induction boundaries or impede) are properties utilized in electrical designing to depict the electrical conduct of straight electrical organizations. These Y-boundaries are utilized in Y-grids (induction frameworks) to compute the approaching and active voltages and flows of an organization.

Y-boundaries are otherwise called 'Hamper Boundaries', as they are determined under open-circuit conditions. Y boundaries of a two-port organization are determined by induction boundaries, and these terms might be utilized conversely in these models. While breaking down Z boundaries (otherwise called impedance boundaries),

we express voltage in the term of current by the accompanying conditions.

$$V_1 = Z_{11}I_1 + Z_{12}I_1,\ V_2 = Z_{21}I_1 + Z_{22}I_2$$

We can address current as far as voltage when we ascertain the permission boundaries of a two port organizations. Then, at that point, we will address the current-voltage relations as:

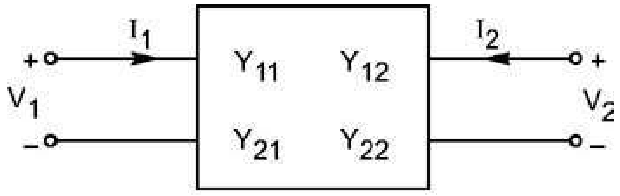
$$I_1 = Y_{11}\ V_1 + Y_{12}V_2,\ I_2 = Y_{21}V_1 + Y_{22}V_2$$

In lattice structure, we can compose it as

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

$$(10.3) I_1 = Y_{11}V_1 + Y_{12}V_2$$

$$(10.4) I_2 = Y_{21}V_1 + Y_{22}V_2$$



(*a*)

Condition (a)

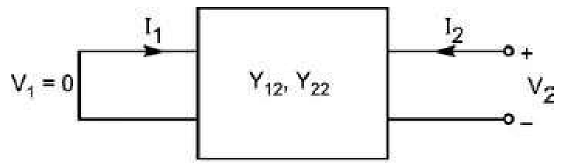
Now if $V_1 = 0$ the input port is short circuit

$$Y_{12} = \frac{I_1}{V_2}\bigg|_{V_1=0}$$

NOTES

and

$$Y_{22} = \frac{I_2}{V_2}\bigg|_{V_1=0}$$

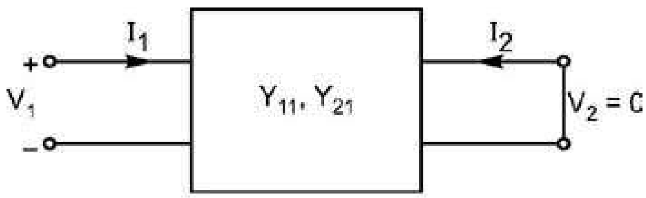


Condition (b)

if $V_2 = 0$ the output port is short circuit

$$Y_{11} = \frac{I_1}{V_1}\bigg|_{V_2=0}$$

$$Y_{21} = \frac{I_2}{V_1}\bigg|_{V_2=0}$$



Condition (c)

Fig. 1.9 *Determination of Y-Parameters*

[Y] boundaries are known as short circuit permission boundaries on the grounds that in both the cases, one-port is short (i.e., $V_1 = 0$ and $V_2 = 0$). The constants Y_{11}, Y_{12}, Y_{21}, and Y_{22} are known as the permission or Y-boundaries of the two-port organization. Condition (a) and Condition (b) are known as the standard condition or characterizing conditions of the Y-boundaries.

Standard equation or defining equations of the Y-parameters.

Y_{11} and Y_{22} are driving-point admittance at port 1 and port 2.

Y_{12} and Y_{21} are known as transfer admittance.

Y_{11} is known as short circuit input admittance.

Y_{12} is known as short circuit reverse transfer admittance.

Y_{21} is known as short circuit forward transfer admittance.

Similarly, Y_{22} is known as short circuit output admittance.

Hybrid Parameters

In the event that the info current I_1 and result voltage V_2 are taken as free factors, the ward factors V_1 and I_2 can be composed as

$$\begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2 \end{bmatrix}$$

NOTES

Where h_{11}, h_{12}, h_{21}, h_{22} are called as hybrid parameters.

$$h_{11} = \left.\frac{V_1}{I_1}\right|_{V_2=0}$$

Input impedence with o/p port short circuited

$$h_{12} = \left.\frac{V_1}{V_2}\right|_{I_1=0}$$

Reverse voltage transfer ratio with i/p circuited port open

$$h_{21} = \left.\frac{I_2}{I_1}\right|_{V_2=0}$$

Forward voltage transfer ratio with o/p port short circuited

$$h_{22} = \left.\frac{I_2}{V_2}\right|_{I_1=0}$$

output impedance with i/p port open circuited

The Hybrid Model for Two Port Network: Based on the definition of hybrid parameters the mathematical model for two pert networks known as H-parameter model can be developed. The hybrid equations can be written as:

$$V_1 = h_i I_1 + h_r V_2$$

$$I_2 = h_f I_1 + h_o V_2$$

(The following convenient alternative subscript notation is recommended by the **IEEE Standards**:

i =11= input　　*o* = 22 = output　　*f* = 21 = forward transfer

r = 12 = reverse transfer)

We may now utilize the four h boundaries to develop a numerical model of the gadget of shown in Figure 1.10 (a). The mixture circuit for any gadget demonstrated shown in Figure 1.10 (b). We can confirm that the model shown in Figure 1.10 (b) fulfill above conditions by composing Kirchhoff's voltage and current laws for info and result ports.

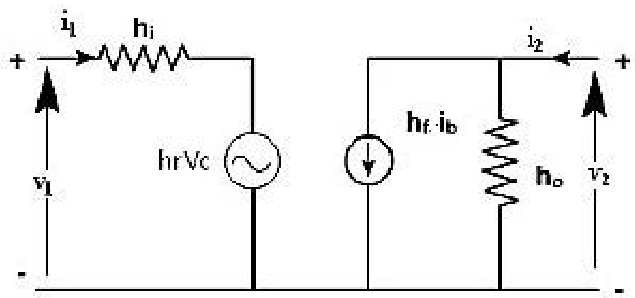


***Fig.** 1.10 (a)*

Assuming these boundaries are determined for a specific setup, then, at that point, postfixes e, b or c are additionally included, for example h_{fe}, h i_b are h boundaries of normal producer and normal authority intensifiers. Utilizing two conditions the summed-up model of the amplifier can be drawn as shown in Figure 1.10 (b).

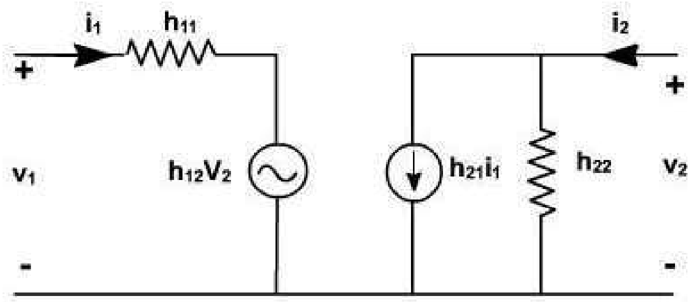


Fig. *1.10 (b)*

Fig. 1.10 *Hybrid Model for Two Port Network*

Transistor Hybrid Model: The hybrid model for a transistor amplifier can be derived as follow:

Allow us to consider C_E setup as shown in Figure 1.11 (a). The factors, I_B, I_C,V_C, and V_B address all out immediate flows and voltages I_B and V_C can be taken as autonomous factors and V_B, I_C as reliant factors.

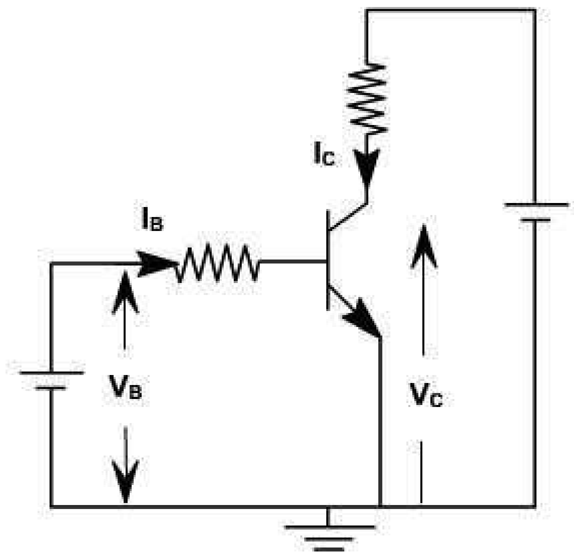


Fig. *1.11 (a)*

$V_B = V_b = f_1 (I_B ,V_C)$

$I_C = f_2 (I_B,V_C)$.

Utilizing Taylor 's series articulation, and dismissing higher request terms we get.

NOTES

$$\Delta v_B = \frac{\partial f_1}{\partial i_B}\bigg|_{v_C} \Delta i_B + \frac{\partial f_1}{\partial v_C}\bigg|_{i_B} \Delta v_C$$

$$\Delta i_C = \frac{\partial f_2}{\partial i_B}\bigg|_{v_C} \Delta i_B + \frac{\partial f_2}{\partial v_C}\bigg|_{i_B} \Delta v_C$$

The incomplete subordinates are taken keeping the authority voltage or base current consistent. The ΔV_B, ΔV_C, ΔI_B, ΔI_C address the little sign (steady) base and authority current and voltage and can be addressed as V_B, I_C, I_B, V_C The model for C_E design is shown in Figure 1.11 (b).

$$\therefore v_b = h_{ie}\, i_B + h_{re}\, v_C$$
$$i_C = h_{fe}\, i_B + h_{oe}\, v_b$$

where

$$h_{ie} = \frac{\partial f_1}{\partial i_B}\bigg|_{v_C} = \frac{\partial v_B}{\partial i_B}\bigg|_{v_C} ; \qquad h_{re} = \frac{\partial f_1}{\partial v_C}\bigg|_{i_B} = \frac{\partial v_B}{\partial v_C}\bigg|_{i_B}$$

$$h_{fe} = \frac{\partial f_2}{\partial i_B}\bigg|_{v_C} = \frac{\partial i_C}{\partial i_B}\bigg|_{v_C} ; \qquad h_{oe} = \frac{\partial f_2}{\partial v_C}\bigg|_{i_B} = \frac{\partial v_B}{\partial v_C}\bigg|_{i_B}$$

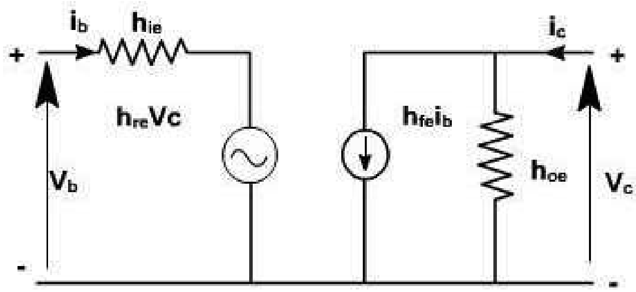


Fig. *1.11 (b)*

To decide the four H-boundaries of semiconductor intensifier, info and result trademark are utilized. Input trademark portrays the connection between input voltage and info current with yield voltage as boundary. The result trademark portrays the connection between yield voltage and result current with input current as boundary. Shown in Figure 1.11 (c) the result attributes of C_E enhancer.

$$h_{fe} = \frac{\partial i_C}{\partial i_B}\bigg|_{V_C} = \frac{i_{C2} - i_{C1}}{i_{b2} - i_{b1}}$$

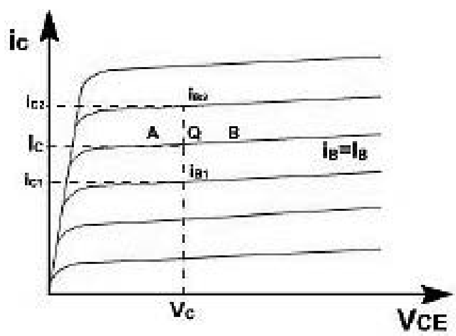


Fig. *1.11 (c)*

NOTES

The current additions are taken around the calm point Q which compares to i_B = IB and to the authority voltage $V_{CE} = V_C$

$$h_{oe} = \left.\frac{\partial i_C}{\partial V_C}\right|_{i_B}$$

The worth of cultivator at the quiet working point is given by the slant of the result trademark at the working point (for example incline of digression AB).

$$h_{ie} = \frac{\partial V_B}{\partial i_B} \approx \left.\frac{\Delta V_B}{\Delta i_B}\right|_{V_C}$$

h_{ie} is the incline of the proper contribution on shown in Figure 1.11 (d), at the working point (incline of digression EF at Q).

$$h_{re} = \frac{\partial V_B}{\partial V_C} = \left.\frac{\Delta V_B}{\Delta V_C}\right|_{I_B} = \frac{V_{B2} - V_{B1}}{V_{C2} - V_{C1}}$$

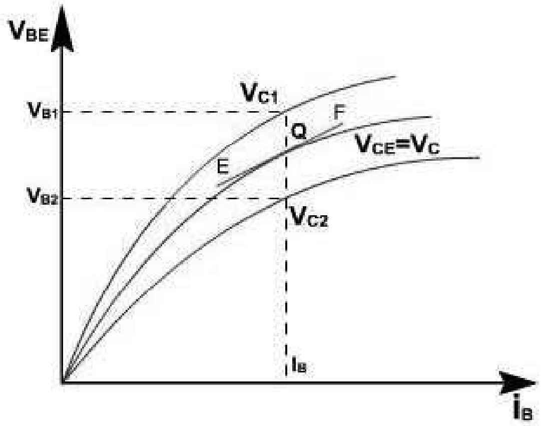


***Fig.** 1.11 (d)*

***Fig. 1.11** Transistor Hybrid Model*

An upward line on the information trademark addresses consistent base current. The boundary hre can be acquired from the proportion $(V_{B2} - V_{B1})$ and $(V_{C2} - V_{C1})$ for at Q.

1.3.2 Conversion of CB to CE and CB to CC Hybrid Parameters

Semiconductor information sheets by and large indicate the semiconductor as far as its H-boundaries for CB association for example h_{ib}, h_{fb}, h_{rb} and h_{ob}. Assuming we need to utilize the semiconductor in CE or CC setup we should change over the given arrangement of boundaries into a bunch of CE or CC boundaries. Inexact change formulae are arranged over leaf:

NOTES

Table

From CB to CE	From CE to CB	From CE to CC
$h_{ie}\ \dfrac{h_{ib}}{1\ \ h_{fb}}$	$h_{ie}\ \dfrac{h_{ie}}{1\ \ h_{fe}}$	$h_{ic} = h_{ie}$
$h_{oe}\ \dfrac{h_{ob}}{1\ \ h_{fb}}$	$h_{ob}\ \dfrac{h_{oe}}{1\ \ h_{fe}}$	$h_{oc} = h_{oe}$
$h_{fe}\ \dfrac{h_{fb}}{1\ \ h_{fb}}$	$h_{fb}\ \dfrac{h_{fe}}{1\ \ h_{fe}}$	$h_{fc} = -(1 + h_{fe})$
$h_{re}\ \dfrac{h_{ib}\,h_{ob}}{1\ \ h_{fb}}\ \ h_{rb}$	$h_{rb}\ \dfrac{h_{ie}\,h_{oe}}{1\ \ h_{fe}}\ \ h_{re}$	$h_{rc} = 1 - h_{re} \cong 1$

If the information current I_1 and result voltage V_2 are taken as free factors, the ward factors V_1 and I_2 can be composed as

Where h_{11}, h_{12}, h_{21}, h_{22} are called as crossover boundaries.

Input impedance with o/p port short circuited

Turn around voltage move proportion with I/p port open circuited

Transistor at Low Frequencies

The semiconductor can be utilized as an enhancing gadget, that is, the result ac power is more noteworthy than the information AC power. The variable that allows an air conditioner power yield more noteworthy than the information ac power is the applied DC power. The intensifier is at first one-sided for the necessary DC voltages and flows. Then, at that point, the air conditioner to be intensified is given as contribution to the intensifier. In the event that the applied ac surpasses the breaking point set by DC level, cutting of the pinnacle area will bring about the result. In this way, appropriate (loyal) intensification configuration necessitates that the DC and ac parts be touchy to one another's prerequisites and restrictions. The superposition hypothesis is pertinent for the investigation and plan of the DC and ac parts of a BJT organization, allowing the detachment of the examination of the DC and ac reactions of the framework.

BJT Transistor modeling:

- A model is a comparable circuit that addresses the AC attributes of the semiconductor.
- A model uses circuit components that inexact the conduct of the semiconductor.
- There are two models normally utilized in little sign AC investigation of a semiconductor:

  i. BC model

  ii. Hybrid equivalent model

Two port device and hybrid model:

For the cross breed comparable model, the boundaries are characterized at a working point.

NOTES

- The quantities hC,hC,hC, and h#C are called cross breed boundaries and are the parts of a little – signal comparable circuit.
- The portrayal of the mixture comparable model will start with the overall two port framework.

1.4 ANALYSIS OF TRANSISTOR AMPLIFIER USING HYBRID MODEL

To shape a semiconductor amplifier it is simply important to interface an outer burden and sign source as demonstrated shown in Figure 1.12 (a) and to inclination the semiconductor appropriately.

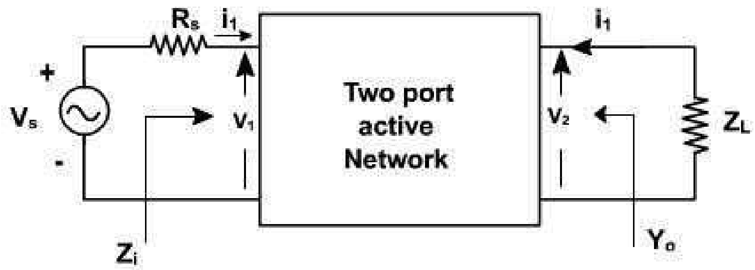


Fig. *1.12 (a)*

Think about the two-port organization of C_E amplifier. R_S is the source opposition and Z_L is the heap impedance h-boundaries are thought to be steady over the working reach. The air conditioner identical circuit is shown in Figure 1.12 (b) (Phaser documentations are utilized expecting sinusoidal voltage input). The amounts of interest are the current addition, input impedance, voltage gain, and result impedance.

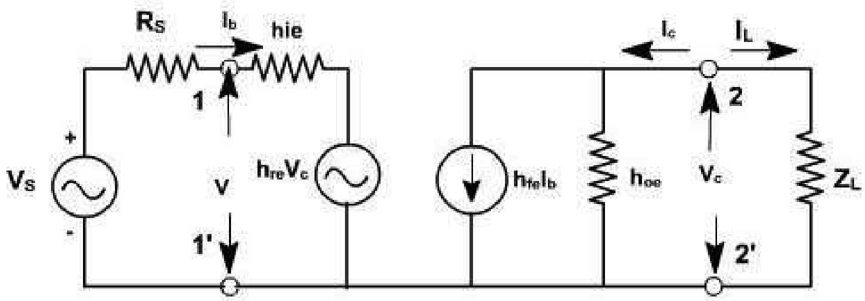


Fig. *1.12 (b)*

Current Gain: For the semiconductor amplifier stage, A_I is characterized as the proportion of result to enter flows.

$$A_I = \frac{I_L}{I_1} = \frac{-I_2}{I_1}$$

Input Impedance: The impedence investigating the enhancer input terminals (1,1') is the information impedence Z_I.

NOTES

$$Z_i = \frac{V_b}{I_b}$$

$$V_b = h_{ie} I_b + h_{re} V_c$$

$$\frac{V_b}{I_b} = h_{ie} + h_{re}\frac{V_c}{I_b}$$

$$= h_{ie} - \frac{h_{re} I_c Z_L}{I_b}$$

$$\therefore Z_i = h_{ie} + h_{re} A_I Z_L$$

$$= h_{ie} - \frac{h_{re}\ h_{fe}\ Z_L}{1 + h_{oe}\ Z_L}$$

$$\therefore Z_i = h_{ie} - \frac{h_{re}\ h_{fe}}{Y_L + h_{oe}} \qquad (\text{since } Y_L = \frac{1}{Z_L})$$

Voltage Gain: The proportion of result voltage to enter voltage gives the addition of the semiconductors.

$$A_v = \frac{V_C}{V_b} = -\frac{I_C Z_L}{V_b}$$

$$\therefore A_v = \frac{I_B A_i Z_L}{V_b} = \frac{A_i Z_L}{Z_i}$$

Output Admittance: It is defined as

$$Y_0 = \frac{I_c}{V_c}\Big|_{V_s} = 0$$

$$I_c = h_{fe} I_b + h_{oe} V_c$$

$$\frac{I_c}{V_c} = h_{fe}\frac{I_b}{V_c} + h_{oe}$$

$$\text{when } V_s = 0, \quad R_s \cdot I_b + h_{ie} \cdot I_b + h_{re} V_c = 0.$$

$$\frac{I_b}{V_c} = -\frac{h_{re}}{R_s + h_{ie}}$$

$$\therefore Y_0 = h_{oe} - \frac{h_{re}\ h_{fe}}{R_s + h_{ie}}$$

Voltage amplification taking into account source impedance (R_S) is given by

$$A_{VS} = \frac{V_c}{V_s} = \frac{V_c}{V_b} * \frac{V_b}{V_s} \qquad \left(V_b = \frac{V_s}{R_s + Z_i} * Z_i\right)$$

$$= A_V \cdot \frac{Z_i}{Z_i + R_s}$$

$$= \frac{A_i\ Z_L}{Z_i + R_s}$$

A_v is the voltage gain for an ideal voltage source ($R_v = 0$).

Consider input source to be a current source I_S in parallel with a resistance R_S as shown in Figure 1.12 (c).

NOTES

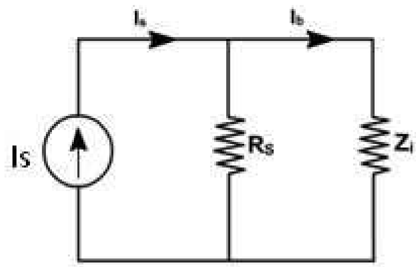


Fig. *1.12 (c)*

In this case, overall current gain A_{IS} is defined as:

$$A_{I_s} = \frac{I_L}{I_s}$$

$$= -\frac{I_c}{I_s}$$

$$= -\frac{I_c}{I_b} * \frac{I_b}{I_s} \qquad \left(I_b = \frac{I_s * R_s}{R_s + Z_i}\right)$$

$$= A_I \cdot \frac{R_s}{R_s + Z_i}$$

$$\text{If } R_s \to \infty, \quad A_{I_s} \to A_I$$

To investigate multistage amplifier the h-boundaries of the semiconductor utilized are gotten from make information sheet. The assembling information sheet generally gives H-boundary in CE design. These boundaries might be changed over into CC and CB esteems. For instance shown in Figure 1.12 (d) h_{rc} as far as CE boundary can be acquired as follows.

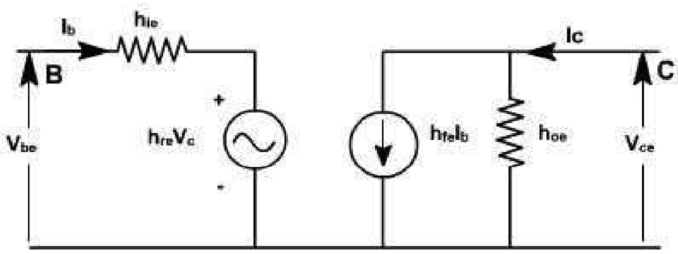


Fig. *1.12 (d)*

Fig. 1.12 *Block Diagram of a Two Port Diagram*

For C_E Transistor Configuration

Vbe = hie Ib + hre Vce

Ic = h fe Ib + hoe Vce

Typical H-parameter values for a transistor

Parameter	CE	CC	CB
hi	1100 Ω	1100 Ω	22 Ω
hr	2.5×10^{-4}	1	3×10^{-4}
hf	50	-51	-0.98
ho	25 μA/V	25 μA/V	0.49 μA/V

The rough crossover recipes for the three associations are recorded underneath. These are relevant when ho and hr is tiny and Rs is extremely huge. The given qualities allude to semiconductor terminals. The upsides of rin(stage) or rin and ro(stage) will rely upon biasing resistors and burden obstruction separately.

Item	*CE*	*CB*	*CC*
r_{in}	h_{ie}	h_{ib}	$h_{ic} + h_{fe}R_L$
r_o	$\frac{1}{h_{oc}}$	$\frac{1}{h_{oB}}$	$\frac{h_{ie}}{h_{fc}}$
A_i	$h_{fe} = \beta$	$-h_{fb} \cong 1$	$-h_{fe} \cong \beta$
A_v	$\frac{h_{ie}R_C}{h_{is}}$	$\frac{f_{fb}}{h_{ib}}R_C$	1

Common Emitter H-Parameter Analysis

The H-boundary likeness the CE circuit shown in Figure 1.13, no producer resistor has been associated.

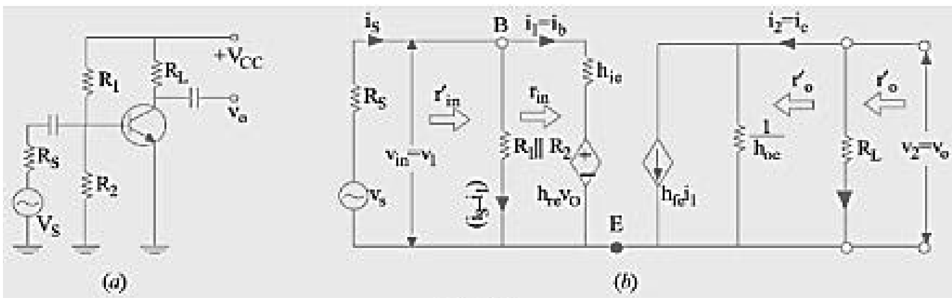


Fig. 1.13

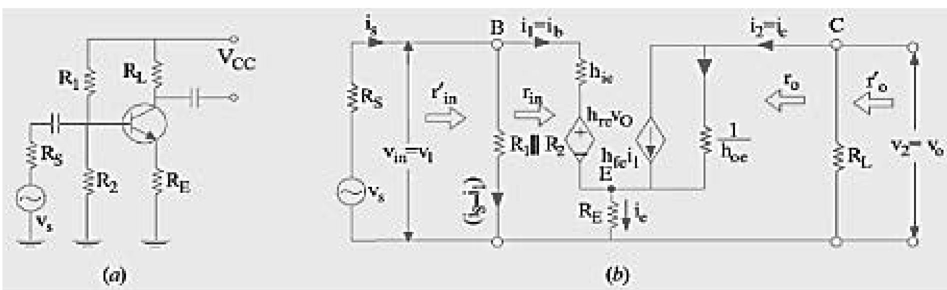


Fig. 1.14

We will presently infer articulations for voltage and current increases for both these circuits.

l. Input Impedance

NOTES

When investigating the base-producer terminals of the semiconductor, h_{ie} in series with $h_{re}V_o$. For a CE circuit, h_{re} is tiny so that $h_{re}V_o$ is unimportant when contrasted with the drop over h_{ie}. Consequently, $r_{in}=h_{ie}$.

Presently, think about the circuit shown in Figure 1.14. Again disregarding $h_{re}V_o$ we have

$$\begin{aligned} v_1 &= h_{ie}i_b + i_e R_E = h_{ie}\, i_b + (i_b + i)\, R_E \\ &= h_{ie}i_b + i_b R_E + h_{fe}\, i_b R_E \qquad ((\quad i_c \quad h_{fe}i_b) \\ &= i_b\,[h_{ie} + R_E(1+h_{fe})] \end{aligned}$$

$$\therefore \qquad r_{in} = r_{in(base)} = \frac{v_1}{i_1} \quad \frac{v_1}{i_b} \quad h_{ie} \quad (1 \quad h_{fe})R_E \ *$$

$$r_{in} \text{ or } r_{in(base)} = R_1 \,||\, R_2 \,||\, r_{in(base)}$$

2. Output Impedance

Looking back into the collector and emitter terminals of the transistor shown in Figure 1.13 (b), $r_o = 1/h_{oe}$.

As seen, r_o' or $r_{o(stage)} = r_o \,||\, R_L = (1/h_o) \,||\, r_L \qquad (\quad r_L \quad R_L)$

Since $1/h_{oe}$ is typically 1 M or so and R_L is usually much smaller, $r_o' \cong R_L = r_L$

3. Voltage Gain

$$A_v \quad \frac{v_2}{v_1} \quad \frac{v_o}{v_{in}}$$

Now, $\quad v_o = -i_c R_L$ and $\qquad v_{in} \cong i_b h_{ie}$

$$\therefore \quad A_v \quad \frac{i_c R_L}{i_b h_{ie}} \quad \frac{i_c}{i_b} \cdot \frac{R_L}{h_{ie}} \quad \frac{h_f R_L}{h_{ie}}$$

Now, consider Fig.13(b). Ignoring $h_{re}v_o$, we have from the input loop of the circuit

$$v_{in} = i_b\,[h_{ie} + R_E(1+h_{fe})] \qquad \text{–proved above}$$

$$\therefore \quad A_v \quad \frac{v_o}{v_{in}} \quad \frac{i_c R_L}{i_b[h_{ie} \quad R_E(1 \quad h_{fe})]} \quad \frac{h_{fe}R_L}{h_{fe} \quad (1 \quad h_{fe})R_E}$$

$$\frac{R_L}{R_E} \qquad \text{— if } (1+h_{fe})\,R_E >> h_{ie}$$

4. Current Gain

$$A_i \quad \frac{i_2}{i_1} \quad \frac{h_{fe}}{1 \quad h_{oe}r_L} \quad h_{fe} \qquad \text{— if } h_{oe}\, r_L << 1$$

$$A_{is} \quad \frac{h_{fe} \cdot R_1 || R_2}{r_{in} \quad R_1 || R_2}$$

NOTES

5. Power Gain

$$A_p = A_v \times A_i$$

Common Collector H-Parameter Analysis

The CC transistor circuit and its h-parameter equivalent are shown in Figure 1.15

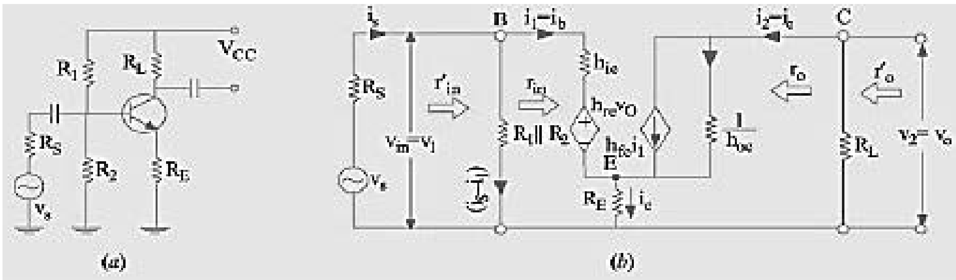


Fig. 1.15

l. Input Impedance

$$v_{in} = i_b h_{ic} + h_{rc} v_o = i_b h_{ic} + v_o = i_b h_{ic} + i_e R_L$$

$$= i_b h_{ic} + h_{fe} i_b R_L = i_b (h_{ic} + h_{fe} R_L) \qquad i_c = h_{fe} i_b$$

$$\therefore \qquad r_{in} \quad \frac{v_{in}}{i_b} \quad h_{ic} \quad h_{fe} R_L$$

As seen, $r_{in(stage)} = r_{in(base)} \| R_1 \| R_2 = r_{in(base)} \| R_B$ where $R_B = R_1 \| R_2$

2. Output Impedance

$$r_o \quad \frac{v_2}{i_2}\Big|_{v_s \; 0} \qquad \frac{v_o}{i_c}\Big|_{v_s \; 0}$$

Now, $\quad i_e \cong i_c = h_{fe} i_b = h_{fc} i_1$

Since $\quad v_s = 0$, i_b is produced by $h_{rc} v_o = v_o$

Hence, considering the input circuit loop, we get

$$i_b \quad \frac{v_0}{h_{ic} \quad (R_S \| R_1 \| R_2)} \quad \frac{v_0}{h_{ic} \quad R_S \| R_B}$$

$$i_c \quad h_{fe}\, i_b \quad \frac{h_{fe}\, v_o}{h_{ic} \quad (R_S \| R_B)} \qquad \text{where } R_B = R_1 \| R_2$$

$$\therefore \qquad r_o \quad \frac{V_o}{i_e} \quad \frac{h_{ic} \quad (R_S \| R_1 \| R_2)}{h_{fe}}$$

Also, r_o' or $\quad r_{o(stage)} = r_o \| R_L$

3. Voltage Gain

$$A_v \quad \frac{v_2}{v_1} \quad \frac{v_o}{v_{in}}$$

Now, $\quad v_o = i_e R_L = h_{fe}\, i_b R_L$ and $i_b = (v_{in} - v_o) / h_{ic}$

$$v_o \quad \frac{h_{fe}\, R_L}{h_{ie}}(v_{in} \quad v_o) \quad \text{or} \quad v_o \; 1 \quad \frac{h_{fe} R_L}{h_{ic}} \quad \frac{h_{fe} R_L v_{in}}{h_{ic}}$$

$$\therefore \quad A_v \quad \frac{v_o}{v_{in}} \quad \frac{h_{fe} R_L / h_{ic}}{1 \quad h_{fe} R_L / h_{ic}} \quad 1$$

4. Current Gain

$$A_i \quad \frac{i_2}{i_1} \quad \frac{i_e}{i_b} \quad h_{fe}\;;\; A_{is} \quad \frac{h_{fe} R_B}{r_{in} \quad R_B} \qquad \text{where } R_B = R_1 \| R_2$$

1.4.1 R-C Coupled CE Amplifier and its Frequency Response

The Resistance Capacitance (RC) coupled amplifiers are relatively inexpensive and face from picking up undesired currents from AC heater leads. These amplifiers have good fidelity over wide frequency range and especially suited to pentode and high-triodes.

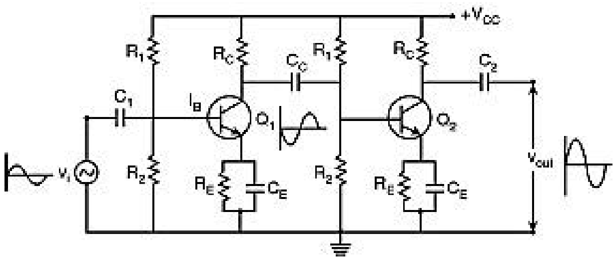


Fig. 1.16 *RC Coupled Amplifier*

Figure 1.16 shows a two stage RC coupled transistor amplifier. The circuit consists of two single stage common-emitter transistor amplifier. A coupling capacitor CC is used to connect the output of first stage of the base (input) of the second stage. The capacitor C_1 is used to couple the input signal to the base of Q_1, while the capacitor C_2 is used to couple the output signal from the collector of Q_2 to the load. The emitter capacitor CE connected at the emitters of Q_1 and Q_2 offer low reactance path to the signal. Without it, the voltage gain of each stage would be lost. The resistors R_1, R_2 and R_C form the biasing and stabilization network. Since the coupling from one stage to next is achieved by a coupling capacitor CC followed by a connection to a shunt resistor, such amplifiers are called *resistance-capacitance coupled amplifiers*.

Operation

When an AC signal is applied to the input of the first stage, it is amplified with a phase reversal by a transistor Q_1 and appears across the collector resistor R_C.

This amplified signal is given to the input of the second stage through a coupling capacitor *CC*. The second stage further amplifies the signal. The amplified output appears across R_C, in the collector of Q_2 with further phase reversal. The output signal of a two stage *RC* coupled amplifier is the twice amplified replica of the input signal. It is in phase with input signal because it has been reversed twice.

The total gain is less than the product of the gains of individual stages. It is because when a second stage is made to follow the first stage, the effective load resistance of first stage is reduced due to the shunting effect of the input resistance of second stage. This reduces the voltages gain of the stage which is loaded by the next stage.

1.4.2 Low and High Frequency Compensation

The frequency response of a typical *RC* coupled amplifier is shown in Figure 1.17. It is evident from the graph that the voltage gain drops off at low frequencies, i.e., frequencies below 50 Hz and at high frequencies, i.e, frequencies above 20 kHz, while it remains constant in the mid-frequency range, i.e., 50 Hz to 20 kHz. This behaviour of the amplifier is explained as shown in Figure 1.17:

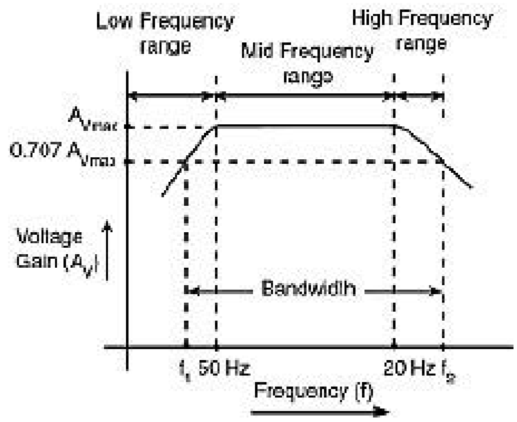


Fig. 1.17 *Frequency Response of R-C Coupled Amplifier*

1. **At low frequencies, i.e., below 50 Hz**

   At low frequencies the coupling capacitor *CC* offers a high reactance (X_C). Hence it will allow only a small part of the signal to pass from one stage to the next stage. In addition to this, the emitter bypass capacitor *CE* cannot shunt the emitter resistor R_E effectively, because of large reactance at low frequencies. As a result of these two factors, the voltage gain drops off at high frequencies.

2. **At high frequencies, i.e., above 20 kHz**

   At high frequencies, the coupling capacitor *CC* after a low reactance (X_C) and it acts as a short circuit. As a result of this, the loading effect of the next stage increases, which reduces the voltage gain. Moreover, at high frequencies, capacitive reactance of base-emitter junction is low which increases the base current. This in turn reduces the current amplification factor β. As a result of these two factors, the voltages gain drops off at high frequencies.

3. **At mid frequencies, i.e., 50 Hz to 20 kHz**

   In the mid frequency range, i.e., from 50 Hz to 20 kHz the effect of coupling capacitor is such that it maintains a constant voltage gain. Thus as the frequency increases, the reactance of capacitor *CC* decreases, which tends to increase the gain. However, at the same time, the reduces. These two factors cancel each other. Thus a constant voltage gain is maintained throughout the mid frequency range.

1.4.3 Concepts of Gain and Bandwidth

Bandwidth of an Amplifier

The limit is set at those frequencies at which the voltage gain reduces to 70.7% of the maximum voltage gain A_{v_m}. These frequencies are known as *cut-off frequencies* of the amplifier and are marked as f_1 and f_2 as shown in Figure 1.17. The frequency f_1 is the *lower cut-off frequency* and the frequency f_2 is the *upper cut-off frequency*. The difference of the two frequencies, that is f_2–f_1 is called the bandwidth of the amplifier. The mid-frequency range of the amplifier is from f_1 to f_2.

$$\textbf{Bandwidth} = (f_2 - f_1)\ \textbf{Hz}$$

Advantages

Advantages of amplifiers are as follows:

1. It requires cheap components like resistors and capacitors. Hence, it is small, light and inexpensive.
2. It has a wide frequency response. The gain is constant over the audio frequency range which is the region of most important for speech, music etc.
3. It provides less frequency distortion.
4. Its overall amplification is higher than that of the other couplings.

Disadvantages

Disadvantages of amplifiers are as follows:

1. The overall gain of the amplifier is comparatively small because of the loading effect of successive stages.
2. *RC* coupled amplifiers have tendency to become noisy with age, especially in moist climates.
3. The impedance matching is poor as the output impedance of *RC* coupled amplifier is several hundred ohms whereas that of a speaker is only a few ohms. Hence, small amount of power will be transferred to the speaker.

Applications

As *RC* coupled amplifiers have excellent audio frequency fidelity over a wide range of frequency, i.e., 50 Hz to 20 kHz, they are widely used as voltage amplifiers. This properly makes it very useful in the initial stages of public address system. However, it may be noted that a coupled amplifier cannot be used as a final stage of the amplifier because of its poor impedance matching characteristics.

NOTES

Advantages of *RC* Coupled Amplifiers Over Single-Stage Amplifier

Advantages of RC coupled amplifiers over single-stage amplifier are as follows:

1. The frequency response of *RC* coupled amplifier is much better over the audio frequency range.
2. The overall amplification is higher than that of single stage amplifier.
3. RC coupled amplifier has maximum nonlinear distortion.
4. It has an excellent audio fidelity over a wide range of frequency.

1.4.4 Cascade Stages

An amplifier is the basic building block of most electronic systems. A single stage amplifier cannot supply enough signal output. For instance, the RF signal at the antenna of a radio receiver is generally in microvolts or millivolts. Audio signal required for a microphone, phonograph or tape recorder is in the order of millivolt range. The voltage or current needed to operate a loudspeaker is, however, much greater than the signal input in the amplifier. The louder the sound we want to hear, the greater audio power output needed.

A single-stage that operates with a low-level signal does not have enough output power. Hence, two or more single-stage of amplification is often employed to achieve greater voltage or current amplification or both. Such an amplifier is called a *multistage amplifier*. Much higher gains can be obtained from the multi-stage amplifiers.

A multi-stage amplifier, using two or more single stage amplifiers is called a *cascade amplifier*. The term cascade denotes the type of connection used for coupling two amplifier stages. A mulit-stage amplifier of *n* stages can be represented as shown in Figure 1.18. Hence the output of the first stage makes the input of the second stage; the output of the second stage makes the input of the third stage and so on.

If $A_{v_1}, A_{v_2}, A_{v_3}, ..., A_{v_n}$ are the individual stage voltage gains, then overall voltage gain is given by

$$A_v = A_{v_1} \times A_{v_2} \times A_{v_3} \times .... \times A_{v_{n-1}} \times A_{v_n}$$

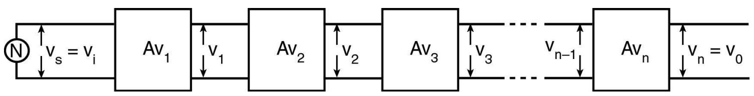


Fig. 1.18 Block Diagram of a Multi-Stage Amplifier Having n Stages

The voltage gain of a single-stage amplifier on the dB scale is given by

$$\textbf{Gain in dB} = 20 \log_{10} \left(\frac{v_0}{v_1}\right)$$

The overall voltage gain of a multistage amplifier is the sum of the decibel voltage gains of the individual stages. That is,

Overall voltage gain in dB = $20 \log_{10} Av$

or $20\log_{10} A_v = 20\log_{10} A_{v_1} + 20\log_{10} A_{v_2} + 20\log_{10} A_{v_3} + ...$

or $\boldsymbol{A_{v\,dB} = A_{vl\,dB} + A_{v2\,dB} + ... + A_{v\,n-1\,dB} + A_{vn\,dB}}$

NOTES

All amplifiers need some kind of coupling network. Even a single-stage amplifier needs coupling to the input source and output load. The multi-stage amplifiers need coupling between their individual stages. This type of coupling is known as *interstage coupling*. It serves the following purposes:

1. It transfers in *AC* coupled amplifier output of one stage to the input of the next stage.
2. It isolates the *DC* conditions of one stage to the next.
2. The coupling network forms a part of the load impedance of the proceeding stage.

Check Your Progress

13. Give a short comparison between Thevenin's and Norton equivalents.
14. What is the admittance parameters?
15. How will you define the conversion of CB to CE and CB to CC hybrid parameters?
16. Define the resistance capacitance.
17. What do you mean by the low and high frequency compensation?
18. Give the two advantages and disadvantages of amplifiers.
19. Define the term multistage amplifier.

1.5 ANSWERS TO 'CHECK YOUR PROGRESS'

1. In immersion locale, both the authority base district and the producer base area are in forward one-sided and substantial current course through the intersection. Furthermore, the area wherein both the intersections of the semiconductor are in switched one-sided is known as the remove locale.
2. The authority current for the C_E circuit is given by $I_C = \beta I_B + (1+\beta) I_{CO}$. The three factors in the situation, β, I_B, and I_{CO} increment with ascend in temperature.
3. Biasing is a course of setting DC working voltage or current to an ideal level across a semiconductor with the goal that AC information can be amplified effectively. It is needed to actuate transistor and keep it from entering saturation or cut-off region.
4. To deliver mutilation free result in amplifier circuits, the inventory voltages and protections build up a bunch of DC voltage V_{CEQ} and I_{CQ} to work the semiconductor in the dynamic area. These voltages and flows are called peaceful qualities which decide the working point or Q point for the semiconductor.
5. The single power hotspot (for instance, a battery) is utilized for both gatherer and base of a semiconductor, albeit separate batteries can likewise be utilized.

NOTES

6. One utilization of fixed inclination is to accomplish rough programmed gain control in the semiconductor by taking care of the base resistor from a DC signal got from the AC result of a later stage.
7. The authority to base predisposition circuit is same as base inclination circuit with the exception of that the base resistor R_B is gotten back to gatherer, rather than to V_{CC} supply as displayed in the figure beneath.
8. The base resistor R_B has its one end associated with base and the other to the gatherer as its name infers. In this circuit, the zero-signal base current is controlled by V_{CB} yet not by V_{CC}.
9. Advantages
   - The circuit is straightforward as it needs just a single resistor.
   - This circuit gives some adjustment, for lesser changes.

   Disadvantages
   - The circuit doesn't give great adjustment.
   - The circuit gives negative criticism.
10. The name voltage divider comes from the voltage divider shaped by R_1 and R_2. The voltage drop across R_2 forward inclinations the base-producer intersection. This causes the base current and subsequently gatherer current stream in the zero sign conditions.
11. The condition for Stability component of this circuit is gotten as stability factor = S = (β + 1) (R_0 + R_3) / R_0 + R_E + βR_E = (β + 1) × 1 + R_0R_E β + 1 + R_0R_E

    Where R_0 = $R_1R_2R_1$ + R_2. Assuming the proportion R_0/R_E is tiny, R_0/R_E can be ignored when contrasted with 1 and the dependability factor becomes

    Stability Factor = S = (β+1) × 1β+1=1.
12. There are two kinds of diode remuneration techniques. They are:
    - Diode pay for shakiness because of V_{BE} variety
    - Diode remuneration for flimsiness because of I_{CO} variety
13. Thévenin Equivalent: A single voltage source and a series impedance can be used to substitute any linear two-terminal circuit.

    Norton Equivalent: A current source and a parallel impedance can be used to substitute any linear two-terminal circuit.
14. Y boundaries (otherwise called induction boundaries or impede) are properties utilized in electrical designing to depict the electrical conduct of straight electrical organizations. These Y-boundaries are utilized in Y-grids (induction frameworks) to compute the approaching and active voltages and flows of an organization.
15. Semiconductor information sheets by and large indicate the semiconductor as far as its H-boundaries for CB association for example h_{ib}, h_{fb}, h_{rb} and h_{ob}. Assuming we need to utilize the semiconductor in CE or CC setup we should change over the given arrangement of boundaries into a bunch of CE or CC boundaries.

16. The Resistance Capacitance (RC) coupled amplifiers are relatively inexpensive and face from picking up undesired currents from *ac* heater leads. These amplifiers have good fidelity over wide frequency range and especially suited to pentode and high-triodes.
17. At low frequencies the coupling capacitor CC offers a high reactance (X_C). Hence it will allow only a small part of the signal to pass from one stage to the next stage.

    At high frequencies, the coupling capacitor CC after a low reactance (X_C) and it acts as a short circuit. As a result of this, the loading effect of the next stage increases, which reduces the voltage gain.
18. Advantages of amplifiers are as follows:
    - It requires cheap components like resistors and capacitors. Hence, it is small, light and inexpensive.
    - It has a wide frequency response. The gain is constant over the audio frequency range which is the region of most important for speech, music etc.

    Disadvantages of amplifiers are as follows:
    - The overall gain of the amplifier is comparatively small because of the loading effect of successive stages.
    - RC coupled amplifiers have tendency to become noisy with age, especially in moist climates
19. A single-stage that operates with a low-level signal does not have enough output power. Hence, two or more single-stage of amplification is often employed to achieve greater voltage or current amplification or both. Such an amplifier is called a multistage amplifier.

1.6 SUMMARY

- In operating point the working point or Q-point in a semiconductor is controlled by the upsides of the I_C (gatherer current) or V_{CE} (authority producer voltage) when no sign is applied to the information. At the point when no sign is given to the info, changes in I_C (gatherer current) and V_{CE} (authority producer voltage) happen around this point, consequently the name working point.
- The authority current for the CE circuit is given by $I_C = \beta I_B + (1+\beta)\, I_{CO}$. The three factors in the situation, β, I_B, and I_{CO} increment with ascend in temperature.
- Negative criticism can be incorporated into the biasing circuit so that expanded authority current prompts diminished base current. Thus, the expanding authority current chokes its source.
- Heat sinks can be utilized that divert additional hotness and keep the base producer temperature from rising.
- Biasing is a course of setting DC working voltage or current to an ideal level across a semiconductor with the goal that AC information can be amplified

NOTES

effectively. It is needed to actuate transistor and keep it from entering saturation or cut-off region.

- The single power hotspot (for instance, a battery) is utilized for both gatherer and base of a semiconductor, albeit separate batteries can likewise be utilized.
- The authority to base predisposition circuit is same as base inclination circuit with the exception of that the base resistor R_B is gotten back to gatherer, rather than to V_{CC} supply as displayed in the figure beneath.
- The base resistor R_B has its one end associated with base and the other to the gatherer as its name infers. In this circuit, the zero-signal base current is controlled by V_{CB} yet not by V_{CC}.
- In a Silicon semiconductor, the progressions in the worth of V_{BE} results in the progressions in I_C. A diode can be utilized in the producer circuit to remunerate the varieties in V_{BE} or I_{CO}.
- An analogous circuit is a theoretical circuit that retains all of the electrical characteristics of a particular circuit in electrical engineering and research.
- A single voltage source and a series impedance can be used to substitute any linear two-terminal circuit.
- A current source and a parallel impedance can be used to substitute any linear two-terminal circuit.
- Y boundaries (otherwise called induction boundaries or impede) are properties utilized in electrical designing to depict the electrical conduct of straight electrical organizations.
- The hybrid model for two port network based on the definition of hybrid parameters the mathematical model for two pert networks known as H-parameter model can be developed.
- Semiconductor information sheets by and large indicate the semiconductor as far as its H-boundaries for CB association for example h_{ib}, h_{fb}, h_{rb} and h_{ob}.
- The semiconductor can be utilized as an enhancing gadget, that is, the result ac power is more noteworthy than the information ac power. The variable that allows an air conditioner power yield more noteworthy than the information ac power is the applied DC power. The intensifier is at first one-sided for the necessary DC voltages and flows.
- A single-stage that operates with a low-level signal does not have enough output power. Hence, two or more single-stage of amplification is often employed to achieve greater voltage or current amplification or both. Such an amplifier is called a multistage amplifier.
- A multi-stage amplifier, using two or more single stage amplifiers is called a cascade amplifier.

1.7KEY TERMS

- **Operating point:** Operating point is a specific point within the operation characteristic of a technical device. This point will be engaged because of the properties of the system and the outside influences and parameters.

- **Thermal runaway:** The authority current for the CE circuit is given by $I_C = \beta I_B + (1+\beta)\, I_{CO}$.
- **Biasing:** Biasing is a course of setting DC working voltage or current to an ideal level across a semiconductor with the goal that AC information can be amplified effectively. It is needed to actuate transistor and keep it from entering saturation or cut-off region.
- **Fixed Bias (base bias):** The single power hotspot (for instance, a battery) is utilized for both gatherer and base of a semiconductor, albeit separate batteries can likewise be utilized.
- **Current Gain:** For the semiconductor amplifier stage, A_I is characterized as the proportion of result to enter flows.

$$A_I = \frac{I_L}{I_1} = \frac{-I_2}{I_1}$$

- **Thévenin Equivalent:** A single voltage source and a series impedance can be used to substitute any linear two-terminal circuit.
- **Norton Equivalent:** A current source and a parallel impedance can be used to substitute any linear two-terminal circuit.
- **Cascade amplifier:** It is a multi-stage amplifier, using two or more single stage amplifiers.

1.8 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What do you mean by the operating point?
2. State the thermal runaways and thermal stability.
3. Define the term biasing.
4. Give two advantages of biasing with collector feedback resistor.
5. Name the two kinds of diode remuneration techniques.
6. What is transistor equivalent circuit?
7. Define the term hybrid parameters.
8. How will you define the current gain?
9. State the low and high frequency compensation.
10. What is cascade stage?

Long-Answer Questions

1. Explain briefly about the operating point with the help of giving examples.
2. Differentiate between the thermal runaways and thermal stability. Give appropriate examples.
3. What do you understand by the biasing? Discuss the four common biasing methods or techniques to bias the transistor.

4. Briefly explain about the biasing with collector feedback resistor with the help of giving examples.
5. Analysis the stabilization against variation in V_{BE} and B-Bias compensation. Give appropriate examples.
6. Illustrate the transistor equivalent circuit with the help of giving examples.
7. Discuss briefly about the V (Admittance) and hybrid parameters. Give appropriate examples.
8. Give the conversion of CB to CE and CB to CC hybrid parameters. Give appropriate examples.
9. Briefly explain about the analysis of transistor amplifier using hybrid model with the help of giving example.
10. Discuss briefly about the low and high frequencies with the help of giving examples. Give the advantages and disadvantages of bandwidth of an amplifier.
11. Explain briefly about the cascade stages with the help of giving examples.

1.9 FURTHER READING

Basavaraj, B. and H.N. Shivashankar. 2004. *Basic Electronics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Black, Uyless D. 1993. *Computer Networks.* New Jersey: Prentice-Hall.

Kumar, A. Anand. 2003. *Fundamentals of Digital Circuits*. New Delhi: Prentice-Hall of India.

Mano, Morris. 1979. *Digital Logic and Computer Design*. New Delhi: Prentice-Hall of India.

Roth, Charles. 2001. *Fundamentals of Logic Design*. Thomson Learning.

Shepherd, Robert D. 1998. *Introduction to Computers and Technology*. EMC/ Paradigm Publishing.

Stallings, William. 1996. *Data and Computer Communications.* New Jersey: Prentice-Hall.

Tanenbaum, Andrew. 2002. *Computer Networks.* New Jersey: Prentice-Hall PTR.

Yarbarough, John M. 1996. *Digital Logic Applications and Design*. Thomson Learning.

Carlson Bruce and Paul B. Crilly. 2010. Communication System. Noida: Tata McGraw-Hill Education.

Lathi B. P. and Zhi Ding. 2010. Modern Digital and Analog Communication Systems. Oxford: Oxford University Press.

Kennedy George and Bernard Davis. 1993. Electronic Communication Systems. London: Glencoe/ McGraw-Hill.

Roddy Dennis and John Coolen. 1977. Electronic Communications. Noida: Pearson Education.

UNIT 2 FEEDBACK CIRCUITS

NOTES

Structure

2.0 INTRODUCTION

Feedback occurs when outputs of a system are routed back as inputs as part of a chain of cause and-effect that forms a circuit or loop. The system can then be said to feed back into itself. A feedback amplifier is also known as closed-loop amplifier ads a closed loop is formed between the input and the output of the amplifier. A Negative-feedback amplifier (or feedback amplifier) is an electronic amplifier that subtracts a fraction of its output from its input, so that negative feedback opposes the original signal. The applied negative feedback can improve its performance (gain stability, linearity, frequency response, step response) and reduces sensitivity to parameter variations due to manufacturing or environment. Because of these advantages, many amplifiers and control systems use negative feedback.

In feedback amplifiers, the emitter follower circuit plays a significant role. A negative current feedback circuit is an emitter follower. In signal generating circuits, this is typically employed as a last stage amplifier. A circuit that creates a continuous, repetitive, alternating waveform without any input is known as an oscillator. Oscillators are devices that convert unidirectional current flow from a DC source into an alternating waveform with the required frequency set by the circuit components. A Colpitts oscillator, invented in 1918 by American engineer Edwin H. Colpitts, is one of a number of designs for LC oscillators, electronic oscillators that use a combination of inductors (L) and capacitors (C) to produce an oscillation at a certain frequency. The distinguishing feature of the Colpitts oscillator is that the feedback for the active device is taken from a voltage divider made of two capacitors in series across the inductor

NOTES

A Uni Junction Transistor (UJT) is a three-terminal semiconductor exchanging gadget. This gadget has a novel trademark that when it is set off, the producer current increments regeneratively until it is restricted by producer power supply. Because of this trademark, the unijunction semiconductor can be utilized in an assortment of utilizations.

A multivibrator is an electronic circuit used to implement a variety of simple two-state devices such as relaxation oscillators, timers, and flip-flops. It consists of two amplifying devices (transistors, vacuum tubes, or other devices) cross-coupled by resistors or capacitors. The first multivibrator circuit, the astable multivibrator oscillator, was invented by Henri Abraham and Eugene Bloch during World War I. They called their circuit a 'Multivibrator' because its output waveform was rich in harmonics.

In this unit you will learn about the feedback in amplifiers, negative feedback and gain stability effect of feedback on input and output impedances and distortions, current and voltage feedback circuits, emitter follower, circuits and working of Hartley, Colpitts, and phase shift oscillators, UJT and its characteristics, transistor as a switch, astable, monostable and bistable miltivibrator.

2.1 OBJECTIVES

After going through this unit, you will be able to:

- Define feedback in amplifiers
- Explain the negative feedback and gain stability effect of feedback on input and output impedances and distortions
- Discuss about the current and voltage feedback circuits
- Understand the basic concept of emitter follower
- Analysis the circuits and working of Hartley, Colpitts, and phase shift oscillators
- Elaborate on the UJT and its characteristics
- Interpret the transistor as a switch
- Learn about the astable, monostable and bistable miltivibrator

2.2 FEEDBACK IN AMPLIFIERS

The basic block diagram of a feedback amplifier is shown in Figure 2.1. A feed-back amplifier is also known as closed-loop amplifier as a closed loop is formed between the input and the output of the amplifier. Essentially two blocks are there:

1. A basic or internal amplifier
2. A feedback circuit.

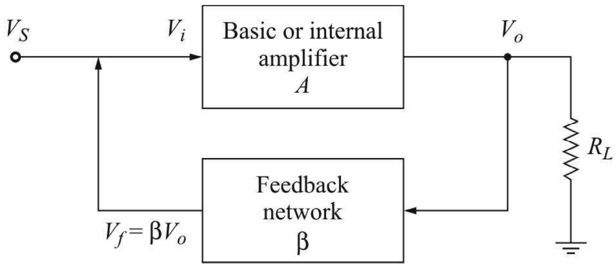


Fig. 2.1 *Block diagram of a feedback amplifier*

NOTES

In a feedback amplifier, the output signal is sampled and feedback to the input to form an error signal that drives the amplifier.

In order to derive the basic feedback equation, certain assumptions are made:

1. Reverse transmission from the amplifier output to the input is zero, and
2. Forward transmission through the feedback network is zero.

The feedback network contains passive circuit elements like resistors, inductors or capacitors and active elements like transistors.

Open loop gain or Transfer gain of the Internal Amplifier: The transfer gain of the basic or internal amplifier is

$$A = \frac{V_O}{V_i}$$

where V_i is the input voltage and V_O is the output voltage of the amplifier.

Feedback Factor

Feedback factor is also known as feedback ratio or reverse transfer ratio or reverse transmission factor and this determines the fraction of the output signal that is added to or subtracted from the externally applied input signal voltage V_S by the feedback network. This is given by $\beta = \frac{V_f}{V_O}$ where the feedback voltage $V_f = \beta V_O$.

The input voltage of the internal amplifier is thus

$$V_i = V_S \pm V_f$$

The positive sign holds for positive feedback and the negative sign for negative feedback.

Closed-loop gain or Transfer gain of the Feedback Amplifier: Closed-loop gain of the feedback amplifier is the ratio of the output voltage V_O and the external input signal voltage V_S and is given by

$$A_f = \frac{V_O}{V_S} \quad ...(2.1)$$

If $V_i = V_S - V_f$, i.e., for negative feedback and considering open loop, we have,

$$V_O = AV_i = A(V_S - V_f) \quad ...(2.2)$$

Now $V_f = \beta V_O$, so from equation (2.2) we get

$$V_o = A(V_S - \beta V_O)$$

or $$V_O + A\beta V_O = AV_S$$

or $$V_O(1 + A\beta) = AV_S$$

Now, $$A_f = \frac{V_O}{V_S} = \frac{A}{1 + A\beta} \quad ...(2.3)$$

NOTES

If we consider open loop, i.e., in the absence of any feedback $\beta = 0$.

$\therefore \quad A_f = A$

Thus, A is the transfer gain without feedback.

Loop Gain: Loop gain is given by the term $-A\beta$ as it gives the product of the gains of the branches making up the loop.

It is also called return ratio or loop transmission. The difference between unity and loop gain, i.e., $(1 + A\beta)$ is called the return difference.

Positive Feedback: Figure 2.2 shows the block diagram of a positive feedback circuit. In this circuit the input signal V_S and the feedback signal V_f are combined in the mixer network to obtain the difference signal V_i which is applied as the input voltage to the basic amplifier. If the feedback signal (V_f) is added in phase with the input signal (V_S), then the feedback is called positive feedback.

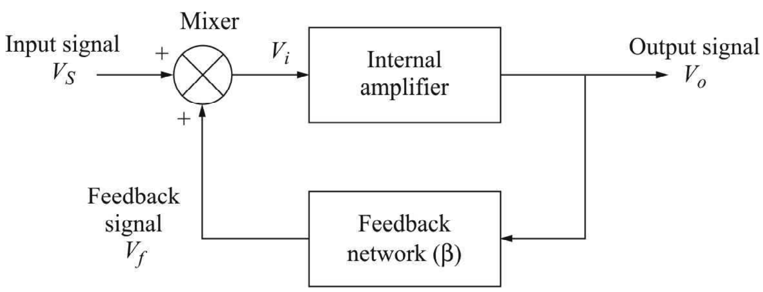


Fig. 2.2 *Block diagram of positive feedback circuit*

For positive feedback

$$|1 + A\beta| < 1, \text{ hence we get}$$
$$|A_f| > |A|$$

Negative Feedback: As shown in Figure 2.3 the block diagram of negative feedback circuit. In this type of circuit, the feedback signal (V_f) and the input signal (V_S) are out of phase with each other.

Here $|1 + A\beta| > 1$, hence we get

$$|A_f| < |A|$$

If $|1 + A\beta| = 0$, then $|A_f| = \infty$ which implies that the amplifier gives an output signal even when the input signal is not present and then the amplifier becomes an oscillator.

In decibel (dB), the amount of feedback introduced into an amplifier is given by

$$F(\text{in dB}) = 20\log_{10}\left|\frac{A_f}{A}\right| = 20\log_{10}\left|\frac{1}{1 + A\beta}\right|$$

F is negative for negative feeback since $|1 + A\beta| > 1$ and is positive for positive feedback since then $1 + A\beta| < 1$.

NOTES

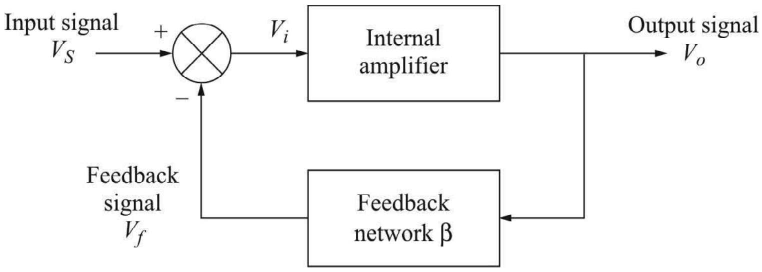


***Fig. 2.3** Block diagram of negative feedback circuit*

2.3 NEGATIVE FEEDBACK AND GAIN STABILITY EFFECT OF FEEDBACK ON INPUT, OUTPUT IMPEDANCES AND DISTORTIONS

Analysis of Negative Feedback Amplifier Circuits

Voltage-series feedback circuit: Illustration of voltage-series negative feedback connection is shown in the emitter follower circuit of shown in Figure 2.4. An emitter follower circuit with self-bias arrangement and a resistance R_C in the collector circuit is shown in Fig. 2.4.

Shown in Figure 2.4, the output voltage V_o across the emitter resistance R_E is totally fedback to the input. So, the feedback fraction β is unity here. The base resistance R_B forward biases the emitter-base junction. The capacitor C_{in} decouples the signal source from the supply voltage V_{CC} and acts like a short circuit for AC. The output voltage is the sampled signal. Since the sampled signal is a voltage, the feedback is a voltage or shunt feedback. Series mixing is obtained by returning the feedback signal voltage in series with the external input signal voltage. Hence, the feedback in the emitter follower circuit is voltage-series or voltage-voltage or shunt-series feedback.

The signal voltage V_S is the input voltage V_i. The output voltage V_o is also the feedback voltage in series with the input voltage. As $V_f = V_o$, we have,

$$\beta = \frac{V_f}{V_o} = 1 \qquad ...(2.4)$$

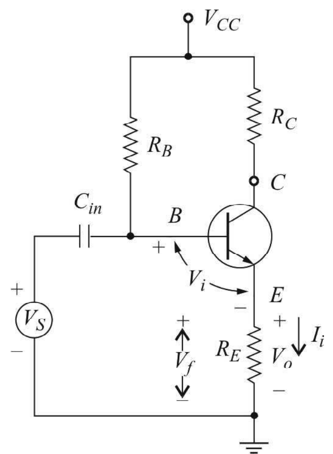


***Fig. 2.4** Emitter follower circuit*

The voltage gain of the circuit without feedback ($V_f = 0$) is given by

$$A = \frac{V_o}{V_S} = \frac{h_{fe} I_i R_E}{V_S} = \frac{h_{fe}(V_S / h_{ie}) R_E}{V_S} = \frac{h_{fe} R_E}{h_{ie}} \quad ...(2.5)$$

NOTES

Here h_{fe} is the forward current gain and $h_{fe} I_i$ is the current source. Again h_{ie} is the input resistance

$$I_i = \frac{V_s}{h_{ie}}$$

Using Equation $A_f = \frac{V_0}{V_s} = \frac{A}{1 + A\beta}$ and Equations (2.4) and (2.5), we get

$$A_f = \frac{A}{1 + A\beta} = \frac{h_{fe} R_E / h_{ie}}{1 + (h_{fe} R_E / h_{ie})} = \frac{h_{fe} R_E}{h_{ie} + h_{fe} R_E} \quad ...(2.6)$$

Here $\beta = 1$ from Equation (2.4)

For $h_{fe} R_E >> h_{ie}$, $A_f \approx 1$.

Input impedance is given by

$$Z_{if} = h_{ie} + (1 + h_{fe}) R_E \quad ...(2.7)$$

Output impedance is expressed as

$$Z_{of} = \frac{R_S + h_{ie}}{1 + h_{fe}} \quad ...(2.8)$$

where R_S is the source resistance.

Thus, input impedance increases by a factor of $(1 + h_{fe}) R_E$.

A practical voltage-series feedback amplifier circuit using a JFET is shown in Figure 2.5 below.

There is a feedback network of resistors R_1 and R_2 whose function is to feedback a part of output voltage V_o. The feedback voltage V_f is connected in series with the source voltage V_S. Their difference being the input voltage V_i. Without feedback, the amplifier gain is given by

$$A = \frac{V_o}{V_i} = -g_m R_L \quad ...(2.9)$$

where g_m is the transconductance of JFET and R_L is the load resistance which is the equivalent resistance of parallel combination of R_D, $(R_1 + R_2)$ and R_o as shown in Figure 2.5.

NOTES

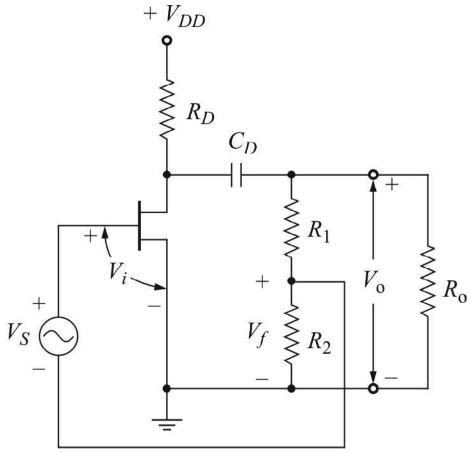


Fig. 2.5 *Voltage-series feedback circuit using JFET*

Thus $$R_L = R_D \parallel (R_1 + R_2) \parallel R_o$$

Here, $$V_f = \frac{V_o R_2}{R_1 + R_2}$$

and the feedback fraction β is given by

$$\beta = \frac{V_f}{V_o} = \frac{R_2}{R_1 + R_2} \quad ...(2.10)$$

Using Equations $A_f = \frac{V_0}{V_s} = \frac{A}{1 + A\beta}$, Equations (2.9) and (2.10), voltage gain with feedback is expressed as:

$$A_f = \frac{-g_m R_L}{1 + \left(g_m \frac{R_2 R_L}{R_1 + R_2} \right)} \quad ...(2.11)$$

Voltage-Shunt Feedback Circuit: Shown in Figure 2.6 below shows a single-stage voltage shunt feedback amplifier circuit using BJT.

Figure 2.6(*a*) shows a CE amplifier circuit where the resistor R_f connected from collector to base provides feedback. This type of arrangement is used to stabilize the operating point against temperature variations. Figure 2.6(*b*) shows the amplifier and feedback blocks separately. In the circuit, the output voltage V_o is much greater than the input voltage V_i and is 180º out of phase with V_i. Here,

$$I_f = \frac{V_i - V_o}{R_f} = -\frac{V_o}{R_f} = \beta V_o$$

$$\therefore \quad \beta = -\frac{1}{R_f} \quad ...(2.12)$$

NOTES

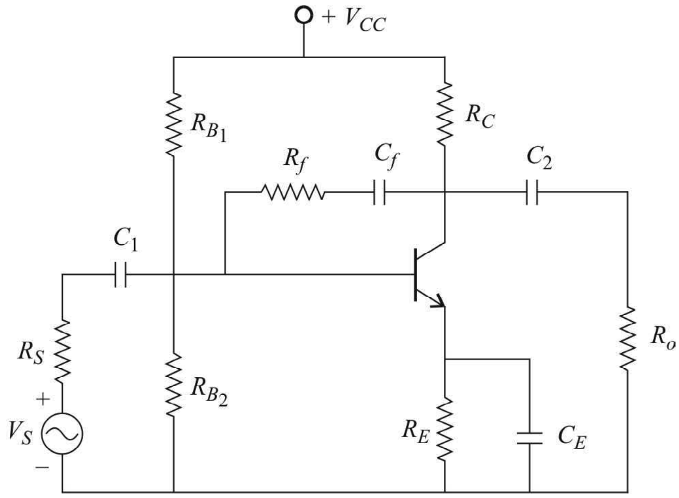


Fig. 2.6 (*a*) *Circuit diagram*

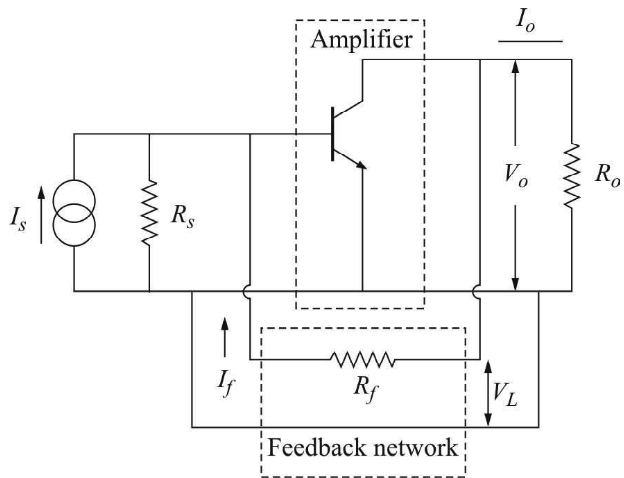


Fig. 2.6 *(b) AC equivalent circuit*

In the circuit shown in Figure 2.6(*a*), the feedback current is proportional to the output voltage, hence this is an illustration of voltage-shunt feedback amplifier. The approximate value of voltage gain A of CE amplifier without feedback is given by

$$A = -\frac{h_{fe}R_i}{h_{ie}} \quad ...(2.13)$$

where $R_L = R_o \| R_c$. Once the values of A and β are known, using above equation, the gain of voltage-shunt feedback amplifier can be obtained as

$$A_f = -\frac{h_{fe}R_L R_f}{h_{fe}R_L + R_f h_{ie}} \quad ...(2.14)$$

Current-Series Feedback Circuit: Current feedback can be established by placing a resistor in the emitter terminal of a CE amplifier as shown in Figure 2.7(*a*) below.

NOTES

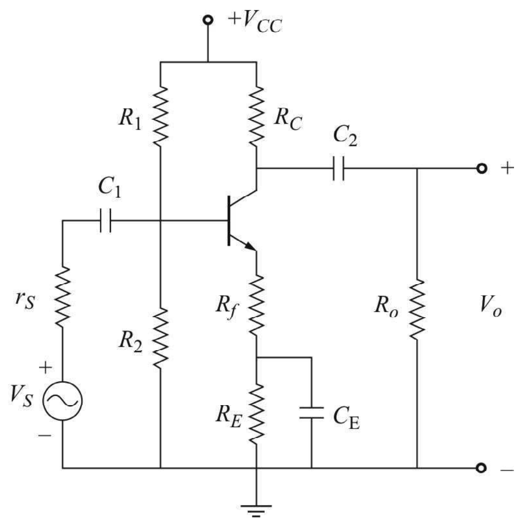


Fig. 2.7 (*a*) *Circuit diagram of current-series feedback amplifier using BJT*

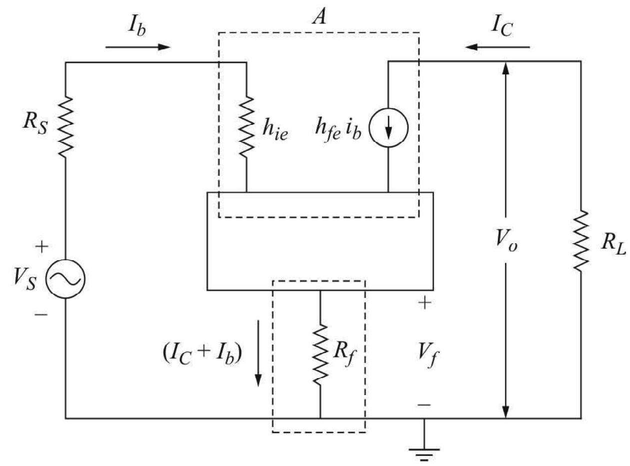


***Fig. 2.7** (b) Approximate AC equivalent circuit*

The purpose of emitter bypass capacitor placed across R_E is to provide DC bias stabilization but there is no AC feedback.

If C_E is removed, an AC voltage is developed across R_E which helps in reducing the input voltage between base and emitter and consequently the output voltage is reduced. For more flexibility in the design of bias stabilization and negative feedback, the resistor R_E consists of two parts: in one part of R_E, AC signal is bypassed through C_E and the other part of R_E that is R_f is used for AC feedback at the input.

From the simplified AC equivalent circuit shown in figure 2.7(*b*), output voltage V_o is given by

$$V_o = -I_C R_L \quad \text{where } R_L = R_C \parallel R_o$$

The feedback voltage across R_f is given by

$$V_f = (I_C + I_b)\, R_f \approx I_C R_f$$

$(\because I_b << I_c)$

Now, $V_f = \beta V_o$, so the feedback fraction β is given by

$$\beta = \frac{V_f}{V_o} = -\frac{R_f}{R_L} \quad ...(2.15)$$

NOTES

For a *CE* amplifier without feedback, the simplified expression of gain is given by Equation (2.13).

Using Equations (2.13) and (2.15), the gain with feedback is given by

$$A_f = \frac{A}{1 + A\beta}$$

$$= \frac{-h_{fe}R_L / h_{ie}}{1 + (-R_f / R_L)(-h_{fe}R_L / h_{ie})}$$

$$= -\frac{h_{fe}R_L}{h_{ie} + (h_{fe}R_f)} \quad ...(2.16)$$

Current-Shunt Feedback Circuit: An illustration of current-shunt feedback topology is shown in Figure 2.8. A current-shunt feedback amplifier is basically a current amplifier.

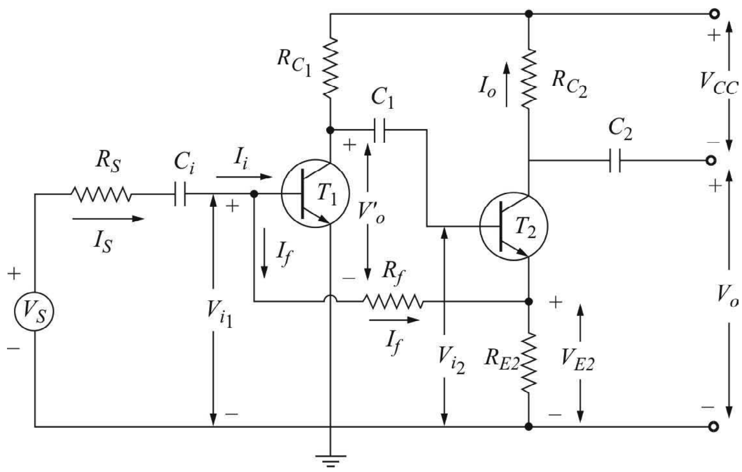


Fig. 2.8 *Current shunt feedback circuit*

In this circuit, two transistors T_1 and T_2 are arranged in cascade. The resistance R_f which is connected between the emitter of T_2 and the base of T_1 provides the feedback path.

The output voltage V_o' of the first stage is much larger than the input voltage V_{i1} and is 180° out of phase with it. The voltage V_{E2} across R_{E2} is slightly less than V_o' and is in phase with it.

In this circuit V_{i2} is much larger than V_{i1} because of voltage gain of the first CE amplifier stage. Again, due to emitter follower action, the voltage at the emitter of Q_2, (V_{E2}), is slightly smaller than V_{i2} and they are in phase. So, V_{E2} and V_{i1} are 180º out of phase with each other and also V_{E2} is greater than V_{i1} in magnitude. I_S increases due to increase of input voltage and hence I_f increases. As a result $I_i (= I_S - I_f)$ is smaller than that it would be without feedback. This action is the characteristic of negative feedback amplifier.

NOTES

The base current of T_1 is negligible compared to the collector current of T_2. Current through the feedback resistor is given by

$$I_f = \frac{V_{i1} - (-V_{E2})}{R_f} \cong \frac{V_{E2}}{R_f} \quad ...(2.17)$$

Since $V_{E2} >> V_{i1}$ and V_{E2} and V_{i1} are at 180º out of phase with each other,

$$V_{E2} = I_o R_{E2} \quad ...(2.18)$$

Using equations (2.17) and (2.18), we get

$$I_f = \frac{I_o R_{E2}}{R_f} \cong \beta I_o, \quad \text{where } \beta = \frac{R_{E2}}{R_f} \quad ...(2.19)$$

Equation (2.18) shows that the feedback current is proportional to the output current I_o. Hence, the circuit is an illustration of current-shunt feedback.

Table 2.1: *Comparison of Different Feedback Systems.*

Feedback Topology	*Voltage-series (series-shunt)*	*Voltage-shunt (shunt-shunt)*	*Current-series (series-series)*	*Current-shunt (shunt-series)*
Sampled signal	Voltage (parallel sampling)	Voltage (parallel sampling)	Current (series sampling)	Current (series sampling)
Feedback signal	Voltage (series mixing)	Current (shunt mixing)	Voltage (series mixing)	Current (shunt mixing)
Gain without feedback (A)	V_o/V_i	V_o/I_i	I_o/V_i	I_o/I_i
Feedback fraction (β)	V_f/V_o	I_f/V_o	V_f/I_o	I_f/I_o
Gain with feedback (A_f)	V_o/V_S	V_o/I_S	I_o/V_S	I_o/I_S
Input resistance (R_{if})	Increases	Decreases	Increases	Decreases
Output resistance (R_{of})	Decreases	Decreases	Increases	Increases
Bandwidth	Increases	Increases	Increases	Increases
Distortion and noise	Decreases	Decreases	Decreases	Decreases
Amplifier type	Voltage amplifier	Transresistance amplifier	Transconduc-tance ampflier	Current amplifier

Effects of Negative Feedback: The effects of negative feedback on the amplifier characteristics are the following:

1. Gain is reduced and stabilized with respect to the variations in transistor parameters like h_{fe}.
2. Non-linear distortion becomes less resulting in improvement of the signal handling capacity of the amplifier.
3. Bandwidth of the amplifier increases and frequency distortion becomes less.

NOTES

4. Phase distortion decreases.
5. Input impedance can be changed by combination of the feedback signal with the externally applied input signal. Extracting the feedback signal from the output also changes the output impedance of the amplifier.
6. Noise level is lowered.

1. Gain Stability: Overall gain of the negative feedback amplifier is

$$A_f = \frac{A}{1 + A\beta} \qquad ...(2.20)$$

where $|1 + A\beta| > 1$. If $A\beta >> 1$, Equation (2.20) gives

$$A_f \approx \frac{1}{\beta} \qquad ...(2.21)$$

Thus, the feedback network mainly determines the overall gain of the amplifier.

Gain stability improves as the feedback network generally consists of stable passive elements. Supply voltage variations, changes of parameters of the active device, ageing and temperature changes will have significant effects on A instead of A_f.

If the loop gain in a negative feedback amplifier is not large, stability of the gain can be found by

$$\frac{dA_f}{A_f} = \frac{1}{1 + A\beta} \cdot \frac{dA}{A} \qquad ...(2.22)$$

Sensitivity S is defined as the ratio of the fractional change in the overall gain to the fractional change in the gain of the internal amplifier which is given as

$$S = \left(\frac{dA_f}{A_f}\right) \Big/ \left(\frac{dA}{A}\right) = \frac{1}{1 + A\beta} \qquad ...(2.23)$$

Desensitivity D is given by the reciprocal of sensitivity S, i.e.,

$$D = \frac{1}{S} = 1 + A\beta \qquad ...(2.24)$$

Equations (2.22) and (2.24) give

$$\frac{dA_f}{A_f} = \frac{1}{D}\left(\frac{dA}{A}\right) \qquad ...(2.25)$$

The percentage change in A_f is less than that in A since $|D| > 1$. Thus the overall gain stability improves.

2. Decrease in Nonlinear Distortion: The output signal gets distorted due to non-linearity in the transfer characteristic of the amplifier when a large-amplitude signal is applied to the input of the amplifier.

Let V_{OS} and V_D represent the signal and distortion component of the output voltage respectively. The input signal voltage is assumed to be sinusoidal and the

NOTES

distortion voltage to be the second harmonic. Let V_S be the applied input signal voltage. Then in the absence of feedback, $V_{OS} = AV_S$. Non-linearity of the active device gives rise to the distortion in the output voltage which is equivalent to an effective distortion voltage V_D/A applied at the input of a distortionless amplifier of gain A. Negative feedback decreases the values of both V_{OS} and V_D. To maintain V_{OS} at its previous value, input signal amplitude is increased from V_S to V_S' which is given by

$$V_S' = V_S(1 + A\beta)$$

Since the output voltage remains constant, the input voltage of the internal amplifier is $V_i = V_S$.

Let V_{Df} be the distortion component of the output voltage in the presence of negative feedback. V_{Df} can be obtained from the overall gain $\left(\frac{A}{1+A\beta}\right)$ and input distortion voltage $\left(\frac{V_D}{A}\right)$ and is given by

$$V_{Df} = \frac{V_D}{A} \times \frac{A}{1+A\beta} = \frac{V_D}{1+A\beta} \quad ...(2.26)$$

3. Effect on Bandwidth and Frequency Distortion: We know that the low frequency gain of one stage of an RC coupled amplifier is

$$A_l = \frac{A_m}{1 - j(f_l / f)} \quad ...(2.27)$$

where A_m = mid-frequency gain

f_l = lower half-power frequency

The high frequency gain is given by

$$A_h = \frac{A_m}{1 + j(f / f_h)} \quad ...(2.28)$$

where f_h = upper half-power frequency with feedback. Thus the overall gains at low and high frequencies are

$$A_{lf} = \frac{A_l}{1 + A_l\beta} \quad ...(2.29)$$

and

$$A_{hf} = \frac{A_h}{1 + A_h\beta} \quad ...(2.30)$$

From Equations (2.29) and (2.30), we obtain

$$A_{lf} = \frac{A_m}{1 + A_m\beta - j(f_l / f)} = \frac{A_{mf}}{1 - j(f_{lf} / f)} \quad ..(2.31)$$

where $A_{mf} = \frac{A_m}{(1 + A_m\beta)}$ is known as the mid-frequency gain with feedback and $f_{lf} = \frac{f_l}{(1 + A_m\beta)}$ is known as the lower half-power frequency with feedback.

NOTES

Negative feedback indicates $|1 + A_m\beta| > 1$ which further implies that $f_{lf} < f_l$. Similarly, from Equations (2.30) and (2.28) we get

$$A_{hf} = \frac{A_{mf}}{1 + j(f / f_{hf})} \quad ...(2.32)$$

where $f_{hf} = f_h (1 + A_m\beta)$ is the upper half-power frequency with negative feedback. This shows that $f_{hf} > f_h$.

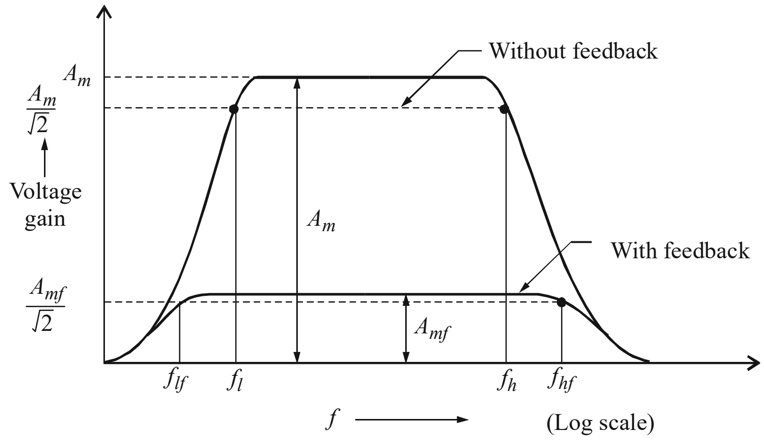


Fig. 2.9 *Plot of magnitude of voltage gain versus frequency for single stage of an RC coupled amplifier with and without feedback*

Figure 2.9 shows the plot of magnitude of the gain with frequency of an RC coupled amplifier with and without feedback.

Thus, the effect of negative feedback on the lower half-power frequency and the upper half-power frequency can be easily obtained. The lower half-power frequency decreases and the upper half-power frequency increases due to feedback.

Now bandwidth is given by the difference between the two half-power frequencies which is consequently enhanced, thus resulting in the reduction of frequency distortion.

4. Effect on Phase Distortion: The gain A of the amplifier without feedback is given by

$$A = |A| < \theta \quad ...(2.33)$$

where A is a complex quantity consisting of magnitude part $|A|$ and phase part $<\theta$. The load or the coupling impedance of the amplifier contains a reactive part. Gain with feedback is given by

$$A_f = \frac{A}{1 + A\beta} \quad ...(2.34)$$

Feedback fraction β can also be real or complex quantity. Considering β to be real for the sake of simplicity we get

$$A_f = \frac{|A| < \theta}{1 + \beta |A| < \theta}$$

$$= \frac{|A| < \theta}{1 + \beta |A| (\cos\theta + j \sin\theta)}$$

$$= \frac{|A| < \theta}{|B| < \phi} \quad ...(2.35)$$

where $|B| = \sqrt{(1+\beta|A|\cos\theta)^2 + (\beta|A|\sin\theta)^2}$ ...(2.36)

and $$\phi = \tan^{-1}\frac{\beta|A|\sin\theta}{1+\beta|A|\cos\theta}$$ ...(2.37)

NOTES

Now, A_f can be written as

$$A_f = |A_f| < \theta_f$$ ...(2.38)

hence Equation (2.35) gives

$$|A_f| = \frac{|A|}{|B|}$$ ...(2.39)

and $$\theta_f = \theta - \phi$$ ...(2.40)

Hence, the phase angle of the gain θ is decreased by angle ϕ due to negative feedback.

Thus, negative feedback reduces the phase distortion since the phase angle by which the output signal leads the input signal decreases.

5. Effect on Input and Output Impedances

(*i*) Effect on Input Impedance: Series-mixing feedback circuit as shown in Fig. (2.10) is used here.

The internal impedance of the internal amplifier is

$$Z_i = \frac{V_i}{I_i}$$ ...(2.41)

The input impedance of the feedback amplifier is

$$Z_{if} = \frac{V_S}{I_i}$$ ...(2.42)

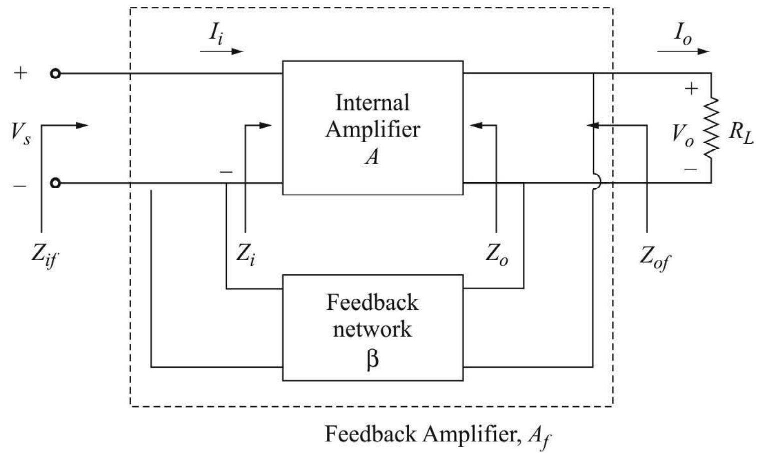


Fig. 2.10 *Voltage series feedback network used for impedance calculation*

The externally applied input signal voltage is

$$V_S = V_i + V_f = V_i + \beta V_o = V_i + A\beta V_i$$ ...(2.43)

Since $V_o = AV_i$, where A is the voltage gain of the internal amplifier including the load resistance R_L.

Equation (2.43) gives

$$V_S = V_i(1 + A\beta)$$ ...(2.44)

Hence, equations (2.42) and (2.41) give

$$Z_{if} = \frac{V_i(1 + A\beta)}{I_i} = Z_i(1 + A\beta) \qquad ...(2.45)$$

Negative feedback implies $|1 + A\beta| > 1$.

NOTES

Therefore, it is evident from Equation (2.45) that the input impedance of the amplifier increases for series-mixing negative feedback.

Let us consider the shunt-mixing feedback shown in Fig. 2.11.

From this figure, the current gain of the internal amplifier is $A = \frac{I_o}{I_i}$.

The input impedance of the internal amplifier is

$$Z_i = \frac{V_S}{I_i} \qquad ...(2.46)$$

The input impedance of the feedback amplifier is

$$Z_{if} = \frac{V_S}{I_S} \qquad ...(2.47)$$

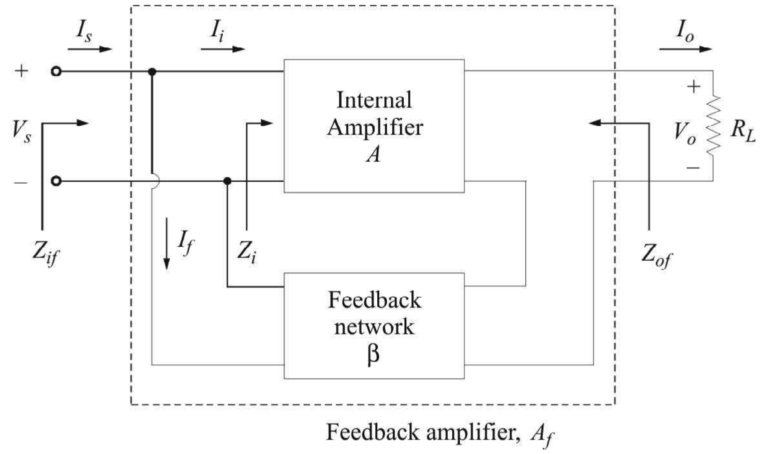


Fig. 2.11 *Calculation of impedances for current-shunt feedback network*

From Kirchoff's current law we have

$$I_S = I_i + I_f \qquad ...(2.48)$$

where I_f is the feedback current. If β is the reverse transmission factor,

$$I_f = \beta I_o = A\beta I_i \qquad ...(2.49)$$

Since $I_o = AI_i$, where A is the current gain of the internal amplifier including the load resistance R_L. Using Equations (2.49) and (2.48), we get

$$I_S = I_i(1 + A\beta) \qquad ...(2.50)$$

Equations (2.47) and (2.46) give

$$Z_{if} = \frac{V_S}{I_i(1 + A\beta)} = \frac{Z_i}{1 + A\beta}$$

The input impedance of a negative feedback amplifier with shunt mixing decreases with feedback since $|1 + A\beta|$ is greater than one for negative feedback.

6. Lowering of Output Noise: Let the noise voltage at the input of an amplifier of voltage gain A in the absence of feedback be V_n.

Then noise voltage of the amplifier output is

$$V_{on} = AV_n \qquad ...(2.51)$$

NOTES

In the presence of feedback with feedback fraction β, the gain reduces to

$$A_f = \frac{A}{1 + A\beta}$$

and then the noise voltage at the output becomes

$$V_{onf} = \frac{AV_n}{1 + A\beta} = \frac{V_{on}}{1 + A\beta} \quad ...(2.52)$$

With negative feedback $|1 + A\beta| > 1$, therefore

$$V_{onf} < V_{on}$$

Thus, the output noise level is reduced. But the signal-to-noise ratio at the amplifier output does not improve since the negative feedback reduces the signal as well as the noise by the same factor.

For negative feedback, $|1 + A\beta| > 1$, hence the distortion voltage is reduced at the output by a factor $1/(1 + A\beta)$. Signal component of the output voltage remains unchanged because the amplitude of the applied input signal is enhanced by $(1 + A\beta)$ which can be achieved by using a preamplifier. Thus, the dynamic range or the signal handling capacity of the amplifier increases.

2.4 CURRENT AND VOLTAGE FEEDBACK CIRCUITS

There are four feedback topologies:

1. Voltage-series feedback (voltage sampling-series mixing)
2. Voltage-shunt feedback (voltage sampling-shunt mixing)
3. Current-series feedback (current sampling-series mixing)
4. Current-shunt feedback (current sampling-shunt mixing)

1. **Voltage-Series (series-shunt) Feedback:** In this type of feedback topology (Refer Figure 2.12), a part of the output voltage V_o is sampled through the feedback network and is connected to the input voltage V_S in series opposition.

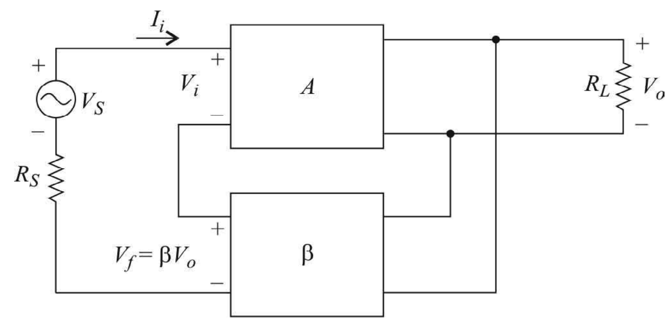


Fig. 2.12 Block diagram of voltage-series feedback network

Voltage Gain: Voltage gain of the basic amplifier in the absence of feedback (i.e., $V_f = 0$) is given by,

$$A = \frac{V_o}{V_S} = \frac{V_o}{V_i} \quad ...(2.53)$$

With the application of feedback signal, $V_f = \beta V_o$, at the input of the amplifier,

$$V_i = V_S - V_f = V_S - \beta V_o \quad ...(2.54)$$

NOTES

From Equations (2.53) and (2.54) we get,

$$V_o = AV_i = AV_S - A\beta V_o$$

Overall voltage gain with feedback A_f is obtained as

$$A_f = \frac{V_o}{V_S} = \frac{A}{1 + A\beta} \qquad ...(2.55)$$

The voltage gain thus reduces by a factor of $(1 + A\beta)$

Input Impedance: From shown in Figure 2.12, the input impedance of the amplifier without feedback (Z_i) is given by

$$Z_i = \frac{V_i}{I_i} = \frac{V_S - (A\beta V_i)}{I_i}$$

or, $$V_S = Z_i I_i + A\beta V_i = Z_i I_i + A\beta I_i Z_i$$

or, $$V_S = Z_i I_i (1 + A\beta) \qquad ...(2.56)$$

Input impedance with feedback (Z_{if}) is obtained from Equation (2.56) as

$$Z_{if} = \frac{V_S}{I_i} = Z_i(1 + A\beta) \qquad ...(2.57)$$

Thus, input impedance with series feedback is given by the value of input impedance without feedback multiplied by the factor $(1 + A\beta)$ and this applies to both voltage-series and current-series feedback topology. Higher input impedance is obtained by series mixing of the feedback signal.

Output Impedance: The output impedance can be obtained by applying a voltage source V at the output that will result in a current I with source signal shorted, i.e., $V_S = 0$.

Let Z_o be the output impedance of the amplifier in the absence of feedback, then voltage $V = IZ_o + AV_i$.

For $$V_S = 0,\ V_i = -V_f$$

Therefore, $$V = IZ_o - AV_f = IZ_o - A\beta V \qquad ...(2.58)$$

or, $$V + A\beta V = IZ_o$$

In the presence of feedback, the output impedance Z_{of} is given by

$$Z_{of} = \frac{V}{I} = \frac{Z_o}{1 + A\beta} \qquad ...(2.59)$$

Thus, the output impedance of the amplifier with feedback is reduced by a factor $(1 + A\beta)$.

The same behaviour is observed for voltage-shunt feedback amplifier. Hence, low-output impedance is observed for voltage-sampling topology. A voltage amplifier should always have high-input impedance and low-output impedance.

2. **Voltage-Shunt (shunt-shunt) Feedback:** Block diagram of voltage-shunt feedback topology is shown in Figure 2.13. Here, a part of the output voltage V_o is fed back in shunt with the input current signal I_S. For this topology, the gain of the feedback amplifier has the dimension of impedance and the

feedback fraction has the dimension of admittance as is clear from the equations for A and β given below.

Voltage-sampling shunt-mixing type feedback topology has a current signal as input and a voltage signal as output.

NOTES

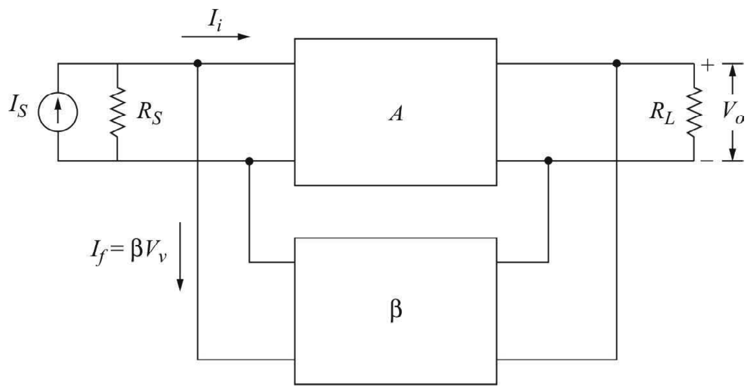


Fig. 2.13 Block diagram of voltage shunt feedback network

Here, gain without feedback

$$A = \frac{V_o}{I_i} \qquad \text{(dimension of impedance)}$$

and feedback fraction

$$\beta = \frac{I_f}{V_o}$$

(dimension of admittance)

When feedback is applied,

$$I_S = I_i + I_f = I_i + \beta V_o$$

$$I_S = I_i + A\beta I_i = I_i(1 + A\beta) \qquad \text{...(2.60)}$$

From equation (2.60), gain with feedback is given by

$$A_f = \frac{V_o}{I_S} = \frac{AI_i}{I_i(1+A\beta)} = \frac{A}{1+A\beta} \text{ (dimension of impedance) ...(2.61)}$$

Equation (2.61) implies that the amplifier gain without feedback is reduced by a factor of $(1 + A\beta)$ after the application of feedback.

Input Impedance: The input impedance (Z_{if}) of voltage-shunt feedback amplifier can be found out from shown in Figure 2.13 and Equation (2.61) as

$$Z_{if} = \frac{V_i}{I_S} = \frac{V_i}{I_i + I_f} = \frac{V_i}{I_i(1+A\beta)}$$

or

$$Z_{if} = \frac{Z_i}{1+A\beta} \qquad \text{...(2.62)}$$

Since input impedance of the amplifier without feedback is given by,

$$Z_i = \frac{V_i}{I_i}$$

Thus, input impedance of the amplifier reduces by a factor of $(1 + A\beta)$ in presence of negative feedback.

Output Impedance: Voltage-shunt feedback topology shown in Fig. 2.13 uses voltage sampling which is in shunt connection and is similar to voltage-series feedback configuration as shown in Figure 2.12. Therefore, the output impedance is given by

NOTES

$$Z_{of} = \frac{V}{I} = \frac{Z_o}{1 + A\beta} \qquad ...(2.63)$$

Thus, when negative feedback is applied, the output impedance of the amplifier reduces by a factor of $(1 + A\beta)$.

3. **Current-Series (series-series) Feedback:** The block diagram shown in Figure 2.14 shows an amplifier with current-series feedback.

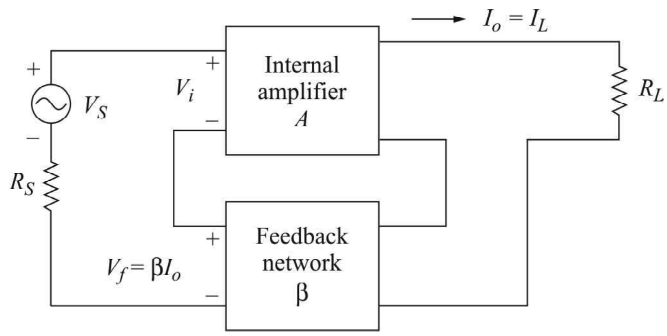


Fig. 2.14 *Block diagram of current-series feedback network*

In this topology, a sample of voltage proportional to the output current is fedback negatively to the input in series with the voltage source. Here, the input is a voltage and the output is a current, hence this topology is known as current-sampling series-mixing or series-series feedback configuration. Here, gain without feedback (A) and feedback fraction (β) are given by,

and
$$\left.\begin{aligned} A &= \frac{I_o}{V_i}\,(\text{dimension of conductance}) \\ \beta &= \frac{V_f}{I_o}\,(\text{dimension of resistance}) \end{aligned}\right\} \qquad ...\ (2.64)$$

From Figure (2.14) and Equation (2.64), the gain with feedback is given by

$$A_f = \frac{I_o}{V_S} = \frac{AV_i}{V_i + V_f} = \frac{AV_i}{V_i + A\beta V_i}$$

(dimension of conductance)

Thus, $$A_f = \frac{A}{1 + A\beta} \qquad ...\ (2.65)$$

Equation (2.65) implies that the gain is reduced by a factor of $(1 + A\beta)$ in presence of negative feedback.

Input Impedance: Current-series feedback topology uses series mixing circuit at the input similar to voltage-series feedback configuration shown in Figure 2.12. Thus the input impedance is given by,

$$Z_{if} = \frac{V_S}{I_i} = Z_i(1 + A\beta) \qquad ...(2.66)$$

Hence, input impedance of the current-series feedback configuration increases by a factor of $(1 + A\beta)$.

Output Impedance: The current-series feedback topology determines the output impedance in a similar way as used in voltage-series feedback configuration. A signal source V is applied at the output with V_S short circuited, resulting in a current I. Output impedance is given by the ratio of V to I.

NOTES

When $V_S = 0$, from as shown Figure 2.14, $V_i = V_f$

$$\therefore \qquad I = \frac{V}{Z_o} - AV_i = \frac{V}{Z_o} - AV_f = \frac{V}{Z_o} - A\beta I$$

or, $\qquad Z_o(1 + A\beta)I = V$

Thus, output impedance with feedback Z_{of} is given by

$$Z_{of} = \frac{V}{I} = Z_o(1 + A\beta) \qquad ...(2.67)$$

So, the output impedance increases by a factor of $(1 + A\beta)$.

4. **Current-Shunt (shunt-series) Feedback:** The block diagram of current-shunt feedback topology is shown in Figure 2.15. Here, the input signal is current and the feedback signal should be current so that it may be mixed in shunt with the current source. The output signal is current and the feedback network samples the current. The direction of feedback current I_f is such that it subtracts from I_S.

   The feedback network gain of the basic amplifier (A_i) and the feedback fraction (β) are given by

$$\left.\begin{aligned} A_i &= \frac{I_o}{I_i}\text{ (dimensionless)} \\ \beta &= \frac{I_f}{I_o}\text{ (dimensionless)} \end{aligned}\right\} \qquad ...(2.68)$$

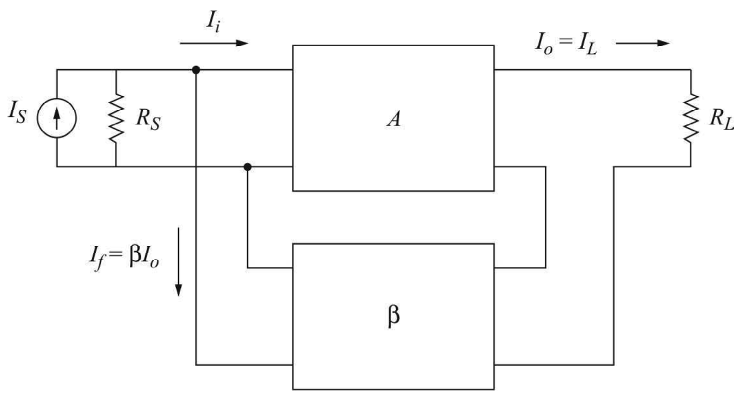


Fig. 2.15 *Block diagram of current-shunt feedback network*

Using Equation (2.68), amplifier gain with feedback A_{if} is given by

$$A_{if} = \frac{I_o}{I_S} = \frac{A_i I_i}{I_i + I_f} = \frac{A_i I_i}{I_i + A_i\beta I_i} = \frac{A_i}{1 + A_i\beta} \qquad ...(2.69)$$

Thus, the current gain of the feedback amplifier reduces by a factor of $(1 + A_i\beta)$.

Input Impedance: Shunt mixing is used in this feedback topology shown in voltage-shunt feedback configuration. Hence input impedance of this feedback topology is given by

$$Z_{if} = \frac{Z_i}{1 + A_i\beta} \qquad ...(2.70)$$

NOTES

Thus, in presence of negative feedback, the input impedance reduces by a factor of $(1 + A_i\beta)$.

Output Impedance: Current sampling is used here and so the feedback network is connected in series with the output of the basic amplifier. The series connection at the output results in increase of the output impedance with feedback by a factor $(1 + A_i\beta)$ and is given by

$$Z_{of} = Z_o(1 + A_i\beta) \qquad ...(2.71)$$

Check Your Progress

1. What is feedback in amplifier?
2. Define the term feedback factor.
3. How current-series feedback circuit can be established?
4. Give the two negative feedback on the amplifier characteristics.
5. State the gain stability.
6. What do you mean by the decrease in nonlinear distortion?
7. Give the name of four feedback topologies.
8. Write a short note on input impedance.
9. How will you define the current-series (series-series) feedback?

2.5 EMITTER FOLLOWER

In feedback amplifiers, the emitter follower circuit plays a significant role. A negative current feedback circuit is an emitter follower. In signal generating circuits, this is typically employed as a last stage amplifier.

The following are some of the most critical characteristics of an emitter follower:

- It has a high input impedance
- Its output impedance is low.
- It's a great circuit for matching impedances.

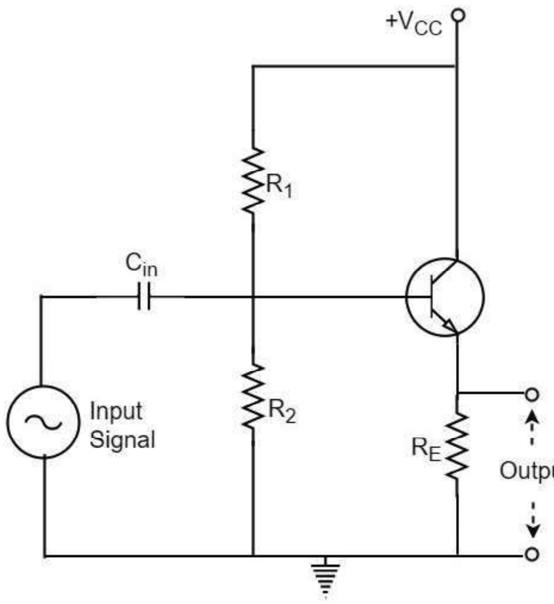
The emitter follower circuit can be used in a variety of ways thanks to all of these advantageous characteristics. This circuit is a current amplifier with no voltage gain.

Construction of Emitter Followers

An emitter follower circuit is virtually identical in construction to a standard amplifier. The key distinction is that the load R_L is not present at the collector terminal of the circuit, instead it is present at the emitter terminal. As a result, the output comes from the emitter terminal rather than the collector terminal.

Biasing is accomplished using either a base resistor or a potential divider. The circuit diagram of an emitter follower is shown in the diagram below.

NOTES



Operation of Emitter Followers

The output voltage V_o across R_E in the emitter section is generated by the input signal voltage applied between the base and the emitter.

Therefore, $V_o = I_E R_E$

Through feedback, the entire output current is applied to the input.

Hence, $V_f = V_o$

This emitter follower circuit is a current feedback circuit because the output voltage created across RL is proportional to the emitter current. As a result,

$\beta = V_{f/} V_o = 1$

It should also be remembered that the transistor's input signal voltage (V_i) is equal to the difference between V_s and V_o, i.e.

$V_i = V_s V_o$.

As a result, the feedback has been negative.

Characteristics of Emitter Follower

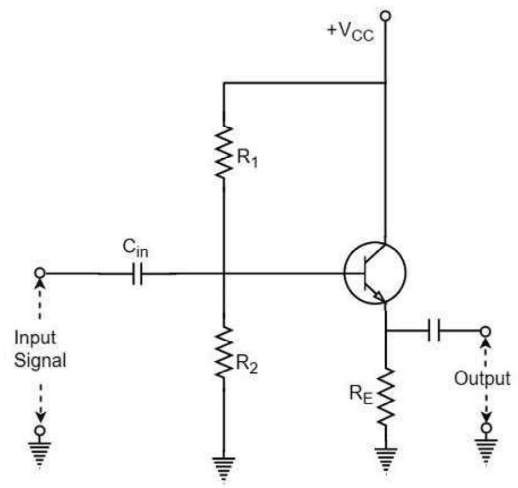
The emitter follower has the following major characteristics:

- There is no voltage gain. In fact, the voltage gain is almost a factor of one.
- Exceptionally high current and power gain.
- Low output impedance and high input impedance.
- The AC voltages at the input and output are in phase.

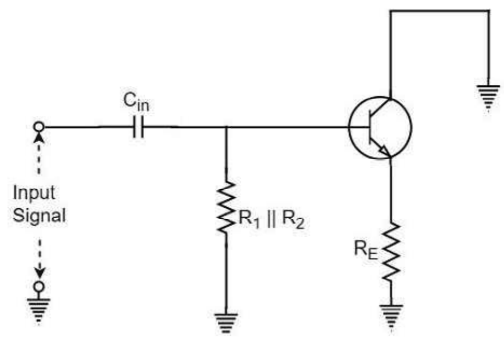
Voltage Gain of Emitter Follower

Because the Emitter Follower circuit is so common, let's try to figure out how to calculate the voltage gain of one. The circuit for our emitter follower is as follows:

NOTES



Because the emitter by pass capacitor is removed, an AC equivalent circuit of the above circuit would look like the one below.

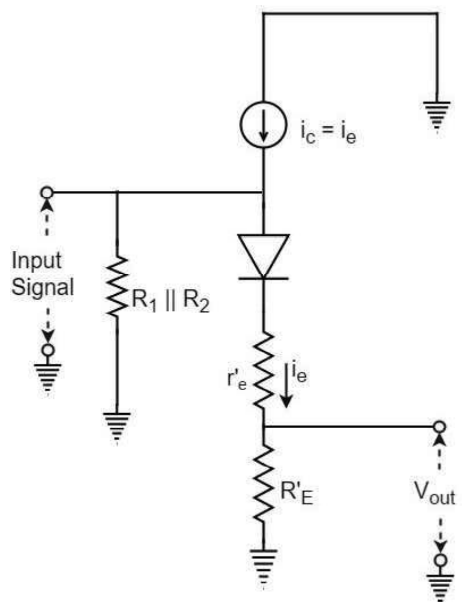


The emitter circuit's AC resistance, r_E, is calculated as

$r_E = r^2_E + R_E$.

where, $r^2_E = 25\text{mV}/I_E$

The preceding chart can be replaced with the following figure to find the amplifier's voltage and gain.



It's important to note that the input voltage is applied across the emitter circuit's ac resistance ($r'_E + R_E$). The output voltage V_{out} will be $V_{out} = i_e R_E$ assuming the emitter diode is perfect.

$V_{in} = i_e(r'_E + R_E)$ is the input voltage V_{in}.

NOTES

As a result, the emitter follower's voltage gain is.

$$A_V = V_{out} / V_{in} = i_e R_E i_e / (r'_e + R_E) = R_E / (r'_e + R_E)$$

Or

$$A_V = R_E / (R2_E + R_E)$$

Some applications are having,

$$R_E >> r2_e$$

So, $A_V \approx 1$. In practise, an emitter follower's voltage gain is between 0.8 and 0.999.

2.6 CIRCUITS AND WORKING OF HARTLEY, COLPITTS, AND PHASE SHIFT OSCILLATORS

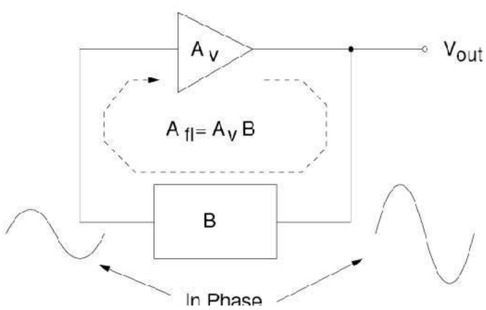
Oscillator: A circuit that creates a continuous, repetitive, alternating waveform without any input is known as an oscillator. Oscillators are devices that convert unidirectional current flow from a DC source into an alternating waveform with the required frequency set by the circuit components.

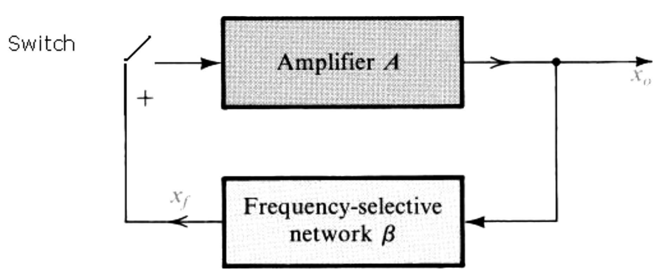
An oscillator's basic operation is as follows:

- If the following criteria are met, an amplifier with positive feedback produces oscillations:

  The loop gain (the product of the amplifier's gain and the feedback network's gain) is unity, and the overall phase shift in the loop is zero.

- A circuit is called a sinusoidal oscillator if the output signal is sinusoidal.





NOTES

There are no oscillations when the switch at the amplifier input is open. Assume that the circuit is given a voltage V_i and that the switch is closed.

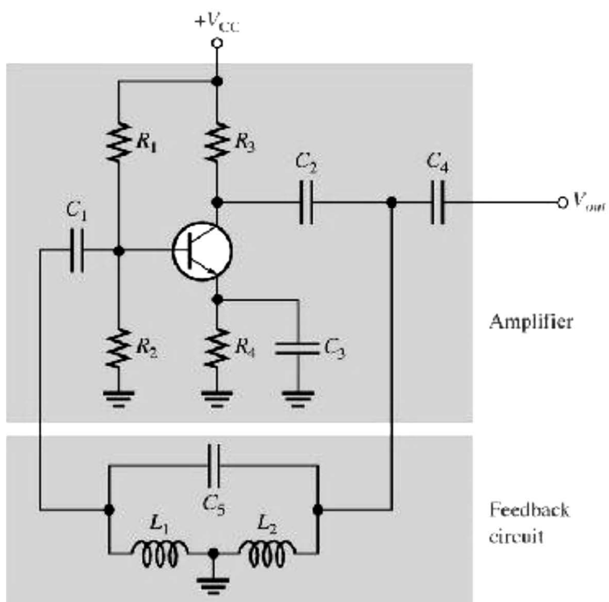
- $\beta V_o = V_f$ is sent back to the circuit as a result of $V_o = A_V V_i$. If we set $V_f = V_i$, the output will continue to exist even if the circuit's input voltage is removed.
- $V_o = A_V V_i$
- $\beta V_o = V_f$
- $\beta A_V V_i = V_f$

If V_f must be the same as V_i, then $\beta A_V = 1$ is obvious from the previous equation. By shutting the switch and eliminating the input in the preceding block diagram, we can produce oscillations at the output if $\beta A_V = 1$, where β AV stands for Loop gain. The fed back signal is in phase with the input signal, which is referred to as positive feedback. This indicates that the signal undergoes no phase change while in the loop.

To obtain persistent oscillations, the preceding criterion, as well as the unity loop gain, must be met. The 'Barkhausen Criterion' refers to these circumstances. Noting the denominator in the basic equation $A_f = A / (1 + \beta A)$ is another way of viewing how the feedback circuit operates as an oscillator.

The denominator becomes 0 and the gain with feedback A_f becomes infinite when $\beta A = -1$ or magnitude 1 at a phase angle of 180°. As a result, even without an input signal, an infinitesimal signal (noise voltage) can produce a quantifiable output voltage, and the circuit operates as an oscillator.

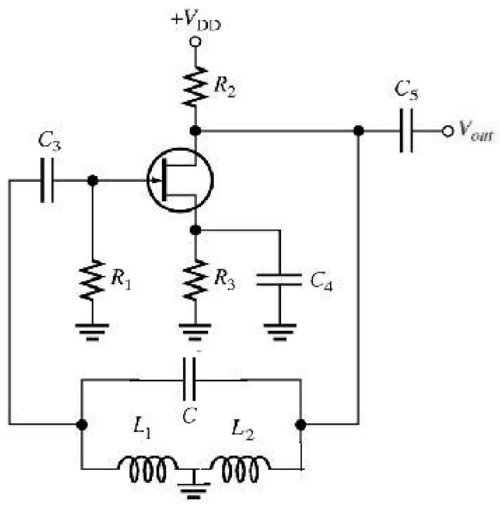
Hartley Oscillator:



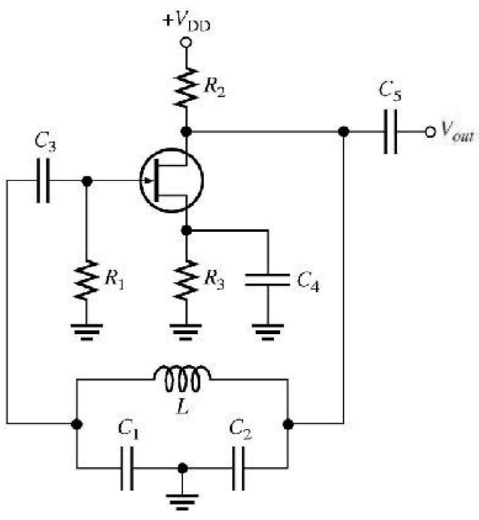
The Hartley oscillator resembles the Colpitts oscillator. Two inductors and one capacitor make up the tank circuit. The resonant frequency is calculated in the same way.

$L_T = L_1 + L_2 + 2M$, where M is mutual coupling, $f = 1/2L_TC$

NOTES



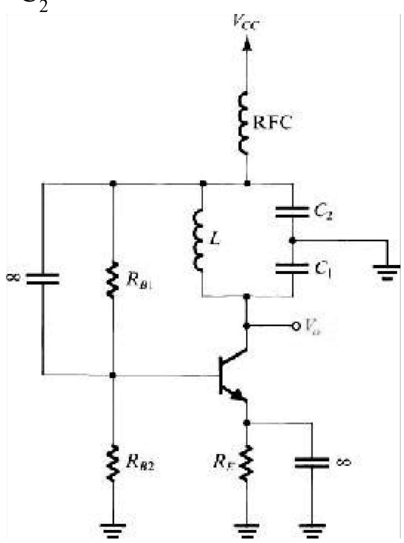
Colpitts Oscillator:



In the feedback loop of the Colpitts oscillator, a tank circuit (LC) is used. The formula below can be used to calculate the resonant frequency. An FET is a superior choice for the active device because the input impedance impacts the Q.

$f_r = 1/2LC_T$

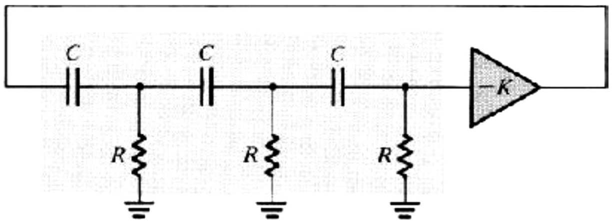
$C_T = C_1C_2 / C_1 + C_2$



NOTES

An Op amp is a type of amplifier that is used to An LC feedback network can be used to set the Oscillator frequency in a Colpitts Oscillator circuit, where the Op amp provides the fundamental amplification and the Oscillator frequency is set by an LC feedback network.

Phase Shift Oscillator:

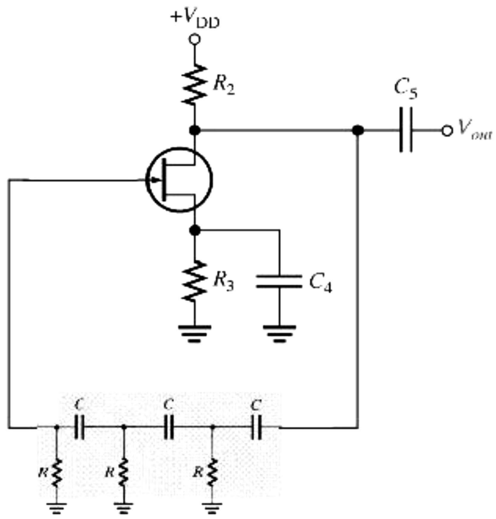


The phase shift oscillator employs three RC circuits to create 180° phase shift, which, when combined with the 180° phase shift of the op-amp, gives the necessary feedback to keep oscillations going.

• To keep the oscillations going, the gain must be at least 29. This type's resonance frequency is comparable to that of any RC circuit oscillator:

FET phase shift oscillator with f_r = 1/2″ð 6RC

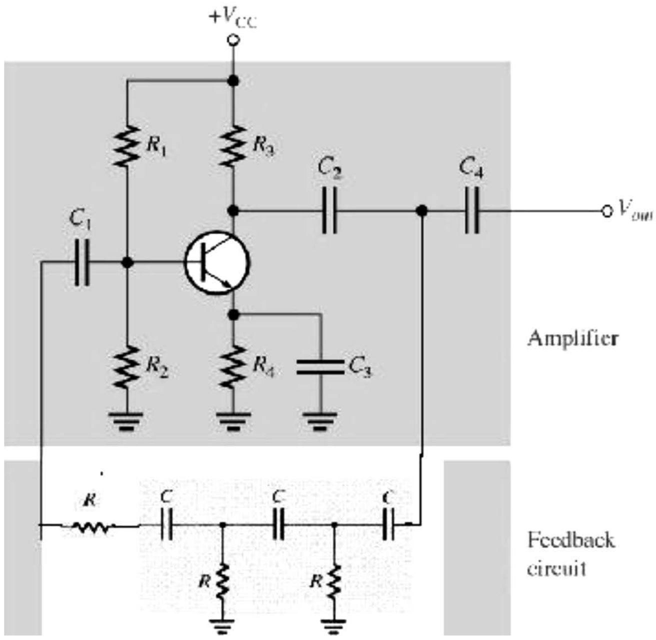
FET Phase Shift Oscillator



- A capacitor bypassed source resistor R_S and a drain bias resistor R_D are used to self-bias the amplifier stage. The gm and rd are FET device parameters of interest.
- $|A| = gmR_L$, where $R_L = (R_D r_d / R_D + r_d)$
- We can suppose that the amplifier's input impedance is infinite at the operating frequency.
- This is a reasonable estimate if the oscillator's operating frequency is low enough to ignore FET capacitive impedances.
- The output impedance of the amplifier stage provided by R_L should be small in comparison to the impedance seen while looking into the feedback network to avoid attenuation due to loading.

RC Phase shift Oscillator - BJT version

NOTES



On the off chance that a semiconductor is utilized as the dynamic component of the enhancer stage, the result of the criticism network is stacked apparently by the generally low info opposition (h_{ie}) of the semiconductor. A producer – devotee input stage followed by a typical producer speaker stage could be used. If a solitary semiconductor stage is wanted, the utilization of voltage – shunt criticism is more appropriate. Here, the criticism signal is coupled through the criticism resistor R' in series with the intensifier stage input opposition (R_i).

$f = (1/2RC)\ [1/6 + 4(RC / R)]$

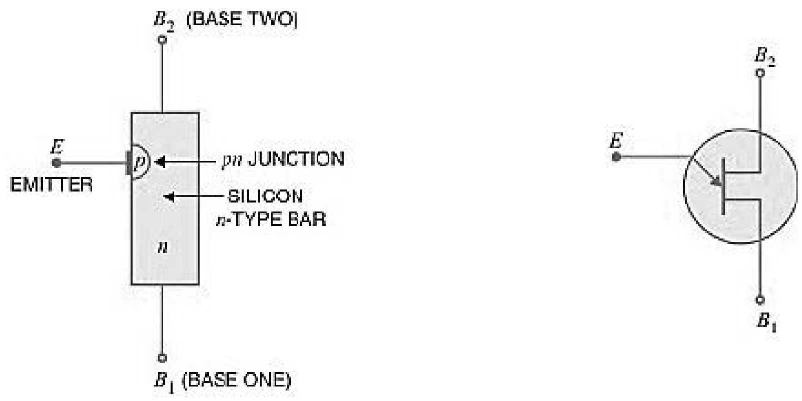
$h_{fe} > 23 + 29\ (R/RC) + 4\ (RC / R)$

2.7 UJT AND ITS CHARACTERISTICS

Uni Junction Transistor (UJT): A unijunction semiconductor (truncated as UJT) is a three-terminal semiconductor exchanging gadget. This gadget has a novel trademark that when it is set off, the producer current increments regeneratively until it is restricted by producer power supply. Because of this trademark, the unijunction semiconductor can be utilized in an assortment of utilizations, e.g., exchanging, beat generator, saw-tooth generator and so forth.

Construction of UJT: As shown in Figure (i) The essential *structure of a unijunction semiconductor. It comprises of a n-type silicon bar with an electrical association on each end. The prompts these associations are called base leads base-one B_1 and base two B_2. Part way along the bar between the two bases, closer to B_2 than B_1, a *pn* intersection is framed between a type producer and the bar. The lead to this intersection is known as the producer lead E. As shown in Figure (ii) The image of unijunction semiconductor. Note that producer is shown nearer to B_2 than B_1.

NOTES



(i) (ii)

(i) Since the gadget has one pn intersection and three leads, it is regularly called a unijunction semiconductor (uni implies single).

(ii) With only one *pn*-intersection, the gadget is actually a type of diode. Since the two base terminals are taken from one part of the diode, this gadget is additionally called double-based diode.

(iii) The producer is vigorously doped having many openings. The *n* district, nonetheless, is daintily doped. Hence, the opposition between the base terminals is exceptionally high (5 to 10 kΩ) when producer lead is open.

Operation of UJT

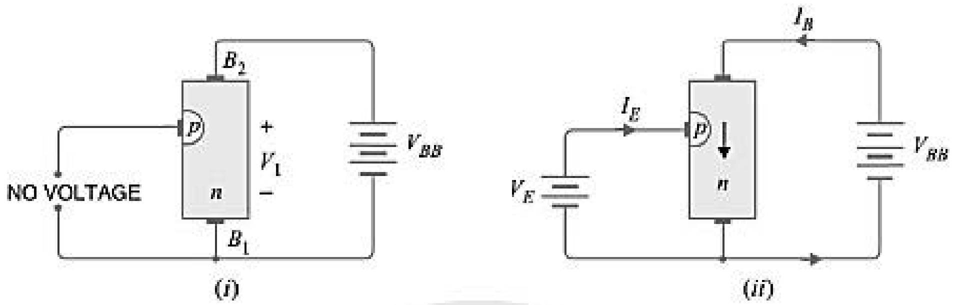


Fig. 2.16 *Shows the Basic Circuit Operation of a Unijunction Transistor.*

The gadget has ordinarily B_2 positive with respect to B_1.

(i) If voltage V_{BB} is applied somewhere in the range of B_2 and B_1 with producer open (Refer Figure 2.16(i)). A voltage slope is set up along the n-type bar. Since the producer is found closer to B_2, more than **half of V_{BB} shows up between the producer and B_1. The voltage V_1 among producer and B_1 builds up an opposite inclination on the *pn* intersection and the producer current is cut off. Obviously, a little spillage current streams from B_2 to producer because of minority transporters.

(ii) Assuming that a positive voltage is applied at the producer (Refer Figure 2.16 (ii)), the *pn* intersection will stay switch one-sided inasmuch as the info voltage is under V_1. In case the information voltage to the producer surpasses V_1, the *pn* intersection becomes *forward one-sided. Under these conditions, openings are infused from *p*-type material into the *n*-type bar. These openings are repulsed by sure B_2 terminal and they are drawn in towards B_1 terminal of the bar. This aggregation of openings in the producer to B_1 area brings about the reduction of opposition in this segment of the bar. The outcome is that inside voltage drop from producer to B_1 is diminished and subsequently the producer current I_E increments. As more openings are infused, a state of immersion will ultimately be reached. Now, the producer current is restricted by producer power supply as it were. The gadget is currently in the ON state.

NOTES

(iii) If a negative heartbeat is applied to the producer, the *pn* intersection is converse one-sided and the producer current is cut off. The gadget is then supposed to be in the OFF state.

Characteristics of UJT

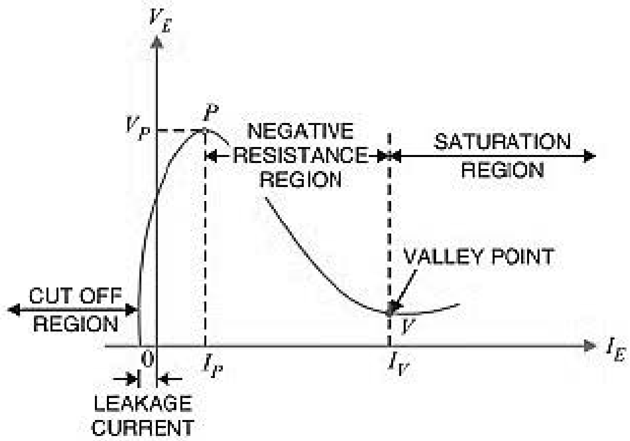


Fig. 2.17 *Characteristics of UJT*

Figure 2.17 shows at a given voltage V_{BB} between the bases, the curve between emitter voltage (V_E) and emitter current (I_E) of a UJT. This is referred to as UJT's emitter characteristic. From the qualities, the following points can be noted:

(i) As V_E rises from zero in the cut-off zone, a small leakage current flows from terminal B_2 to the emitter. The minority carriers in the reverse biased diode are responsible for this current.

(ii) Forward I_E begins to flow above a particular value of V_E, increasing until the peak voltage V_P and current I_P are attained at point P.

(iii) An attempt to enhance V_E is followed by a sharp increase in emitter current I_E and a corresponding fall in V_E after the peak point P. Because V_E reduces as I_E increases, this represents a negative resistance part of the curve. As a result, the device features a negative resistance zone that is stable enough to be used in various applications, such as trigger circuits, sawtooth generators, and timing circuits.

Advantages of UJT

The UJT was presented in 1948 however didn't open up until 1952. From that point forward, the gadget has accomplished extraordinary notoriety because of the accompanying reasons:

(i) It is a minimal expense gadget.

(ii) It has amazing qualities.

(iii) It is a low-power engrossing gadget under ordinary working conditions.

NOTES

Applications of UJT

Due to above reasons, this gadget is being utilized in an assortment of uses. A couple incorporate oscillators, trigger circuits, saw-tooth generators, bistable organization and so on.

The UJT is exceptionally well known today basically because of its high exchanging speed.

A couple of select utilizations of the UJT are as per the following:

(i) It is utilized to trigger SCRs and TRIACs

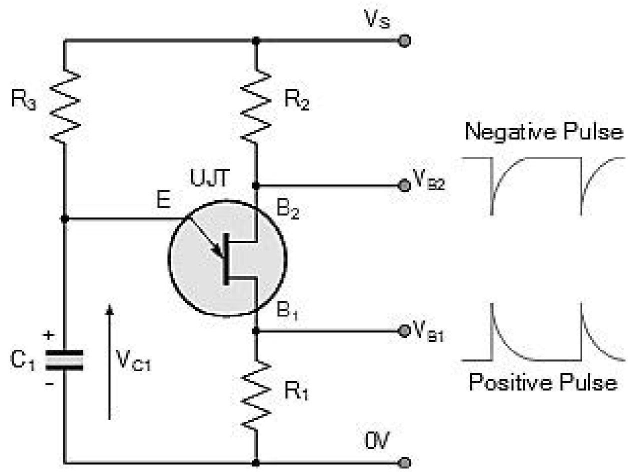
(ii) It is utilized in non-sinusoidal oscillators

(iii) It is utilized in stage control and timing circuits

(iv) It is utilized in saw tooth generators

(v) It is utilized in oscillator circuit plan.

2.7.1 UJT as Relaxation Oscillators



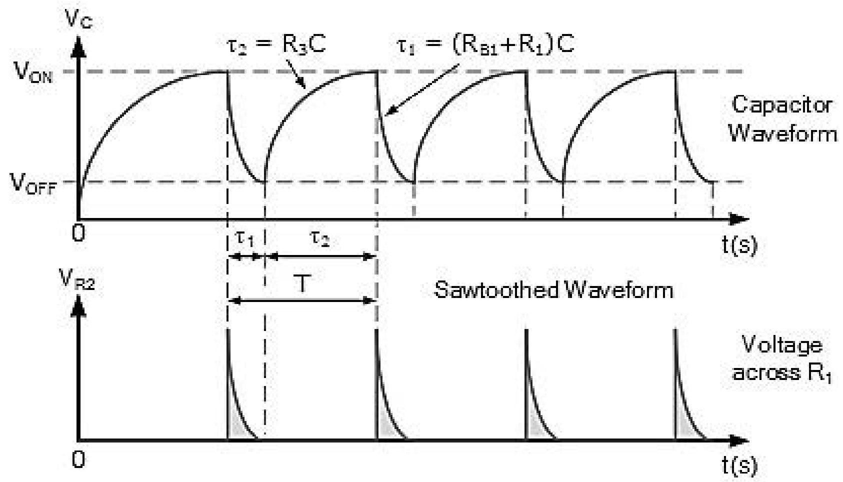
At the point when a voltage (Vs) is first and foremost applied, the unijunction semiconductor is 'OFF' and the capacitor C_1 is completely released yet starts to energize dramatically through resistor R_3. As the Emitter of the UJT is associated with the capacitor, when the charging voltage Vc across the capacitor becomes more prominent than the diode volt drop esteem, the *pn* intersection acts as an ordinary diode and becomes forward one-sided setting off the UJT into conduction. The unijunction semiconductor is 'ON'. Now the Emitter to B_1 impedance implodes as the Emitter goes into a low impedance immersed state with the progression of Emitter current through R_1 occurring.

As the ohmic worth of resistor R_1 is exceptionally low, the capacitor releases quickly through the UJT and a quick rising voltage beat shows up across R_1. Additionally, in light of the fact that the capacitor releases more rapidly through the UJT than it does energizing through resistor R_3, the releasing time is much not exactly the charging time as the capacitor releases through the low opposition UJT.

At the point when the voltage across the capacitor diminishes beneath the holding point of the *pn* intersection (V_{OFF}), the UJT turns 'OFF' and no current streams into the Emitter intersection so by and by the capacitor energizes through resistor R_3 and this charging and releasing cycle among V_{ON} and V_{OFF} is continually rehashed while there is a stockpile voltage, Vs applied.

NOTES

UJT Oscillator Waveforms



Then, at that point, we can see that the unijunction oscillator ceaselessly switches 'ON' and 'OFF' with next to no criticism. The recurrence of activity of the oscillator is straightforwardly impacted by the worth of the charging obstruction R_3, in series with the capacitor C_1 and the worth of η. The result beat shape created from the $Base_1$ (B_1) terminal is that of a saw tooth waveform and to manage the time-frame, you just need to change the ohmic worth of opposition, R_3 since it sets the RC time steady for charging the capacitor.

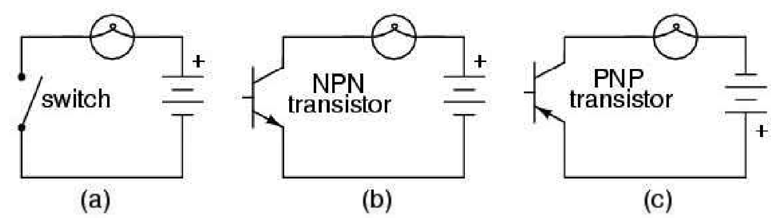
The time-frame, T of the saw-toothed waveform will be given as the charging time in addition to the releasing season of the capacitor. As the release time, τ_1 is for the most part extremely short in contrast with the bigger RC charging time, τ_2 the time span of swaying is pretty much identical to $T\tau_2$. The recurrence of wavering is along these lines given by $f = 1/T$.

2.8 TRANSISTOR AS A SWITCH

Since a semiconductor's authority current is relatively restricted by its base current, it tends to be utilized as a kind of current-controlled switch. A somewhat little progression of electrons sent through the foundation of the semiconductor can apply command over a lot bigger progression of electrons through the gatherer. Assume we had a light that we needed to turn on and off with a switch. Such a circuit would be very basic below as shown in Figure (a). For delineation, how about we embed a semiconductor instead of the change to show how it can handle the progression of electrons through the light. Recall that the controlled current through a semiconductor should go among gatherer and producer. Since it is the

NOTES

current through the light that we need to control, we should situate the gatherer and producer of our semiconductor where the two contacts of the switch were. We should likewise ensure that the light's present will move against the course of the producer bolt image to guarantee that the semiconductor's intersection predisposition will be right as below shown in Figure (b).



(a) Mechanical switch, (b) NPN semiconductor switch, (c) PNP semiconductor switch.

A PNP semiconductor could likewise have been decided for the work. Its application is above shown in Figure (c).

The decision among NPN and PNP is truly subjective. The only thing that is important is that the appropriate current bearings are kept up with for right intersection biasing (electron stream conflicting with the semiconductor image's bolt).

Returning to the NPN semiconductor in our model circuit, we are confronted with the need to add something all the more so we can have base current. Without an association with the base wire of the semiconductor, base current will be zero, and the semiconductor can't turn on, bringing about a light that is consistently off. Recall that for a NPN semiconductor, base current should comprise of electrons moving from producer to base

(Against the producer bolt image, actually like the light current). Maybe the least complex thing to do is interface a switch between the base and authority wires of the semiconductor as shown in Figure beneath (a).

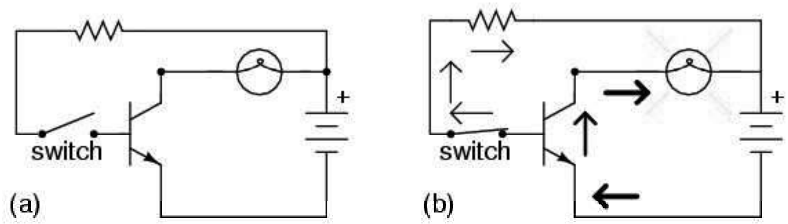


Fig. 2.18 *Semiconductor (a) Cutoff, light off; (b) Soaked, light on.*

Assuming the switch is open as in (Refer Figure 2.18(a)), the base wire of the semiconductor will be left 'Drifting' (not associated with anything) and there will be no current through it. In this express, the semiconductor is supposed to be cutoff. Assuming that the switch is shut as in

(Refer Figure 2.18(b)), in any case, electrons will actually want to move from the producer through to the foundation of the semiconductor, through the switch and up to the left half of the light, back to the positive side of the battery. This base current will empower a lot bigger progression of electrons from the producer through to the authority, accordingly illuminating the light. In this condition of greatest circuit current, the semiconductor is supposed to be soaked.

NOTES

Obviously, it might appear to be inconsequential to utilize a semiconductor in this ability to control the light. All things considered, we're actually utilizing a switch in the circuit, right? Assuming that we're actually utilizing a change to control the light - - if by some stroke of good luck in a roundabout way - - then, at that point, why bother having a semiconductor to control the current? Why not simply return to our unique circuit and utilize the switch straightforwardly to control the light current?

Two focuses can be made here, really. First is the way that when utilized thusly, the switch contacts need just handle what minimal base current is important to turn the semiconductor on; the actual semiconductor handles the vast majority of the light's current. This might be a significant benefit in case the switch has a low current rating: a little switch might be utilized to control a moderately high-current burden. More significant, the current-controlling conduct of the semiconductor empowers us to utilize something totally unique to wind down the light on or. As shown in Figure 2.19 underneath, where a couple of sun based cells gives 1 V to defeat the 0.7 V_{BE} of the semiconductor to cause base current stream, which thusly controls the light.

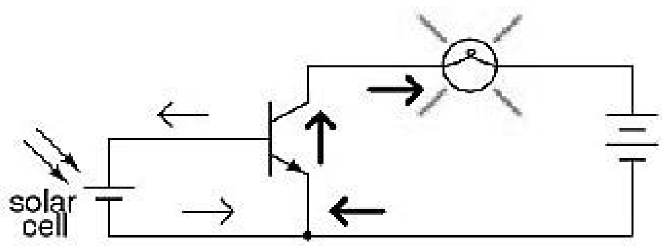


Fig. 2.19 *Solar cell serves as light sensor.*

2.9 ASTABLE, MONOSTABLE AND BISTABLE MULTIVIBRATOR

An electronic gadget that delivers a non-sinusoidal waveform as its result is known as a Multivibrator. The created non-sinusoidal waveforms are fundamentally a square wave, rectangular wave, a three-sided wave, sawtooth wave, or slope wave and so on

It is a 2 phase RC coupled enhancer that works in two modes. The modes are essentially named as conditions of the multivibrator.

Definition of Multivibrator

A multivibrator is an electronic circuit used to execute an assortment of straightforward two state frameworks like oscillators, clocks and flip-flops. It is portrayed by two intensifying gadgets (semiconductors, electron tubes or different gadgets) cross coupled by resistors or capacitors. The name 'Multivabrator' was at first applied to the free running oscillator adaptation of the circuit since its result waveform was wealthy in sounds. Multivibrator are a gathering of regenerative circuits that are broadly utilized in planning application.

NOTES

Types of Multivibrators

1. Astable Multivibrator
2. Mono Stable Miltivibrator
3. Bistable Multivibrator

1. Astable Multivibrator

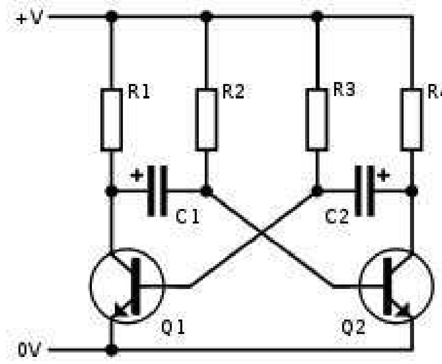


Fig. 2.20 *Astable Multivibrator*

It is known as free running multivibrator, it is likewise called as square wave generator. The circuit has two semi-stable expresses (no steady state). Accordingly, three is a wavering between two state and no outer signs are needed to create change in state. Astable circuits are utilized to create square wave. For instance, clock generators in advanced framework.

2. Mono Stable Multivibrator

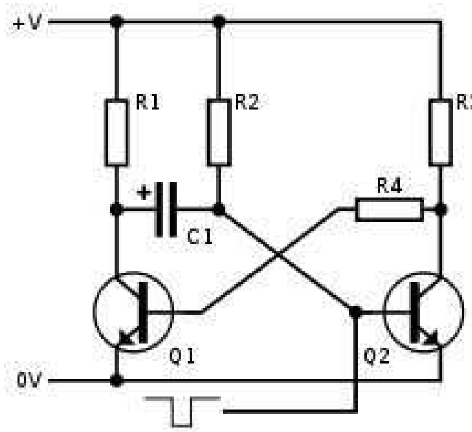


Fig. 2.21 *Mono Stable Miltivibrator*

The mono steady or one - shot multivibrator produces a sign beat of indicated span to every outer trigger. As its name suggests just one stable state exist. Utilization of each trigger makes a change the semi stable state. The circuit stays in the semi stable state for a proper timespan and afterward returns to its unique stable state. Indeed, an inward trigger sign is created which delivers the change to the steady state. Generally, the charging and releasing of capacitor gives this inward trigger sign.

3. Bistable Multivibrator

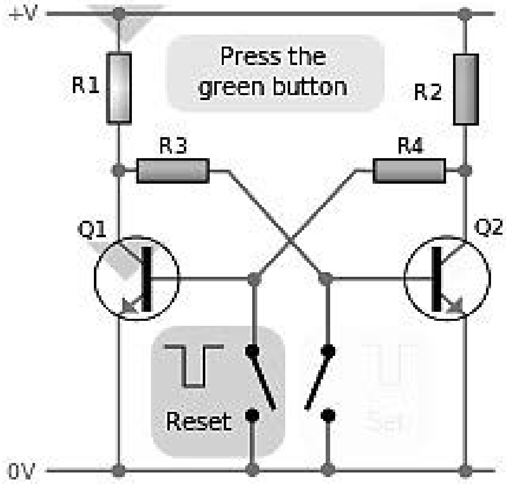


Fig. 2.22 *Bistable Multivibrator*

NOTES

It is called as bistable lock of flip failures. Schmitt trigger is another bistable circuit significant property of a bistable multivibrator is that it keeps a given result voltage level except if an outer sign is applied. Use of outer sign causes a difference in state and this result level is kept up with endlessly until a subsequent trigger is applied. Accordingly, a bistable multivibrator requires two outer triggers before it gets back to its underlying state.

Check Your Progress

10. What is emitter follower?
11. Give some of the most critical characteristics of an emitter follower.
12. Define the term oscillator.
13. What do you mean by the Hartley oscillator?
14. How will you define the phase shift oscillator?
15. Give a short note on Uni Junction Transistor (UJT).
16. State the transistor as a switch.
17. What is multivibrator?
18. How many types of multivibrators?

2.10 ANSWERS TO 'CHECK YOUR PROGRESS'

1. A feedback amplifier is also known as closed-loop amplifier as a closed loop is formed between the input and the output of the amplifier.
2. Feedback factor is also known as feedback ratio or reverse transfer ratio or reverse transmission factor and this determines the fraction of the output signal that is added to or subtracted from the externally applied input signal voltage V_S by the feedback network. This is given by $\beta = \frac{V_f}{V_O}$ where the feedback voltage $V_f = \beta V_O$.

NOTES

3. Current feedback can be established by placing a resistor in the emitter terminal of a CE amplifier.
4. The effects of negative feedback on the amplifier characteristics are the following:
    - Gain is reduced and stabilized with respect to the variations in transis tor parameters like h_{fe}.
    - Non-linear distortion becomes less resulting in improvement of the signal handling capacity of the amplifier.
5. Gain Stability: Overall gain of the negative feedback amplifier is

$$A_f = \frac{A}{1 + A\beta}$$

6. The output signal gets distorted due to non-linearity in the transfer characteristic of the amplifier when a large-amplitude signal is applied to the input of the amplifier.
7. There are four feedback topologies:
    - Voltage-series feedback (voltage sampling-series mixing)
    - Voltage-shunt feedback (voltage sampling-shunt mixing)
    - Current-series feedback (current sampling-series mixing)
    - Current-shunt feedback (current sampling-shunt mixing)
8. Input impedance with series feedback is given by the value of input impedance without feedback multiplied by the factor $(1 + A\beta)$ and this applies to both voltage-series and current-series feedback topology. Higher input impedance is obtained by series mixing of the feedback signal.
9. A sample of voltage proportional to the output current is fedback negatively to the input in series with the voltage source. Here, the input is a voltage and the output is a current, hence this topology is known as current-sampling series-mixing or series-series feedback configuration.
10. In feedback amplifiers, the emitter follower circuit plays a significant role. A negative current feedback circuit is an emitter follower. In signal generating circuits, this is typically employed as a last stage amplifier.
11. The following are some of the most critical characteristics of an emitter follower:
    - It has a high input impedance.
    - Its output impedance is low.
    - It's a great circuit for matching impedances.

    The emitter follower circuit can be used in a variety of ways thanks to all of these advantageous characteristics. This circuit is a current amplifier with no voltage gain.

12. A circuit that creates a continuous, repetitive, alternating waveform without any input is known as an oscillator. Oscillators are devices that convert unidirectional current flow from a DC source into an alternating waveform with the required frequency set by the circuit components.
13. The Hartley oscillator resembles the Colpitts oscillator. Two inductors and one capacitor make up the tank circuit. The resonant frequency is calculated in the same way.

    LT= $L_1 + L_2$ + 2M, where M is mutual coupling, f = 1/2LTC
14. The phase shift oscillator employs three RC circuits to create 180o phase shift, which, when combined with the 180o phase shift of the op-amp, gives the necessary feedback to keep oscillations going. To keep the oscillations going, the gain must be at least 29.
15. A unijunction semiconductor (truncated as UJT) is a three-terminal semiconductor exchanging gadget. This gadget has a novel trademark that when it is set off, the producer current increments regeneratively until it is restricted by producer power supply. Because of this trademark, the unijunction semiconductor can be utilized in an assortment of utilizations, e.g., exchanging, beat generator, saw-tooth generator and so forth.
16. A semiconductor's authority current is relatively restricted by its base current, it tends to be utilized as a kind of current-controlled switch. A somewhat little progression of electrons sent through the foundation of the semiconductor can apply command over a lot bigger progression of electrons through the gatherer.
17. A multivibrator is an electronic circuit used to execute an assortment of straightforward two state frameworks like oscillators, clocks and flip-flops. It is portrayed by two intensifying gadgets (semiconductors, electron tubes or different gadgets) cross coupled by resistors or capacitors.
18. Multivibrators are three types:
    - Astable multivibrator
    - Mono stable miltivibrator
    - Bistable multivibrator

2.11 SUMMARY

- A feedback amplifier is also known as closed-loop amplifier as a closed loop is formed between the input and the output of the amplifier.
- In a feedback amplifier, the output signal is sampled and feedback to the input to form an error signal that drives the amplifier.
- Closed-loop gain of the feedback amplifier is the ratio of the output voltage V_O and the external input signal voltage V_S
- Loop gain is given by the term $-A\beta$ as it gives the product of the gains of the branches making up the loop.

NOTES

- The capacitor C_{in} decouples the signal source from the supply voltage V_{CC} and acts like a short circuit for AC.
- Current feedback can be established by placing a resistor in the emitter terminal of a CE amplifier.
- A current-shunt feedback amplifier is basically a current amplifier.
- Gain is reduced and stabilized with respect to the variations in transistor parameters like h_{fe}.
- Non-linear distortion becomes less resulting in improvement of the signal handling capacity of the amplifier.
- Input impedance with series feedback is given by the value of input impedance without feedback multiplied by the factor $(1 + A\beta)$ and this applies to both voltage-series and current-series feedback topology. Higher input impedance is obtained by series mixing of the feedback signal.
- A sample of voltage proportional to the output current is fedback negatively to the input in series with the voltage source. Here, the input is a voltage and the output is a current, hence this topology is known as current-sampling series-mixing or series-series feedback configuration.
- The emitter follower circuit plays a significant role. A negative current feedback circuit is an emitter follower. In signal generating circuits, this is typically employed as a last stage amplifier.
- The emitter follower circuit can be used in a variety of ways thanks to all of these advantageous characteristics. This circuit is a current amplifier with no voltage gain.
- An emitter follower circuit is virtually identical in construction to a standard amplifier. The key distinction is that the load RL is not present at the collector terminal of the circuit, instead it is present at the emitter terminal.
- The output voltage Vo across R_E in the emitter section is generated by the input signal voltage applied between the base and the emitter.
- A circuit that creates a continuous, repetitive, alternating waveform without any input is known as an oscillator. Oscillators are devices that convert unidirectional current flow from a DC source into an alternating waveform with the required frequency set by the circuit components.
- The Hartley oscillator resembles the Colpitts oscillator. Two inductors and one capacitor make up the tank circuit. The resonant frequency is calculated in the same way.
- The phase shift oscillator employs three RC circuits to create 180o phase shift, which, when combined with the 180o phase shift of the op-amp, gives the necessary feedback to keep oscillations going. To keep the oscillations going, the gain must be at least 29.
- A unijunction semiconductor (truncated as UJT) is a three-terminal semiconductor exchanging gadget. This gadget has a novel trademark that

when it is set off, the producer current increments regeneratively until it is restricted by producer power supply. Because of this trademark, the unijunction semiconductor can be utilized in an assortment of utilizations.

- A semiconductor's authority current is relatively restricted by its base current, it tends to be utilized as a kind of current-controlled switch. A somewhat little progression of electrons sent through the foundation of the semiconductor can apply command over a lot bigger progression of electrons through the gatherer.
- A multivibrator is an electronic circuit used to execute an assortment of straightforward two state frameworks like oscillators, clocks and flip-flops. It is portrayed by two intensifying gadgets (semiconductors, electron tubes or different gadgets) cross coupled by resistors or capacitors.
- Astable multivibrator is known as free running multivibrator, it is likewise called as square wave generator. The circuit has two semi-stable expresses (no steady state). Accordingly, three is a wavering between two state and no outer signs are needed to create change in state. Astable circuits are utilized to create square wave. For instance, clock generators in advanced framework.

NOTES

2.12 KEY TERMS

- **Feed back amplifier:** A feedback amplifier is also known as closed-loop amplifier as a closed loop is formed between the input and the output of the amplifier.
- **Loop gain:** Loop gain is given by the term $-A\beta$ as it gives the product of the gains of the branches making up the loop.
- **Gain stability:** Overall gain of the negative feedback amplifier is

$$A_f = \frac{A}{1 + A\beta}$$

- **Emitter follower:** The emitter follower circuit plays a significant role. A negative current feedback circuit is an emitter follower. In signal generating circuits, this is typically employed as a last stage amplifier.
- **Oscillator:** A circuit that creates a continuous, repetitive, alternating waveform without any input is known as an oscillator. Oscillators are devices that convert unidirectional current flow from a DC source into an alternating waveform with the required frequency set by the circuit components.
- **Uni Junction Transistor (UJT):** A unijunction semiconductor (truncated as UJT) is a three-terminal semiconductor exchanging gadget. This gadget has a novel trademark that when it is set off, the producer current increments regeneratively until it is restricted by producer power supply. Because of this trademark, the unijunction semiconductor can be utilized in an assortment of utilizations.
- **Multivibrator:** A multivibrator is an electronic circuit used to execute an assortment of straightforward two state frameworks like oscillators, clocks

and flip-flops. It is portrayed by two intensifying gadgets (semiconductors, electron tubes or different gadgets) cross coupled by resistors or capacitors.

NOTES

2.13 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is feedback in amplifiers?
2. Define voltage-series feedback circuit.
3. Name the four feedback topologies.
4. What do you understand by the emitter follower?
5. Define the term Colpitts oscillator.
6. Write a short note on Uni Junction Transistor (UJT).
7. Differentiate between NPN semiconductor switch and PNP semiconductor switch.
8. Give the definition of multivibrator.
9. What is astable multivibrator?

Long-Answer Questions

1. Briefly discuss about the closed-loop gain or transfer gain of the feedback amplifier. Give appropriate examples.
2. Analysis the negative feedback amplifier circuits with the help of giving examples.
3. Discuss the four feedback topologies with the help of relevant examples.
4. What is emitter follower? Discuss some of the most critical characteristics of an emitter follower.
5. Explain briefly about the oscillator's basic operation with the help of giving examples.
6. What do you understand by the uni junction transistor? Give the construction process of UJI.
7. Illustrate the mechanical switch, NPN semiconductor switch and PNP semiconductor switch. Give appropriate examples.
8. Discuss briefly about the types of multivibrator with the help of giving examples.

2.14 FURTHER READING

Basavaraj, B. and H.N. Shivashankar. 2004. *Basic Electronics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Black, Uyless D. 1993. *Computer Networks.* New Jersey: Prentice-Hall.

NOTES

Kumar, A. Anand. 2003. *Fundamentals of Digital Circuits*. New Delhi: Prentice-Hall of India.

Mano, Morris. 1979. *Digital Logic and Computer Design*. New Delhi: Prentice-Hall of India.

Roth, Charles. 2001. *Fundamentals of Logic Design*. Thomson Learning.

Shepherd, Robert D. 1998. *Introduction to Computers and Technology*. EMC/ Paradigm Publishing.

Stallings, William. 1996. *Data and Computer Communications.* New Jersey: Prentice-Hall.

Tanenbaum, Andrew. 2002. *Computer Networks.* New Jersey: Prentice-Hall PTR.

Yarbarough, John M. 1996. *Digital Logic Applications and Design*. Thomson Learning.

Carlson Bruce and Paul B. Crilly. 2010. Communication System. Noida: Tata McGraw-Hill Education.

Lathi B. P. and Zhi Ding. 2010. Modern Digital and Analog Communication Systems. Oxford: Oxford University Press.

Kennedy George and Bernard Davis. 1993. Electronic Communication Systems. London: Glencoe/ McGraw-Hill.

Roddy Dennis and John Coolen. 1977. Electronic Communications. Noida: Pearson Education.

UNIT 3 OPERATIONAL AMPLIFIER

NOTES

Structure

3.0 INTRODUCTION

A differential amplifier is a type of electronic amplifier that amplifies the difference between two input voltages but suppresses any voltage common to the two inputs. Single amplifiers are usually implemented by either adding the appropriate feedback resistors to a standard Op-amp, or with a dedicated integrated circuit containing internal feedback resistors. It is also a common sub-component of larger integrated circuits handling analog signals.

An operational amplifier is a DC-coupled high-gain electronic voltage amplifier with a differential input and, usually, a single-ended output. In this configuration, an Op-amp produces an output potential (relative to circuit ground) that is typically 100,000 times larger than the potential difference between its input terminals. The popularity of the Op-amp as a building block in analog circuits is due to its versatility. By using negative feedback, the characteristics of an Op-amp circuit, its gain, input and output impedance, bandwidth, etc., are determined by external components and have little dependence on temperature coefficients or engineering tolerance in the Op-amp itself. Adding speaker or a snake is utilized to total two sign voltages. Voltage snake circuit is a basic circuit that empowers you to add a few signals together. It has wide assortment of utilizations in electronic circuits. A subtractor is an electrical circuit that produces a result that is the same as the difference between the applied information sources. This section delves into

NOTES

the operation of an amp-based subtractor circuit. A differentiator is an electronic circuit that creates a result equivalent to the main subordinate of its feedback. This part examines about the operation amp-based differentiator exhaustively. An integrator is an electronic circuit that creates a result that is the mix of the applied info. This segment examines about the operation amp-based integrator. Now the Op-amp increases its output voltage to compensate for the VBE drop. The circuit is actually a buffered non-inverting amplifier driven by a constant input voltage. It keeps up this constant voltage across the constant sense resistor.

The comparator is an electronic dynamic circuit that employs a functional enhancer with a very high addition in its open-circle form. A square wave generator is an electronic circuit which creates square wave. This segment examines about operation amplifier based square wave generators. An electronic circuit that generates a three-sided wave is known as a three-sided wave generator.

A voltage multiplier is a circuit that generates a d.c. voltage that is equivalent to a number of different discrete input voltages. At least two pinnacle identifiers or rectifiers are included. Voltage multipliers were used to track down applications in circuits that required high voltage with low current, such as picture tubes in TV receivers, oscilloscopes, and so on. A voltage multiplier is an electrical circuit that changes over AC electrical power from a lower voltage to a higher DC voltage through capacitors and diodes joined into an organization. A wave forming circuit is the one which can be utilized to change the state of a waveform from rotating current or direct current. The three terminal voltage controllers are of two sorts: Fixed and movable voltage controllers. In fixed voltage controllers we have positive voltage controllers and negative voltage controllers. The 78XX series is a progression of fixed positive voltage controllers and the 79XX series is a progression of fixed negative voltage controllers.

In this unit, you will learn about the differential amplifier, circuits and working of operational amplifier, Op-amp parameters, inverting and non-inverting Op-amp amplifiers, use of IC as adder, subtractor, differentiator and integrator, Op-amp as constant current source, comparator, square and triangular wave generator, voltage multipliers circuits, wave shaping circuits, clipping, clamping, differentiating, and integrating circuits, voltage regulated power supply, regulation sensitivity and stability factors, over voltage and short circuit performance, three terminal IC regulated power supply circuits for positive and negative voltages.

3.1 OBJECTIVES

After going through this unit, you will be able to:

- Explain differential amplifier
- Explain the circuits and working of operational amplifier
- Discuss about the use of IC as adder, subtractor, differentiator and integrator
- Describe the Op-amp as constant current source
- Explain the comparator, square and triangular wave generator
- Discuss the voltage multipliers circuits

- Discuss the basic concept of wave shaping circuits
- Illustrate the clipping, clamping, differentiating, and integrating circuits
- Interpret the voltage regulated power supply
- Elaborate on the regulation sensitivity and stability factors
- Analyse the over voltage and short circuit performance
- Explain the three terminal IC regulated power supply circuits for positive and negative voltages

NOTES

3.2 DIFFERENTIAL AMPLIFIER

The differential amplifier is the most widely used circuit-building block in analog integrated circuits as it is invariably used as the input stage of every Op-amp to produce high-voltage gain. One of the basic characteristics of the differential amplifier is that it is DC coupled and avoids large capacitors which is impracticable to realise in ICs. Also, the high-speed logic circuit family or the Emitter-Coupled Logic (ECL) is formed by the BJT differential amplifier.

This section deals with the discussion of differential amplifier which forms the basis for understanding the operational amplifier. Analysis of different circuit configurations of differential amplifier is described in this section. Differential amplifier is a small signal amplifier which is generally used as a voltage amplifier rather than as a current amplifier or power amplifier.

Circuit Configurations: The block diagram of differential amplifier is shown in the Figure 3.1 below:

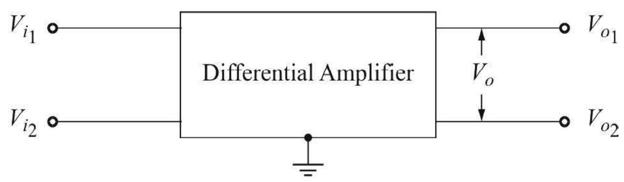


Fig. 3.1 *Block diagram of differential amplifier*

Here, V_{i1} and V_{i2} are the signals applied to two input terminals and V_{o1} and V_{o2} are the output signals respectively. The input signals are applied at V_{i1} and V_{i2} with respect to ground and outputs are taken at V_{o1} and V_{o2} with respect to ground. V_o is the differential output which is obtained as the difference of the two output signals V_{o1} and V_{o2}.

A differential amplifier is classified into various topologies according to the number of input singals used and the way an output voltage is measured.

The four differential amplifier topologies are:

1. Dual-input, balanced-output differential amplifier.
2. Dual-input, unbalanced-output differential amplifier.
3. Single-input, balanced-output differential amplifier.
4. Single-input, unbalanced-output differential amplifier.

If two input signals are applied at the two input terminals, then the configuration is said to be dual-input, otherwise it is a single-input configuration. At the output

NOTES

side, if the output voltage is measured between the two output terminals, it is called balanced output, whereas if the output voltage is measured at one of the output terminals with respect to ground then the output is referred to as unbalanced output.

Let us consider the case of a dual-input unbalanced-output differential amplifier which is similar to an Op-amp configuration.

In an ideal situation,

$$V_o = A_d(V_{i1} - V_{i2}) \quad ...(3.1)$$

where A_d is the differential gain of the amplifier and V_{i1} and V_{i2} are the input signals. If a common signal is applied to both the inputs, that is if $V_{i1} = V_{i2}$, then $V_o = 0$, which means the signal cancels at the output. The output of the amplifier not only depends on the difference of V_{i1} and V_{i2} but also on the average level.

Now let us consider two cases:

1. $V_{i1} = +10$ mV, $V_{i2} = -10$ mV and
2. $V_{i1} = 110$ mV, $V_{i2} = 90$ mV

The difference signal is 20 mV in both the cases but in the second case the average signal is 100 mV where it is zero in the first case. This defines the two signals, one is the difference-mode signal (V_d) and the other is the common-mode signal (V_c).

$$V_d = V_{i1} - V_{i2} \quad \text{and} \quad V_c = \frac{V_{i_1} + V_{i_2}}{2} \quad ...(3.2)$$

V_{i1} and V_{i2} can be expressed in terms of V_c and V_d as:

$$V_{i1} = V_c + \frac{V_d}{2} \quad \text{and} \quad V_{i2} = V_c - \frac{V_d}{2} \quad ...(3.3)$$

Applying the principle of linear superposition, the output voltage V_o can be experssed in terms of V_d and V_c.

When $V_{i1} \neq 0$ and $V_{i2} = 0$ (grounded), $V_o = A_1 V_{i1}$ when $V_{i1} = 0$ (grounded) and $V_{i2} \neq 0$, $V_o = A_2 V_{i2}$.

Now when both V_{i1} and V_{i2} are present, applying superposition principle we get

$$V_o = A_1 V_{i1} + A_2 V_{i2} \quad ...(3.4)$$

Substituting the values of V_{i1} and V_{i2} from Equation (3.3), we get

$$V_o = A_c V_c + A_d V_d \quad ...(3.5)$$

where $\quad A_c = (A_1 + A_2) \quad \text{and} \quad A_d = \frac{A_1 - A_2}{2} \quad ...(3.6)$

Equations (3.3) and (3.4) may be used to find differential-mode gain A_d and common-mode gain A_c of any differential amplifier and the output will be given by Equation (3.5).

NOTES

3.2.1 Circuits and Working of Operational Amplifier

The circuit of a general-purpose monolithic IC Op-amp is shown in Figure 3.2 (*a*). The Darlington connection formed by the two *n-p-n* transistors provides high input impedance. There are two differential gain stages, the output is taken from a single-ended emitter follower. The transistors T_1 and T_2, and T_3 and T_4 are identical. The resistors R_1 and R_2 are also equal. The errors due to thermal drift are reduced since the corresponding components of each pair are affected almost equally as they are closely spaced in the IC chip.

Let the inverting input terminal '*a*' be grounded and a small positive voltage v_1 be applied to the non-inverting input terminal '*b*'. This decreases the potential of the point Q so that T_4 will conduct more and thus increasing the potential at the point N. Now T_5 will conduct less, so that the output voltage at '*c*' becomes more positive. Thus, the signal appearing at '*c*' is of the same polarity as the signal appearing at '*b*' but of increased amplitude. If A_1 is the voltage gain, the output voltage is $v_o = A_1 v_{i1}$. Here, A_1 is positive as v_o and v_{i1} have the same polarity.

Now, let '*b*' be grounded and a small positive voltage v_2 be aplied to '*a*'. The point P becomes less positive, so that T_3 conducts more. As a result there is a drop of potential at the point Q and T_4 conducts less. The potential of N thus becomes more negative so that T_5 conducts more. Thus, the potential of the output terminal '*c*' falls. So, a signal applied at '*a*' appears in an amplified form with reversed polarity at '*c*'. If A_2 is the voltage gain, the output voltage $v_o = A_2 v_{i2}$. Now A_2 will be negative since v_o and v_{i2} have opposite polarity.

Now v_{i1} and v_{i2} are simultaneously applied to '*b*' and '*a*' respectively. Then the output voltage is given by $v_o = A_1 v_{i1} + A_2 v_{i2}$ which is Equation (3.4). Ideally, $A_1 = -A_2$, so we obtain $v_o = 0$ for $v_1 = v_2$.

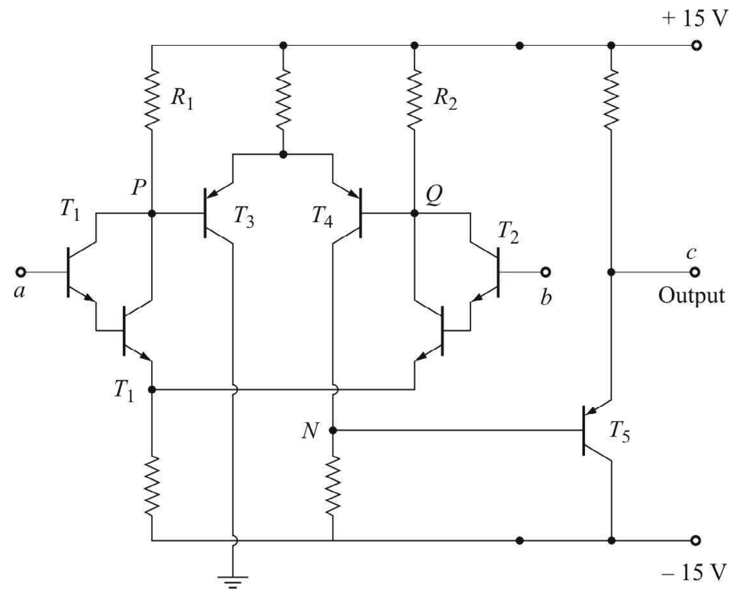
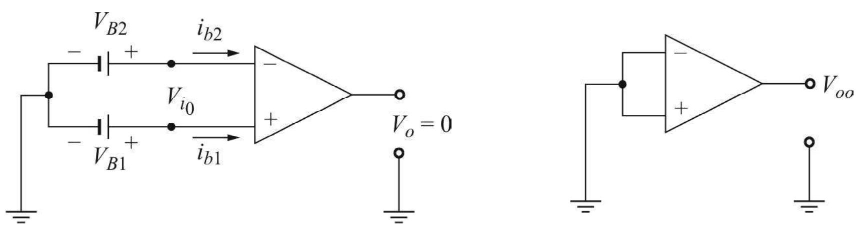


Fig. 3.2 *(a) Circuit of an IC Op-amp*

NOTES

3.2.2 Op-amp Parameters

1. **Offset Error Voltages and Current:** It is desired that the DC voltage at the output of ideal Op-amp is zero for equal input voltages v_{i1} and v_{i2} is, an ideal Op-amp is perfectly balanced or $v_o = 0$ for $v_{i1} = v_{i2}$. But in practice, because of unequal amount of current drawn by the input transistors of the first differential amplifier due to imbalance in the circuit caused by the mismatch of the built-in transistors following the inverting and the non-inverting terminals, the output voltage will not be zero.
2. **Input Offset Voltage:** Input offset voltage is the voltage that must be applied between the two input terminals of an Op-amp to nullify the output. Typical value of input offset voltage of μA 741 Op-amp is 1 mV, [Refer Figure 3.3 (*a*)].
3. **Input Offset Current:** Input offset current is the difference between the currents into the inverting and non-inverting input terminals of a balanced amplifier [Fig. 3.3 (*a*)]. Its typical value is 20 nA for a μA 741 Op-amp. Therefore, input offset current $i_{io} = i_{b1} - i_{b2}$, when $v_o = 0$. Improvement of matching between input terminals reduces this value and it depends on improvement of technology.



(a) Input offset voltage *(b) Output offset voltage*

Fig. 3.3

4. **Input Bias Current:** Input bias current is the average of the input currents that flow into the inverting and non-inverting input terminals of an Op-amp [Refer Figure 3.3 (*a*)].

   Therefore, input bias current

   $$i_B = \frac{i_{b1} + i_{b2}}{2}, \quad \text{when } v_o = 0.$$

   Typical values of input bias current for a μA 741 IC is 80 nA.
5. **Output Offset Voltage:** Output offset voltage V_{oo} is the voltage at the output terminals when the two input terminals are grounded. [Refer Figure 3.3 (*b*)].

   Arrangements are there in practical Op-amps to balance the offset voltage.
6. **Gain:** Ideal Op-amp has infinite open-loop differential gain. But practical Op-amps have open-loop gains that vary considerably according to the type of Op-amp used and ranges from 25,000 to 3,00,000. If v_1 is the input and v_2 is the output voltage of the Op-amp, the gain in decibels is

   $$\text{dB} = 20 \log_{10} \frac{V_2}{V_1}.$$

NOTES

7. **Input Resistance:** Ideal Op-amp has infinite input resistance. But practical Op-amps have input resistance ranging from 250 kΩ and 40 MΩ for the Op-amps with bipolar transistor input and 10^{12} Ω for the Op-amps with Field-Effect Transistor (FET) input.
8. **Output Resistance:** Op-amp is basically a voltage amplifier, therefore, its output resistance should be as low as possible. Practical Op-amps have output resistance of the order of 100 Ω.
9. **Slew rate:** The slew rate of an Op-amp is defined as the maximum rate at which its output voltage can vary. It is expressed in volts per micro second (V/μs) 741 Op-amp has a slew rate of 0.5 V/μs which means that the output of 741 Op-amp can change by 0.5 V every micro-second.

Thus, the maximum permissible output voltage of the Op-amp when a signal at a particular frequency is applied at its input is determined by slew rate and if we try to increase the output voltage even more than the maximum permissible value, the signal waveform gets distorted. This fact is shown in Figure 3.4.

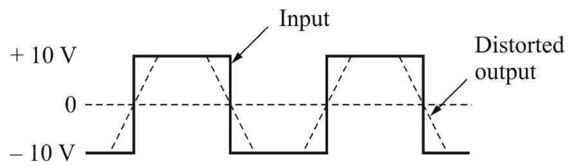


Fig. 3.4 *Illustrating the effect of inadequate slew rate*

The figure shows the input (square waveform) and output waveforms of an Op-amp with inadequate slew rate. Typically slew rate for Op-amps ranges from 0.3 – 12 V/μs. Slew rate can also be used to determine the maximum operating frequency of the Op-amp as follows:

$$f_{\max} = \frac{\text{Slew rate}}{2\pi V_{\text{peak}}}$$

where V_{peak} is the maximum or peak value of the input applied signal.

10. **Bandwidth:** The open-loop voltage gain of an Op-amp is not constant at all frequencies but drops at high frequencies due to capacitive effects. Manufacturers specify the gain-frequency characteristics of Op-amps in two ways:
    (*i*) The bandwidth for large signals over which less than 5% distortion is obtained.
    (*ii*) The frequency at which the gain falls to unity.

The bandwidth of an amplifier is defined as the range of frequencies for which the gain remains constant. The gain-bandwidth product of an Op-amp is always a constant.

As shown in Figure 3.5 shows that the voltage gain is constant (approximately 106 dB) at frequencies below 10 Hz, but the gain falls off at the constant rate of 6 dB octave at higher frequencies. Its value becomes 0 dB (i.e., 1) at a frequency of 1 MHz. This is known as unity-gain frequency (funity). Other Op-amps have similar gain-frequency characteristics but with different values of low frequency gain and unity-gain frequency.

NOTES

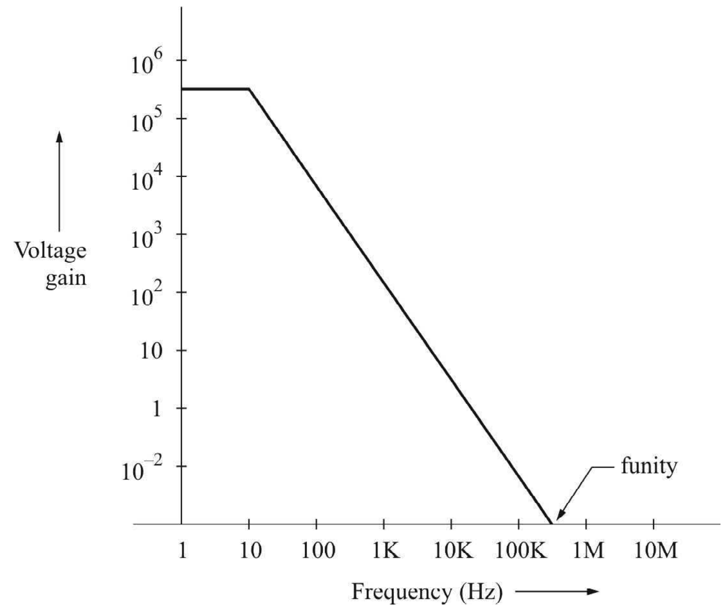


Fig. 3.5 *Open-loop voltage gain of Op-amp as a function of frequency*

11. **Input Capacitance:** Input capacitance is the equivalent capacitance that can be measured at either the inverting or non-inverting terminal with the other terminal connected to ground. The typical value of input capacitance is 1.4 pF for the μA 741 Op-amp.
12. **Transient Response:** Transient response is an important consideration for selecting an Op-amp for AC applications. Steady-state response is the response of a network after it attains a fixed value and is independent of time, but transient response is time-variant. The main characteristics of transient response are 'Rise Time' and 'Percent of Overshoot'. The rise time (also known as transition time) is defined as the time required by the output to go from 10% to 90% of its final value. Overshoot is the maximum amount by which the output deviates from the steady-state value and is generally expressed as a percentage. The rise time is 0.3 μsec and overshoot is 5% for μA 741 Op-amp. Higher bandwidth is achieved for a smaller value of rise time since the rise time varies inversely as the unity gain bandwidth of the Op-amp.
13. **Power Consumption:** The amount of quiescent power that must be consumed by the Op-amp for proper operation of the device is known as power consumption. Typical value of power consumption is 50 mW for a μA 741 Op-amp.
14. **Applications of Op-amp:** In the Op-amp applications described below, all voltages are measured with respect to ground. Balanced DC supplies (such as +15 V and –15 V) with respect to ground are required in order to energize the Op-amp circuits and are connected externally to the proper pins of the Op-amp and are not shown in the schematic circuits discussed.

NOTES

3.2.3 Inverting and Non-Inverting OP-amp

(A) Inverting Amplifier

The circuit diagram of an inverting amplifier using Op-amp is shown in Figure 3.6. The source voltage V_S is conneced in series with a resistance R_1 at the inverting input terminal of the Op-amp. A feedback resistor R_f is connected between the output and the inverting input terminal. This type of connection provides a negative feedback causing a decrease in the output signal as any increase in the output signal results in a feedback signal into the inverting input. This configuration is a voltage-shunt feedback amplifer. The non-inverting terminal (positive) is connected to ground.

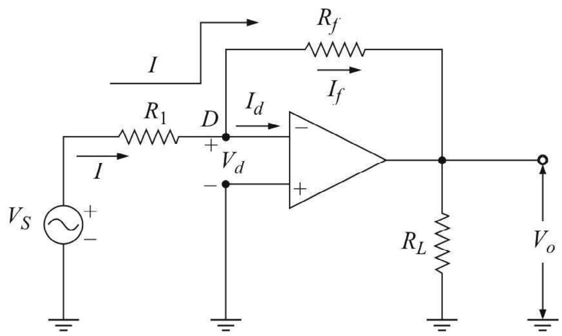


Fig. 3.6 *Circuit diagram of an inverting amplifier*

The input and output voltages are V_S and V_o respectively. Let V_d be the voltage at the inverting input terminal. As the open loop gain A of the Op-amp is very high and the output voltage V_o is finite due to negative feedback, we have, $V_d = \frac{V_o}{A}$ which tends to zero as $|A|$ tends to infinty. Therefore, the inverting input terminal is practically at ground potential. Thus, though the point D is not actually connected to ground. It is held virtually at ground potential independent of the magnitudes of V_S and V_o.

Difference between 'Actual Ground' and 'Virtual Ground': When a terminal is actually grounded, any amount of current can flow to ground through the terminal. Thus, actual ground serves as a 'Sink' for infinite current.

But no current can flow into the Op-amp through the virtual ground as the input impedance of the Op-amp is infinite. So, a virtual ground cannot serve as a sink for current.

Now since the point D is at virtual ground,

$$I_d = 0 \quad \text{and} \quad V_d = 0$$

So, we can write

$$I_d = I + I_f \qquad ...(3.7)$$

As $\quad I_d = 0 \quad$ (at virtual ground)

$$I = -I_f$$

or, $$\frac{V_s - V_d}{R_1} = \frac{V_d - V_o}{R_f}$$

As the point D is at virtual ground, voltage at point D is zero, i.e., $V_d = 0$.

So, $$\frac{V_s}{R_1} = -\frac{V_o}{R_f}$$

Thus, the voltage gain of the inverting amplifier is given by

$$A_V^{\text{INV}} = \frac{V_o}{V_s} = -\frac{R_f}{R_1} \qquad \text{...(3.8)}$$

NOTES

The negative sign in the gain expression indicates that the input and output signals are out of phase by 180°.

Due to this phase inversion, the amplifier shown in Figure 3.6 is called an inverting amplifier. From Equation (3.8) it is clear that the gain of the inverting amplifer is decided by selecting a ratio of feedback resistance R_f to the input resistance R_1. If may be noted that the inverting amplifier gain is independent of the load resistance R_L and any other parameter of the Op-amp. The gain can be fixed to any value even less than 1 and can be maintained accurately. The inverting amplifier is more applicable than non-inverting amplifier because of the unique nature of gain expression.

The input resistance of the amplifier system is

$$R_{\text{in}} = \frac{V_s}{I} = \frac{V_s}{(V_s - V_d)/R_1} \approx R_1 \qquad \text{...(3.9)}$$

(using the equation $I = \frac{V_s - V_d}{R_1}$ and noting that $V_d \approx 0$). R_{in} refers to the entire amplifier system and not to the Op-amp which has an infinte input impedance. The output resistance of the inverting amplifier is very small.

(B) Non-inverting Amplifier

The circuit diagram of a non-inverting amplifier using Op-amp is shown in Fig. 3.7.

The feedback resistance R_f is connected between the output and inverting input terminal and resistance R_1 is connected between inverting terminal and ground. The input signal V_s is connected to the non-inverting input terminal and a load R_L is connected at the output. The feedback circuit is composed of two resistors R_f and R_1. The feedback voltage is 180° out-of-phase with the input voltage, hence the feedback is negative. The circuit configuration is a voltage-series negative feedback amplifier.

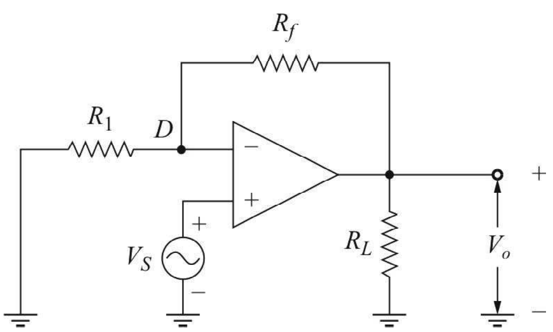


Fig. 3.7 *Circuit diagram of non-inverting amplifier*

The current flowing into the Op-amp is negligible since the input impedance is very large.

Now, applying Kirchhoff's current loop at the point D, we obtain

$$\frac{V_o - V_s}{R_f} = \frac{V_s}{R_1} \qquad \text{...(3.10)}$$

The non-inverting gain is given by

$$A_V = \frac{V_o}{V_s} = \frac{R_1 + R_f}{R_1}$$

or, $$A_V = 1 + \frac{R_f}{R_1} \quad \text{...(3.11)}$$

NOTES

From the Equation (3.11), it is clear that the voltage gain is greater than unity by a factor $\frac{R_f}{R_1}$.

The gain of the non-inverting amplifier is positive which means that input and output voltages are in phase and hence the name non-inverting amplifier. The gain is determined by the ratio of two resistors R_1 and R_f and is independent of source resistance, load resistance and parameters of the Op-Amp as long as assumptions of ideal Op-Amp are valid. The gain of this amplifier is very stable and its minimum value is 1 (when $R_f = 0$).

3.2.4 Use of 741 IC as Adder, Subtractor, Differentiator, and Integrator

A functional enhancer, often known as an operation amp or operation Amp, is a built-in circuit that is mostly used to do simple computations. It has a very high voltage gain, which is generally requested for 105. (100dB). Despite the fact that they were designed to perform numerical tasks such as expansion, deduction, mix, separation, and so on (hence the name Operational Amplifiers), they can also be used as an enhancer and for some, different capacities such as channels, comparators, and so on by utilising outside parts such as resistors and capacitors to make a necessary input instrument. In this post, we'll look at the IC 741 Op-amp, which is one of the most widely used Op-Amp ICs.

IC 741 Op-amp (Operational Amplifier)

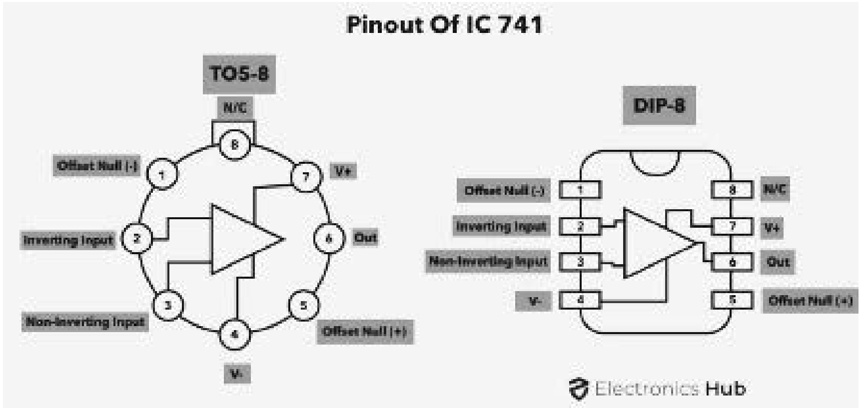
The 741 Op-amp IC is a well-designed integrated circuit that includes a versatile Operational Amplifier. Fairchild Semiconductors created it for the first time in 1963. This functional speaker IC includes 7 useable pins, 4 pins for data collection, and 1 result pin, as indicated by the number 741. The IC 741 Op-Amp can provide high voltage acquisition and operate over a wide range of voltages, making it the best choice for integrators, adding intensifiers, and general criticism applications. It also has a hamper and internal recurring remuneration circuit built in. This Op-Amp IC is available in the following structural factors:

- 8-pin DIP package
- TO5-8 metal can bundle
- 8-pin SOIC package

Pinout of IC 741 Op-amp and their Functions

The underneath figure outlines the pin designs and inward square graph of IC 741 of every 8 pin DIP and TO5-8 metal can bundle.

NOTES



- **Power Supply (Pin4 and Pin7):** The positive voltage supply terminal is Pin7, while the negative voltage supply terminal is Pin4. These pins provide power to the 741 IC, which allows it to function. Between these two pins, the voltage might be anywhere between 5 and 18 volts.
- **Pin6 (Output):** This is the IC 741's outcome pin. The voltage at this pin is determined by the signals at the data pins as well as the criticism system in use. Assuming that the result is expected to be high, this means that the voltage at the result is equal to the positive stock voltage. Assuming that the result is supposed to be low, this means that the voltage at the result is the same as the negative stock voltage.
- **Pins 2 and 3 (Input):** These are the IC's input pins. Pin2 is the non-altering input, while Pin3 is the reversing information. The result signal remains low if the voltage at Pin2 is higher than the voltage at Pin3, i.e. the voltage at the modifying input is higher. Essentially, the result goes high if the voltage at Pin3 is greater than the voltage at Pin2, i.e., the voltage at non-upsetting information is high.
- **Pin1 and Pin5 (Offset Null):** Due to the high addition provided by the 741 Op-Amp, even minor voltage differences at the upsetting and non-rearranging inputs, resulting by errors in the assembling cycle or external aggravations, might have an impact on the output. An offset voltage can be added at pins 1 and 5 to negate this effect, which is usually done with a potentiometer.
- **Pin8 (N/C):** No circuit in the 741 IC is connected to this pin. It's only a false lead that's used to fill the void in typical 8-pin bundles.

Specifications

The following are the IC 741 essential determinations:

- **Power Supply:** A minimum voltage of 5V is required, and it can endure up to 18V.
- **Voltage Gain:** 200,000 at low frequencies (200 V/mV) • Input Impedance: Approximately 2 M • **Yield Impedance:** Approximately 75
- **Maximum Output Current:** 20 mA • Maximum Suggested Output Load: greater than 2 K • Input Offset: 2 mV to 6 mV

0.5V/S slew rate (It is the rate at which an Op-Amp can distinguish voltage changes)

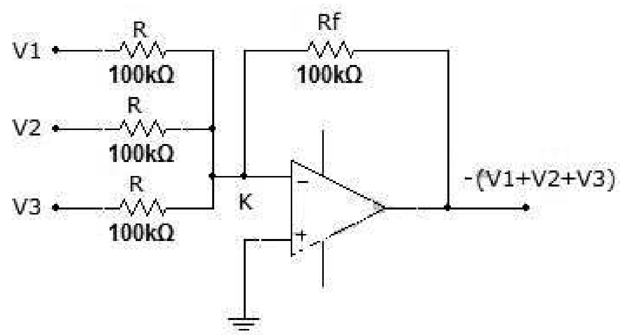
IC 741 is a voltage enhancer that comes close to being optimal due to its high information impedance and small result impedance.

NOTES

Adder Circuit Using Op-Amp 741

Adding speaker or a snake is utilized to total two sign voltages. Voltage snake circuit is a basic circuit that empowers you to add a few signals together. It has wide assortment of utilizations in electronic circuits. For instance, on an accuracy speaker, you might have to add a little voltage to drop the offset mistake of the operation amp itself. A sound blender is one more genuine instance of adding waveforms (sounds) together from various channels (vocals, instruments) prior to conveying the joined message to a recorder.

You can change the addition or add one more contribution without screwing up with the increases of different data sources. Simply recollect that the transforming adding intensifier circuit rearranges the info signals. That is not no joking matter. In case you want the contrary extremity, you should simply to put an altering stage previously or after the late spring.



- The information voltages V1, V2, and V3 are given to the viper circuit here.
- Because yield is the number of contributions with a sign shift, this is a reverse adding speaker.
- You can run one 'Modifying Speaker' with solidarity gain alongside this circuit to create a non-transforming snake.
- The outcome of this snake circuit is – (V1+V2+V3).
- Consider how much current flows through the information resistors.
- The current flowing through input resistor Rf is then given by the sum of these three flows, according to Kirchof's present law.
- Due to the fact that the point 'K' serves as a virtual ground point, this current will flow through the input resistor Rf. As a result, the voltage drop at Rf is given by -ve sign, and the operating amp related in upsetting mode is expected.

NOTES

- This circuit is known as an adding 'Intensifier' since it can give gain. By changing the worth of Rf the increase can be changed.
- Then, at that point, the result becomes

$$I_1 = \frac{V_1}{R},$$

$$I_2 = \frac{V_2}{R},$$

$$I_3 = \frac{V_3}{R}$$

$$I_K = I_1 + I_2 + I_3$$

$$= \frac{V_1}{R} + \frac{V_2}{R} + \frac{V_3}{R}$$

$$= \frac{1}{R}(V_1 + V_2 + V_3)$$

$$V_O = -I_K R$$

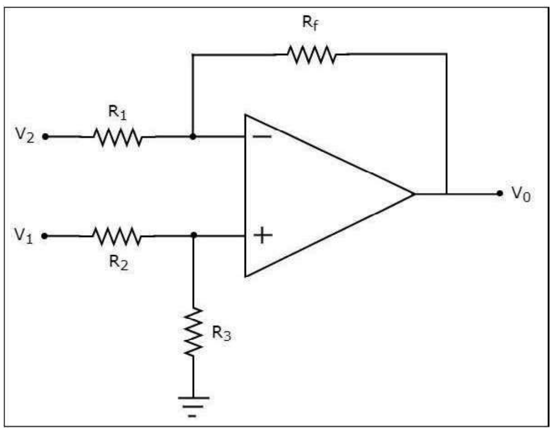
$$= -\frac{1}{R}(V_1 + V_2 + V_3) R$$

$$= -(V_1 + V_2 + V_3)$$

$$V_O = -\frac{R_F}{R}(V_1 + V_2 + V_3)$$

Subtractor Utilizing 741 IC

A subtractor is an electrical circuit that produces a result that is the same as the difference between the applied information sources. This section delves into the operation of an amp-based subtractor circuit. The difference between the information voltages applied at its transforming and non-rearranging terminals is equivalent to the result of an operation amp based subtractor. It's also known as a distinction intensifier because the end result is improved. The circuit diagram for an operation amp based subtractor is shown in the diagram below:



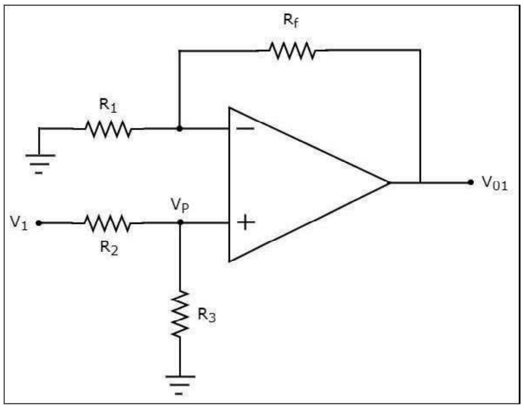
Presently, let us find the articulation for yield voltage V_0 of the above circuit utilizing superposition hypothesis utilizing the accompanying advances –

Step 1

NOTES

First and foremost, let us work out the result voltage V_{01} by considering just V_1.

For this, dispose of V_2 by making it hamper. Then, at that point, we acquire the adjusted circuit outline as displayed in the accompanying figure "



Presently, utilizing the voltage division guideline, ascertain the voltage at the non-modifying input terminal of the operation amp.

$=> V_p = V_1 (R_3 R_2 + R_3)$

Presently, the above circuit resembles a non-upsetting enhancer having input voltage V_p. Thusly, the result voltage V_{01} of above circuit will be

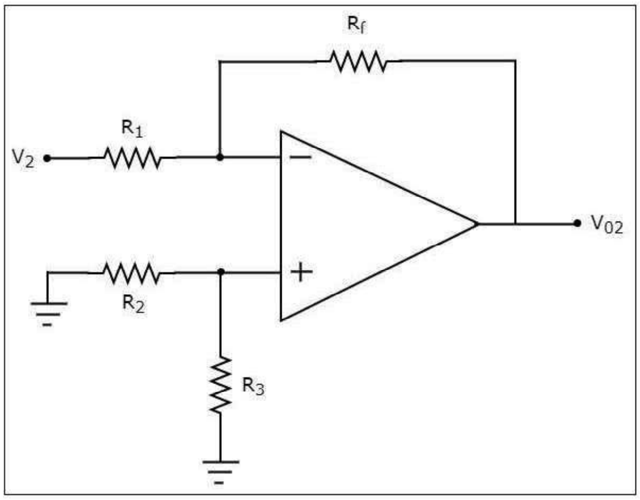
$V_{01} = V_p(1 + R_f R_1)$

Substitute, the worth of V_p in above condition, we acquire the result voltage V_{01} by considering just V_1, as–

$V_{01} = V_1(R_3 R_2 + R_3) (1 + R_f R_1)$

Step 2

In this progression, let us track down the result voltage, V_{02} by considering just V_2. Like that in the above advance, dispose of V_1 by making it hamper. The adjusted circuit chart is displayed in the accompanying figure.



NOTES

It can see that the voltage at the non-altering input terminal of the operation amp will be zero volts. That is to say, the above circuit is essentially a reversing operation amp. In this way, the result voltage V_{02} of above circuit will be–

$V_{02} = (-R_f R_1)\ V_2$

Step 3

In this progression, we will acquire the result voltage V_0 of the subtractor circuit by adding the result voltages got in Step1 and Step2. Numerically, it tends to be composed as

$V_0 = V_{01} + V_{02}$

Subbing the upsides of V_{01} and V_{02} in the above condition, we get–

$V_0 = V_1(R_3 R_2 + R_3)\ (1 + R_f R_1) + (-R_f R_1)V_2$

$=> V_0 = V\ (R_3 R_2 + R_3)\ (1 + R_f R_1) - (R_f R_1)V_2$

In the event that $R_f = R_1 = R_2 = R_3 = R$, the result voltage $V_0 V_0$ will be

$V_0 = V_1(R_R + R)(1 + R_R) - (R_R)V_2$

$=> V_0 = V_1(R_2 R_2) - (1)V_2$

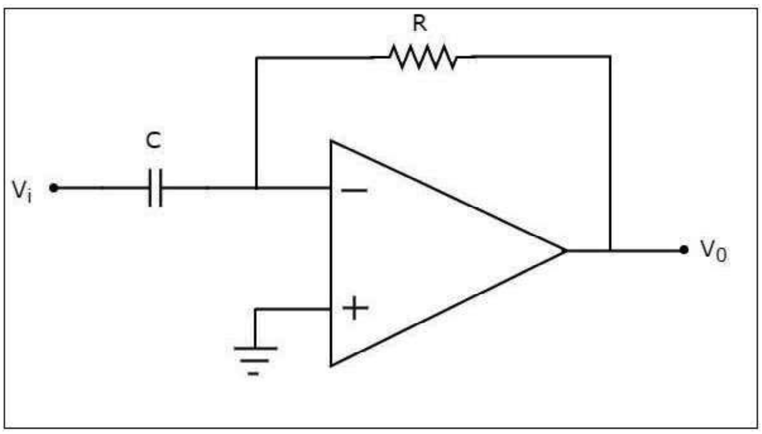
$V_0 = V_1 - V_2$

Along these lines, the operation amp based subtractor circuit talked about above will create a result, which is the distinction of two info voltages V_1 and V_2, when every one of the resistors present in the circuit are of same worth.

OP-amp as Differentiator

A differentiator is an electronic circuit that creates a result equivalent to the main subordinate of its feedback. This part examines about the operation amp-based differentiator exhaustively.

An operation amp-based differentiator delivers a result, which is equivalent to the differential of information voltage that is applied to its rearranging terminal. The circuit graph of an operation amp-based differentiator is displayed in the accompanying figure–



In the above circuit, the non-reversing input terminal of the operation amp is associated with ground. That implies zero volts is applied to its non-reversing input terminal.

As per the virtual short idea, the voltage at the altering input terminal of OP-amp will be equivalent to the voltage present at its non-upsetting information terminal. Thus, the voltage at the altering input terminal of operation amp will be zero volts.

NOTES

The nodal condition at the altering input terminal's hub is–

$C_d(0–V_i)/d_t+0–V_0/R=0$

$–C_dV_i/d_t=V_0/R$

$=>V_0 = –RC_dV_i/d_t$

On the off chance that RC=1sec, the result voltage V_0 will be–

$V_0 = –dV_i/d_t$

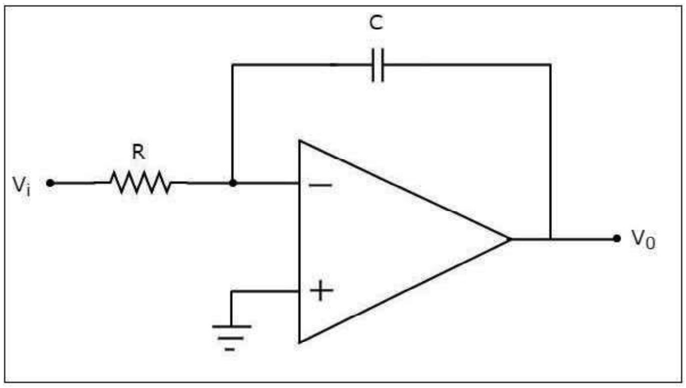
Consequently, the operation amp-based differentiator circuit displayed above will deliver a result, which is the differential of info voltage V_i, when the sizes of impedances of resistor and capacitor are proportional to one another.

Note that the result voltage V_0 is having a negative sign, which shows that there exists a 1800 stage contrast between the info and the result.

Integrator

An integrator is an electronic circuit that creates a result that is the mix of the applied info. This segment examines about the operation amp-based integrator.

An operation amp-based integrator creates a result, which is an essential of the information voltage applied to its modifying terminal. The circuit graph of an operation amp-based integrator is displayed in the accompanying figure–



In the circuit displayed over, the non-transforming input terminal of the operation amp is associated with ground. That implies zero volts is applied to its non-upsetting info terminal.

As per virtual short idea, the voltage at the reversing input terminal of operation amp will be equivalent to the voltage present at its non-rearranging input terminal. Along these lines, the voltage at the modifying input terminal of operation amp will be zero volts.

The nodal condition at the altering input terminal is–

$0–V_i/R+C_d(0–V_0)/d_t = 0$

$=>–V_i/R = C_dV_0/d_t$

=>$dV_0/d_t = –V_i/RC$

=>$dV_0=(–V_i/RC)\ d_t$

Incorporating the two sides of the situation displayed above, we get–

NOTES

$+–dV_0=+–(–V_i/RC)dt$

=>$V_0= –1/RC+ –V_t d_t$

On the off chance that RC=1sec, the result voltage, V_0 will be –

$V_0= + –V_i d_t$

In this way, the operation amp based integrator circuit examined above will create a result, which is the basic of information voltage V_i, when the greatness of impedances of resistor and capacitor are proportional to one another.

Note: The result voltage, V_0 is having a negative sign, which shows that there exists 1800 stage distinction between the information and the result.

3.2.5 Op-amp as Constant Current Source

The circuit diagram of dual-input balanced-output difference amplifier as shown in Figure 3.8 sets up the DC emitter current by using a combination of R_E and V_{EE}.

This arrangement can be replaced by using a constant current bias circuit which provides better current stabilization and stable operting point. This type of arrangement is shown in Figure 3.8. A constant current transistor (T_3) circuit is used in place of the emitter resistor R_E. The resistors R_1, R_2 and R_E set up the dc collector current in transistor T_3.

Neglecting the base loading effect, the voltage at the base of T_3 is given by

$$V_{B3} = \frac{-R_2 V_{EE}}{R_1 + R_2} \quad \text{and} \quad V_{E3} = V_{B3} - V_{BE3}$$

Therefore,

$$I_{E3} \cong I_{C3} = \frac{V_{E3} - (-V_{EE})}{R_E}$$

$$= \frac{V_{EE} - [R_2 V_{EE}/(R_1 + R_2)] - V_{BE3}}{R_E} \qquad ...(3.12)$$

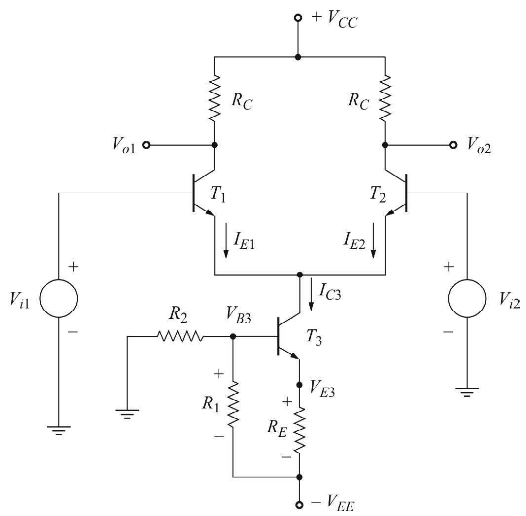


Fig. 3.8 Differential amplifier circuit using a constant current bias

The collector current of transistor T_3, i.e., I_{C3} is divided into two halves with each half flowing through the emitter of transistors T_1 and T_2 due to symmetry of the circuit.

NOTES

$$I_{E1} = I_{E2} \cong \frac{I_{C3}}{2}$$

$$= \frac{V_{EE} - [R_2 V_{EE}/(R_1 + R_2)] - V_{BE3}}{2R_E} \quad ...(3.13)$$

The collector current I_{C3} is fixed due to no signal injection into either the emitter or base of T_3. Thus, the transistor T_3 acts as a soure of constant emitter current for transistors T_1 and T_2 of the differential amplifier.

The constant current bias provides constant emitter current and also a very high source resistance since an ideal DC current source has infinite source resistance at the emiter which improves the CMRR of the amplifier. So, all the performance equations of the differential amplifer with emitter bias will also be applicable to the differential amplifier with constant current bias. Using diodes instead of R_1 improves the thermal stability of a constant current transister T_3.

Current mirror is another example of constant current source in which the output current is forced to be equal the input current. The current mirror is a special case of constant current bias, and therefore, can be used to set up constant emitter currents in differential amplifier stages.

The *n-p-n* transistors Q_1 and Q_2 make up the differential amplifier and the *p-n-p* transistors Q_3 and Q_4 make up the current mirror.

The current mirror acts as the collector load and provides a high effective collector load resistance, increasing the gain, thus producing a gain of 5000 or more with no load but decreases with loading. It is used in feedback loops and as a comparator.

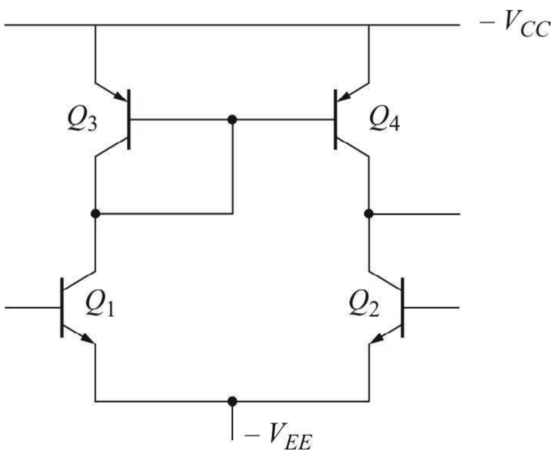


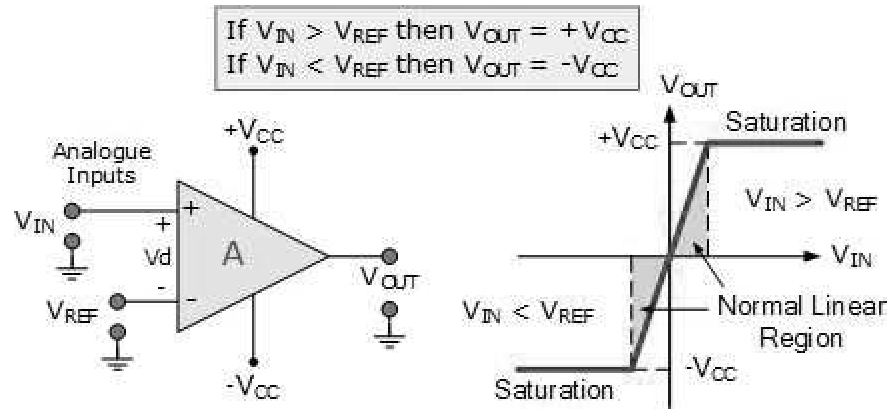
Fig. 3.9 *Circuit configuration of a typical current mirror circuit used in an Op-amp*

The current mirror circuit is commonly used in integrated amplifiers such as differential and operational amplifiers as it uses fewer components than constant current bias circuit because of its simplicity and ease of fabrication in I_C form.

3.2.6 Comparator, Square and Triangular Wave Generator

NOTES

1. Op-amp Comparator Circuit



The comparator is an electronic dynamic circuit that employs a functional enhancer with a very high addition in its open-circle form, i.e., no criticism resistor.

The Op-amp comparator compares two simple voltage levels, or a predefined reference value, VREF, and generates a result signal based on the voltage correlation. The operation amp voltage comparator, at the end of the day, considers the sizes of two voltage inputs and determines which is the larger of the two.

We've shown in previous instructional activities how the functional enhancer can be used with negative input to adjust the magnitude of its result signal in the straight district while performing a variety of functions. We've also shown that the open-circle gain AO of a conventional functional intensifier describes its outcome voltage, which is given by the articulation: $V_{OUT} = AO(V+ - V-)$, where V+ and V- are the individual voltages at the noninverting and rearranging terminals.

Voltage comparators, on the other hand, use either positive criticism or no input at all (open-circle mode) to switch their result between two saturated states, because the intensifier's voltage gain is fundamentally similar to AVO in the openloop mode. The result from the comparator then swings either entirely to its positive stockpile rail, +Vcc, or completely to its negative stockpile rail, - Vcc, on the usage of changing information signal that passes some preset edge esteem because of this high open circle gain.

The open-circle operation amp comparator is a basic circuit that operates in a non-straight location as variations in the two simple data sources, V+ and V, cause it to act like a computerised bistable device as setting off causes it to have two outputs. The voltage comparator is basically a 1-bit simple to computerised converter at that moment, because the information signal is simple but the output acts carefully.

Allows first to expect that V_{IN} is not exactly the DC voltage level at V_{REF}, (V_{IN} V_{REF}), in the operation amp comparator circuit above. The result will be LOW and at the negative stock voltage, - Vcc, resulting in a negative immersion of the result because the comparator's non-altering (positive) contribution is not exactly the reversing (negative) input.

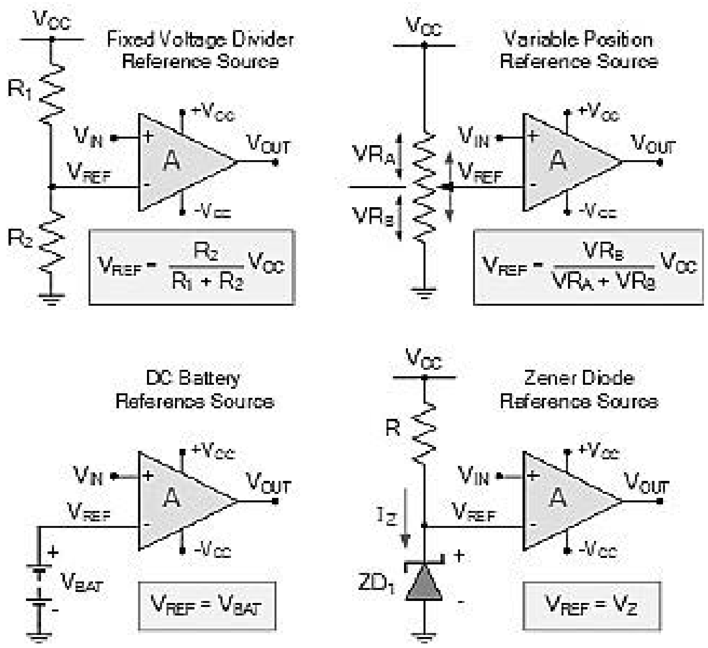
If we now increase the information voltage, V_{IN}, so that its value is greater than the reference voltage V_{REF} on the rearranging input, the result voltage swiftly shifts HIGH to the positive stock voltage, +Vcc, causing the result to be positive immersed. Assuming we reduce the info voltage V_{IN} to a value that is somewhat less than the reference voltage, the operation amp's result returns to its negative immersion voltage, which serves as a limit indicator.

Then we can see that the operation amp voltage comparator is a device whose result is dependent on the value of the information voltage, V_{IN} concerning some DC voltage level, as the result is HIGH when the voltage on the non-reversing input is greater than the voltage on the upsetting information, and LOW when the non-reversing input is not exactly the altering input voltage. This condition holds true whether the info signal is linked with the comparator's rearranging or non-transforming contribution.

We can also see that the result voltage's value is completely dependent on the OP-Amp's power supply voltage. Because of the operating amp's enormous open-circle gain, the size of its result voltage might theoretically be infinite in two ways, (). Regardless, and for obvious reasons, it is limited by the operation amps supply rails, which give V_{OUT} = +Vcc or V_{OUT} = -Vcc.

The essential functioning amp comparator generates a positive or negative voltage yield by comparing its feedback voltage to a fixed DC reference voltage, as we previously stated. A resistive voltage divider is typically used to set the information reference voltage of a comparator, but a battery supply, zener diode, or potentiometer for a variable reference voltage might all be used as shown.

Comparator Reference Voltages



In principle the comparators reference voltage can be set to be anyplace among 0v and the inventory voltage however there are useful restrictions on the real voltage range contingent upon the operation amp comparator being gadget utilized.

2. Square and Triangular Wave Generator

A waveform generator is an electronic circuit, which creates a standard wave. There are two kinds of operation amp-based waveform generators–

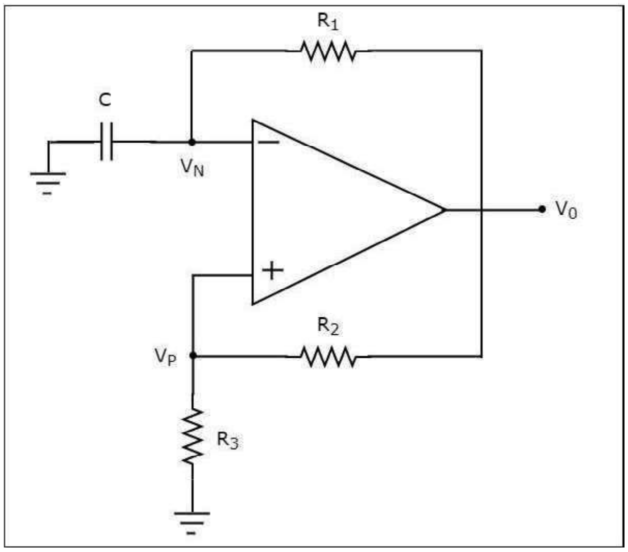
NOTES

- Square wave generator
- Three-sided wave generator

Square Wave Generator

A square wave generator is an electronic circuit which creates square wave. This segment examines about operation amp based square wave generators.

The circuit chart of an operation amp based square wave generator is displayed in the accompanying figure



Notice how the resistor R_1 is connected between the operation amp's modifying input terminal and the result of the operation amp in the circuit diagram above. The resistor R_1 is used in the negative criticism in this manner. In essence, the resistor R_2 is connected to the operating amp's noninverting input terminal and the result. The resistor R_2 is used in this way as a positive input resistor.

Between the changing input terminal of the operation amp and ground lies a capacitor C. As a result, the voltage across capacitor C will be the information voltage at the operating amp's converting terminal. A resistor R_3 is also connected between the operation amp's non-upsetting info terminal and ground. As a result, the voltage across resistor R_3 will be the information voltage at the operating amp's non-transforming terminal.

The activity of a square wave generator is clarified underneath–

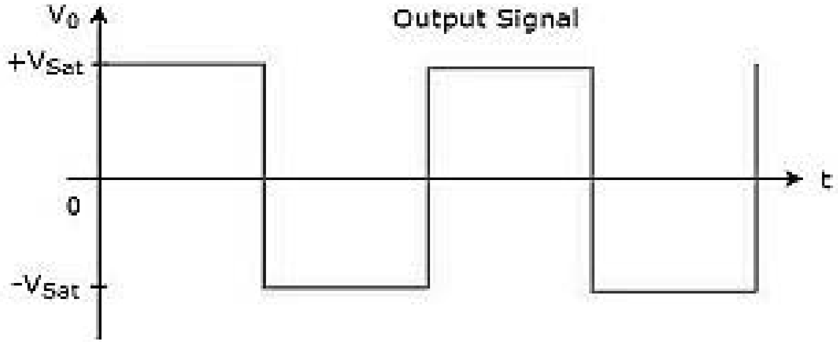
- Accept that no charge is initially stored in the capacitor. At that time, the voltage present at the operating amp's upsetting terminal is zero volts. However, there is some balanced voltage at the operating amp's non-altering terminal. As a result, the value present at the output of the aforementioned circuit will be $+V_{sat}$.

NOTES

- Currently, the capacitor C is being charged via a resistor R_1. When the voltage across the capacitor C is only slightly higher than the voltage (positive worth) across resistor R_3, the value at the output of the aforesaid circuit changes to V_{sat}.
- When the result of the aforesaid circuit is V_{sat}, the capacitor C begins to release through a resistor R_1. When the voltage across capacitor C reaches only slightly less than (worse) the voltage (negative worth) across capacitor A, the worth present at the result of the aforesaid circuit will change to $+V_{sat}$.

Hence, the circuit displayed in the above chart will create a square wave at the result as displayed in the accompanying figure –

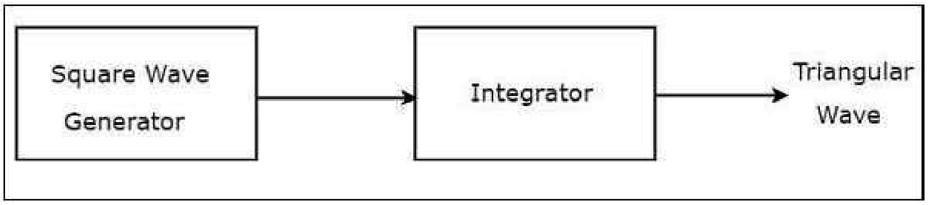
From the above figure we can see that the result of square wave generator will have one of the two qualities: $+V_{sat}$ and $–V_{sat}$. Thus, the result stays at one incentive for some term and afterward changes to one more worth and stays there for some span. Thusly, it proceeds.



The output of the square wave generator will have one of two values: $+V_{sat}$ or $–V_{sat}$, as shown in the diagram above. As a result, the output remains at one value for a period of time before transitioning to another value and remaining there for another period of time. It goes on in this manner.

Triangular Wave Generator

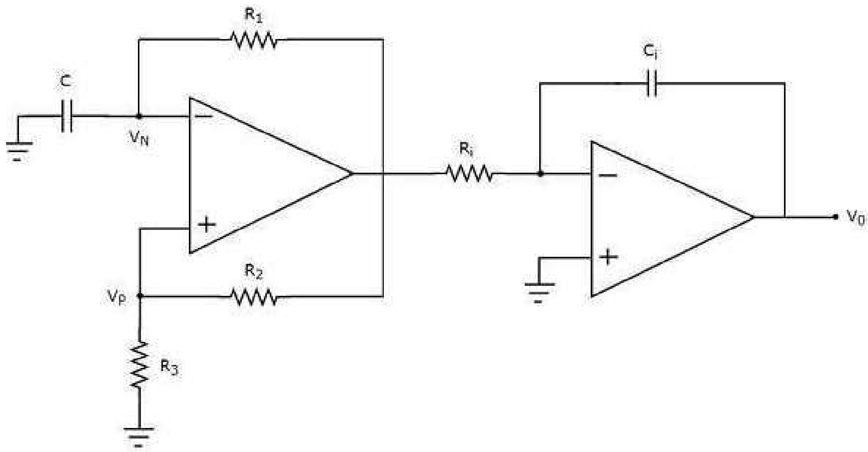
An electronic circuit that generates a three-sided wave is known as a three-sided wave generator. The preceding figure shows the square graph of a three-sided wave generator –



A three-sided wave generator's square chart consists mostly of two squares: a square wave generator and an integrator. These two squares have fallen apart. That is, the outcome of the square wave generator is used as an integrator contribution. It's worth noting that a square wave's joining is only a three-sided wave.

The circuit diagram for a three-sided wave generator based on operation amps is shown in the adjacent figure –

NOTES



The circuit graphs of a square wave generator and an integrator have been effectively seen. By replacing the squares in the square graph of a three-sided wave generator with individual circuit outlines, we generated the above circuit chart of an operating amp based three-sided wave generator.

Check Your Progress

1. Define the term differential amplifier.
2. What are the four differential amplifier configurations?
3. Write a short note on slew rate.
4. How manufactures specify the gain-frequency characteristics of Op-amps in two ways?
5. What is the difference between actual ground and virtual ground?
6. How will you define the adder circuit using Op-amp 741?
7. Define the term waveform generator.

3.3 VOLTAGE MULTIPLIERS CIRCUITS

Introduction: A voltage multiplier is a circuit that generates a d.c. voltage that is equivalent to a number of different discrete input voltages. At least two pinnacle identifiers or rectifiers are included. Voltage multipliers were used to track down applications in circuits that required high voltage with low current, such as picture tubes in TV receivers, oscilloscopes, and so on.

A voltage multiplier is an electrical circuit that changes over AC electrical power from a lower voltage to a higher DC voltage through capacitors and diodes joined into an organization.

Contingent upon the result voltage, multipliers can be of various kinds:

- Voltage Doublers
- Voltage Drunkards
- Voltage Quadrupler

Voltage Doublers

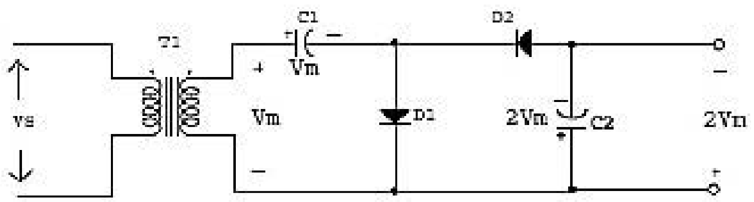
NOTES

A Voltage doubler creates a DC voltage twice the rms worth of the information a.c. voltage. Voltage doubler can be of two sorts:

- Half wave voltage doubler
- Full wave voltage doubler

Half Wave Voltage Doubler

The circuit for a half wave voltage doubler is shown in Figure 3.10 (a) D_1 behaves and D_2 is cut off during the positive half pattern of the optional voltage diode. Capacitor C_1 is currently charged to the pinnacle modified voltage Vm, as shown in the figure. The optional voltage is in series with the voltage across the capacitor C_1 during the negative half cycle. C_2 will attempt to charge to 2Vm (Vm of the information and Vm of the capacitor C_1) in this manner. The voltage across the capacitor C_2 will be equivalent to 2Vm after a few cycles. (Refer Figure 3.10) We can aggregate the voltages around the external circle since diode D_2 is a short during the negative half-cycle (and diode D_1 is open). For example, Vm VC_1 VC_2 = 0 or Vm Vm VC_2 = 0 from which VC_2 = 2VM



Half Wave Voltage Doubler

Capacitor C_1 in the circuit will release in the negative half cycle. It starts charging again in the positive half cycle. As a result, the half wave voltage doubler provides one-half cycle voltage to the heap. As a result, the half-wave voltage doubler's guideline is weak.

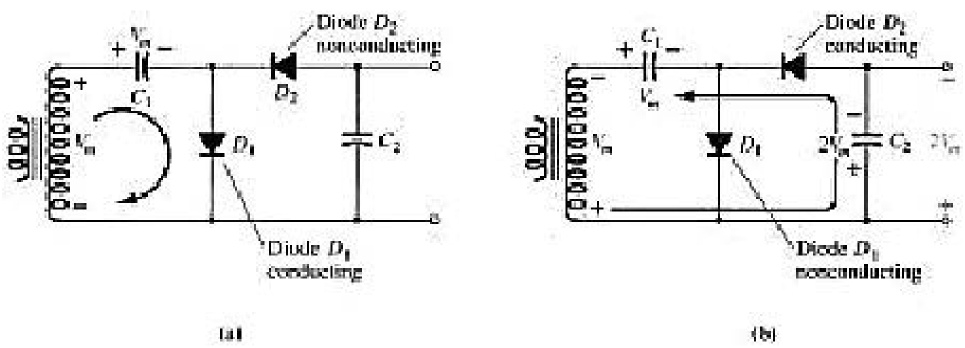


Fig. 3.10 Half Wave Voltage Doubler During (a) Positive and (b) Negative half patterns of information

Full Wave Voltage Doubler

Figure 3.10 (b) shows another voltage doubler circuit termed a full wave voltage doubler. The optional voltage diode D_1 conducts during the positive half pattern, charging the capacitor C_1 to the maximum voltage Vm. Diode D_2 is currently non-directing. Diode D_2 conducts during the negative half cycle, charging capacitor C_2

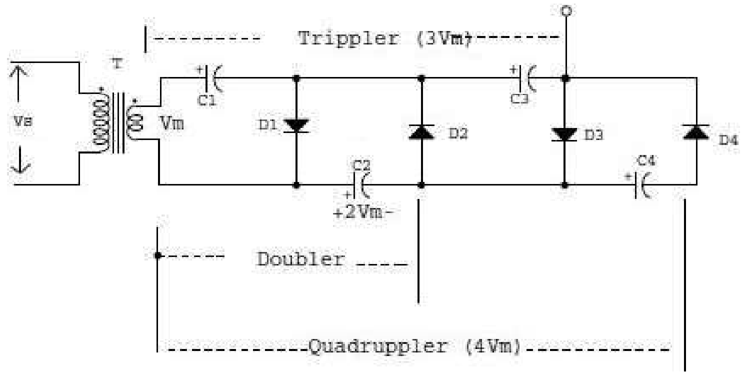
NOTES

to Vm with the extremity as checked, while diode D_2 is non-directing. Because the two capacitors C_1 and C_2 are connected in series, the final voltage is around 2Vm. Because one of the result capacitors is charged per half pattern of the information voltage, this circuit is called a full wave voltage doubler.

Voltage Trippler/Quadruppler

A voltage trippler/quadruppler generates a DC voltage three/four times that of the information na.c. voltage. The voltage trippler/quadruppler circuit shown in the diagram is an enhancement of the half wave voltage doubler.

The diode D_1 conducts during the positive half cycle, charging the capacitor C_1 to Vm with the extreme indicated. The diode D_2 conducts charging the capacitor C_2 to 2Vm during the primary negative half cycle. During the next positive half cycle, the diode D_3 conducts in place of D_1, charging the capacitor C_1 while the voltage across the capacitor C_2 charges capacitor C_3 to the same value of 2Vm. The diodes D_2 and D_4 behave differently on the succeeding negative half cycle, and capacitor C_3 charges C_4 to 2Vm. As a result, the voltage across C_2 is 2Vm, 3Vm across C_1 and C_3, and 4Vm between C_2 and C_4.



Assuming extra part of diode and capacitor are utilized, every capacitor will be charged to 2Vm. Estimating from the highest point of the transformer winding will give odd products of Vm at the result, though estimating from the lower part of the transformer, the result voltage will give even products of the pinnacle voltage Vm.

3.3.1 Wave Shaping Circuits

A wave forming circuit is the one which can be utilized to change the state of a waveform from rotating current or direct current.

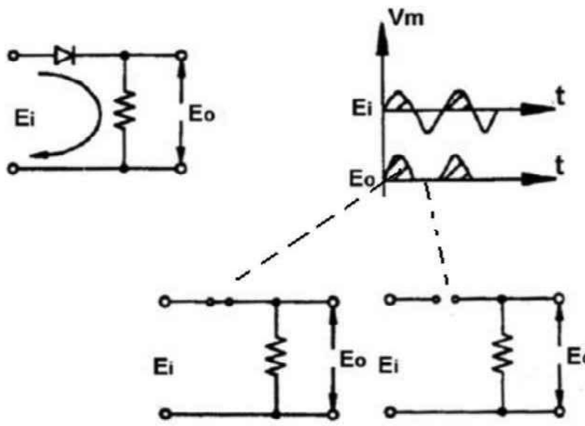
For instance, a trimmer circuit is utilized to keep the waveform voltage from surpassing the foreordained voltage without influencing the excess piece of the waveform. This is only waveshaping.

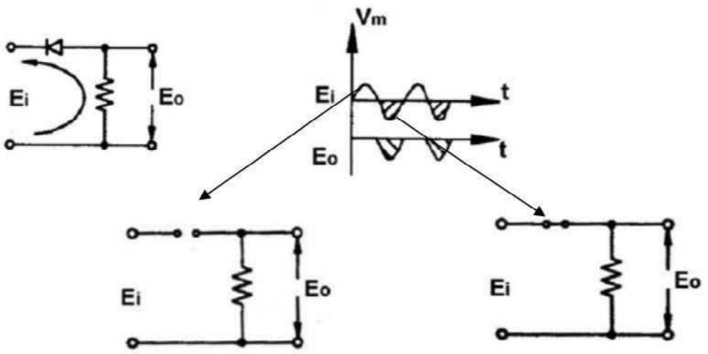
3.3.2 Clipping, Clamping, Differentiating, and Integrating Circuits

Diode Clippers

- The Diode Clipper, otherwise called a Diode Limiter, is a wave molding circuit that takes an information waveform and clasps or cuts off its top half, base half or the two parts together.

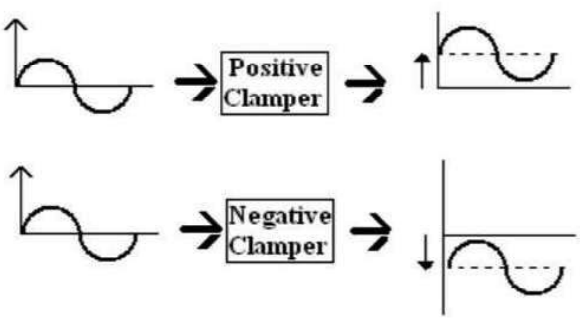
- There are two sorts of trimmer circuits, the series and equal diode cutting circuits.
- Series Diode Clipping Circuit
- In these sorts of circuits, the diode is associated between the information and result voltage terminals.
- As the accompanying figure uncovers, the negative pattern of the information voltage can be cut off by this sort of series trimmers.
- Opposite of the diode pins respects a positive cycle cutting circuit





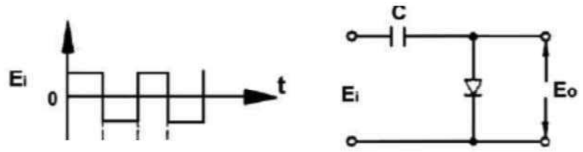
Clamper Circuits

- Clamper Circuits, or momentarily clampers are utilized to change the D.C. level of a sign to an ideal worth.



- Being unique in relation to trimmers, bracing circuits utilizes a capacitor and a diode association. At the point when diode is in its on express, the result voltage equivalents to diode drop voltage (in a perfect world zero) or more the voltage source, assuming any.

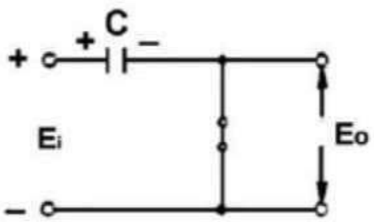
NOTES



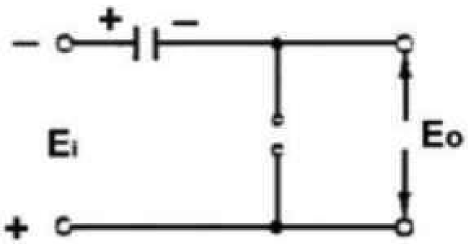
Typical Clamping Circuit

- As you probably are aware, this circuit, truth be told, is a series R-C circuit.
- The obstruction of diode (A few ohms over its drop voltage) and the little capacitance respect a modest consistent for this circuit.
- This implies that the capacitor will quickly be charged if any info voltage, that is sufficient to swtich on the diode, is applied.
- The diode will lead during the positive pattern of the info sign and result voltage will be in a perfect world zero (Practically speaking this voltage rises to ~0.6 V).

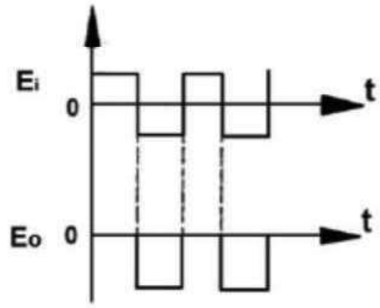
Diode Conducts During Positive Cycle



- Note that during positive cycle the capacitor is quickly accused in converse extremity of the info voltage. After progress to negative cycle, the diode becomes to its off state.
- For this situation, the result voltage equivalents to the amount of the info voltage and the voltage across the terminals of the capacitor which have a similar extremity with one another.
- $\mathbf{E_0 = -(|E_i| + |E_c|)}$
- Diode is switched off during negative cycle



- The subsequent sign after a total cycle is displayed beneath.

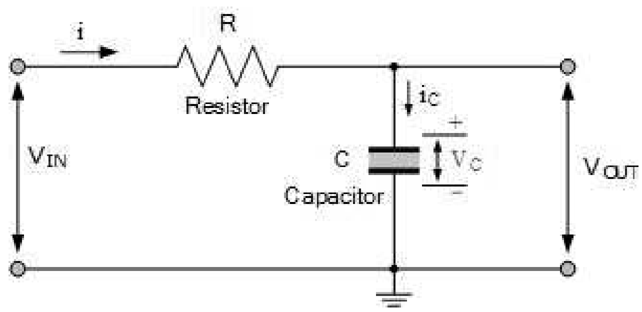


NOTES

- By this interaction, the information signal is moved to negative DC esteem (its most extreme worth is in a perfect world zero) with next to no adjustment of its sufficiency preferably.

RC Integrator

- The RC integrator is a series associated Resistor-Capacitor network that delivers a result signal which compares to the numerical course of combination.
- For an inactive RC integrator circuit, the information is associated with an obstruction while the result voltage is taken from across a capacitor being the specific inverse to the RC Differentiator Circuit. The capacitor energizes when the info is high and releases when the information is low.
- An inactive RC network is just a resistor in series with a capacitor that is a proper obstruction in series with a capacitor that has a recurrence dependant reactance which diminishes as the recurrence across its plates increments.
- Accordingly at low frequencies the reactance, Xc of the capacitor is high while at high frequencies its reactance is low because of the standard capacitive reactance equation of $Xc = 1/(2\pi f C)$.
- Then, at that point, assuming the information signal is a sine wave, a RC integrator will basically go about as a basic Low Pass Channel (LPF) with a cut-off or corner recurrence that compares to the RC time consistent (tau, τ) of the series organization and whose result is diminished over this remove recurrence point. Consequently, when taken care of with an unadulterated sine wave a RC integrator goes about as a detached low pass channel.
- Consequently, the pace of charging or releasing relies upon the RC time consistent, $\tau = RC$.



NOTES

- For a RC integrator circuit, the info signal is applied to the obstruction with the result taken across the capacitor, then, at that point, V_{OUT} rises to V_C.
- As the capacitor is a recurrence dependant component, the measure of charge that is set up across the plates is equivalent to the time area necessary of the current.
- That is it requires some investment for the capacitor to completely charge as the capacitor can not charge promptly just charge dramatically.
- Along these lines the capacitor current can be composed as:

$$i_{C(t)} = C\frac{dV_{C(t)}}{dt}$$

- This fundamental condition above of i_C = C(dVc/dt) can likewise be communicated as the quick pace of progress of charge, Q regarding time providing us with the accompanying standard condition of: i_C = dQ/dt where the charge Q = Cx Vc, that is capacitance times voltage.
- The rate at which the capacitor charges (or releases) is straightforwardly corresponding to the measure of the obstruction and capacitance giving the time consistent of the circuit.
- Since capacitance is equivalent to Q/Vc where electrical charge, Q is the progression of a flow (I) over the long run (t), that is the result of I x t in coulombs, and from Ohms law we realize that voltage (V) is equivalent to I x R, subbing these into the situation for the RC time consistent gives:

RC Time Constant

$$RC = R\frac{Q}{V} = R\frac{i\times T}{i\times R} = \cancel{R}\frac{\cancel{i}\times T}{\cancel{i}\times \cancel{R}} = T$$
$$\therefore T = RC$$

Capacitor Voltage

- The capacitors current can be communicated as the pace of progress of charge, Q regarding time.
- In this manner, from an essential standard of differential analytics, the subsidiary of Q as for time is dQ/dt and as I = dQ/dt we get the accompanying relationship of:
- Q = +– idt (the charge Q on the capacitor at any moment on schedule)
- Since the info is associated with the resistor, a similar current, I should go through both the resistor and the capacitor ($i_R = i_C$) delivering a V_R voltage drop across the resistor so the current, (I) coursing through this series RC network is given as:

$$i(t) = \frac{V_{IN}}{R} = \frac{V_R}{R} = C\frac{dV}{dt}$$
$$V_{OUT} = V_C = \frac{Q}{C} = \frac{\int i dt}{C} = \frac{1}{C}\int i(t)dt$$

As $I = V_{IN}/R$, subbing and improving to settle for V_{OUT} as an element of time gives:

$$V_{OUT} = \frac{1}{C}\int\left(\frac{V_{IN}}{R}\right)dt = \frac{1}{RC}\int V_{IN}\,dt$$

NOTES

- Then, at that point, expecting the underlying charge on the capacitor is zero, that is $V_{OUT} = 0$, and the information voltage V_{IN} is consistent, the result voltage, V_{OUT} is communicated in the time area as:

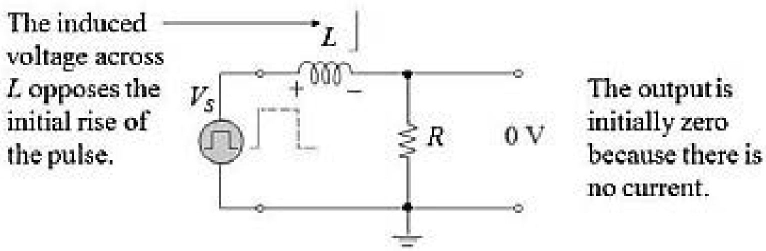
Capacitor Charging

$$V_{OUT} = \frac{1}{RC}\int_0^t V_{IN(t)}\,dt$$

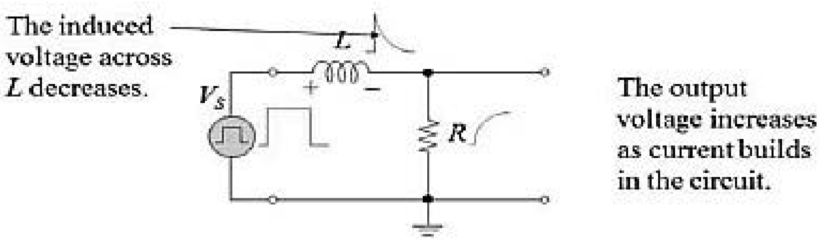
$$V_{C(t)} = V\left(1 - e^{-\left(\frac{t}{RC}\right)}\right)$$

The RL Integrator

- Like the RC integrator, a RL integrator is a circuit that approximates the numerical course of joining.
- Under identical conditions, the waveforms resemble the RC integrator.
- For a RL circuit, t = L/R.
- An essential RL integrator circuit is a resistor in series with an inductor and the source. The result is taken across the resistor.
- At the point when the beat generator yield goes high, a voltage promptly shows up across the inductor as per Lenz's law.
- The quick current is zero, so the resistor voltage is at first zero.

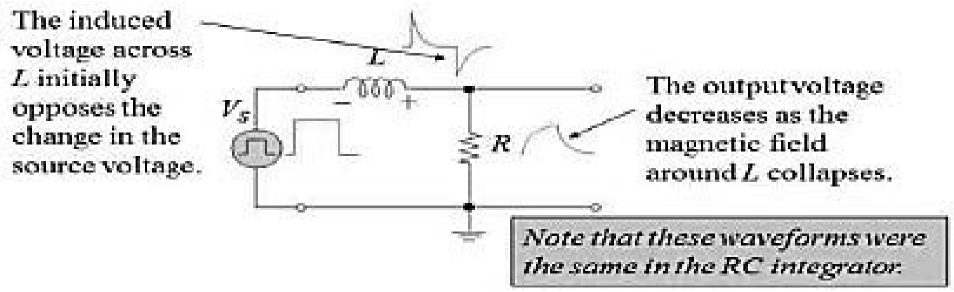


- At the highest point of the info beat, the inductor voltage diminishes dramatically and current increments.
- Subsequently, the voltage across the resistor increments dramatically. As on account of the RC integrator, the result will be 63% of the last worth in 1t.



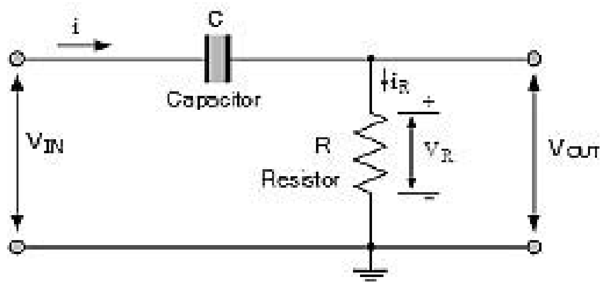
NOTES

- At the point when the beat goes low, an opposite voltage is initiated across L restricting the change.
- The inductor voltage at first is a negative voltage that is equivalent and inverse to the generator; then, at that point, it dramatically increments.



RC Differentiator

- The aloof RC differentiator is a series associated Resistor-Capacitor network that delivers a result signal which compares to the numerical course of separation.
- For an aloof RC differentiator circuit, the info is associated with a capacitor while the result voltage is taken from across an obstruction being the specific inverse to the RC Integrator Circuit.
- An aloof RC differentiator is just a capacitance in series with an obstruction that is a recurrence subordinate gadget which has reactance in series with a decent opposition (the inverse to an integrator).
- Actually like the integrator circuit, the result voltage relies upon the circuit's RC time steady and info recurrence.
- In this manner at low information frequencies the reactance, Xc of the capacitor is high obstructing any d.c. voltage or gradually differing input signals.
- While at high information frequencies the capacitors reactance is low permitting quickly changing heartbeats to pass straightforwardly from the contribution to the result.
- This is on the grounds that the proportion of the capacitive reactance (Xc) to obstruction (R) is diverse for various frequencies and the lower the recurrence the less result.
- So for a given time frame steady, as the recurrence of the info beats expands, the result beats increasingly more look like the information beats in shape.
- At the point when taken care of with an unadulterated sine wave a RC differentiator circuit goes about as a basic detached high pass channel because of the standard capacitive reactance equation of $Xc = 1/(2\pi f C)$.



- For a RC differentiator circuit, the information signal is applied aside of the capacitor with the result taken across the resistor, then, at that point, V_{OUT} approaches V_R.

NOTES

Capacitor Current

$$i_{(t)} = \frac{dQ}{dt} = \frac{d(C \times dV_C)}{dt} = C\frac{dV_C}{dt} = C\frac{dV_{IN}}{dt}$$

Accordingly the capacitor current can be composed as:

$$i_{(t)} = C\frac{dV_{IN(t)}}{dt}$$

As V_{OUT} approaches V_R where V_R as per ohms law is equivalent as well: i_R x R. The current that courses through the capacitor should likewise move through the obstruction as they are both associated together in series. In this manner:

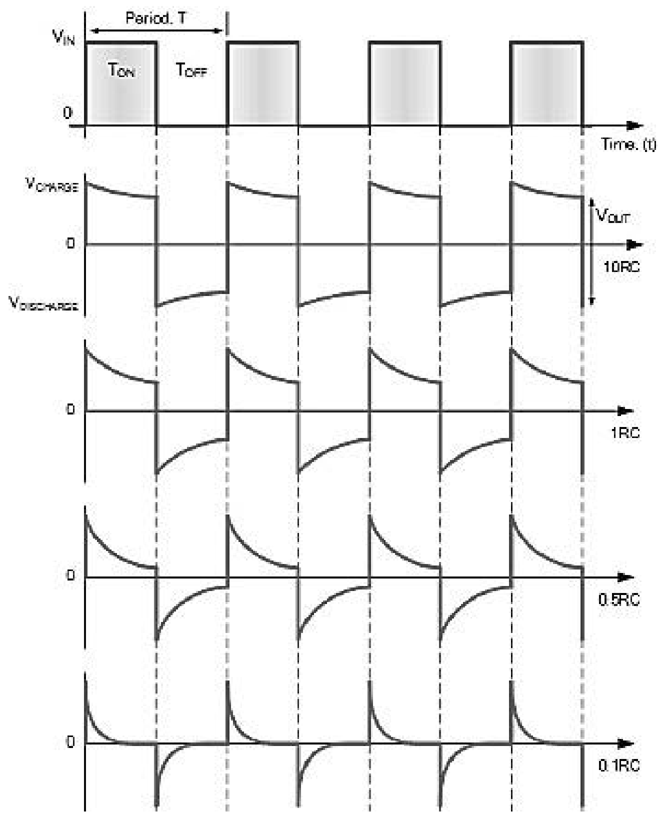
$$V_{OUT} = V_R = R \times i_R \quad i_C = C\frac{dV_{IN}}{dt}$$

As $i_R = i_C$, therefore: $V_{OUT} = RC\frac{dV_{IN}}{dt}$

$$V_{OUT} = RC\frac{dV_{IN}}{dt}$$

Then, at that point, we can see that the result voltage, V_{OUT} is the subsidiary of the information voltage, V_{IN} which is weighted by the consistent of RC. Where RC addresses the time consistent, τ of the series circuit.

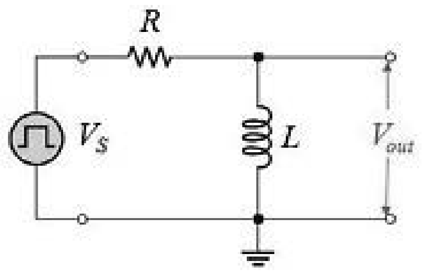
RC Differentiator Output Waveforms



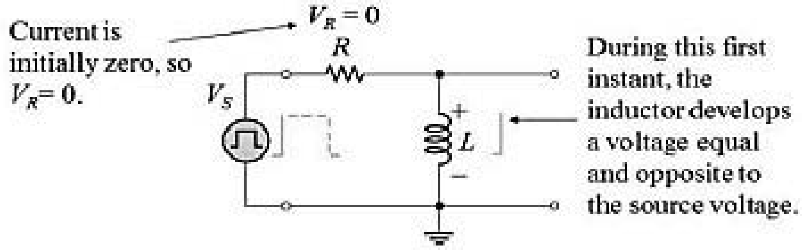
NOTES

The RL Differentiator

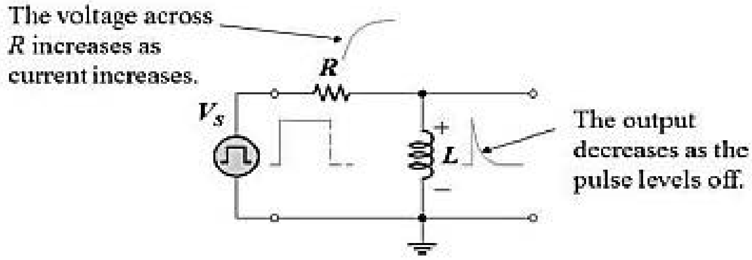
- A RL differentiator is additionally a circuit that approximates the numerical course of separation.
- It can create a result that is the pace of progress of the contribution under specific conditions.
- An essential RL differentiator circuit is an inductor in series with a resistor and the source.
- The result is taken across the inductor.



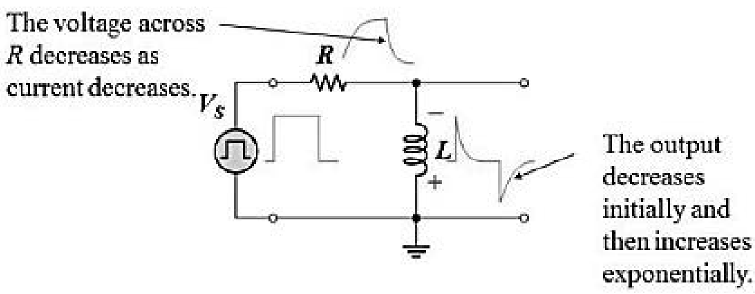
- At the point when a heartbeat generator is associated with the contribution of a RL differentiator, the inductor has a voltage initiated across it that goes against the source; at first, no current is in the circuit.



- After the underlying edge has passed, current expansions in the circuit.
- In the long run, the current arrives at a consistent state esteem given by Ohm's law.



- Then, the falling edge of the beat causes a (negative) voltage to be initiated across the inductor that goes against the change.
- The current reductions as the attractive field breakdowns.



- A use of an integrator is to produce a period delay.

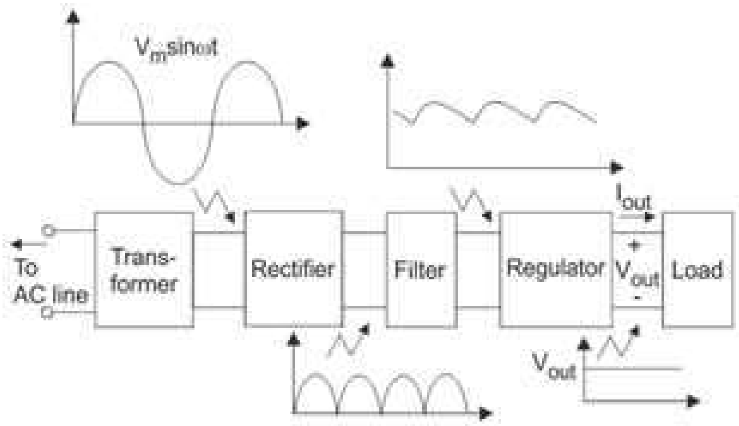
  The voltage at B ascends as the capacitor charges until the limit circuit identifies that the capacitor has arrived at a foreordained level.

NOTES

3.3.3 Voltage Regulated Power Supply

Today pretty much every electronic gadget needs a DC supply for its smooth activity and they should be worked inside specific power supply limits. This necessary DC voltage or DC supply is gotten from single stage ac mains.

A managed power supply can change over unregulated an AC (exchanging current or voltage) to a steady DC (direct current or voltage). A controlled power supply is utilized to guarantee that the result stays consistent regardless of whether the information changes. A directed DC power supply is likewise called as a straight power supply, it is an implanted circuit and comprises of different squares. The directed power supply will acknowledge an AC information and give a steady DC yield. Figure beneath shows the square outline of an average managed DC power supply.



Components of typical linear power supply

The essential structure squares of a managed DC power supply are as per the following:

1. A Advance Down Transformer
2. A Rectifier
3. A DC Channel
4. A Controller

Step Down Transformer

A stage down transformer will venture down the voltage from the air conditioner mains to the necessary voltage level. The turn's proportion of the transformer is so changed, for example, to get the necessary voltage esteem. The result of the transformer is given as a contribution to the rectifier circuit.

Rectification

Rectifier is an electronic circuit comprising of diodes which does the correction cycle. Amendment is the method involved with changing over an exchanging voltage or current into Comparing Direct (DC) amount. The contribution to a rectifier is ac though its result is unidirectional throbbing DC. Normally, a full wave rectifier or an extension rectifier is utilized to correct both the half patterns of the air conditioner supply (full wave amendment). Figure 3.11 shows beneath shows a full wave span rectifier.

NOTES

DC Filtration

The amended voltage from the rectifier is a throbbing DC voltage having extremely high wave content. Yet, this isn't we need, we need an unadulterated wave free DC.

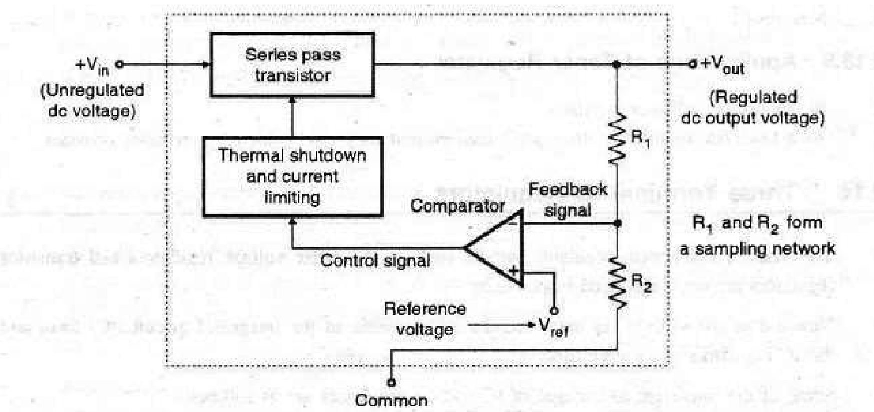
Regulation

This is the last square in a directed DC power supply. The result voltage or current will change or vary when there is change in the contribution from ac mains or because of progress in load current at the result of the directed power supply or because of different variables like temperature changes. This issue can be killed by utilizing a controller. A controller will keep up with the result consistent in any event, when changes at the information or some other changes happen. Semiconductor series controller, Fixed and variable IC controllers or a zener diode worked in the zener area can be utilized relying upon their applications. IC resembles 78XX and 79XX are utilized to acquired fixed upsides of voltages at the result. waveform. Subsequently a channel is utilized. Various kinds of channels are utilized, for example, capacitor channel, LC channel, Choke input channel, π type channel.

Regulated Power Supply

The voltage controllers which are manufactured on a solitary silicon chip are called solid or IC controllers. Benefits of these controllers are little size, simple to utilize, superior execution and minimal expense. Solid voltage controllers are three terminal gadgets having three terminals indicated as info, result and normal terminals. These controllers are accessible in fixed or variable result voltage, positive or negative result voltage.

Block Diagram of Three Pin Regulator



Functional block diagram of a three pin IC voltage regulator

Fig. 3.11 *Block Diagram of a Three Pin IC Voltage Regulator*

Figure 3.11 shows worked on block graph of three pin controller. The three terminals are V_{in} (input voltage), V_{out} (yield voltage) and normal terminal. The dc unregulated voltage is applied to the terminal V_{in}, supply ground to normal and managed yield is taken at terminal V_{out}.

NOTES

The square outline shown in Figure 3.11 is fundamentally a series controller. The two resistors R_1 and R_2 structure the examining organization to deliver the criticism signal relative to the result voltage. V_{ref} is the inside produced voltage reference voltage. This reference voltage is contrasted and the criticism signal by the comparator to deliver the control voltage. This control signal is moved through the warm closure and current restricting square to the series pass semiconductor. The series pass semiconductor is working as the control component. The voltage across the control component is changed by the control sign to get a steady result voltage.

The warm closure and current restricting square give the insurance against expanded inward temperature or over current. On the off chance that the inward temperature surpasses a foreordained worth, then, at that point, the warm closure square won't permit the control sign to go through to the series pass semiconductor. This will wind down the chip consequently. The current restricting circuit ensures the controller against over current.

3.3.4 Regulation Sensitivity and Stability Factors

The stability factor S is defined as the change in collector current in relation to reverse saturation current while keeping and V_{BE} constant. This can be expressed in the following way:

The Thermal Stability Factor: $S = \partial Ic / \partial Ico |Vbe, \beta$

Regulation sensitivity: A voltage generation circuit that generates a regulated voltage and a load current; a sensing circuit that communicates with the voltage generation circuit and senses the load current's peak magnitude; the sensing circuit stores a peak signal based on the peak magnitude of the load current.

3.3.5 Over Voltage and Short Circuit Performance (Transistorised)

Short-circuit protection is provided by this circuit. If the output is ground-shorted. The regulator will be turned off, leaving R_4 as the sole source of power.

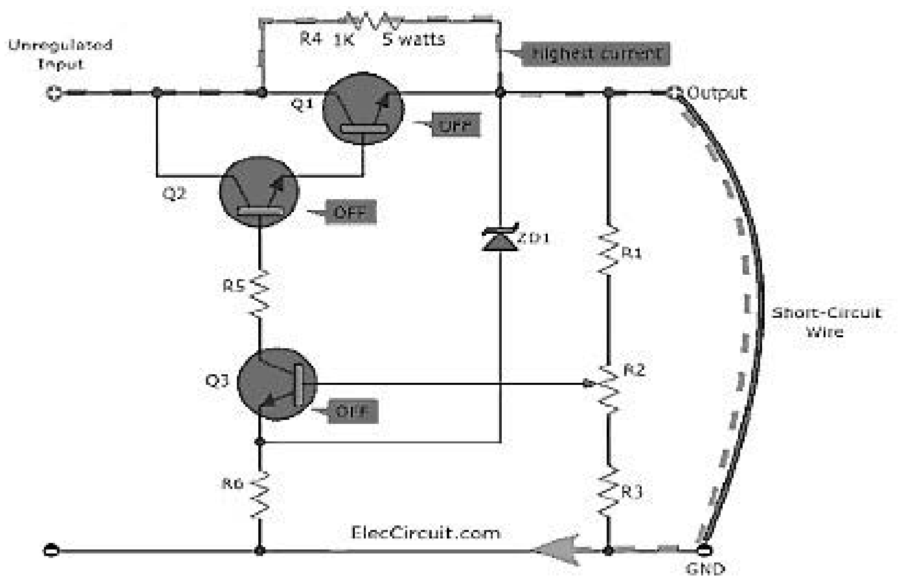


Fig. 3.12 (a)

NOTES

There will be no current going through R_1, R_2, or R_3 in this state. As a result, there is no voltage between them. There is no current flowing through the base of the Q_3 transistor. It is not a current conductor. Finally, both Q_2 and Q_1 do not run. As a result, we should utilize a 5 watt wire-wound resistor for R_4. Despite the fact that it does not produce any current, the power source is operational. The R_4 is just required for starting up once more. To start the circuit, it requires give roughly 1V into the output. When the set output pot (R_2) detects around 50% of the output voltage, Q_3 is turned on.

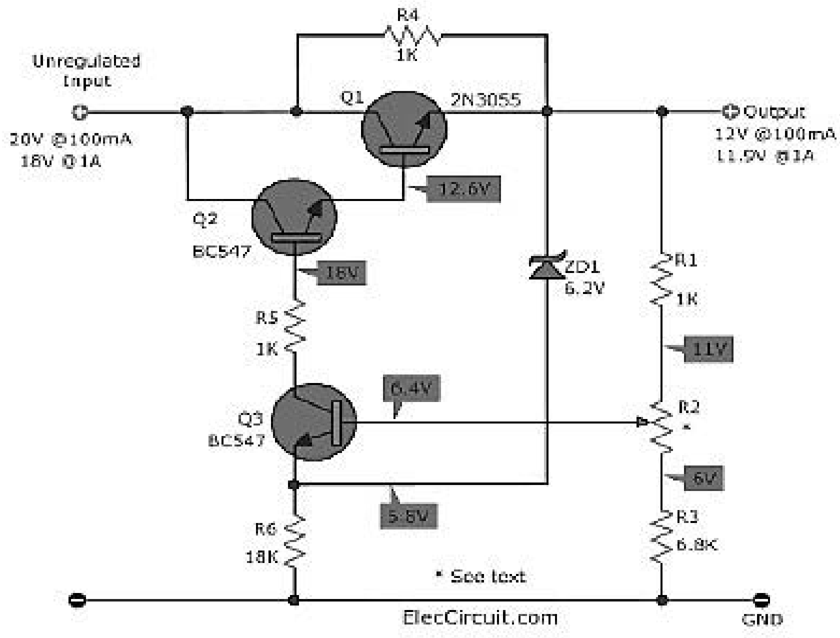


***Fig. 3.12** (b)*

***Fig. 3.12** Simple Transistor Intercom Circuit*

The power driver transistor Q_2 will then be turned on. It will also activate the power regulator transistor Q_1. With the output at 12V, this condition will improve and stabilise. And a load current of around 100mA. The output voltage drops to 11.9V when the load increases to 1Amp. This essentially raises Q1's base-emitter voltage to 0.7V, causing it to switch on faster. The input voltage may drop by about 2 volts. However, Q1's base voltage will remain constant at 12.6V.

Protection Overload

The two preceding circuits have around two-thirds of the advantages of a decent power source. Smoothing and regulation are provided by them. Short-circuit protection is also present in the second circuit. Also, have a look at the third circuit. Overload protection is a crucial function to include in a power supply. Now we'll look at the feature that limits the power transformer and power transistor's maximum current rating. The output will reach a peak of up to 10 times the typical current in a short period of time. This leads to overheating and component damage. The output will reach a peak of up to 10 times the typical current in a short period of time. This leads to overheating and component damage. Power supplies that deliver more than 1 Amp should always have overload protection. Not only that, but it also decreases the possibility of a fire breaking out and severe damage to the equipment being delivered. We can divide them into two groups.

NOTES

Fuse and Circuit Breaker

The simplest is using a circuit breaker or fuse (cheap) in the output. This will blow when the current rise approx. 30% over the recommended max. When it runs, the circuit will not reset itself. We need a manual reset or replace the fuse. The obvious disadvantage with this is the inconvenience of physically changing the fuse. Another way we can use electronic overload is better.

3.3.6 Three Terminal IC Regulated Power Supply Circuits for Positive and Negative Voltages

The three terminal voltage controllers are of two sorts: Fixed and movable voltage controllers. In fixed voltage controllers we have positive voltage controllers and negative voltage controllers. The 78XX series is a progression of fixed positive voltage controllers and the 79XX series is a progression of fixed negative voltage controllers.

78XX Series

78Xx series are three terminal, positive fixed voltage controllers. Here XX show the result voltage. These controllers are accessible in seven diverse result voltages like 5, 6, 8, 12, 15, 18 and 24 volts. For example IC 7812 methods +12V.

IC Number	7805	7806	7808	7812	7815	7818	7824
Output Voltage	5V	6V	8V	12V	15V	18V	24V

79XX Series

79Xx series are three terminal, negative fixed voltage controllers. Here XX show the result voltage. These controllers are accessible in nine diverse result voltages, for example, 5, 6, 8, 12, 15, 18 and 24 volts notwithstanding this series gives - 2V and - 5.2V. For example IC 7812 methods +12V.

IC Number	7905	7906	7908	7912	7915	7918	7924
Output Voltage	-5V	-6V	-8V	-12V	-15V	-18V	-24V

A Positive 12 V Power supply using IC 7812

The circuit outline of a 12 V directed power supply utilizing IC 7812 is shown in Figure 3.13.

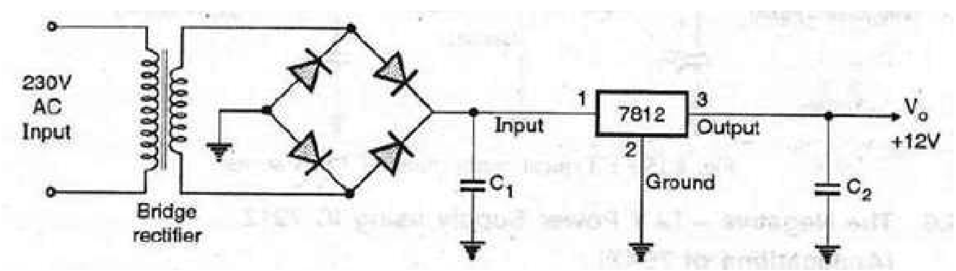


Fig. 3.13 *Positive 12 V Power supply using IC 7812*

NOTES

The extension rectifier and capacitor input channel produce an unregulated dc voltage which is applied at the Input terminal of IC 7812. As the base drop out voltage is 2v for Ic 7812, the voltage applied at the Input terminal ought to be basically 14V. C1 is the channel capacitor and C2 is associated at the result terminal to work on the transient reaction of the controller.

A Negative 12 V Power supply using IC 7912

The circuit graph of a - 12 V managed power supply utilizing IC 7912 is shown in Figure 3.14.

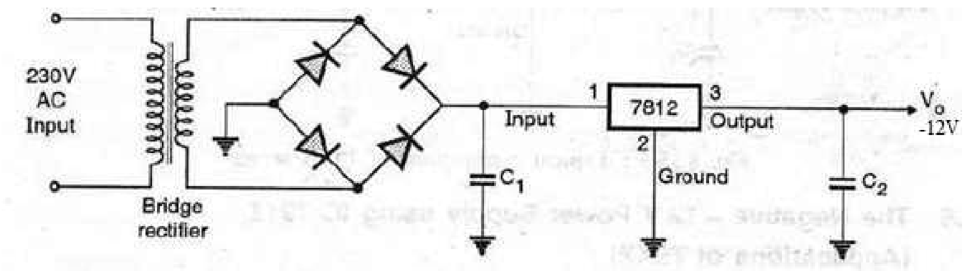


Fig. 3.14 *A Negative 12 V Power supply using IC 7912*

A full wave rectifier and capacitor channel Ci delivers the unregulated negative dc contribution to the controller IC. At the result of 7912, we get a directed result of - 12V.

Check Your Progress

8. What are voltage multipliers used for?
9. State the wave shaping circuits.
10. What are the four primary components of a power supply?
11. How will you expressed the thermal stability factor?
12. Define the term regulation sensitivity.
13. What do you understand by the fuse and circuit breaker?
14. Write a short note on 78Xx series.
15. What is 79XX series?

3.4 ANSWERS TO 'CHECK YOUR PROGRESS'

1. The differential amplifier is the most widely used circuit-building block in analog integrated circuits as it is invariably used as the input stage of every Op-amp to produce high-voltage gain.
2. The four differential amplifier topologies are:
   - Dual-input, balanced-output differential amplifier.
   - Dual-input, unbalanced-output differential amplifier.
   - Single-input, balanced-output differential amplifier.
   - Single-input, unbalanced-output differential amplifier.

NOTES

3. The slew rate of an Op-amp is defined as the maximum rate at which its output voltage can vary. It is expressed in volts per micro second (V/μs) 741 Op-amp has a slew rate of 0.5 V/μs which means that the output of 741 Op-amp can change by 0.5 V every micro-second.
4. Manufacturers specify the gain-frequency characteristics of Op-amps in two ways:
   (*i*) The bandwidth for large signals over which less than 5% distortion is obtained.
   (*ii*)The frequency at which the gain falls to unity.
5. When a terminal is actually grounded, any amount of current can flow to ground through the terminal. Thus, actual ground serves as a 'Sink' for infinite current.
   But no current can flow into the Op-amp through the virtual ground as the input impedance of the Op-amp is infinite. So, a virtual ground cannot serve as a sink for current.
6. Adding speaker or a snake is utilized to total two sign voltages. Voltage snake circuit is a basic circuit that empowers you to add a few signals together. It has wide assortment of utilizations in electronic circuits.
7. A waveform generator is an electronic circuit, which creates a standard wave. There are two kinds of operation amp-based waveform generators:
   - Square wave generator
   - Three-sided wave generator
8. Voltage multipliers were used to track down applications in circuits that required high voltage with low current, such as picture tubes in TV receivers, oscilloscopes, and so on.
9. A wave forming circuit is the one which can be utilized to change the state of a waveform from rotating current or direct current.
10. The essential structure squares of a managed DC power supply are as per the following:
    - A Advance Down Transformer
    - A Rectifier
    - A DC Channel
    - A Controller
11. The thermal stability factor: S= ∂Ic /∂Ico |Vbe, β
12. A voltage generation circuit that generates a regulated voltage and a load current; a sensing circuit that communicates with the voltage generation circuit and senses the load current's peak magnitude; the sensing circuit stores a peak signal based on the peak magnitude of the load current.
13. The simplest is using a circuit breaker or fuse (cheap) in the output. This will blow when the current rise approx. 30% over the recommended max. When it runs, the circuit will not reset itself. We need a manual reset or replace the

NOTES

fuse. The obvious disadvantage with this is the inconvenience of physically changing the fuse. Another way we can use electronic overload is better.

14. 78XX series are three terminal, positive fixed voltage controllers. Here XX show the result voltage. These controllers are accessible in seven diverse result voltages like 5, 6, 8, 12, 15, 18 and 24 volts.
15. 79XX series are three terminal, negative fixed voltage controllers. Here XX show the result voltage. These controllers are accessible in nine diverse result voltages, for example, 5, 6, 8, 12, 15, 18 and 24 volts notwithstanding this series gives - 2 and - 5.2V.

3.5 SUMMARY

- The differential amplifier is the most widely used circuit-building block in analog integrated circuits as it is invariably used as the input stage of every Op-amp to produce high-voltage gain.
- The high-speed logic circuit family or the Emitter-Coupled Logic (ECL) is formed by the BJT differential amplifier.
- If two input signals are applied at the two input terminals, then the configuration is said to be dual-input, otherwise it is a single-input configuration.
- Input Offset Voltage: Input offset voltage is the voltage that must be applied between the two input terminals of an Op-amp to nullify the output. Typical value of input offset voltage of μA 741 Op-amp is 1 mV
- Ideal Op-amp has infinite input resistance. But practical Op-amps have input resistance ranging from 250 kΩ and 40 MΩ for the Op-amps with bipolar transistor input and 10^{12} Ω for the Op-amps with Field-Effect Transistor (FET) input.
- Op-amp is basically a voltage amplifier, therefore, its output resistance should be as low as possible. Practical Op-amps have output resistance of the order of 100 Ω.
- The slew rate of an Op-amp is defined as the maximum rate at which its output voltage can vary. It is expressed in volts per micro second (V/μs) 741 Op-amp has a slew rate of 0.5 V/μs which means that the output of 741 Op-amp can change by 0.5 V every micro-second.
- The open-loop voltage gain of an Op-amp is not constant at all frequencies but drops at high frequencies due to capacitive effects.
- When a terminal is actually grounded, any amount of current can flow to ground through the terminal. Thus, actual ground serves as a ‘Sink’ for infinite current.
- But no current can flow into the Op-amp through the virtual ground as the input impedance of the Op-amp is infinite. So, a virtual ground cannot serve as a sink for current.

NOTES

- Adding speaker or a snake is utilized to total two sign voltages. Voltage snake circuit is a basic circuit that empowers you to add a few signals together. It has wide assortment of utilizations in electronic circuits.
- A square wave generator is an electronic circuit which creates square wave. This segment examines about operation amp based square wave generators.
- Triangular wave generator is an electronic circuit that generates a three-sided wave is known as a three-sided wave generator.
- A voltage multiplier is a circuit that generates a d.c. voltage that is equivalent to a number of different discrete input voltages. At least two pinnacle identifiers or rectifiers are included.
- Voltage multipliers were used to track down applications in circuits that required high voltage with low current, such as picture tubes in TV receivers, oscilloscopes, and so on.
- A wave forming circuit is the one which can be utilized to change the state of a waveform from rotating current or direct current.
- A stage down transformer will venture down the voltage from the air conditioner mains to the necessary voltage level. The turn's proportion of the transformer is so changed, for example, to get the necessary voltage esteem. The result of the transformer is given as a contribution to the rectifier circuit.
- Rectifier is an electronic circuit comprising of diodes which does the correction cycle. Amendment is the method involved with changing over an exchanging voltage or current into Comparing Direct (DC) amount.
- The amended voltage from the rectifier is a throbbing DC voltage having extremely high wave content. Yet, this isn't we need, we need an unadulterated wave free DC.
- A voltage generation circuit that generates a regulated voltage and a load current; a sensing circuit that communicates with the voltage generation circuit and senses the load current's peak magnitude; the sensing circuit stores a peak signal based on the peak magnitude of the load current.
- 78XX series are three terminal, positive fixed voltage controllers. Here XX show the result voltage. These controllers are accessible in seven diverse result voltages like 5, 6, 8, 12, 15, 18 and 24 volts.
- 79XX series are three terminal, negative fixed voltage controllers. Here XX show the result voltage. These controllers are accessible in nine diverse result voltages, for example, 5, 6, 8, 12, 15, 18 and 24 volts notwithstanding this series gives - 2 and - 5.2V.

3.6 KEY TERMS

- **Differential amplifier:** The differential amplifier is the most widely used circuit-building block in analog integrated circuits as it is invariably used as the input stage of every Op-amp to produce high-voltage gain.

NOTES

- **Input offset voltage:** Input offset voltage is the voltage that must be applied between the two input terminals of an Op-amp to nullify the output. Typical value of input offset voltage of μA 741 Op-amp is 1 mV.
- **Slew rate:** The slew rate of an Op-amp is defined as the maximum rate at which its output voltage can vary. It is expressed in volts per micro second (V/μs) 741 Op-amp has a slew rate of 0.5 V/μs which means that the output of 741 Op-amp can change by 0.5 V every micro-second.
- **Bandwidth:** The open-loop voltage gain of an Op-amp is not constant at all frequencies but drops at high frequencies due to capacitive effects.
- **Square wave generator:** A square wave generator is an electronic circuit which creates square wave. This segment examines about operation amp based square wave generators.
- **Triangular wave generator:** Triangular wave generator is an electronic circuit that generates a three-sided wave is known as a three-sided wave generator.
- **Wave shaping circuits:** A wave forming circuit is the one which can be utilized to change the state of a waveform from rotating current or direct current.
- **DC filtration:** The amended voltage from the rectifier is a throbbing DC voltage having extremely high wave content. Yet, this isn't we need, we need an unadulterated wave free DC.
- **78XX series:** 78XX series are three terminal, positive fixed voltage controllers. Here XX show the result voltage. These controllers are accessible in seven diverse result voltages like 5, 6, 8, 12, 15, 18 and 24 volts.
- **79XX series:** 79XX series are three terminal, negative fixed voltage controllers. Here XX show the result voltage. These controllers are accessible in nine diverse result voltages, for example, 5, 6, 8, 12, 15, 18 and 24 volts notwithstanding this series gives - 2 and - 5.2V.

3.7 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What is differential amplifier?
2. What are the Op-amp parameters?
3. Define inverting and non-inverting Op-amp.
4. State the adder circuit using Op-amp 741.
5. How will you define the square wave generator?
6. What do you understand by the voltage multiplier circuit?
7. Define the terms clipping, clamping, differentiating, and integrating circuits.
8. What is rectification?

9. Write a short note on regulation sensitivity.
10. How will you define the over voltage and short circuit performance?
11. Define 78XX series and 79XX series.

NOTES

Long-Answer Questions

1. What do you understand by differential amplifier? Discuss about the four differential amplifier topologies.
2. Discuss briefly about the operational amplifier with the help of giving examples.
3. Briefly explain about the inverting and non-inverting Op-amp. Give appropriate examples.
4. Explain about the use of 741 IC as adder, subtractor, differentiator and integrator with the help of giving examples.
5. Differentiate between the comparator, square and triangular wave generator. Give appropriate examples.
6. Illustrate the voltage multiplier circuit with the help of giving examples.
7. Discuss about the Clipping, clamping, differentiating, and integrating circuits. Give appropriate examples.
8. What are the essential structure squares of a managed DC power supply? Discuss with the help of giving examples.
9. Explain briefly about the regulation sensitivity and stability factors with the help of giving examples.
10. Analysis the over voltage and short circuit performance with the help of giving examples.
11. Illustrate the three terminal IC regulated power supply circuits for positive and negative voltages. Give appropriate examples

3.8 FURTHER READING

Basavaraj, B. and H.N. Shivashankar. 2004. *Basic Electronics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Black, Uyless D. 1993. *Computer Networks.* New Jersey: Prentice-Hall.

Kumar, A. Anand. 2003. *Fundamentals of Digital Circuits*. New Delhi: Prentice-Hall of India.

Mano, Morris. 1979. *Digital Logic and Computer Design*. New Delhi: Prentice-Hall of India.

Roth, Charles. 2001. *Fundamentals of Logic Design*. Thomson Learning.

Shepherd, Robert D. 1998. *Introduction to Computers and Technology*. EMC/ Paradigm Publishing.

Stallings, William. 1996. *Data and Computer Communications.* New Jersey: Prentice-Hall.

NOTES

Tanenbaum, Andrew. 2002. *Computer Networks.* New Jersey: Prentice-Hall PTR.

Yarbarough, John M. 1996. *Digital Logic Applications and Design*. Thomson Learning.

Carlson Bruce and Paul B. Crilly. 2010. Communication System. Noida: Tata McGraw-Hill Education.

Lathi B. P. and Zhi Ding. 2010. Modern Digital and Analog Communication Systems. Oxford: Oxford University Press.

Kennedy George and Bernard Davis. 1993. Electronic Communication Systems. London: Glencoe/ McGraw-Hill.

Roddy Dennis and John Coolen. 1977. Electronic Communications. Noida: Pearson Education.

UNIT 4 COMMUNICATION ELECTRONICS

NOTES

Structure

4.0 INTRODUCTION

Modulation is the process of varying one or more properties of a periodic waveform, called the carrier signal, with a separate signal called the modulation signal that typically contains information to be transmitted. Modulation plays a key role in communication systems since it is employed for encoding information digitally in the analog world. It is very important to modulate signals prior to transmitting them to the receiver section for larger distance transfer, accurate data transfer and low-noise data reception. During modulation, the frequency of the message signal gets raised to a range that will make it more useful for transmission. Let us look at some reasons that make modulation importance for a communication system. The modulation technique referred to as Amplitude Modulation (AM) is employed in the field of electronic communication, mainly for the purpose of information transmission by using a radio carrier wave.

NOTES

In telecommunication, Carson's bandwidth rule defines the approximate bandwidth requirements of communications system components for a carrier signal that is frequency modulated by a continuous or broad spectrum of frequencies rather than a single frequency. Carson's rule does not apply well when the modulating signal contains discontinuities, such as a square wave. Carson's rule originates from John Renshaw Carson's 1922 paper. Generation of DSB-SC Modulation of AM Wave is to be understood that generally a useful modulation signal m(t) will not comprise a single sine wave. DSB-SC modulation of AM wave are several forms of Amplitude Modulation (AM) generation, the normal AM or the Double-Side-Band (DSB) is generated with 'Switching-Modulator' or the 'Square-Law' modulator and the Double-Side-Band-Suppressed Carrier (DSB-SC) with the balanced modulator or the 'Ring-Modulator'.

Double-Sideband Suppressed-Carrier Transmission (DSB-SC) is the type of transmission where the frequencies that are produced by AM are spaced symmetrically below and above the carrier frequency. At the same time, the carrier level is reduced to the lowest practical level, with ideally it being suppressed completely. As opposed to AM, in the case of DSB-SC modulation, the wave carrier is not transmitted; so, most of the power gets distributed between the sidebands, which implies an increase of the cover in DSB-SC, compared to AM, for the same power used. Employed in the case of radio data systems, the DSB-SC transmission is a special case of double-sideband reduced carrier transmission. The Weaver method is also referred to as the '3rd Method' or the 'Third Method'. This could be because the first method is considered to be the super het method and the second method the direct conversion phasing method.

A single-sideband modulation is a type of modulation that is used to send data over radio waves, such as an audio transmission. This modulation is used in radio communications to change the AM signal by utilising additional transmitter power and bandwidth (amplitude modulation). Many single sideband radio communication devices, such as SSB Tx, SSB Rx, and SSB transceiver, are available on the market. Vestigial Sideband (VSB) is an amplitude modulation technique in which a portion of the signal is called a vestige and is modulated with one sideband. Both bands are unnecessary for transmission because they are a waste of time. FDM (Frequency Division Multiplexing) is a multiplexing technology that involves mixing numerous signals over a shared channel. Signals of multiple frequencies are merged for simultaneous transmission in FDM.

In this unit, you will learn about the types of modulation analysis and production of AM and FM wave, generation of DSB-SC modulation of AM waves, demodulation of AM waves generation of DSB-SC waves, coherent detection of DSB-SC waves, SSB modulation, generation and detection of SSB waves vestigial sideband modulation and frequencies division multiplexing.

4.1 OBJECTIVES

After going through this unit, you will be able to:

- Describe the types of modulation analysis and production of AM and FM wave

- Explain the generation of DSB-SC modulation of AM wave
- Discuss the coherent defection of DSB-SC wave
- Understand the basic concept of SSB techniques
- Explain VSB modulation
- Analyse the Frequency Division Multiplexing (FDM)

NOTES

4.2 TYPES OF MODULATION ANALYSIS AND PRODUCTION OF AM AND FM WAVE

Modulation plays a key role in communication systems since it is employed for encoding information digitally in the analog world. It is very important to modulate signals prior to transmitting them to the receiver section for larger distance transfer, accurate data transfer and low-noise data reception.

During modulation, the frequency of the message signal gets raised to a range that will make it more useful for transmission. Let us look at some reasons that make modulation importance for a communication system.

- In the case of signal transmission, multiplexers are used to simultaneously transmit signals, arising from several different sources, via a common channel through. These signals, in case they are transmitted simultaneously with certain bandwidth, tend to create interference. For preventing such a situation, speech signals are modulated to various carrier frequencies in order for the receiver to tune them to desired bandwidth of own choice within the range of transmission.
- One more reason of the technical type is the size of the antenna. The size of the antenna is inversely proportional to the frequency of the signal which is being radiated. The order of the antenna aperture size is at least one by tenth of the wavelength of the signal. Its size is not practicable if the signal is 5 KHz; therefore, raising frequency by modulating process will certainly reduce the height of the antenna.
- Due to the fact that low frequency signals cannot travel long distance, it is essential to employ modulation for transferring signals over large distances.
- Modulation plays a key role in allocating more channels for users and in increasing noise immunity.

Modulation refers to a process of mixing a signal with a sinusoid in such a way that a new signal is created. The process of modulation employees two signals which are the modulating signal and the carrier signal. The modulating signal refers to the baseband signal/information signal. These are the band of frequencies representing the original signal, usually has low frequency and is the one that is to be sent or transmitted to the receiver. On the other hand, the carrier is always some signal which a high frequency sinusoidal signal and is the carrier signal. Mostly, the frequency of the carrier signal is higher than that of the original baseband signal and therefore it can travel faster and longer distances.

The process of modulation has some such carrier wave parameter (like phase, frequency and amplitude) which varied in accordance with the modulating

NOTES

signal. Then, the signal thus modulated gets transmitted by the transmitter. The received modulated signal is received by the receiver which demodulates it and gets back the original information signal. Modulation and demodulation are exactly opposite of each other. During the modulation process, the carrier wave behaves as carrier that carries the information signal from the transmitter to receiver. The new signal produced during the process of modulation, conceivably possesses some specific benefits as compared with a signal that is un-modulated. As shown in Figure 4.1 illustrates the process of modulation in a communication system.

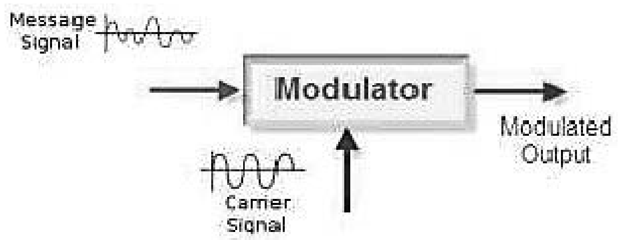


Fig. 4.1 *Modulation in a Communication System*

Modulation refers to the process of taking the wave which is to be transmitted and modifying its characteristics with the help of superimposing the message signal on the high frequency signal. During such a process voice, video and other data signals modify high frequency signals – also known as carrier wave. The carrier wave could wither be AC or DC or pulse chain basis the application which is being employed. Generally use is made of a high frequency sine wave for a carrier wave signal.

It is possible to broadly categories the modulation techniques into two major categories: Analog and digital or pulse modulation. In both the analog and digital modulation techniques, the baseband information is converted to radio frequency signals, but in analog modulation these RF communication signals are continuous range of values, whereas in digital modulation these are prearranged discrete states.

4.2.1 Types of Modulation

There are three basic types of modulation which are as follows.

1. Amplitude Modulation

Amplitude modulation is a form of modulation in which there is modulation or changing of the carrier signal's amplitude proportionate to the message signal. In this form of modulation, no change is made to the phase or to the frequency.

2. Frequency Modulation

Frequency modulation is a form of modulation in which there is the modulation of the carrier signal's frequency proportionate to the message signal. This type of modulation does not in any way modify the phase or the amplitude.

3. Phase Modulation

Phase modulation is a form of modulation in which there is the modulation of the carrier signal's phase in accordance with the low frequency of the message signal. Let us understand all of them in detail.

Analog Modulation

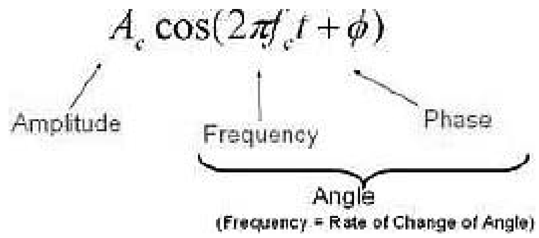


Fig. 4.2 Analog Modulation

NOTES

During analog modulation, the carrier wave is a continuously varying sine wave that modulates the data/message signal. The below given illustration depicts the general function of a sinusoidal wave. Here, it is possible to alter 3 parameters for achieving modulation and their parameters are phase, frequency and amplitude. Modulations that fall in the category of analog modulation are:

- Amplitude Modulation (AM)
- Frequency Modulation (FM)
- Phase Modulation (PM)

In the case of amplitude modulation, there is the varying of the carrier wave's amplitude in proportion to the message signal, while phase and frequency are kept constant. The illustration given below depicts such a modulated signal as well as its spectrum consists of lower frequency band, upper frequency band and carrier frequency components. In such a modulation more power and more bandwidth is required. Also, in such a modulation, filtering is extremely difficult.

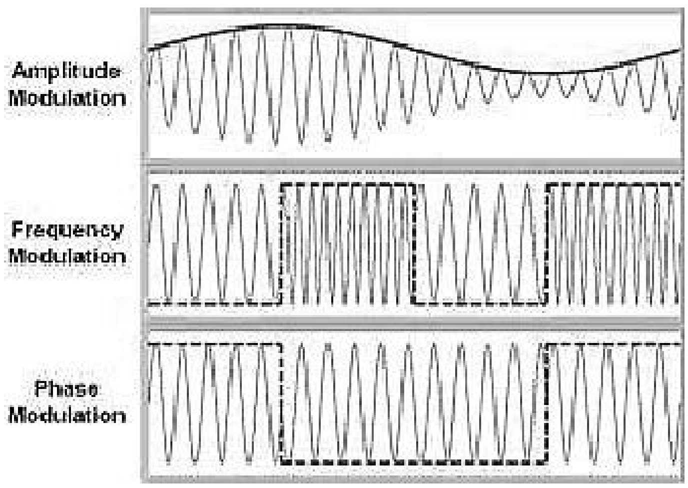


Fig. 4.3 Waveform of Analog Modulations

In the case of Frequency Modulation (FM) the frequency of the carrier is varied in proportion to the message or data signal with the amplitude and phase being kept constant. FM provides more suppression of noise at the expense of bandwidth in FM than does AM. Its use is made in such applications as telemetry seismic prospecting, radar and radio. The bandwidths and efficiency are dependent on maximum modulating frequency and modulation index.

In the case of phase modulation, the carrier phase is varied in accordance with the data signal. With phase modulation, a change in phase will affect frequency, due to which such modulation also falls in the category of frequency modulation.

NOTES

There is greater noise sensitivity in Analog Modulation (AM, FM and PM). In case noise enters a system, the noise will persist all the way to the end receiver. To overcome this problem, digital modulation technique is employed.

Digital Modulation

Digital modulation technique provides efficient and clear communication. Its key advantages over analog modulation are high noise immunity, available bandwidth and permissible power. During digital modulation, there is the conversion of a message signal from analog to digital message, which is then modulated with the help of a carrier wave.

The carrier wave is keyed or switched on and off to create pulses such that the signal is modulated. Just as in the case of analog, in the case of digital too, parameters like amplitude, frequency and phase variation of the carrier wave go to decide the type of modulation.

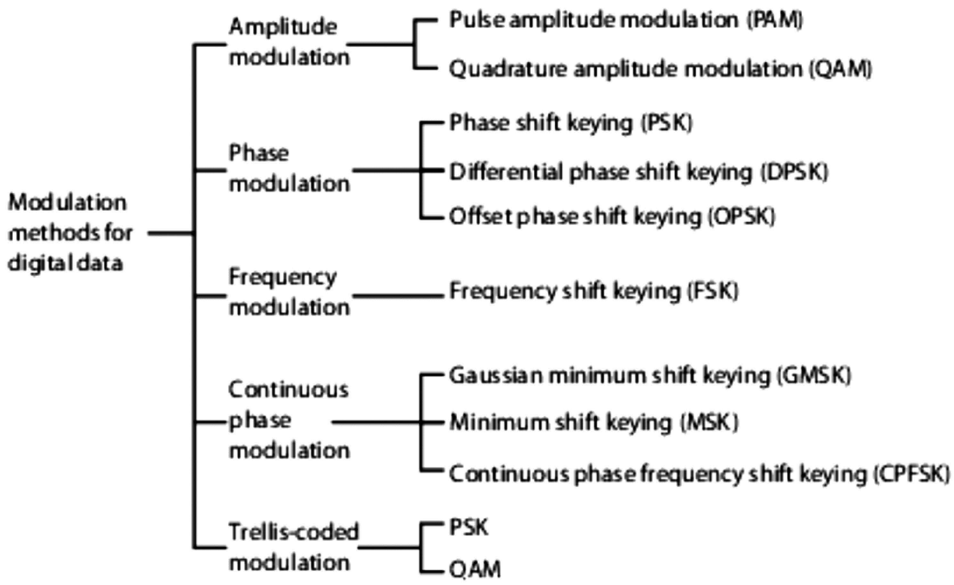


Fig. 4.4 *Modulation Methods of Digital Data*

There are various types of digital modulation categorized on the basis of the type of application and signal that is used, like Amplitude Shift Keying, Frequency Shift Keying, Phase Shift Keying, Differential Phase Shift Keying, Quadrature Phase Shift Keying, Minimum Shift Keying, Gaussian Minimum Shift Keying, and Orthogonal Frequency Division Multiplexing. When amplitude shift keying is used, it modifies the carrier wave's amplitude on the basis of the message signal or the base band signal, which is in digital format. Its use is made in the case of low-band requirements and it has high noise sensitivity.

NOTES

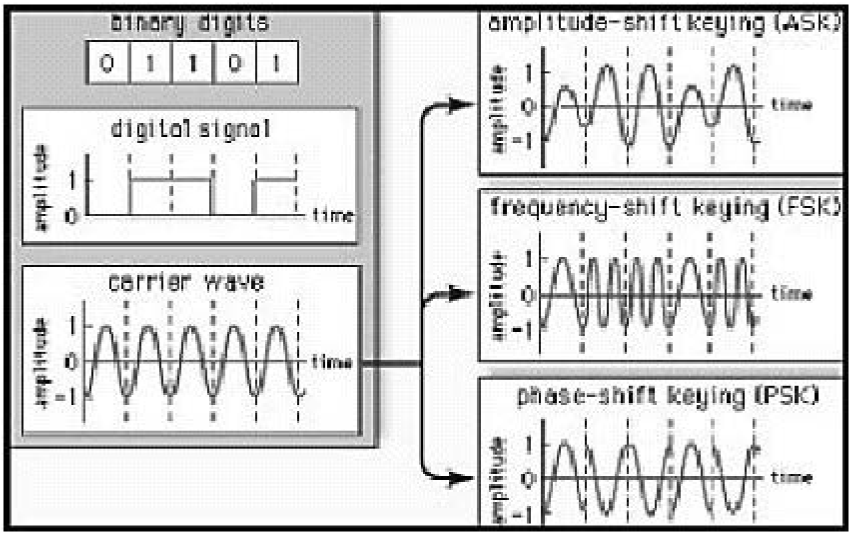


Fig. 4.5 *Digital Modulation*

In frequency shift keying, the frequency of the carrier wave is varied for each symbol in the digital data. It requires larger bandwidths. Similarly, the phase shift keying changes the phase of the carrier for each symbol and it not very noise sensitive.

4.2.2 Need of Modulation

Why should modulation be employed when it is possible to directly transmit a baseband signal? Modulation is required because there are several limitations associated with baseband transmission that can be overcome using modulation. There is a translation of the baseband signal during the modulation process which converts the low frequency signal to a high frequency signal. The shift in frequency is in proportion to the carrier's frequency.

The following are the advantages that modulation provides:

• Reduction in the height of antenna

In the case of radio signal transmission, the height of the antenna has to be multiple of $\lambda/4$ (λ refers to the wavelength).

$$\lambda = c/f$$

Where,

c refers to the velocity of light.

f refers to the frequency of the signal to be transmitted.

So, the calculation of the minimum antenna height required to transmit a baseband signal of f = 10 kHz is:

$$\text{Minimum antenna height} = \frac{\lambda}{4} = \frac{c}{4f} = \frac{3\times10^{8}}{4\times10\times10^{3}} = 7500 \text{ meters i.e. } 7.5 \text{ km}$$

It is practically not possible to install and antenna of this height.

NOTES

With the modulated signal at f = 1 MHz, it is possible to give the minimum antenna height with:

$$\text{Minimum antenna height} = \frac{\lambda}{4} = \frac{c}{4f} = \frac{3\times10^{8}}{4\times10\times10^{6}} = 75 \text{ meters}$$

Practically, it is easy to install this antenna. So, with modulation it is possible to reduce the antenna's height to one which is feasible to install.

• Prevents mixing of signal

In case of transmission of baseband sound signals without modulation by more than one transmitter, the signals will all fall in the same frequency range - 0 to 20 kHz. This will result in all of the signals getting intermixed and leaving the receiver unable to separate one from the other.

On the other hand, if every baseband sound signal is used to modulate a different carrier, each one will fall in a separate frequency domain (different channels) slot and will prevent signal mixing.

• Increase in range of communication

A baseband signal is a low frequency signal. Low frequency signals when transmitted are unable to go long distances getting heavily attenuated.

Attenuation causes a reduction with increase in frequency of the transmitted signal, enabling the signal to travel longer distances.

The frequency, of the signal to be transmitted, is increased when there is modulation of the signal. This leads to the communication range being increased.

• Enables multiplexing

With multiplexing, it becomes possible to simultaneously transmit two or more signals over the same communication channel. This can only be achieved with modulation.

With multiplexing, several signals can make use of the same channel at the same time. To take an example, several different TV channels can use the same frequency range without one interfering or getting mixed with another. It even allows different frequency signals can be transmitted at the same time.

• Improved reception quality

Frequency Modulation (FM) along with digital communication techniques like PCM, there is significant decrease in the effect of noise. Due to this, reception quality becomes much improved.

4.2.3 Amplitude Modulation

The modulation technique referred to as Amplitude Modulation (AM) is employed in the field of electronic communication, mainly for the purpose of information transmission by using a radio carrier wave. For implementation of amplitude modulation, the carrier wave's amplitude (signal strength) needs to be varied proportionate to the waveform which is to be transmitted. For example, the specific waveform may, for instance, correspond to the sounds to be reproduced by a loudspeaker, or the intensity of light in the pixels of a television. AM is a technique which is directly in contrast with both technique of frequency modulation (FM) where the carrier signal's frequency is varied, and phase modulation where there is the varying of the phase of the carrier signal.

The earliest method of modulation was the AM technique for the purpose of transmitting of voice by radio. Its development happed in the initial twenty years of the 20th century, starting with the 1900 when the radiotelephone experiments were conducted by Reginald Fessenden and Landell de Moura. Several communications make use of AM even today. Some examples being, portable two-way radios, VHF aircraft radio, Citizen's Band Radio, and in computer modems (in the form of QAM). "AM" is often used to refer to mediumwave AM radio broadcasting.

Amplitude Modulation: Forms

In the field of telecommunications and electronics, modulation implies the varying of a certain aspect of a higher frequency continuous wave carrier signal with an information-bearing modulation waveform, live a video signal that depicts an image or audio signal that depicts sound, so that the carrier will "Carry" the information. When the same gets to the defined destination, demodulation is used for the extraction of the information signal from the modulated carrier.

In the case of the amplitude modulation, that which is varied is the carrier oscillations strength or amplitude. Let us take an example. In the case AM radio communication, prior to its transmission, the amplitude of a continuous wave radio-frequency signal (sinusoidal carrier wave) is modulated by an audio waveform. The audio waveform modifies the amplitude of the carrier wave and determines the *envelope* of the waveform. When the frequency domain is taken into account, amplitude modulation results in a signal which has its power concentrated at the carrier frequency and two adjacent sidebands. Both of the sidebands are individually equal in bandwidth to that of the modulating signal, and are mirror image of the other. So, at times the standard AM is also referred to as 'double-sideband amplitude modulation' (DSB-AM) so that it can be told apart from sophisticated modulation methods which too are based on AM.

A disadvantage that plagues each and every technique of amplitude modulation, and not just standard AM, lies with the receiver as it amplifies and detects electromagnetic interference and noise in equal proportion to the signal. If there is to be an increase in the received signal-to-noise ratio, for example by a factor of 10 (which is 10 decibel improvement), it will need that there be an increase in the transmitter power by a factor of 10. This stands in contrast to digital audio and to Frequency Modulation (FM) in which the effect of such noise post demodulation gets reduced greatly till the reduced remains well above the reception threshold. This is the reason why AM broadcast is preferred for broadcasts and voice communication but is not sought for the purpose of high fidelity broadcasting or for music.

NOTES

AM suffers for the disadvantage of power usage inefficiency. It is in the carrier signal that a minimum of two-thirds of the power gets concentrated. No part of the original information which it being transmitted is present in the carrier signal. Nevertheless, it being present creates a simple means of demodulation which employs envelope detection, providing a frequency and phase reference to extract the modulation from the sidebands. Certain AM based modulation systems require a lower transmitter power via a complete or partial removal of the carrier component, nevertheless these signals require costlier and more complex receivers. A copy might be regenerated of the carrier frequency (generally moved to the intermediate frequency) by the receiver, with a "pilot" carrier (in reduced-carrier transmission or DSB-RC) which has been greatly reduced, for use with the process of demodulation. When the carrier is totally eliminated as it is in the case of double-sideband suppressed-carrier transmission, it is still possible to accomplish carrier regeneration through employing Costas phase-locked loop. Nevertheless, this will not work in the case of single-sideband suppressed-carrier transmission (SSB-SC), causing the "Donald Duck" sound from such receivers when slightly detuned. However, there is wide usage of single sideband in amateur radio and other voice communications since it is power efficient as well as bandwidth efficient (reducing RF bandwidth by half in comparison to standard AM). Furthermore, in case of broadcasting of the short wave and medium wave, standard AM with the full carrier enables reception by employing inexpensive receivers. The extra power cost are absorbed by the broadcaster to greatly increase potential audience.

The standard AM's carrier provides an amplitude reference which is lost in single or double-sideband suppressed-carrier transmission. In the receiver, the Automatic Gain Control (AGC) responds to the carrier so that the reproduced audio level stays in a fixed proportion to the original modulation. Furthermore, when there is suppressed-carrier transmission, there will not be any transmitted power when there is a pause in the modulation, so the AGC must respond to peaks of the transmitted power during peaks in the modulation. Such a situation will typically have what is referred to as *fast attack, slow decay* circuit that holds the AGC level for a second or more following peaks of this type, in between the program's short pauses or syllables. In the case of communications radios, such a thing is extremely acceptable, since the compression of the audio helps with intelligibility. On the other hand, it is not at all favored for music or normal broadcast programming, where a faithful reproduction of the original program, including its varying modulation levels, is expected.

AM's trivial form that is employed for the purpose of binary data transmission is called on-off keying. On-off keying is the most basic and simple form of *amplitude-shift keying*, and in this technology the zeros and ones are depicted with the absence or presence of a carrier. Use is made of on-off keying by radio amateurs for the transmission of Morse code and here it is called continuous wave

(CW) operation, despite the fact that this transmission is does not remain "Continuous", In the strictest sense." Quadrature amplitude modulation is a form of AM which is much more complex and it is used more commonly for digital data, while making more efficient use of the available bandwidth.

ITU designations

International Telecommunication Union (ITU), in the year 1982, fixed the various types of amplitude modulation. The following table depicts these amplitude modulations.

Table 4.1 *Amplitude Modulation with their Designation*

Designation	Description
R3E	single-sideband reduced-carrier
Lincompex	linked compressor and expander
J3E	single-sideband suppressed-carrier
H3E	single-sideband full-carrier
C3F	vestigial-sideband
B8E	independent-sideband emission
A3E	double-sideband a full-carrier - basic AM scheme

Standard AM: Simplified Analysis

Let us look at a carrier wave (sine wave) whose amplitude will be depicted as A and frequency by f_c as given below:

$$c(t) = A{\cdot}\sin(2\pi f_c t)$$

Consider that the modulation waveform is represented by $m(t)$. in our example, the modulation to be just a sine wave whose frequency is f_m, which is a frequency that is way lower (for example an audio frequency) than f_c:

$$m(t) = M{\cdot}\cos(2\pi f_m t + \phi)$$

In the above equation, M represents the modulation's amplitude. In the example, it is assumed that $M < 1$ and in which case $(1 + m(t))$ will never be negative. With $M > 1$, there will be overmodulation and reconstruction of message signal from the transmitted signal shall cause loss of the original signal. The occurrence of amplitude modulation happens in the case where the carrier $c(t)$ is multiplied by the positive quantity $(1+m(t))$:

$$\begin{aligned} y(t) &= [1 + m(t)]{\cdot}c(\mathrm{t}) \\ &= [1 + M{\cdot}\cos(2\pi f_m t + \phi)]{\cdot}A{\cdot}\sin(2\pi f_c t) \end{aligned}$$

In our example, M is identical to the modulation index, which is explained below. With $M = 0.5$ the amplitude modulated signal $y(t)$ thus corresponds to the top graph (labelled "50% Modulation") in the following illustration.

NOTES

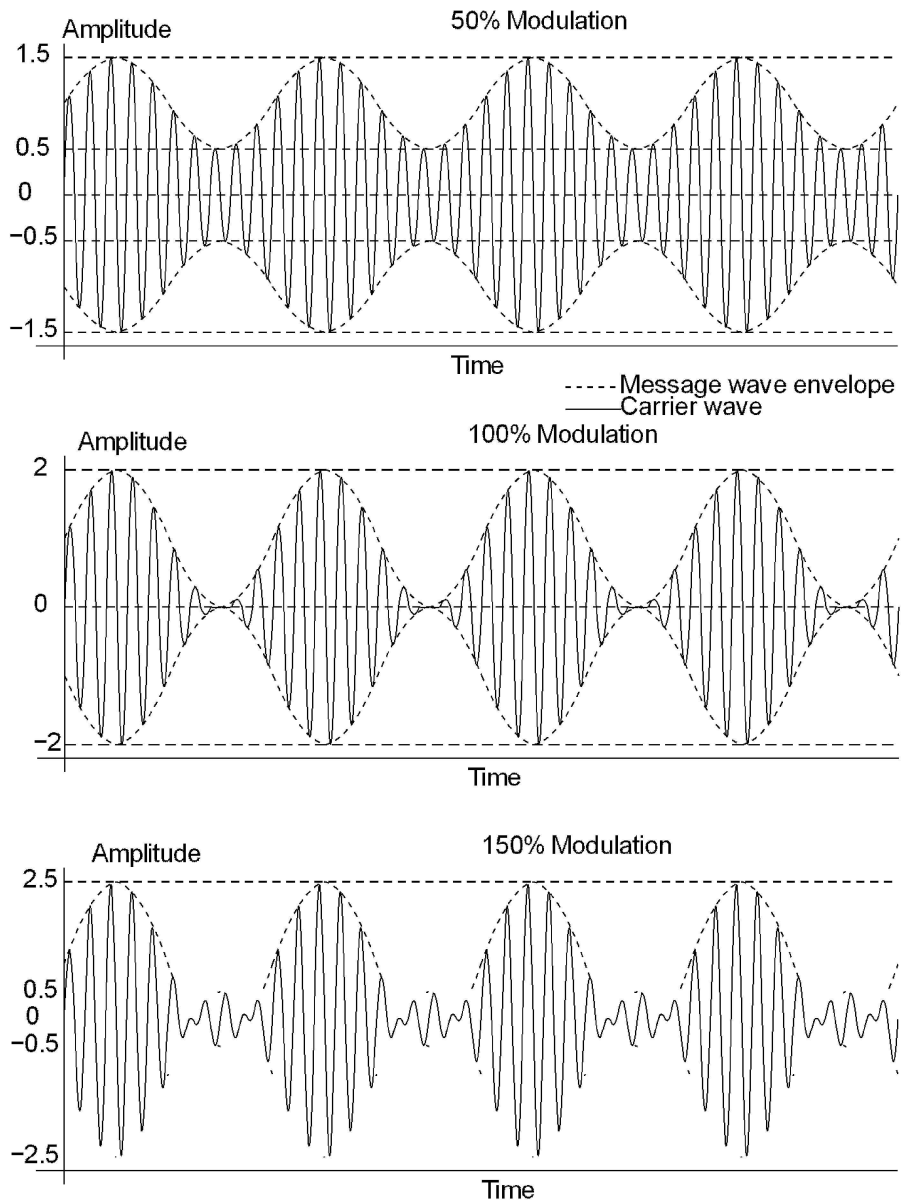


Fig. 4.6 *Modulation Depth, the Unmodulated Carrier has An amplitude of 1*

If use is made of prosthaphaeresis identities, it is possible to depict $y(t)$ as being the sum of three sine waves in the following manner.

$$y(t) = A \cdot \sin(2\pi f_c t) + \frac{AM}{2}[\sin(2\pi(f_c + f_m)t + \phi) + \sin(2\pi(f_c - f_m)t - \phi)$$

Thus, there are three components of the modulated signal which are: the carrier wave $c(t)$ which is unchanged, and two pure sine waves (known as sidebands) with frequencies slightly above and below the carrier frequency f_c.

4.2.4 Frequency Modulation

It is possible for a signal to be carried by either an AM or a frequency modulation radio wave. Figure 4.7 shows a signal in AM and FM wave.

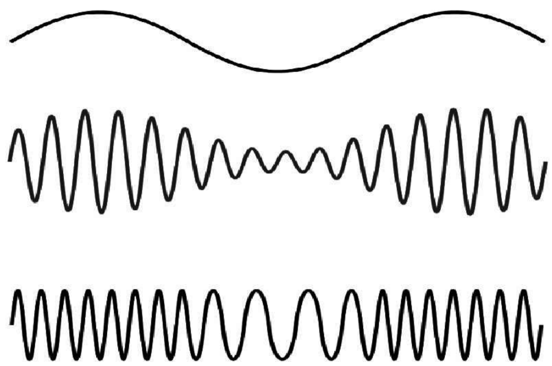


Fig. 4.7 A Signal Represented in AM and FM Radio Wave

In comparison to amplitude modulation, there is more superior noise (RFI) rejection in FM. In signal processing and telecommunications, FM refers to the encoding of information in a carrier wave through varying the wave's instantaneous frequency. This is in contrast to AM where the carrier wave's amplitude varies but the frequency remains constant.

In analog FM, for example in FM radio broadcasting an audio signal (like music or voice), the instantaneous frequency deviation, the difference between the frequency of the carrier and its center frequency, will be proportional to the modulating signal.

FM can be used to encode and transmit digital data through shifting the frequency of the carrier among a set of predefined frequencies that represent digits (as an example it can be considered that one frequency depicts a binary 1, another depicts a binary 0). This modulation technique is referred as Frequency-Shift Keying (FSK). FSK is widely used in fax modems and other modems and it can even be employed for transmitting Morse code. FSK is also used in radioteletype. There is wide use of frequency modulation in FM radio broadcasting. Frequency modulation is also used in the following:

- Certain video-transmission systems
- Magnetic tape-recording systems
- Monitoring of newborns for seizures via EEG
- Music synthesis
- Radar
- Seismic prospecting
- Telemetry
- Two-way radio systems

The advantage of using frequency modulation in radio transmission is that it has a larger signal-to-noise ratio because of which it rejects radio frequency interference better than an equal power Amplitude Modulation (AM) signal. That is why a majority of radio music is broadcasted in the FM mode via FM radio.

There is a close relationship between frequency modulation and phase modulation. Mostly, phase modulation is being used as an intermediate step to arrive at frequency modulation. When seen in mathematical terms, phase modulation as well as frequency modulation are seen to be a special case of Quadrature Amplitude Modulation (QAM).

Theory

NOTES

If the information to be transmitted (i.e. the baseband signal) is $x_m(t)$ and the sinusoidal carrier is $x_c(t) = A_c \cos(2\pi f_c t)$, with f_c being the base frequency of the carrier, and A_c being the amplitude of the carrier, the carrier is combined with the baseband data signal by the modulator to obtain the transmitted signal:

$$\begin{aligned} y(t) &= A_c \cos\left(2\pi\int_0^t f(r)dr\right) \\ &= A_c \cos\left(2\pi\int_0^t [f_c + f_\Delta x_m(r)dr\right) \\ &= A_c \cos\left(2\pi f_c t + 2\pi f_\Delta \int_0^t x_m(r)dr\right) \end{aligned}$$

Where:

$f\Delta = K_f A_m$

K_f represents the frequency modulator's sensitivity and

A_m is baseband signal or modulating signal's amplitude.

$f(r)$ is the oscillator's instantaneous frequency

$f\Delta$ is the frequency deviation that shows maximum shift away from f_c in one direction, having $x_m(t)$ is limited to the range ± 1.

Despite the fact that a majority of the signal's energy lies within $f_c \pm f_\Delta$, Fourier analysis can show that there is need for a wide range of frequencies for precise representation of an FM signal. An actual FM signal's frequency spectrum contains components that extend infinitely, but their amplitude decreases and higher-order components are often neglected in practical design problems.

Sinusoidal Baseband Signal

Mathematically, it is possible to approximate the baseband modulated signal by a sinusoidal continuous wave signal with a frequency f_m. This method is referred to as Single-tone Modulation. Such a signal has the following integrals:

$$\int_0^t x_m(r)dr = \frac{A_m \sin(2\pi f_m t)}{2\pi f_m}$$

The above mentioned expression for $y(t)$ can be simplified as follows:

$$y(t) = A_c \cos\left(2\pi f_c t + \frac{A_m f_\Delta}{f_m}\sin(2\pi f_m t\right)$$

Where the modulating sinusoid's amplitude A_m is gets represented by the peak deviation f_Δ. Bessel functions can be used to represent the harmonic distribution of a sine wave carrier modulated by a sinusoidal signal of this type. It gives the basis for gaining a mathematical understanding of frequency modulation in the frequency domain.

Modulation Index

The modulation index is indicative of the amount by which the modulated variable varies around its unmodulated level. It relates to carrier frequency variations:

$$h = \frac{\Delta f}{f_m} = \frac{f_\Delta \,|\, x_m(t) \,|}{f_m}$$

Where,

- f_m is highest frequency component present in the modulating signal $x_m(t)$
- Δf is the peak frequency-deviation

In case of sine wave modulation, the modulation index is the ratio of the carrier wave's peak frequency deviation to the modulating sine wave's frequency. It is referred to as Narrowband FM when bandwidth is approximately $2f_m$. Sometimes, modulation index $h < 0.3$ rad is taken to be Narrowband FM while otherwise it is seen as being Wideband FM.

In the case of modulation systems that are digital (like Binary Frequency Shift Keying (BFSK)), in which the carrier is modulated by a binary signal, the modulation index is given by:

$$h = \frac{\Delta f}{f_m} = \frac{\Delta f}{\frac{1}{2T_s}} = 2\Delta f T_s$$

Where,

- T_s represents the symbol period
- $f_m = \frac{1}{2T_s}$ is the highest *fundamental* of modulating binary waveform.

With digital modulation, the transmission of the carrier f_c never happens. Instead, there is the transmission of one of two frequencies i.e. $f_c + \Delta f$ or $f_c - \Delta f$. The basis for the transmission is the modulating signal's binary state - 0 or 1.

If $h \gg 1$, the modulation will be referred to as *wideband FM*. This has a bandwidth of about $2f_\Delta$. The signal-to-noise ratio can be improved using the more bandwidth of wideband FM. For example, the signal-to-noise ratio will see an eight-fold improvement on keeping f_m constant and doubling the value of Δf.

In case of a tone-modulated FM wave, increasing the modulation index and keeping the modulation frequency constant, there is an increase in the (non-negligible) bandwidth of the FM signal and the spacing between spectra remains the same. While there is a decrease in the strength of some spectral components and there is an increase in some others. If there is an increasing of the modulation frequency while the frequency deviation is maintained as constant, there is an increase in the spacing between spectra.

There are two ways to classify frequency modulation:

- Narrowband – when the change in the carrier frequency is about the same as the signal frequency
- Wideband - when the change in the carrier frequency is much higher (modulation index > 1) than the signal frequency

Bessel functions

When the carrier is modulated by a single sine wave, it is possible to use Bessel Function of the first kind to calculate the resulting frequency spectrum, as a function

NOTES

of the modulation index and the sideband number. For the different modulation indices of FM signals, the carrier and sideband amplitudes are illustrated. In the case of the modulation index's specific values, the carrier amplitude becomes zero and all the signal power is in the sidebands.

The count of the sidebands is doubled since they are on both sides of the carrier. Their count is then multiplied by the modulating frequency to arrive at the bandwidth. Consider an example, a 3 kHz deviation modulated by a 2.2 kHz audio tone produces a modulation index of 1.36. Assume that only those sidebands are considered that have a relative amplitude of at least 0.01. As provided in the table, this modulation index creates 3 sidebands which when doubled provide (6 * 2.2 kHz) or a 13.2 kHz required bandwidth.

Modulation index	Sideband amplitude																
	Carrier	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
0.00	1.00																
0.25	0.98	0.12															
0.5	0.94	0.24	0.03														
1.0	0.77	0.44	0.11	0.02													
1.5	0.51	0.56	0.23	0.06	0.01												
2.0	0.22	0.58	0.35	0.13	0.03												
2.41	0	0.52	0.43	0.20	0.06	0.02											
2.5	−0.05	0.50	0.45	0.22	0.07	0.02	0.01										
3.0	−0.26	0.34	0.49	0.31	0.13	0.04	0.01										
4.0	−0.40	−0.07	0.36	0.43	0.28	0.13	0.05	0.02									
5.0	−0.18	−0.33	0.05	0.36	0.39	0.26	0.13	0.05	0.02								
5.53	0	−0.34	−0.13	0.25	0.40	0.32	0.19	0.09	0.03	0.01							
6.0	0.15	−0.28	−0.24	0.11	0.36	0.36	0.25	0.13	0.06	0.02							
7.0	0.30	0.00	−0.30	−0.17	0.16	0.35	0.34	0.23	0.13	0.06	0.02						
8.0	0.17	0.23	−0.11	−0.29	−0.10	0.19	0.34	0.32	0.22	0.13	0.05	0.03					
8.65	0	0.27	0.06	−0.24	−0.23	0.03	0.26	0.34	0.26	0.18	0.10	0.05	0.02				
9.0	−0.09	0.25	0.14	−0.18	−0.27	−0.06	0.20	0.33	0.31	0.21	0.12	0.06	0.03	0.01			
10.0	−0.25	0.04	0.25	0.06	−0.22	−0.23	−0.01	0.22	0.32	0.29	0.21	0.12	0.06	0.03	0.01		
12.0	0.05	−0.22	−0.08	0.20	0.18	−0.07	−0.24	−0.17	0.05	0.23	0.30	0.27	0.20	0.12	0.07	0.03	0.01

Carson's Rule

This rule states that approximately all (~98 percent) of the power of a FM signal is within a bandwidth $B_T B_T$.

$$B_T = 2(\Delta f + f_m)$$
$$= 2f_m(\beta + 1)$$

Where,

- Δf is peak deviation of the instantaneous frequency ($f(t)$) from the center carrier frequency
- β is modulation index (ratio of frequency deviation to highest frequency in the modulating signal)
- f_m is the highest frequency in the modulating signal.

Sinusoidal signals is the only condition for the application of Carson's rule.

$$B_T = 2(\Delta f + W)$$
$$= 2W(D + 1)$$

W represents the modulating signal's highest frequency but it is by nature non-sinusoidal

D represents the deviation ratio which the ratio of frequency deviation to highest frequency of modulating non-sinusoidal signal.

NOTES

Reduction of Noise

In comparison to AM, with FM it is possible to obtain improved Signal-to-Noise Ratio (SNR). Further, when looking at SNR, in comparison to an optimum AM scheme, typically FM provides poorer SNR below a certain level of signal referred to as noise threshold, but above a higher level (which is the full improvement/full quieting threshold) there is great improvement in SNR over AM. It is deviation and modulation level which decide the improvement. In the case of typical channels of voice communications, the typical improvements will be 5-15 dB. It is possible to attain even more improvement by implementing FM broadcasting by employing wider deviation. Generally such other additional techniques, like pre-emphasis of higher audio frequencies with corresponding de-emphasis in the receiver, are employed for improving the overall SNR in FM circuits. Due to the fact that there is constant amplitude FM signals, the FM receivers generally come with limiters which eliminate the AM noise, which makes the SNR even better.

Implementation

Modulation

It is possible to attain the generation of FM signals via indirect as well as direct frequency modulation as given follow:

- To achieve direct FM modulation, the message has to be fed directly into the input of a voltage-controlled oscillator.
- To achieve indirect FM modulation, the message signal needs to be integrated to create a phase-modulated signal. This signal is then employed for modulating a crystal-controlled oscillator, and the result which is thus attained is passed through a frequency multiplier so that an FM signal is generated. Which modulation of this type, there is the generation of narrow band FM, creating wide band FM later and due to this fact this form of modulation is called indirect FM modulation.

Demodulation

There are several varied FM detector circuits. One of the common ways in which the information signal is recovered is by employing a Foster-Seeley discriminator. It is possible to make use of a phase-locked loop as an FM demodulator. A slope detection makes use of a tuned circuit, with its resonant frequency slightly offset from the carrier, for demodulating an FM signal. With the rise and fall in the frequency, there is a changing amplitude of response provided by the tuned circuit, converting FM to AM. It could be that AM receivers detect some FM transmissions by this means, even though it does not give any efficient way of detection for FM broadcasts.

Applications

Let us now examine the application of FM in different technologies.

NOTES

Magnetic tape

Analog VCR systems (including VHS) make use of FM at intermediate frequencies for the recording of the luminance (black and white) parts of the video signal. Generally, the recording of the chrominance component gets carried out in the form of a conventional AM signal, by making use of the higher-frequency FM signal as bias. For the luminance ("Black and White") component of video to be recorded to (and for retrieving video from) magnetic tape without distortion the only method that is feasible is FM. There is a huge range of frequency components associated with video signals, such as starting from just a few hertz to several megahertz, which become extremely wide for equalizers to handle because of the electronic noise that exists below –60 dB. With FM, the tape remains at saturation level which acts as a means of noise reduction; it is possible for a limiter to mask playback output variations, while pre-echo and print-through are removed by the FM capture effect. If there is the addition of a continuous pilot-tone in the signal, it becomes possible to assist timebase correction and have mechanical jitter under control.

Such FM systems are unusual, since they carry a ratio of carrier to maximum modulation frequency of lower than 2. When this is compared with FM audio broadcasting, it is seen that in FM audio broadcasting the ratio is around 10,000. It is essential to design the system in a way that will bring down that this unwanted output to a level that is acceptable.

Sound

Use of FM is also made at audio frequencies for synthesizing of sound. This technique is called FM synthesis, made popular with early digital synthesizers, it is now a standard feature in several generations of personal computer sound cards.

Radio

An American electrical engineer, Edwin Howard Armstrong (1890–1954) is considered to be the inventor of wideband FM radio. The Wideband FM (WFM) needs a wider signal bandwidth in comparison to AM by an equivalent modulating signal; it turns the signal more robust against interference and noise. Frequency modulation better handles signal-amplitude-fading phenomena than AM. Due to these several reasons, there was the clear choice of using FM as the modulation standard for high frequency, high fidelity radio transmission, reason for the term FM radio. A special detector is employed by FM receivers for the purpose of FM signals and this produces a phenomenon called the *capture effect*, where the tuner will "Capture" the stronger of two stations on the same frequency and ignore the weaker of the two. Nevertheless, it is possible that lack of selectivity or even frequency drift lead to a one station being overtaken by another on an adjacent channel.

It is possible to use FM signal to transmit a stereo signal. This is achieved by employing multiplexing and demultiplexing prior to and post the FM process. For the monaural as well as the stereo processes, the process of FM modulation

NOTES

and demodulation is the same. Use can be made of a high-efficiency radio-frequency switching amplifier for the transmission of FM signals and also for other constant-amplitude signals. For a given signal strength (which is measured at the receiver antenna), less battery power is used by the switching amplifiers than linear amplifiers as their cost is also less than of a linear amplifier, an advantage for FM over other modulation methods which have to make use of linear amplifiers (like QAM and AM).

Common use is made of FM at VHF radio frequencies for broadcasting speech and music as high-fidelity broadcasts. FM is also use to broadcast sound for analog TV. In the case of both commercial and amateur radio settings, use is made of narrowband FM for the purpose of voice communications. In the case of broadcast services that consider audio fidelity as being important to the service, use is generally made of wideband FM. In the case of two-way radio, use is made of Narrowband FM (NBFM) for conserving bandwidth for land mobile, marine mobile and other radio services.

4.2.5 FM Noise Triangle

The FM noise triangle, is the triangular distribution of noise in FM. Noise is the unwanted deviation in the carrier frequency. In the case of FM, phase modulation and amplitude modulation is produced by noise signal. The magnitude of frequency deviation is dependent upon the noise's relative amplitude respective to the carrier. Here is how one can look vectorially at the carrier signal and the noise signal. When the noise signal gets superimposed upon the carrier, there is the addition of the amplitude of the noise with the amplitude of the carrier. This creates the amplitude modulation. As shown in Figure 4.8, the amplitude carrier's maximum deviation is V_n.

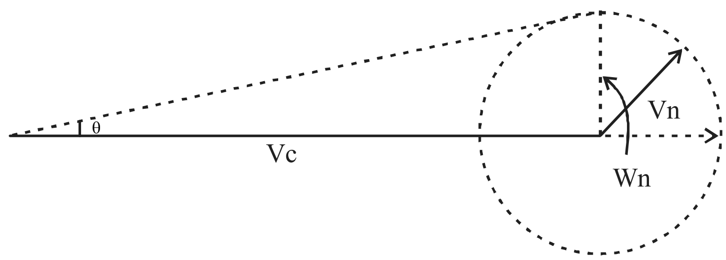


Fig. 4.8 *The Vector Effect on the Carrier of Noise*

Simultaneously, the phase angle is constantly being changed by the noise vector with respect to the carrier signal V_c which will change the phase deviation (θ) of FM wave (phase deviation ¸ of FM changes because of noise). So, the carrier is modulated by noise both as far as phase and amplitude are concerned.

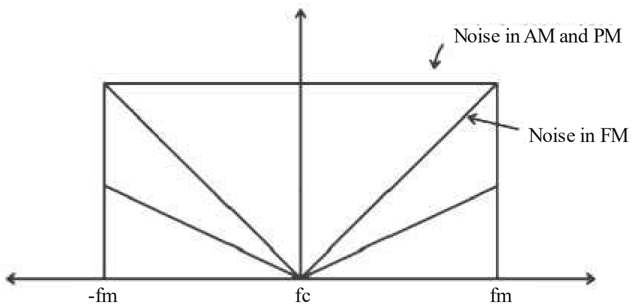


Fig. 4.9 *Noise Triangle*

NOTES

The noise triangle in Figure 4.9 shows that in the case of both PM and AM, noise stays constant for the entire audio range. This is due to the fact that because of the modulation index are independent of modulating frequency. While in case of FM, increase in the modulating frequency will also results in increase of noise. So, in FM, at higher frequencies, there is more effect of noise.

Emphasis

Normally, before certain processes are performed, for example recording on tape or phonographic record or transmission over cable, there is boosting of the input frequency range which is most susceptible to noise. This procedure is called "pre-emphasis". It is performed prior to or "pre-" the process the signal that the signal has to go through. After receiving the signal from the recording, the reverse transformation is applied in order to verify that the output reproduces the same input as received. This reverse transformation is called as de-emphasis. During the de-emphasis stage, any noise added by transmission or record/playback, to the frequency range previously boosted is now attenuated.

There is the emphasizing of the high-frequency signal components so that a more equal modulation index is created for the transmitted frequency spectrum, which leads to a better signal-to-noise ratio for the entire frequency range. There is common use of emphasis for FM broadcasting and the creation of vinyl records.

Let us look briefly at pre-emphasis and de-emphasis.

Pre-emphasis

When electronic audio signals are processed, the term pre-emphasis implies a process which is created for increasing (in the limits of a frequency band) the magnitude of some (usually higher) frequencies with respect to the magnitude of other (generally lower) frequencies in order to improve the overall signal-to-noise ratio by minimizing the adverse effects of such phenomena as attenuation distortion or saturation of recording media in subsequent parts of the system. The operation which is the exact opposite of this is referred to as de-emphasis, and the entire system is referred to as emphasis. In order to perform pre-emphasis, a pre-emphasis network is used. A pre-emphasis network is just a calibrated filter. In this process, the decision of the frequency response is done by the use of special time constants. This value can be used to calculate the cutoff frequency. One common use of pre-emphasis is in telecommunications, besides displaying the spectrograms of speech signals, FM broadcasting transmissions, record cutting and digital audio recording. For example the Dolby noise-reduction system for magnetic tapes and the RIAA equalization curve of the vinyl records of 24 rpm and 33 rpm.

Pre-emphasis is used in high speed digital transmission to improve the quality of the signal at output of data. When signals are transmitted at high data rates, there is possibility of introduction of distortions by transmission medium, and pre-emphasis is used for the distortion of the transmitted signal so that this distortion can be corrected. If pre-emphasis is properly implemented, then a received signal is produced which is much more like the original signal or the desired signal, and it enables higher frequencies to be used as well as the creation of fewer bit errors.

Pre-emphasis is used with phase modulation or frequency modulation transmitters for the purpose of equalizing of the modulating signal drive power for the deviation ratio. The process of receiver demodulation has a reciprocal network referred to as a de-emphasis network which is used for restoring the original signal power distribution.

NOTES

De-emphasis

The process of de-emphasis stands complementary to the process of pre-emphasis. "The process of de-emphasis is a system process which is designed for the purpose of decreasing (for a band of frequencies), the magnitude of some (generally higher) frequencies with respect to the magnitude of other (generally lower) frequencies in order to improve the overall signal-to-noise ratio by minimizing the adverse effects of such phenomena as attenuation distortion or saturation of recording media in subsequent parts of the system". The frequency response curve is used to calculate the cutoff frequency and dictated by the special time constants.

With data transmission of the serial type, de-emphasis refers to the reduction of the level of all bits except the first one after a transition. It leads to the low frequency content being de-emphasized while the high frequency content, because of transition, being emphasized in comparison to the low frequency content. It being a form of transmitter equalization; also compensates for losses that occur over the channel, which will be greater for higher frequencies. The serial data standards like SAS, SATA and PCI Express need use of de-emphasis by the transmitted signals.

4.2.6 Comparison of Narrow Band FM and Wide Band FM

FM systems can be classified into two categories:

1. Narrow band FM
2. Wide band FM/Broadband FM

Let us look at how one differs from the other.

Narrow Band FM

The FM wave that has a small bandwidth is referred to as narrow band FM. The Narrow band FM has a modulation index (m_f) which is small in comparison to one radian. Therefore, a narrow band FM's spectrum comprises the carrier and the upper and lower sidebands.

The values of the j coefficients for the small values of m_f is given below.

$$J_0(m_f) = 1,$$
$$J_1(m_f) = m_{f/2}$$
$$J_n(m_f) = 0 \text{ for } n > 1$$

Mathematically, a narrow band FM wave is represented as:

$$e_{FM}(t) = s(t) = \underbrace{E_c \sin w_C t}_{\text{Carrier}} + \frac{m_f E_c}{2} \sin \underbrace{(w_c + w_m)t}_{\text{USB}} - \frac{m_f E_c}{2} \sin \underbrace{(w_c + w_m)t}_{\text{LSB}}$$

The (-) sign associated with the LSB depicts a 180° phase shift. In the practical sense, in the narrow band FM systems the m_f is less than 1. The maximum frequency deviation that is permissible is about 5 kHz. It is in FM mobile communications that this system is used, like radio-taxi, ambulance, police wireless, to name a few.

NOTES

Narrow band FM Analysis

The following expression depicts the instantaneous frequency of the FM wave.

$$f_i = f_c + k_f x(t)$$

Where,

- $x(t)$ represents the modulating signal
- $k_f x(t)$ depicts the frequency deviation
- k_f(constant) controls the deviation

In case of small k_f values, there is a small frequency deviation and at the same time the FM signal's spectrum will have a narrow band. Due to this, it is referred to as narrow band FM.

Consider the following FM wave's expression:

$$s(t) = E_c \cos[2\pi f_c t + 2\pi k_f \int x(t)dt]$$

Expressing it in terms of w, we have

$$s(t) = E_c \cos[w_c t + 2\pi k_f \int x(t)dt]$$

We can represent this in the exponential manner as under:

$$s(t) = E_c \cos\theta(t) = E_c e^{j\theta(t)}$$

This has been written by considering only the real part of $E_c e^{j\theta(t)}$

Therefore,

$$s(t) = E_c e^{j\theta(t)} = E_c e^{j[\cos w_c t + k_f \int x(t)dt]}$$

Let $\int x(t)dt = g(t)$

Thus,

$$s(t) = E_c e^{j[\cos w_c t + k_f g(t)]}$$

If $k_f g(t) << 1$ for all values (which is the case for narrow band FM), then, the expression for FM will be

$$\hat{s}(t) = E_c[1 + jk_f g(t)]e^{jw_c t}$$

Also,

$$s(t) = R_c[\hat{s}(t)]\underbrace{E_c \cos w_c t}_{\text{Carrier}} - \underbrace{E_c k_f g(t) \sin w_c t}_{\text{Side band}}$$

Figure 4.10 shows the generation of narrow band FM with the use of a balanced modulator.

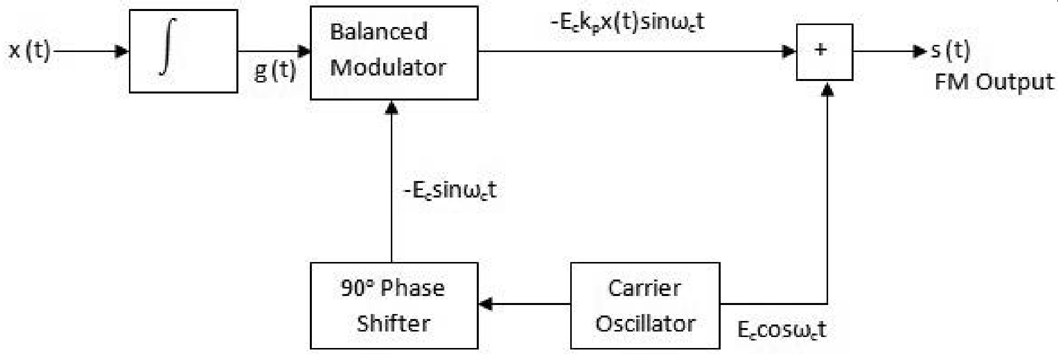


Fig. 4.10 *Narrow band FM generation*

Wideband FM

In the case of modulation index m_f large values, ideally the FM wave will comprise a carrier plus an infinite number of sidebands which systematically located around that carrier. An FM wave of this type has an infinite bandwidth due to which it is referred to as wideband FM.

For a wideband FM, the modulation index will be higher than 1. The maximum deviation that is permitted is 75 kHz. This is employed for entertainment broadcasting applications like FM radio.

Wideband FM wave Frequency Spectrum

In the case of a wideband FM, its expression is a complex one due to the fact that it is the sine of sine function. Only the Bessel functions can help to solve this equation. The following is the equation for wideband FM wave obtained by employing the Bessel functions.

$$
\begin{aligned}
eFM = s(t) = E_c\{J_0(m_f)\sin\omega_c t + J_1(m_f)[\sin(\omega_c+\omega_m)t - \sin((\omega_c-\omega_m)t] \\
+ J_2(m_f)[\sin(\omega_c+2\omega_m)t - \sin((\omega_c-2\omega_m)t + J_3(m_f)[\sin(\omega_c+3\omega_m)t \\
- \sin((\omega_c-3\omega_m)t + J_4(m_f)[\sin(\omega_c+4\omega_m)t - \sin((\omega_c-4\omega_m)t \ldots\ldots\ldots(1)
\end{aligned}
$$

From this equation, it is possible to draw the following conclusions:

- The FM wave has a carrier. Equation (1)'s first term represents this carrier.
- Ideally, there are an infinite number of sidebands in an FM wave ideally consists of infinite number of sidebands. Other than the first term, the rest are all sidebands.
- The J coefficients decide the sidebands' and the carrier's amplitudes.
- The J coefficients' values are dependent on the modulation index m_f, and it is the modulation index which decides how many sideband components have significant amplitudes.
- Few J coefficients can be negative. So, there is a phase shift of a 180 for that specific pair of sidebands.
- The carrier component does not remain constant. The $J_0(m_f)$ varies and with it the carrier's amplitude also varies. Nevertheless, the FM's amplitude stays constant.

NOTES

- In case of certain modulation index values, there will be a complete disappearance of the carrier component. Such values are known as Eigen values.
- The total power that is transmitted will always remain constant in the case of FM. This is independent of modulation index. It is because the FM signal's amplitude E_c remains constant at all times. The power that is transmitted is depicted in the following manner:

$$P_t = \frac{\left(\frac{E_c}{\sqrt{2}}\right)^2}{R} = \frac{E_c^2}{2R}$$

Where,

$$E_c = \text{peak amplitude of FM wave}$$

Hence:

$$P_t = \frac{E_c^2}{2R} \text{ If } R = 1\Omega$$

4.2.7 Generation and Detection of FM

FM signal generation can be performed in the following two ways:

- Direct Method
- Indirect Method

FM signal generation using direct method

$$f_i = f_c + k_f m(t)$$

In the case of a direct FM system, there is a direct varying of the instantaneous frequency with the information signal. To vary the frequency of the carrier, an Oscillator is to be employed and its components that can be varied determine its resonant frequency. The modulating signal amplitude thus changes the oscillator frequency.

FM signal generation indirect method

$$x(t) = A_c \cos(2\pi f_c t + \theta(t))$$

$$\theta(t) = 2\pi k_p m(t)$$

$$\theta(t) = 2\pi k_f \int_0^t m(\tau)d\tau$$

Angle modulation comprises PM and FM and both are interrelated, when one changes, so does the other. The information signal frequency also deviates the carrier frequency in PM. Frequency modulation is produced by phase modulation. Due to the fact that there is varying of the phase shift amount, it seems that the frequency is changed.

Due to the fact that PM produces FM, the former is known as indirect FM. First, there is the integration of the information signal and then it is used for phase modulating crystal-controlled oscillator, which gives frequency stability. For ensuring minimum distortion in the phase modulator, it is essential to keep the modulation index small, which leads to a *narrow-band FM-signal*.

A frequency multiplier is used to multiply the narrow-band FM signal in frequency to achieve the desired wide-band FM signal. Narrow band can be converted to wide band with the help of a frequency multiplier. The new wave form's frequency deviation will be "*M*" times the frequency deviation of the old wave form, and there will be no change in the rate at which the instantaneous frequency varies.

NOTES

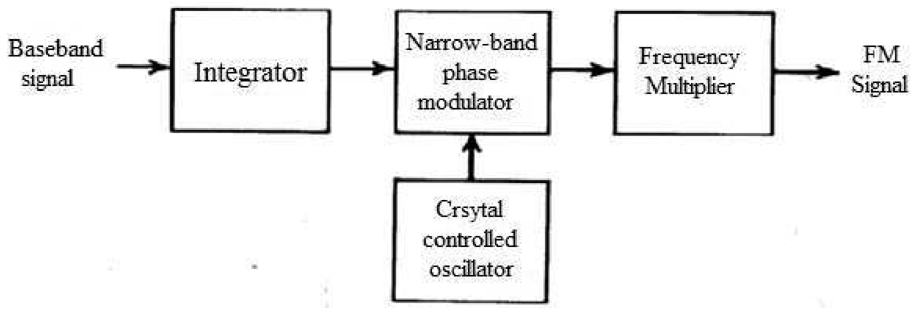


Fig. 4.11 *Block diagram of the indirect method of generating a wide band FM signal*

4.2.8 Discriminator

An ideal FM discriminator comprises of slope circuits (differentiator) which is followed by an envelope detector. A limiter is inserted before the differentiator to ensure a constant amplitude of the input signal. Following expression is derived from the property of Fourier Transform.

$$F\left\{\frac{d}{dt}x(t)\right\} = j2\pi f x(f)$$

The differentiator's transfer function will be:

$$H(f) = j2\pi f$$

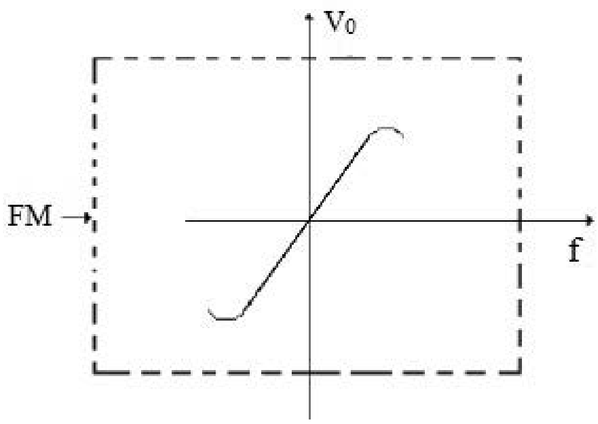


Fig. 4.12 *Circuit symbol of frequency discriminator*

Since the slope discriminator has poor linearity, it is used rarely. The following illustration is of a balanced frequency discriminator which in comparison to a single slope detector gives better linearity.

NOTES

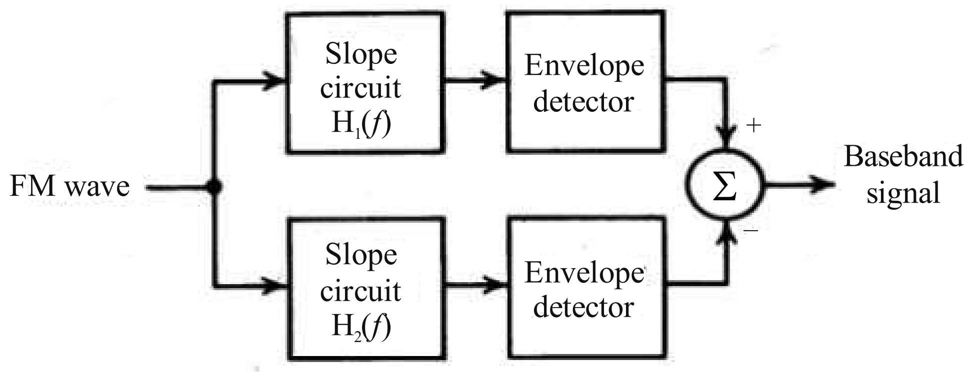


Fig. 4.13 *Balanced frequency discriminator*

In Figure 4.13, the lower and upper resonant filter sections are tuned below and above the unmodulated carrier frequency f_c signal, f_1 and f_2.

Consider that each of the two filters has a high quality factor or Q-factor. A resonant circuit's Q-factor is a measure of the goodness of the entire circuit.

The separation of the two resonant frequencies determines the linearity of the useful portion of the total response, whose center is f_c.

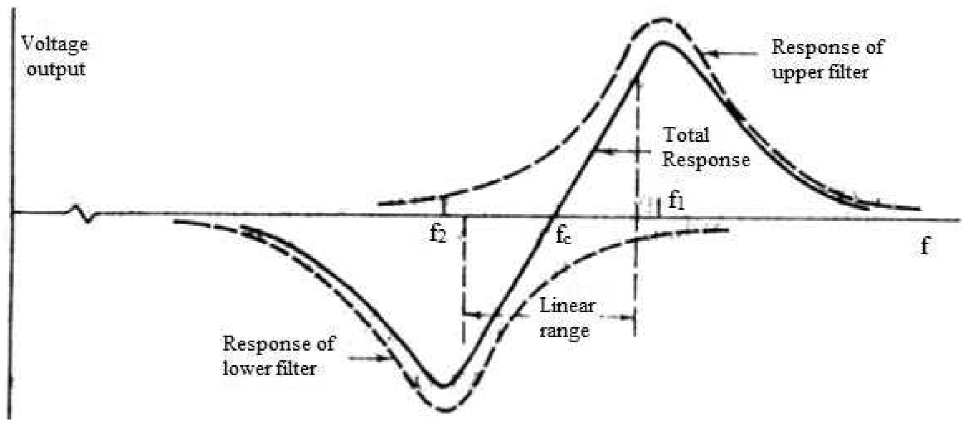


Fig. 4.14 *Response Curve of the Upper and Lower Resonant Filter*

Check Your Progress

1. Define the term modulation.
2. What do you mean by the digital modulation?
3. How will you define the sinusoidal baseband signal?
4. Name the two ways to classify frequency modulation.
5. State the Bessel functions.
6. Write the name of applications of demodulation.
7. What is narrow band FM?

NOTES

4.3 GENERATION OF DSB-SC MODULATION OF AM WAVE

It is to be understood that generally a useful modulation signal m(t) will not comprise a single sine wave, as shown in the above example. Nevertheless, using the Fourier decomposition principle, it is possible to express m(t) to be the sum of a several sine waves of varied frequencies, amplitudes, and phases. Performing multiplication of 1+ m(t) with c(t) (as above) will give the result which comprises a sum of sine waves. Once more, the c(t) carrier is found present with no change at all, while every frequency component of m at f_i there are two sidebands at frequencies $f_c + f_i$ and f_c - f_i. The collection of the former frequencies that lies below the carrier frequency is referred to as the lower sideband, and those above form the upper sideband. When looked at in a bit different manner, it is possible to consider the modulation m(t) as comprising an equal mix of negative and positive frequency components (as results from a formal Fourier transform of a real valued quantity) as shown in the top of following illustration. It is also possible to look at the sidebands to be that modulation m(t) which have only been shifted in frequency by f_c as shown at the bottom right of following illustration (formally, the modulated signal even comprised identical components at negative frequencies, which is depicted at the bottom left of the following illustration for the purpose of completeness).

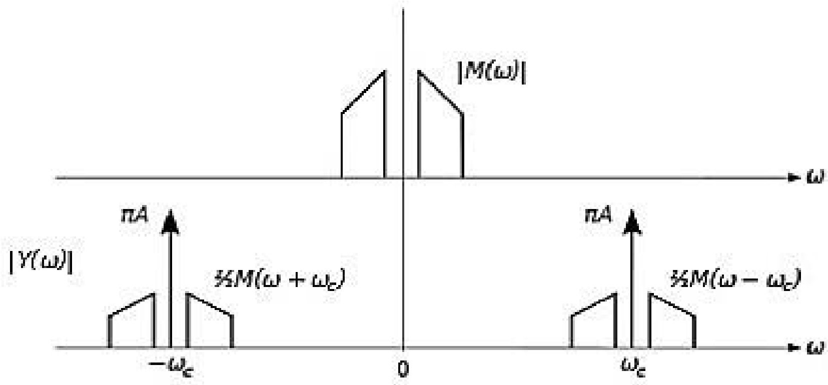


Fig. 4.15 *Double-sided spectra of Baseband and AM Signals*

If only the short-term spectrum of modulation is considered, then it is possible to plot the frequency content (on the horizontal axis) as a function of time (on the vertical axis) as in the following image. With the varying of the modulation frequency content, at any point in time there is an upper sideband generated according to those frequencies shifted above the carrier frequency, and the same content mirror-imaged in the lower sideband below the carrier frequency. The carrier is always constant, and its power is always higher than the total power of the sideband.

Power and spectrum efficiency

In AM transmission the RF bandwidth is two times the bandwidth of the modulating/ baseband signal. This is so because both the lower and upper sidebands around the carrier frequency come with a bandwidth which is as wide as the highest modulating frequency. While an AM signal has a narrower bandwidth than one employing Frequency Modulation (FM), its bandwidth is still two times as wide as single-sideband techniques. In this way, it is possible to see it as spectrally inefficient. So, within a frequency bandit is possible to accommodate only half as many channels or transmissions. Due to this analog television makes use of a variant of single-sideband called vestigial sideband so that the required channel spacing can be decreased. Standard AM can also be made better with suppression/ reduction of the modulated spectrum's carrier component.

NOTES

In the above Figure 4.15, it is seen in the spike that lies between the sidebands; even if the sine wave modulation is at 100% which is full modulation, the carrier component's power will remain two times that present in the sidebands, and even then it will not be carrying any information that is unique. Thus there is a great advantage in efficiency in reducing or totally suppressing the carrier, either in conjunction with elimination of one sideband (single-sideband suppressed-carrier transmission) or with both sidebands remaining (double sideband suppressed carrier). Though the suppressed carrier transmissions provide efficiency of transmitter power, they only work with receivers that are more sophisticated and make use of synchronous detection and regeneration of the carrier frequency. Due to this, there is continued wide usage of standard AM, more so in the field of broadcast transmission, so that inexpensive receivers (that employ envelope detection) can be used. Television (analog) which has a largely suppressed lower sideband, also has enough carrier power to use envelope detection. On the other hand, in case of such communications systems in which the receivers as well as the transmitters are optimizable, the suppression of the carrier and the one sideband display a net advantage and are used frequently.

Modulation Index

The AM modulation index is a measure based on the ratio of the modulation excursions of the RF signal to the level of the unmodulated carrier. Following is how it is, therefore, defined.

$$h = \frac{\text{peak value of } m(t)}{A} = \frac{M}{A}$$

The modulation amplitude is the peak (be it negative or positive) change in the RF amplitude from its unmodulated value. Generally, we express the modulation index in the form of a percentage, and it may be displayed on a meter which is attached with the AM transmitter. So, if the carrier amplitude will be varying by 100% above (and below) its unmodulated level. In the case of, it varies by 50%. In case when there is 100% modulation, it is possible that at times the wave amplitude reach zero, and it will imply full modulation using standard AM and at times this is the target (so that highest possible signal-to-noise ratio can be attained) but this should not be crossed/exceeded. If the modulating signal is increased further than this point (called over modulation), it will result in a standard AM modulator failing, due to the fact that the negative excursions of the wave envelope are incapable of becoming lower than zero, leading to distortion/clipping of the modulation that is received. Typically, in transmitters there is a limiter circuit which prevents overmodulation, and/or a compressor circuit (more so with voice communications) to approach 100%

modulation for maximum intelligibility above the noise. Circuits of this type are also called vogad.

NOTES

Nevertheless, it is possible to have a modulation index exceeding 100%, with no distortion introduced, as with double-sideband reduced-carrier transmission. In that case, negative excursions beyond zero require the carrier phase's reversal. It is not possible to produce this thru employing the efficient high-level (output stage) modulation techniques that are used widely and specifically so in broadcast transmitters that are of high power. Furthermore, such a waveform is produced by a special modulator at a low level and this is followed with a linear amplifier. Also, it is not possible a standard AM receiver that uses an envelope detector to properly demodulating a signal of this kind. What is needed in this case is synchronous detection. Due to this, generally double-sideband transmission does not fall in the category of "AM" despite it forming an RF waveform which is identical to a standard AM till the modulation index stays lower than 100%. Majority of the times, systems of this type try to create a radical reduction of the carrier level compared to the sidebands reduced to zero. Each case of this type displays a loss of vale of the term "Modulation Index" as it refers to the ratio of the modulation amplitude to a rather small (even zero) remaining carrier amplitude.

Modulation methods

The classification of modulation circuit designs falls in two categories: low- level and high-level. This classification is based on whether they modulate in a low-power domain—followed by amplification for transmission—or in the high-power domain of the transmitted signal.

Generation: Low-level

The modern radio systems, generates modulated signals through the process of Digital Signal Processing (DSP). Using DSP, it is possible to have several types of AM with software control, such as Independent Sideband (ISB), SSB suppressed-carrier and DSB with carrier. Calculated digital samples are converted to voltages with a digital to analog converter, typically at a frequency less than the desired RF-output frequency. Following this, the analog signal has to be shifted in frequency amplified linearly for the required power and frequency level (linear amplification is required for the prevention of distortion of modulation). Several Amateur Radio transceivers make use of this low-level method for AM.

Generation: High-level

AM transmitters that are of high power (called high-power AM transmitters for example, the ones employed for AM broadcasting) are based on high-efficiency class-D and class-E power amplifier stages, where supply voltage is varied for attaining modulation.

Even designs that are older (in amateur radio and in broadcast) generate AM through controlling the gain of the transmitter's final amplifier (usually class-C, to attain efficiency). The list given below is of the types for vacuum tube transmitters and even transistors come with similar options.

NOTES

- **Plate Modulation:** In this case, the RF amplifier's plate voltage is modulated with the audio signal. The requirement audio power is just 50 percent of the RF-carrier power.
- **Constant-Current/Heising Modulation:** In this modulation, the RF amplifier plate voltage is fed through a "choke" which is a high-value inductor. The same inductor is also used to feed the AM modulation tube plate, so the modulator tube diverts current from the RF amplifier. In the audio range, the choke performs the role of a constant current source. There is low efficiency of power in such a system.
- **Control Grid Modulation:** It is possible to control the operating bias and gain of the final RF amplifier through varying of the voltage of the control grid. In this method, very little audio power is needed though it is essential to take care to lessen distortion.
- **Clamp Tube/screen Grid Modulation:** It is possible to control the screen-grid bias with a "clamp tube", which will reduces voltage basis the modulation signal. In this system, while maintaining low distortion it is extremely difficult to achieve 100-percent modulation.
- **Doherty Modulation:** In this case, while one of the tube provides power under carrier conditions, another tube operates only for positive modulation peaks. In this system, there is low distortion and good overall efficiency.
- **Outphasing Modulation:** In this method, there is parallel operation of two tubes, though they are just partially out of phase with each other. Since these tubes are differentially phase modulated, when combined, their amplitude is either smaller or greater. In this system, if the adjustment is proper, then there is low distortion and good overall efficiency.
- **Pulse Width Modulation (PWM) or Pulse Duration Modulation (PDM):** In this technique, there is the application of an extremely efficient high voltage power supply which is applied to the tube plate. This supply's output voltage is varied at an audio rate to follow the program.

4.3.1 Frequency spectrum of AM wave

- Amplitude modulation refers to a process that is used to vary the amplitude of high frequency carrier signal in accordance with the amplitude of the low frequency modulating or information signal, while the carrier signal's phase and the frequency are kept constant.
- Consider the carrier voltage to be v_c and the modulating voltage to be v_m, both of which are being represented as follows:

$$v_c = V_c \sin w_c t$$
$$v_m = V_m \sin w_m t$$

- With amplitude modulation, the unmodulated carrier's (v_c) amplitude will be varied in proportion with the instantaneous modulating voltage V_m sin $\acute{E}_{mt}$.
- In case where modulation is not there, the carrier's amplitude will be equal to its unmodulated value and in the case where there is modulation, the

carrier's amplitude will be varied by its instantaneous value (as seen in the illustration that follows).

NOTES

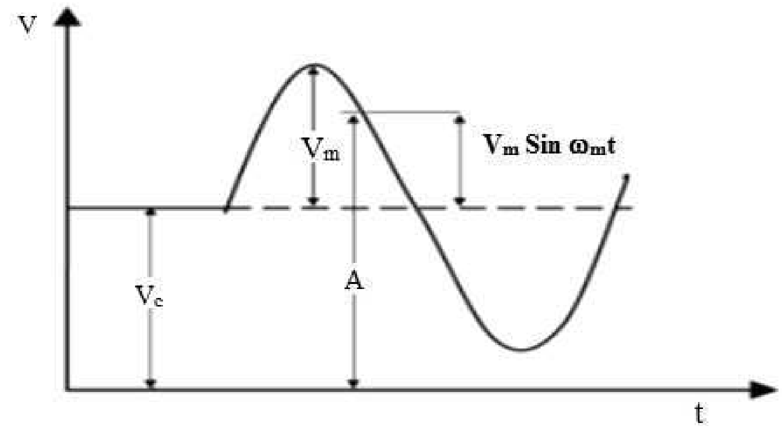


Fig. 4.16 *Amplitude of AM Wave*

- Illustration given below provides an amplitude modulated wave's time domain representation depicting the carrier signal's modulation by the modulating signal.
- The modulating wave as well as the carrier wave in the case or AM wave is sinusoidal in nature yet the modulated wave is not. Following is the amplitude of the AM.

$$v_{AM} = V_c \sin w_c t + \frac{mV_c}{2}\cos(w_c - w_m)t - \frac{mV_c}{2}\cos(w_c + w_m)t$$

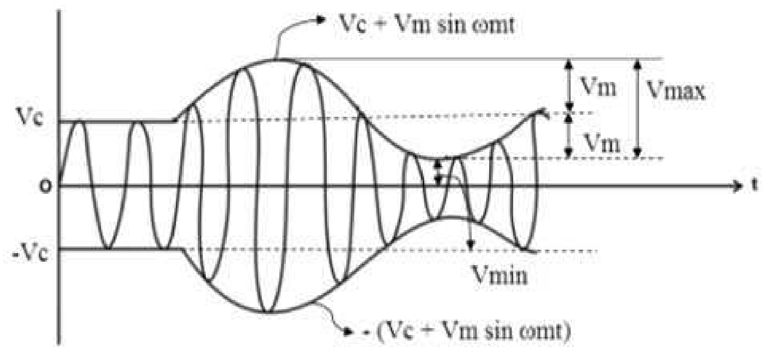


Fig. 4.17 *Time Domain Representation of AM Wave*

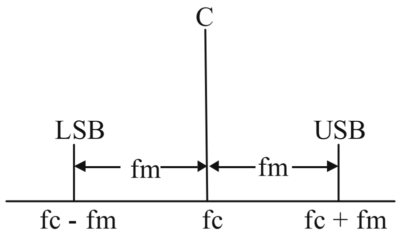
- In light of the above, it is possible to say that besides the original signal two additional sine waves exist, one below and one above the carrier frequency. So, the complete AM signal comprises one carrier wave plus two additional frequencies one on rather side, referred to as side frequencies.
- The frequency below the carrier frequency is referred to as lower sideband and the one above the carrier frequency is referred to as upper sideband.
- The frequency of the Lower Sideband (LSB) is:

  $f_c - f_m$
- The frequency of the Upper Sideband (USB) is:

  $f_c + f_m$

NOTES

- The illustration given below represents the spectrum of amplitude modulated wave.



- The amplitude of the central frequency (the carrier frequency f_c) is the highest. The sideband frequencies which lie on one on each side of the carrier frequency have lower amplitude. The following is used to provide the bandwidth of the amplitude modulated wave:

$$BW = (f_c + f_m) - (f_c - f_m) = 2f_m$$

4.3.2 DSB-SC Modulation of AM Wave

While there are several forms of amplitude modulation (AM) generation, the normal AM or the Double-Side-Band (DSB) is generated with 'Switching-Modulator' or the 'square-law' modulator and the Double-Side-Band-Suppressed Carrier (DSB-SC) with the balanced modulator or the 'ring-modulator'. When there is DSB generation by employing 'switching modulator', it becomes essential that the amplitude of the carrier remain much higher than the amplitude of the message due to the fact that the switching action has to remain dependent on just the carrier, though in the square-law modulator, the total amplitude of the message and carrier to be 'low' enough to operate the active device in the 'square-law' region. So, it must be noted that there are strict conditions for operation associated with the two existing methods of discrete component realization of the AM. For DSB-SC, the 'ring modulator' which is employed is not of the kind with which has operating conditions that are difficult to maintain and strict. With the carrier, the DSB-SC becomes the DSB, and when DSB has the carrier removed it leads to DSB-SC modulation. Nevertheless, it is not possible to use the 'square-law' modulator or the 'Switching-Modulator' for generating DSB-SC. Similarly, the 'ring-modulator' cannot be employed for DSB generation.

4.3.3 Demodulation of AM Waves Generation of DSB-SC Waves

Double-Sideband Suppressed-Carrier Transmission (DSB-SC) is the type of transmission where the frequencies that are produced by AM are spaced symmetrically below and above the carrier frequency. At the same time, the carrier level is reduced to the lowest practical level, with ideally it being suppressed completely.

As opposed to AM, in the case of DSB-SC modulation, the wave carrier is not transmitted; so, most of the power gets distributed between the sidebands, which implies an increase of the cover in DSB-SC, compared to AM, for the same power used.

Employed in the case of radio data systems, the DSB-SC transmission is a special case of double-sideband reduced carrier transmission.

NOTES

DSB-SC Spectrum

Basically, DSB-SC is an amplitude modulation wave sans the carrier, due to which there is a reduction in power waste and also 100% efficiency. In comparison with normal AM transmission (DSB) this is an increase. DSC provides a 33.333% efficiency at the maximum and this is because the carrier uses 2/3rd of the power even though it does not carry any intelligence, and also both the sideband carries identical information. There is hundred percent efficiency with the use of Single Side Band (SSB) Suppressed Carrier.

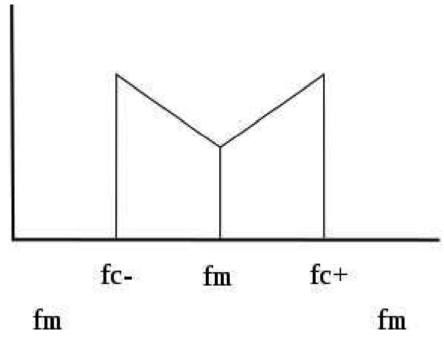


***Fig. 4.18** Spectrum Plot of a DSB-SC Signal*

DSB-SC Generation

A mixer is used to generate DSB-SC. It comprises a message signal multiplied by a carrier signal. Following is a mathematical representation of the process, which makes use of the product-to-sum trigonometric identity.

$$\underbrace{V_m \cos(w_m t)}_{\text{Message}} \times \underbrace{V_c \cos(w_c t)}_{\text{Carrier}} = \underbrace{\frac{V_m V_c}{2}[\cos((w_m + w_c)t) + \cos((w_m - w_c)t)]}_{\text{Message}}$$

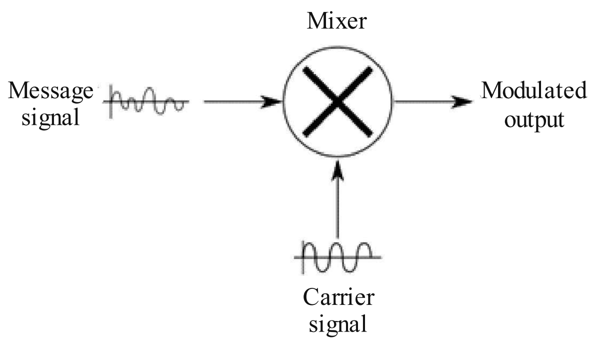


***Fig. 4.19** Mathematical Representation of DSB-SC Generation*

DSB-SC Demodulation

Demodulation is similar to the process of modulation and is accomplished by the multiplication of the DSB-SC signal with the carrier signal. The signal thus obtained is passed through a low pass filter which then created a scaled version of original message signal. In case the modulation index is less than unity, it is possible to use a simple envelope detector, such as AM, to demodulate DSB-SC. Carrier re-insertion has to be carried out in the case of full depth modulation.

$$\overbrace{\frac{V_m V_c}{2}[\cos(w_m + w_c)t) + \cos((w_m - w_c)t)]}^{\text{Modulated Signal}} \times \overbrace{V_c' \text{ os}(w_c t)}^{\text{Carrier}}$$

=

NOTES

$$\left(\frac{1}{2}V_c V_c'\right)\underbrace{V_m \cos(w_m t)}_{\text{original message}} + \frac{1}{4}V_c V_c' V_m[\cos((w_m + 2w_c)t) + \cos((w_m - 2w_c)t)]$$

The above equation depicts that when the carrier signal and the modulated signal are together multiplied, it creates a scaled version of the original message signal and also a second term.

With: $w_c \gg w_m$

the additional second term has a frequency that is much higher compared with the original message. When this signal has passed through a low pass filter, it leads to the removal of the higher frequency component and only the original message remains.

DSB-SC Distortion and Attenuation

To accomplish demodulation, it is essential that the demodulation oscillator's phase and frequency remain exactly the same as that of the modulation oscillator, else there will be distortion and/or attenuation.

The following provides an understanding of this effect:

- Message signal to tranmitted $f(t)$
- Modulation (carrier) signal $V_c \cos(w_c t)$
- Demodulation signal (with frequency and phase the deviations from the modulation signal) $V_c' \cos[(w_c + \Delta w)t + \theta]$

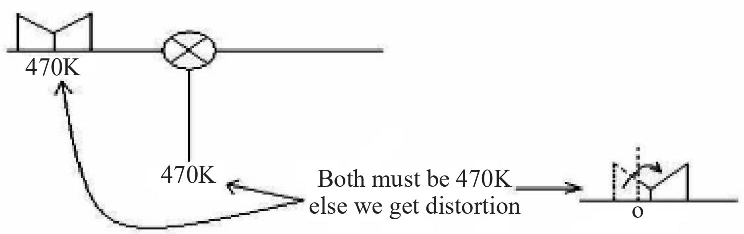
The resultant signal can then be given by

$$f(t) \times V_c \cos(w_c t) \times V_c' \cos[(w_c + \Delta w)t + \theta]$$

$$= \frac{1}{2}V_c V_c' f(t)\cos(\Delta w \cdot t + \theta) + \frac{1}{2}V_c V_c' f(t)\cos[(2w_c + \Delta w)t + \theta]$$

$$\xrightarrow{\text{After low pass filter}} \frac{1}{2}V_c V_c' f(t)\cos(\Delta w \cdot t + \theta)$$

The $\cos(\Delta w \cdot t + \theta)$ terms cause the original message signal to be attenuated and distorted. Specifically so, with frequencies being correct, and the phase being wrong, contribution from θ is a constant attenuation factor, also $\Delta w \cdot t$ depicts a cyclic inversion of the recovered signal, which is a serious form of distortion.

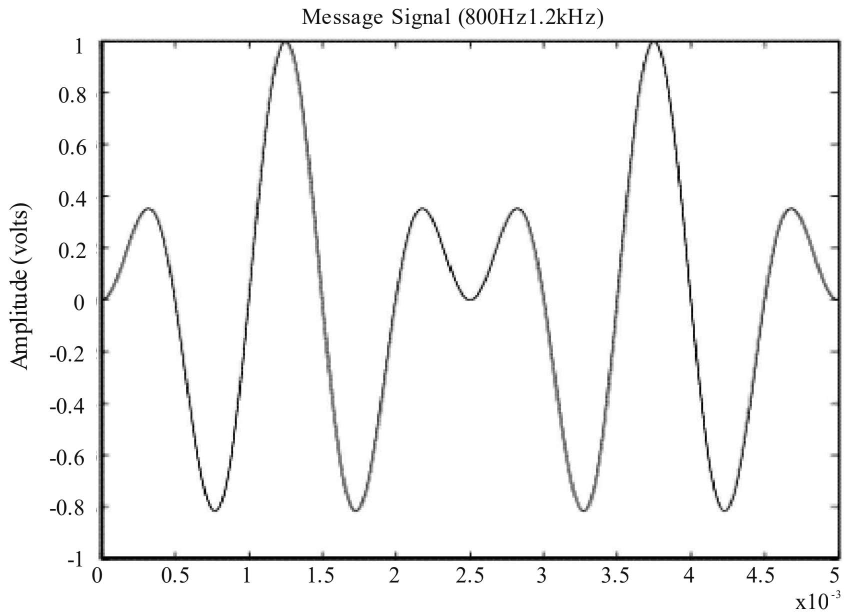


Its Working

The easiest manner of depicting the working of a DSB-SC is graphically. The

NOTES

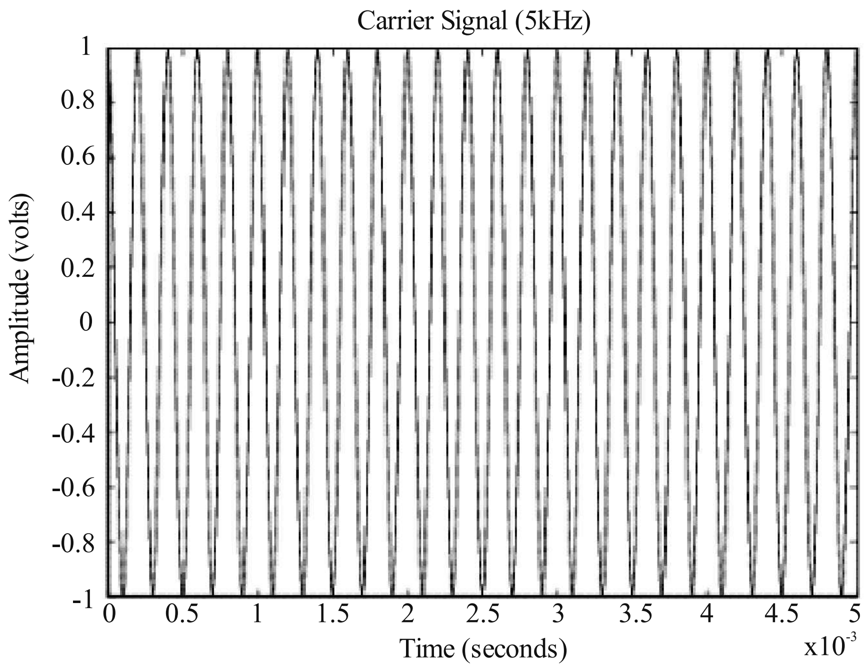
illustration below depicts a message signal that one may wish to modulate onto a carrier, comprising a couple of sinusoidal components with frequencies respectively 800 Hz and 1200 Hz.



In case of this message signal, the equation will be:

$$s(t) = \frac{1}{2}\cos(2\pi 800t) - \frac{1}{2}\cos(2\pi 1200t)$$

Here, the carrier is a plain 5 kHz ($c(t) = \cos(2\pi 5000t)$) sinusoid, shown in the following illustration.

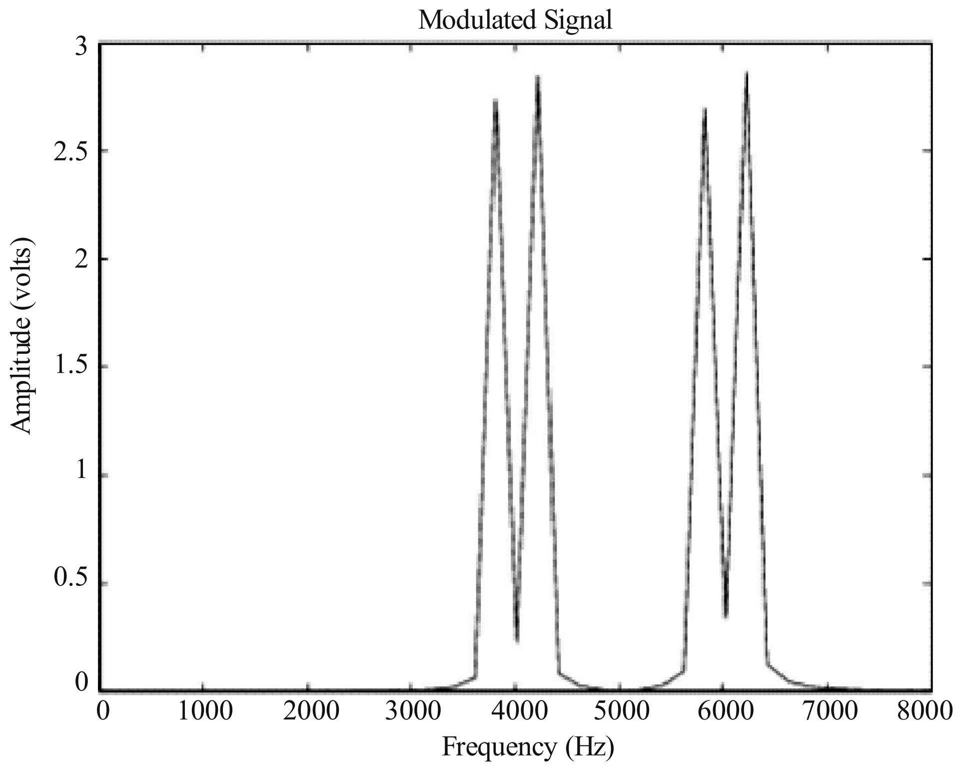


Multiplication in the time domain is used to perform the modulation, which rests in a 5 kHz carrier signal, with varying amplitude in the same manner as the message signal.

$$x(t) = \underbrace{\cos(2\pi 5000t)}_{\text{Carrier}} \times \underbrace{\frac{1}{2}\cos(2\pi 800t) - \frac{1}{2}\cos(2\pi 1200t)}_{\text{Message Signal}}$$

NOTES

The term "Suppressed Carrier" is derived from the carrier signal component being suppressed and not present in the output signal. The output signal's spectrum clearly depicts the absence of the carrier signal component. The illustration below depicts four peaks, the two peaks below 5000 Hz are the Lower Sideband (LSB) and the two peaks above 5000 Hz are the Upper Sideband (USB), and there is no peak at the 5000 Hz mark, which is the frequency of the suppressed carrier.



4.3.4 Balanced Modulator

A method, referred to as the Armstrong method, was patented by Edwin H. Armstrong in 1933 to generate frequency modulation of radio signals. It generates a double sideband suppressed carrier signal, phase shifts this signal, and then reinserts the carrier to produce a frequency modulated signal.

With frequency modulation, there is the generation of high quality audio. Furthermore, frequency modulation greatly decreases the noise on the channel in comparison to amplitude modulation. AM was employed by early broadcasters since it was easier to generate compared with frequency modulation, and receivers for it were simpler to construct. According to electronics theory, a signal that has been frequency modulated has infinite bandwidth and in the case of an amplitude modulated signal, the bandwidth is about twice the highest modulating frequency.

It was understood by Armstrong that in the case of a signal that is frequency modulated the bandwidth would be infinite, and also that just the first few sets of sidebands would be significant and the others being insignificant could be ignored.

An amplitude modulated voice channel bandwidth would be approximately 6 kilohertz; a common frequency modulated voice channel bandwidth could be 15 kilohertz.

Its Working

In the Armstrong method, foremost there is the generating of a carrier signal at a frequency that is extremely low, such as 500 kilohertz. The frequency will be lower than that of the AM broadcast band and much lower than the current FM broadcast band of 88 to 108 megahertz. The carrier signal will then be applied to two stages in the transmitter: a balanced modulator and a mixer.

It is essential to have knowledge of amplitude modulation and its working to be able to understand a balanced modulator and the working of balanced modulator. Amplitude modulation can be described as a method of changing the strength of the carrier (amplitude) in sync with the modulating audio. While there is a change in the power output with modulation, the change is due to the fact that all AM modulator generates two sidebands, one below and one above the carrier. When power enters these 2 sidebands, there is an increase in the power output. Then, the amplitude modulated signal, comprises two sidebands and one constant strength carrier. The information is carried by the sidebands and the carrier carries not relevant additional information. The carrier can be removed at the transmitter and reinserted at the receiver so that the transmitter can divert all of the power to the sidebands.

Sidebands are also generated by a frequency modulator and these are several on either side of the carrier rather than just one on either side of the carrier. The existence of many sidebands on either side of the carrier widens the bandwidth of FM. The power output from an FM transmitter is constant with modulation, therefore, with power going into the sidebands, the power in the carrier decreases.

In a balanced modulator, there is the mixing of the radio frequency carrier and the audio signal along with suppression of the carrier, due to which only the sidebands remain. The balanced modulator’s output is a double sideband suppressed carrier signal which holds every information that is there in the AM signal yet sans the carrier. An AM signal can be generated thru taking the balanced modulator’s output and performing a carrier reinsertion on it.

While using the Armstrong method, the radio frequency carrier signal and the audio signal are both applied to the balanced modulator so that a double sideband suppressed carrier signal is generated. The output signal’s phase then is shifted 90 degrees respective to the original carrier. The balanced modulator output can either lead or lag the carrier’s phase. Then, both the original carrier signal and the double sideband signal are applied to the mixer, and the original carrier—90 degrees out of phase—is reinserted. The mixer’s output is a frequency modulated signal.

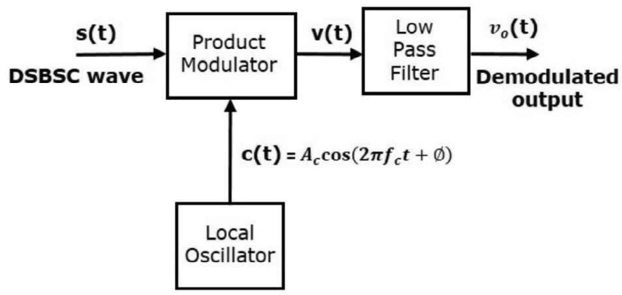
An AM signal will get produced if the carrier is reinserted without using the phase shift. When the carrier is reinserted with the 90 degree phase shift, a PM signal is produced. When intelligence is integrated prior to being applied to the resulting phase modulator, it becomes equivalent to an FM signal.

NOTES

A problem that is associated with the Armstrong method is that to keep distortion at a minimum it is essential to keep the frequency deviation (modulation amount) small. While the maximum deviation is a fraction of 1 kilohertz, in FM broadcast it needs to be 75 kilohertz deviation and a typical FM voice channel it needs to be 5 kilohertz deviation. To resolve this, Armstrong multiplied the signal several times to a higher frequency to obtain the necessary deviation.

4.3.5 Coherent Defection of DSB-SC Wave

The message signal is detected using the same carrier signal that is utilised to generate the DSB-SC signal. As a result, this detection method is known as coherent or synchronous detection. The coherent detector's block diagram is shown below.



By multiplying the message signal with a carrier that has the same frequency and phase as the carrier used in DSB-SC modulation, the message signal can be recovered from the DSB-SC wave. A Low Pass Filter is subsequently applied to the resulting signal. This filter's output is the intended message signal.

Let the DSB-SC wave be

$$s(t) = A_c \cos(2\pi f_c t) m(t)$$

The local oscillator's output is

$$c(t) = A_c \cos(2\pi f_c t + \phi)$$

Where,

The phase difference between the local oscillator signal and the carrier signal used for DSB-SC modulation is ϕ.

The output of the product modulator can be written as shown in the diagram.

$$v(t) = s(t)\, c(t)$$

In the above equation, substitute s(t) and c(t) values.

$$\Rightarrow v(t) = A_c \cos(2\pi f_c t) m(t)\, A_c \cos(2\pi f_c t + \phi)$$

$$= {A_c}^2 \cos(2\pi f_c t) \cos(2\pi f_c t + \phi) m(t)$$

$$= \frac{{A_c}^2}{2} [\cos(4\pi f_c t + \phi) + \cos\phi]\, m(t)$$

$$v(t) = \frac{{A_c}^2}{2} \cos\phi\, m(t) + \frac{{A_c}^2}{2} \cos(4\pi f_c t + \phi) m(t)$$

NOTES

The scaled form of the message signal is the first term in the equation above. By running the above signal through a low pass filter, it can be retrieved.

As a result, the low pass filter's output is

$$v_0 t = \frac{A_c^{\ 2}}{2} \cos \phi m\,(t)$$

The amplitude of the demodulated signal will be at its maximum, where $\phi = 0^p$.

As a result, the local oscillator signal and the carrier signal must be in phase, i.e., there must be no phase difference between them.

The amplitude of the demodulated signal will be zero, where $\phi = \pm 90p$

The quadrature null effect is the name for this effect.

4.4 GENERATION AND DETECTION OF SSB WAVES

The following are the various techniques for generating the SSB waves.

Filter Method

To use the filter method to generate an SSB modulated wave, it is essential for the message signal to satisfy some conditions, which are:

1. There must not be any low frequency content in the message signal.
2. In the spectrum of the message signal, the highest frequency should be much smaller than the carrier frequency.

System Block Diagram

The diagram given below depicts a block diagram of an SSB modulator operating on the frequency discrimination principle.

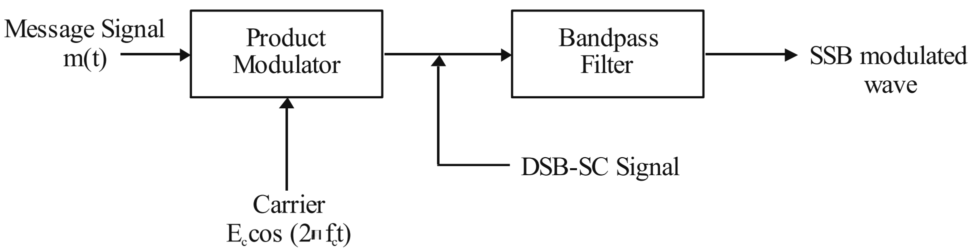


Fig. 4.20 *Block Diagram of an SSB Modulator*

As Shown in the Figure 4.20, the modulator comprises a carrier oscillator, product modulator and bandpass filter made for the purpose of passing the desired sideband. The output of the product modulator is the DSB-SC modulated wave that has just two sidebands. With the bandpass filter, on single sideband will pass and generate the SSB modulated wave at its output

Disadvantage

As shown in the Figure 4.20, the DSB-SC signal at the output of the product modulator comprises both of the two sidebands. As evident from the illustration given below, the difference of frequency between LSB's highest frequency and USB's lowest frequency is extremely minute.

NOTES

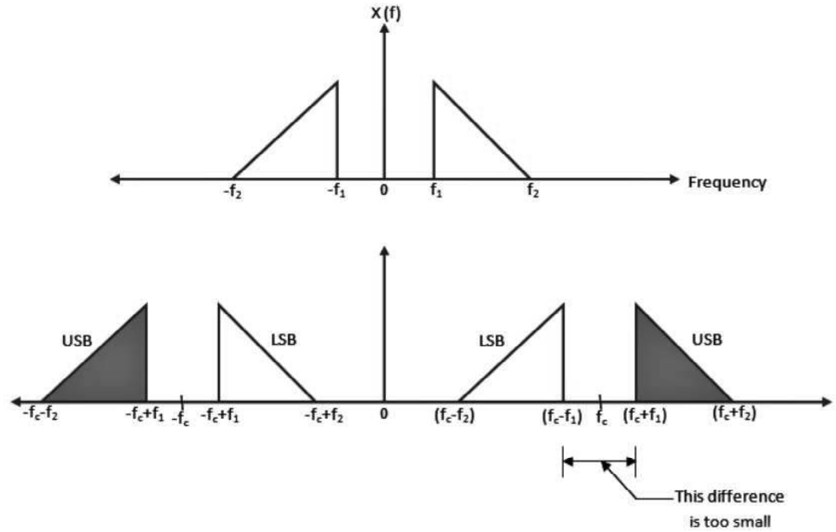


Fig. 4.21 *Spectrum of Message Signal and Product Modulator Output*

This makes the design of the bandpass filter extremely difficult because its frequency response needs to have very sharp change over from attenuation to pass band and vice versa.

Design of bandpass Filter

The conditions given below need to be satisfied by the design of the bandpass filter.

1. Passband of the BPF and the spectrum of the desired SSB modulated wave should occupy the same frequency range.
2. The width of the guard band which separates the passband from stopband be =-twice the lowest frequency component of the message signal. This will be Guard band = 2f1 Hz

4.4.1 Phase Shift Method

As shown in Figure 4.22 represents use of phase shift method for SSB generation. The depicted system is employed for the purpose of suppressing the lower sideband. The system employs two balanced modulators M1 and M2 and two 90o phase shifting networks.

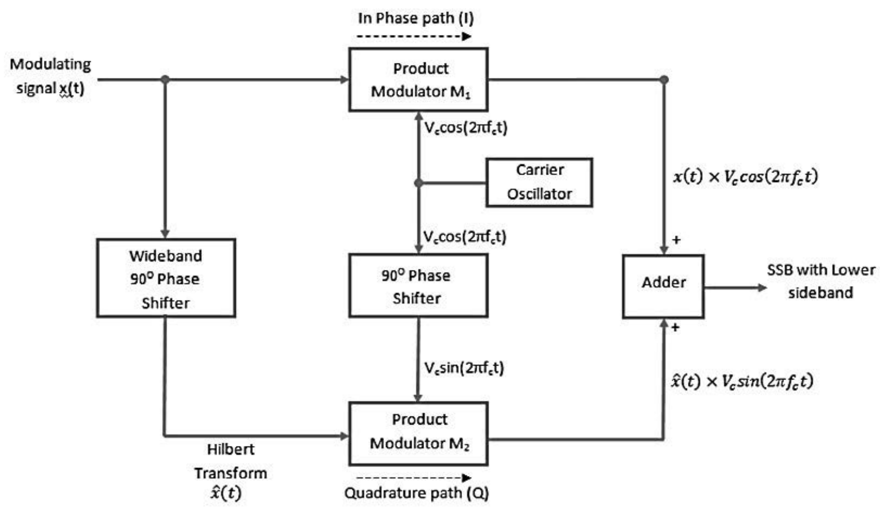


Fig. 4.22 *Phase Shift Method for SSB Signal Generation*

Working Operation

Application of the message signal x(t) is done to the product modulator M_1 and via a 90° phase shifter it is applied to the product modulator M_2.

This, at the output of the wideband 90° phase shifter, provides the Hilbert transform:

NOTES

$$\hat{x}(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{x(k)}{t-k}dk$$

The carrier oscillator's output gets applied as is to the modulator M_1 and it is applied to M_2 post being passed through a 90° phase shifter.

$$\text{Output of } M_1 = x(t) \times V_c \cos(2\pi f_c t)$$

and Output of $M_2 = \hat{x}(t) \times V_c \sin(2\pi f_c t)$

Both M_1 and M_2 outputs are then applied to an adder.

So, the output of the adder is as given below.

$$\text{Adder Output} = x(t) \times V_c \cos(2\pi f_c t) + \hat{x}(t) \times V_c \sin(2\pi f_c t)$$

Or Adder Output $= V_c[x(t) \times \cos(2\pi f_c t) + \hat{x}(t) \times \sin(2\pi f_c t)]$

The above expression, depicts the SSB signal that has rejected the USB and has just a single LSB.

SSB with Upper Sideband (USB)

Given below is a block diagram depicting the use of the phase shift method for the generation of USB with SSB.

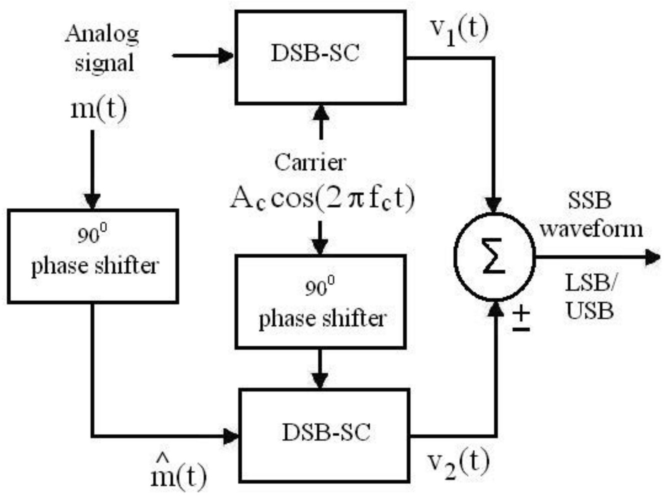


Fig. 4.23 *USB Suppression Phase Shift Method*

While the adder polarities for the in-phase is positive it is negative for the quadrature paths.

Suppression of the upper sideband

It is possible to arrange the blocks as given in the illustration to suppress the LSB and generate the SSB signal consisting of the USB.

Here, the carrier and modulating signals get applied to the upper balanced modulator directly (with no phase shift). Whereas, both these signals are 90° phase shifted and then applied to the lower balanced modulator.

NOTES

Advantages of Phase Shift Method

The following are the advantages of the phase shift method:

1. It is capable of generating the SSB signal at any frequency, so there is no need for the frequency up converter stage.
2. It is capable of employing the low audio frequencies as modulating signal. This cannot be done by using the filter method.
3. This enables easy switching between sidebands.

Disadvantages of Phase Shift Method

1. Foremost drawback is that the extreme criticality of the design of the 90^o phase shifting network for the modulating signal.
2. This network has to provide a correct phase shift of 90^o at all the modulating frequencies which is practically impossible to attain.

4.4.2 Third method/Weavers method

The Weaver method is also referred to as the "3rd method" or the "Third method". This could be because the first method is considered to be the superhet method and the second method the direct conversion phasing method. The Weaver receiver does not need an accurate audio phase shift network; rather, it uses a pair of Low Pass Filters (LPF) in two (90-degree phased) audio signal paths. It uses an audio subcarrier which is located at a point somewhere in the middle of the audio passband, for example at 1.5 kHz all of the phasing is done through an extra set of mixers. The LPF's audio have to be matched closely and this is a task that is greatly easier than building a precise 90-degree audio phase shift in one path. This provides an output without any unwanted sidebands. To explain it better, the unwanted sideband in the traditional sense is determined by the steep roll-off of the audio LPF. A narrow gap exists around the 1.5 kHz sub-carrier where we cannot hear any audio. So, since 1.5 kHz is not useful for the human voice, this gap's loss becomes immaterial.

The "unwanted sideband" that is caused by any mismatches anywhere is actually within the audio passband. It remains mirrored around/folded back around the 1.5 kHz subcarrier. So if we are listening to a wanted carrier signal at 1.0kHz, and there is also an interfering signal 1kHz higher up (equal distance the other side of the 1.5kHz subcarrier), then we will also hear the unwanted signal at some attenuation, depending on how perfectly matched everything is. In the SSB modulation mde, you do not hear anything from any adjacent channel and the attenuated reflection within the same channel goes unnoticed.

This method is a technique which does not employ the use of an accurate 90-degree audio phase shift, and with it no existence of any unwanted sideband interference from adjacent SSB channels. But it is important to also understand that due to the use of 4 mixers it has a high complexity in its circuit.

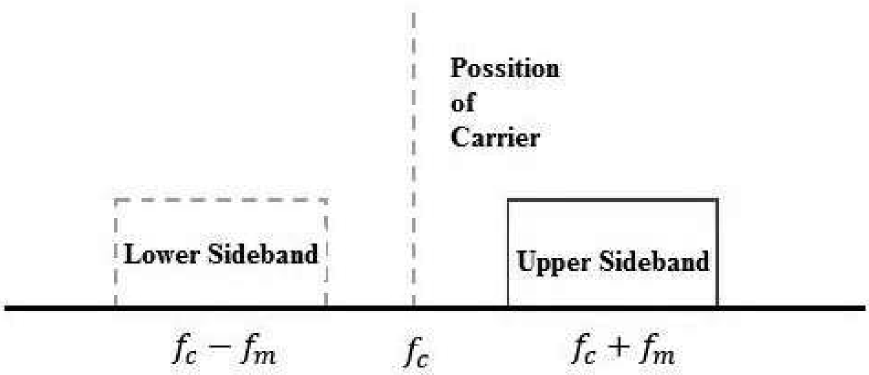
4.4.3 SSB Modulation and Its Applications

NOTES

In the short wave or HF sector of the radio spectrum, modulation such as SSB (single-sideband modulation) is widely employed for radio communications in two ways. This communication employs SSB modulation, which has a wide range of applications, including marine, HF point-to-point transmissions, military, and radio amateurs or radio hams. Because it overcomes various problems of AM (amplitude modulation), this modulation is evolved from AM. In general, this modulation is utilised for a variety of purposes, including voice transmission and radio communication in two directions using an analogue signal.

What is SSB Modulation?

A single-sideband modulation is a type of modulation that is used to send data over radio waves, such as an audio transmission. This modulation is used in radio communications to change the AM signal by utilising additional transmitter power and bandwidth (amplitude modulation). Many single sideband radio communication devices, such as SSB Tx, SSB Rx, and SSB transceiver, are available on the market.



SSB-Lower Sideband (LSB), Upper Sideband (USB), Double Sideband (DSB), Single Sideband Suppressed Carrier (SSB SC), Vestigial Sideband (VSB), and SSB decreased carrier are some of the SSB modulation variations.

Power Measurement of SSB

It's common to be asked to explain an SSB transmission's o/p power. For example, determining the transmitter power utilised for mutual radio communication is critical in order to maximise its efficiency for specific purposes. Because the real o/p power relies on the stage of the modulating signal, power measurement for an SSB transmission is difficult. PEP (peak envelope power) can be used to overcome this.

The RF envelope power for conduction is received here, and the peak level signal can be used at any moment. Watts can be used to compute the peak envelope level for power. These are only the power stages that are connected to 1 Watt, otherwise 1 mW.

NOTES

Advantages and Disadvantages of SSB Modulation

The following are the key benefits of single-sideband modulation.

- BW is smaller than amplitude modulation and DSB signals due to the occupied spectrum space.
- It may be possible to broadcast an additional number of signals.
- It is possible to save energy.
- The transmission of a high signal is possible.
- There is less noise present • There is less signal loss.

The following are the primary drawbacks of single-sideband modulation.

- The single-sideband signal finding and generating method is complicated.
- When the SSB transmitter and receiver have exceptional frequency strength, signal quality suffers.

Applications

The following are some of the applications of SSB modulation.

- This modulation is utilised in situations where power efficiency and bandwidth are important.
- It's utilised for communications in the air, on land, and at sea.
- It is used in military communications such as amateur radio, etc. • It is used in radio and point-to-point communications • It is used in communications such as telemetry, TV, and radar • It is used in communications such as telemetry, TV, and radar.

Thus, on the HF part of the radio spectrum, SSB modulation is utilised to broadcast analogue voice for mutual radio communication. When compared to other types, it is more effective in terms of power and spectrum. As a result, this modulation is an effective choice for mutual radio communication since it reduces development inefficiency.

4.4.4 Vestigial Sideband Modulation

The most commonly utilised modulations in the realm of communication are SSB and DSBSC. The DSBSC, or double sideband suppressed carrier system, is created when the carrier is covered up and the conserved power is distributed to the two sidebands. SSB or SSBSC refers to the process of suppressing one of the sidebands while transmitting only one sideband utilising the carrier (single sideband suppressed carrier system). When a sideband is transferred throughout the filters in SSB, the filter, such as bandpass, cannot function entirely in practise, and some data may be lost. DSBSC, on the other hand, sends two sidebands to transport the waste data. To solve this issue, a method known as VSB modulation was adopted, which combines DSBSC and SSB. The term "vestige" refers to something that is distinct from its derived name.

What is VSB Modulation?

Vestigial Sideband (VSB) is an amplitude modulation technique in which a portion of the signal is called a vestige and is modulated with one sideband.

NOTES

Both bands are unnecessary for transmission because they are a waste of time. However, if only one band is transferred, the data will be lost. As a result, this strategy has evolved. The following diagram depicts a VSB signal.

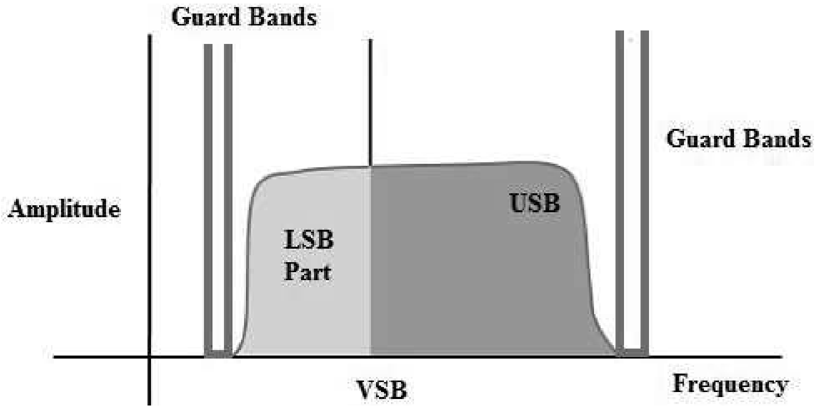


Fig. 4.24 *VSB Modulation*

In addition to the higher sideband, a small portion of the lower sideband is broadcast in this approach.

To prevent invasions, a very little protector band is placed on either side of the vestigial sideband. This modulation is commonly used in television broadcasts. The transmission BW of a VSB modulated wave is equal to the sum of message bandwidth and VSB size.

Advantages

The advantages of VSB modulation include the following.

- The reduction in BW is the key benefit of this modulation. It is similar to SSB in terms of efficiency.
- Extremely effective
- When great accuracy is not required, the filter design is simple.
- The filter limitations will be alleviated due to the transmission permission of a lower sideband section, allowing for the transmission of low-frequency components as well as good phase characteristics without difficulty.
- Practical filters are used to suppress incomplete LSB.The main benefit of this modulation is the decrease in BW. It is approximately efficient like SSB

Disadvantages

The disadvantages of VSB modulation include the following.

- When compared to a single-sided band, the bandwidth is greater (SSB).
- Demodulation is challenging.

VSB Modulation Applications

The following are some of the applications of VSB modulation.

- VSB modulation is the industry standard for TV signal transmission. Because video signals demand a large transmission BW, techniques such as DSB-

NOTES

FC or DSF-SC are used.

- This is a form of amplitude modulation that is mostly utilised for television broadcasting around the world. It is critical to broadcast visual and audio information simultaneously in this broadcast.
- The higher sideband of the video signal and the image carrier are aired without any control during VSB transmission. When a vestige is a fraction of the lower sideband, it is transmitted and the remainder is hidden.
- When considering the use of BW, this is the most appropriate and effective strategy.

This concludes our discussion of VSB modulation. The most common and important application of this modulation is the transmission of television signals.

4.4.5 Frequency Division Multiplexing

FDM (Frequency Division Multiplexing) is a multiplexing technology that involves mixing numerous signals over a shared channel. Signals of multiple frequencies are merged for simultaneous transmission in FDM.

Concept and Process

The complete bandwidth is partitioned into a collection of non-overlapping frequency bands in FDM. One of the sending devices generates and modulates a separate signal on each of these bands. To prevent signal overlapping, the frequency bands are divided by guard bands, which are strips of unused frequencies.

A multiplexer (MUX) is used in the transmitting end to combine the modulated signals. The combined signal is sent through the communication channel, allowing for the transmission of numerous independent data streams at the same time. Demultiplexing is a procedure that extracts individual signals from a mixed signal at the receiving end (DEMUX).

Example

Multiplexing with FDM is conceptualised in the diagram below. It contains four frequency bands, each capable of carrying a signal from a single sender to a single receiver. A frequency band is assigned to each of the four senders. The communication channel multiplexes the four frequency bands and sends them. A demultiplexer on the receiving end regenerates the original four signals as outputs.

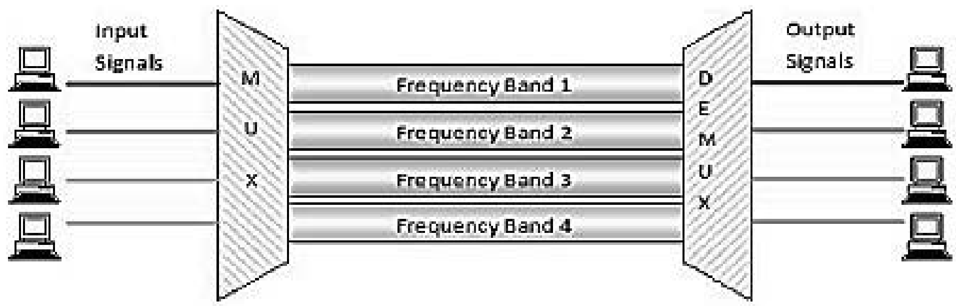


Fig. 4.25 *Frequency Band*

If the frequency bands have a bandwidth of 150 KHz and are separated by 10 KHz guard bands, the communication channel's capacity should be at least 630 KHz (channels: 150 4 + guard bands: 10 3).

NOTES

Uses and Applications

It permits several independent signals created by numerous users to use a single transmission medium, such as a copper or fibre optic cable.

In telephone networks, FDM is commonly used to multiplex calls. It's also suitable for usage in cellular and wireless networks, as well as satellite communications.

Orthogonal Frequency Division Multiplexing

OFDM is a method of signal transmission in which the channel bandwidth is divided into many densely packed sub-carriers or narrowband channels, each of which transmits signals independently using techniques such as QAM (Quadrature Amplitude Modulation). As a result, they don't need guard bands and can make greater use of the available bandwidth.

Check Your Progress

8. What do you mean by the control grid modulation?
9. Define demodulation of AM wave's generation of DSB-SC waves.
10. State the balanced modulator method.
11. What is coherent detector?
12. How will you define the filter method in SSB techniques?
13. What is SSB Modulation?
14. Define the term VSB modulation.
15. How will you define the Frequency Division Multiplexing (FDM)?

4.5 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Modulation plays a key role in communication systems since it is employed for encoding information digitally in the analog world. It is very important to modulate signals prior to transmitting them to the receiver section for larger distance transfer, accurate data transfer and low-noise data reception.
2. Digital modulation technique provides efficient and clear communication. Its key advantages over analog modulation are high noise immunity, available bandwidth and permissible power. During digital modulation, there is the conversion of a message signal from analog to digital message, which is then modulated with the help of a carrier wave.
3. Mathematically, it is possible to approximate the baseband modulated signal by a sinusoidal continuous wave signal with a frequency fm. This method is referred to as single-tone modulation.

NOTES

4. There are two ways to classify frequency modulation:
   - Narrowband: When the change in the carrier frequency is about the same as the signal frequency.
   - Wideband: When the change in the carrier frequency is much higher (modulationindex > 1) than the signal frequency.
5. When the carrier is modulated by a single sine wave, it is possible to use Bessel function of the first kind to calculate the resulting frequency spectrum, as a function of the modulation index and the sideband number.
6. Applications of demodulation are:
   - Magnetic tape
   - Sound
   - Radio
7. The FM wave that has a small bandwidth is referred to as narrow band FM. The Narrow band FM has a modulation index (mf) which is small in comparison to one radian. Therefore, a narrow band FM's spectrum comprises the carrier and the upper and lower sidebands.
8. It is possible to control the operating bias and gain of the final RF amplifier through varying of the voltage of the control grid. In this method, very little audio power is needed though it is essential to take care to lessen distortion.
9. Double-Sideband Suppressed-Carrier Transmission (DSB-SC) is the type of transmission where the frequencies that are produced by AM are spaced symmetrically below and above the carrier frequency. At the same time, the carrier level is reduced to the lowest practical level, with ideally it being suppressed completely.
10. A method, referred to as the Armstrong method, was patented by Edwin H. Armstrong in 1933 to generate frequency modulation of radio signals. It generates a double sideband suppressed carrier signal, phase shifts this signal, and then reinserts the carrier to produce a frequency modulated signal.
11. The message signal is detected using the same carrier signal that is utilised to generate the DSB-SC signal. As a result, this detection method is known as coherent or synchronous detection.
12. To use the filter method to generate an SSB modulated wave, it is essential for the message signal to satisfy some conditions, which are:
    - There must not be any low frequency content in the message signal.
    - In the spectrum of the message signal, the highest frequency should be much smaller than the carrier frequency.
13. A single-sideband modulation is a type of modulation that is used to send data over radio waves, such as an audio transmission. This modulation is used in radio communications to change the AM signal by utilising additional transmitter power and bandwidth (amplitude modulation). Many single sideband radio communication devices, such as SSB Tx, SSB Rx, and SSB transceiver, are available on the market.

14. Vestigial Sideband (VSB) is an amplitude modulation technique in which a portion of the signal is called a vestige and is modulated with one sideband.
15. FDM (Frequency Division Multiplexing) is a multiplexing technology that involves mixing numerous signals over a shared channel. Signals of multiple frequencies are merged for simultaneous transmission in FDM.

4.6 SUMMARY

- Modulation plays a key role in communication systems since it is employed for encoding information digitally in the analog world. It is very important to modulate signals prior to transmitting them to the receiver section for larger distance transfer, accurate data transfer and low-noise data reception.
- Amplitude modulation is a form of modulation in which there is modulation or changing of the carrier signal's amplitude proportionate to the message signal. In this form of modulation, no change is made to the phase or to the frequency.
- Frequency modulation is a form of modulation in which there is the modulation of the carrier signal's frequency proportionate to the message signal. This type of modulation does not in any way modify the phase or the amplitude.
- Phase modulation is a form of modulation in which there is the modulation of the carrier signal's phase in accordance with the low frequency of the message signal. Let us understand all of them in detail.
- Digital modulation technique provides efficient and clear communication. Its key advantages over analog modulation are high noise immunity, available bandwidth and permissible power. During digital modulation, there is the conversion of a message signal from analog to digital message, which is then modulated with the help of a carrier wave.
- The baseband modulated signal by a sinusoidal continuous wave signal with a frequency fm. This method is referred to as single-tone modulation.
- When the carrier is modulated by a single sine wave, it is possible to use Bessel function of the first kind to calculate the resulting frequency spectrum, as a function of the modulation index and the sideband number.
- The FM wave that has a small bandwidth is referred to as narrow band FM. The Narrow band FM has a Modulation Index (MF) which is small in comparison to one radian. Therefore, a narrow band FM's spectrum comprises the carrier and the upper and lower sidebands.
- An ideal FM discriminator comprises of slope circuits (differentiator) which is followed by an envelope detector. A limiter is inserted before the differentiator to ensure a constant amplitude of the input signal.
- The RF amplifier's plate voltage is modulated with the audio signal. The requirement audio power is just 50 percent of the RF-carrier power.
- Control grid modulation is possible to control the operating bias and gain of the final RF amplifier through varying of the voltage of the control grid. In this method, very little audio power is needed though it is essential to take care to lessen distortion.

NOTES

- Pulse duration modulation is the application of an extremely efficient high voltage power supply which is applied to the tube plate. This supply's output voltage is varied at an audio rate to follow the program.
- Balanced modulator method referred to as the Armstrong method, was patented by Edwin H. Armstrong in 1933 to generate frequency modulation of radio signals. It generates a double sideband suppressed carrier signal, phase shifts this signal, and then reinserts the carrier to produce a frequency modulated signal.
- The message signal is detected using the same carrier signal that is utilised to generate the DSB-SC signal. As a result, this detection method is known as coherent or synchronous detection.
- A single-sideband modulation is a type of modulation that is used to send data over radio waves, such as an audio transmission. This modulation is used in radio communications to change the AM signal by utilising additional transmitter power and bandwidth (amplitude modulation). Many single sideband radio communication devices, such as SSB Tx, SSB Rx, and SSB transceiver, are available on the market.
- Vestigial Sideband (VSB) is an amplitude modulation technique in which a portion of the signal is called a vestige and is modulated with one sideband.
- FDM (Frequency Division Multiplexing) is a multiplexing technology that involves mixing numerous signals over a shared channel. Signals of multiple frequencies are merged for simultaneous transmission in FDM.

4.7 KEY TERMS

- **Modulation:** Modulation plays a key role in communication systems since it is employed for encoding information digitally in the analog world. It is very important to modulate signals prior to transmitting them to the receiver section for larger distance transfer, accurate data transfer and low-noise data reception.
- **Digital modulation:** Digital modulation technique provides efficient and clear communication. Its key advantages over analog modulation are high noise immunity, available bandwidth and permissible power. During digital modulation, there is the conversion of a message signal from analog to digital message, which is then modulated with the help of a carrier wave.
- **Bessel function:** When the carrier is modulated by a single sine wave, it is possible to use Bessel function of the first kind to calculate the resulting frequency spectrum, as a function of the modulation index and the sideband number.
- **Plate modulation:** The RF amplifier's plate voltage is modulated with the audio signal. The requirement audio power is just 50 percent of the RF-carrier power.
- **Balanced modulator:** Balanced modulator method referred to as the Armstrong method, was patented by Edwin H. Armstrong in 1933 to generate frequency modulation of radio signals. It generates a double

sideband suppressed carrier signal, phase shifts this signal, and then reinserts the carrier to produce a frequency modulated signal.

- **Coherent detector:** The message signal is detected using the same carrier signal that is utilised to generate the DSB-SC signal. As a result, this detection method is known as coherent or synchronous detection.
- **SSB modulation:** A single-sideband modulation is a type of modulation that is used to send data over radio waves, such as an audio transmission. This modulation is used in radio communications to change the AM signal by utilising additional transmitter power and bandwidth (amplitude modulation). Many single sideband radio communication devices, such as SSB Tx, SSB Rx, and SSB transceiver, are available on the market.
- **VSB modulation:** Vestigial Sideband (VSB) is an amplitude modulation technique in which a portion of the signal is called a vestige and is modulated with one sideband.
- **Frequency Division Multiplexing (FDM):** FDM (Frequency Division Multiplexing) is a multiplexing technology that involves mixing numerous signals over a shared channel. Signals of multiple frequencies are merged for simultaneous transmission in FDM.

NOTES

4.12 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. What do you mean by the modulation?
2. What is the significance of modulation?
3. Define modulation index.
4. State the wide band FM/broadband FM.
5. Define directed and indirected method.
6. How will you define the generation of DSB-SC modulation of AM wave?
7. What do you understand by the coherent detector?
8. Give the techniques for generating the SSB waves.
9. Define SSB modulation.
10. What is VSB modulation?
11. State the Frequency Division Multiplexing (FDM).

Long-Answer Questions

1. Explain briefly about the types of modulation with the help of giving examples.
2. Discuss analog and digital modulation in detail. Give appropriate examples.
3. Differentiate between the narrow band FM and wide band FM/broadband FM. Give appropriate examples.
4. Briefly explain about the FM signal generation can be performed with the help of giving examples.

5. Explain Double-Sideband Suppressed-Carrier transmission (DSB-SC) with balanced modulator.
6. Discuss about the types for vacuum tube transmitters and even transistors come with similar options. Give appropriate examples.
7. What do you mean by the coherent detector? Explain with the help of coherent detector's block diagram.
8. Explain about the techniques for generating the SSB waves with the help of giving examples.
9. What do you understand by the SSB modulation? Give the advantages and disadvantages of SSB modulation.
10. Briefly explain about the VSB modulation. Give appropriate examples.
11. Illustrate the Frequency Division Multiplexing (FDM) with the help of giving examples.

NOTES

4.13 FURTHER READING

Basavaraj, B. and H.N. Shivashankar. 2004. *Basic Electronics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Black, Uyless D. 1993. *Computer Networks.* New Jersey: Prentice-Hall.

Kumar, A. Anand. 2003. *Fundamentals of Digital Circuits*. New Delhi: Prentice-Hall of India.

Mano, Morris. 1979. *Digital Logic and Computer Design*. New Delhi: Prentice-Hall of India.

Roth, Charles. 2001. *Fundamentals of Logic Design*. Thomson Learning.

Shepherd, Robert D. 1998. *Introduction to Computers and Technology*. EMC/ Paradigm Publishing.

Stallings, William. 1996. *Data and Computer Communications.* New Jersey: Prentice-Hall.

Tanenbaum, Andrew. 2002. *Computer Networks.* New Jersey: Prentice-Hall PTR.

Yarbarough, John M. 1996. *Digital Logic Applications and Design*. Thomson Learning.

Carlson Bruce and Paul B. Crilly. 2010. Communication System. Noida: Tata McGraw-Hill Education.

Lathi B. P. and Zhi Ding. 2010. Modern Digital and Analog Communication Systems. Oxford: Oxford University Press.

Kennedy George and Bernard Davis. 1993. Electronic Communication Systems. London: Glencoe/ McGraw-Hill.

Roddy Dennis and John Coolen. 1977. Electronic Communications. Noida: Pearson Education.

UNIT 5 ELECTRONIC DEVICES AND PHOTONIC DEVICES

NOTES

Structure

5.0 INTRODUCTION

The Junction-Gate Field-Effect Transistor (JFET) is one of the simplest types of field-effect transistor. JFETs are three-terminal semiconductor devices that can be used as electronically controlled switches or resistors, or to build amplifiers. A MESFET (Metal-Semiconductor Field-Effect Transistor) is a field-effect transistor semiconductor device similar to a JFET with a Schottky (metal-semiconductor) junction instead of a p-n junction for a gate. The Metal-Oxide-Semiconductor Field-Effect Transistor (MOSFET, MOS-FET, or MOS FET), also known as the Metal-Oxide-Silicon Transistor (MOS transistor, or MOS), is a type of insulated-gate field-effect transistor that is fabricated by the controlled oxidation of a semiconductor, typically silicon. The voltage of the covered gate determines the electrical conductivity of the device; this ability to change conductivity with the amount of applied voltage can be used for amplifying or switching electronic signals. The MOSFET was invented by Mohamed M. Atalla and Dawon Kahng at Bell Labs in 1959, and first presented in 1960.

Microwaves are incredibly short waves, as the name implies. In general, RF energy ranges from DC to infrared, and it is a type of electromagnetic energy. A quick check at the various frequency ranges reveals that the Microwave frequency range includes UHF (Ultra-High Frequency) and SHF (Super High Frequencies), with wave length (λ) ranging from 1 to 100 cm. Microwaves and low-frequency radio waves both work on the same underlying basis. The phenomena can be

NOTES

explained in terms of current flow in a closed electric circuit in this case. A parametric device is one that employs a time-varying or nonlinear reactance (capacitance or inductance). Because capacitance or inductance, which is a reactive property, may be utilized to create capacitive or inductive excitation, the term parametric is derived from the phrase parametric excitation.

Excitations are the transitions of electrons from lower to higher energy states by absorbing external energy. Excitations are the transitions of electrons from lower to higher energy states by absorbing external energy. Electron hole recombination is the passage of electrons from higher energy levels to lower energy ones in semiconductors. Electrons release energy in any form during rest. A Light Dependent Resistor (also called a photoresistor or LDR) is a resistive device whose resistance varies in response to incoming electromagnetic radiation. As a result, they are photosensitive devices. Additionally, they are referred to as photoconductors, photoconductive cells, or just photocells. They are constructed using high-resistance semiconductor materials. A photodetector is a device which absorbs light and converts the optical energy to measurable electric current. Photon detectors follow the principle of photons to electrons conversion. Unlike the thermal detectors, such detectors are based on the rate of absorption of photons rather than on the rate of energy absorption.

Solar cells are photovoltaic devices that turn light into electricity. Photovoltaic devices, such as solar cells, produce voltage when exposed to light. In 1839, Alexander-Edmond Becquerel discovered the photovoltaic effect in a junction created between a platinum electrode and an electrolyte (silver chloride). The photodiode's function is comparable to that of a solar cell (photodetector). An unbiased photodiode connected to a load is used (impedance). Light-emitting diodes are just forward-biased p-n junctions that spontaneously emit light. Radiative recombination of electron-hole pairs in the depletion area causes spontaneous emission (or electroluminescence). LEDs have a large angular bandwidth and a broad spectral bandwidth (20-150 nm).

In this unit, you will learn about the JFET, MOSFET, MESFET, microwave devices, parametric devices, radiative and non-radiative transmitter, LDR, photodiode detectors, solar cells, LED and optical gain and threshold current for lasing.

5.1 OBJECTIVES

After going through this unit, you will be able to:

- Define the term JFET
- Explain about the MOSFET and MESFET
- Describe various microwave devices
- Elaborate on the parametric devices
- Understand the basic concept of radiative and non-radiative transmitter
- Describe the Light Dependent Resistor (LDR)
- Discuss about the photodiode detectors

- Explain about the solar cells
- Elaborate on Light Emitting Diode (LED)
- Illustrate the optical gain and threshold current for lasing

NOTES

5.2 JFET: STRUCTURE, WORKING AND CHARACTERISTICS

The JFET (Junction-Gate Field-Effect Transistor) can be classified into two types: (*i*) *N*-channel JFET where electrons are the majority carriers and (*ii*) *P*-channel JFET where holes are the majority carriers.

Physical Structure and Construction of *N*-Channel JFET

N-channel JFET is constructed with an uniformly dosed *N*-type silicon or gallium arsenide substrate. Metallic contacts at the two ends of the bar are ohmic contacts (terminals) called Source and Drain. *P*-type impurities are diffused on the two sides of the *N*-type bar. The current flows along the length of the bar when a voltage is applied between the two terminals. This current is due to electrons as majority carriers. The basic structure of an *N*-channel JFET is shown in Figure 5.1.

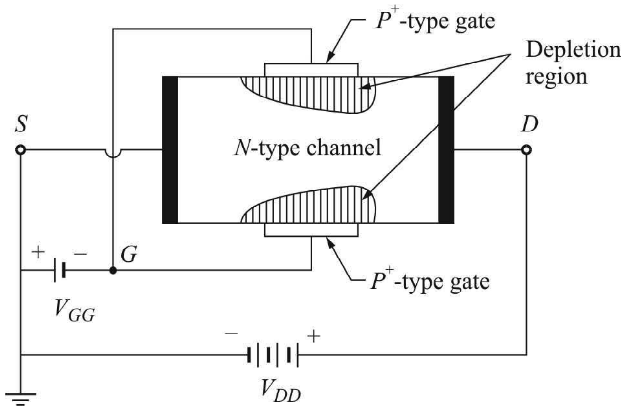


Fig. 5.1 *Schematic diagram and basic structure of an N-channel JFET*

The JFET has three terminals which are known as Source (*S*), Drain (*D*) and Gate (*G*). These terminals are defined as follows:

Source (*S*): In an *N*-channel JFET the terminal which is connected to the negative pole of the DC source providing drain voltage (V_{DD}) and through which the majority carries, i.e., electrons enter into the semiconductor bar is called Source.

Drain (*D*): The terminal of an *N*-channel JFET connected to the positive pole of the drain voltage source (V_{DD}) and through which the majority carriers, i.e., electrons leave the bar is known as Drain.

Gate (*G*): The heavily doped P^+ regions on the two sides of the *N*-type silicon bar forming *P-N* junctions which are joined together are called the gate. A voltage V_{GS} is applied between the gate and source to reverse bias the *P-N* junctions.

Channel: The region of *N*-type silicon bar between the two depletion regions of *P-N* junctions through which the majority carriers, i.e., electrons move from the source to drain due to a potential difference V_{DS} applied between the source and drain is known as Channel.

NOTES

Principle of Operation

The *P-N* junction formed by the *P*-type gate and *N*-type channel of JFET is always made reverse biased by the source V_{GG}. A depletion region devoid of mobile carriers is formed which extend more towards the lightly doped *N*-type bar than towards heavily doped P^+-type gate. The conductivity of depletion region is zero. If the gate to source voltage be zero ($V_{GS} = 0$), the depletion region is uniform and symmetrical about the junction. If V_{GS} be made negative then the depletion layer width increases with increasing reverse bias. Thus the effective channel width through with the current flows in the bar decreases with increasing reverse bias (V_{GS}). At a particular value of $-V_{GS}$, the two depletion regions about the junction would widen to such an extent that they tend to touch each other. Under this situation the effective channel width becomes zero and the channel is said to be cut off.

When V_{GS} is zero and a positive voltage is applied between drain and source, i.e., V_{DS} is positive, the electrons flow from source to drain and the conventional current (I_D) is directed from drain to source. Assuming uniform channel resistance, the voltage drop will gradually increase from source to drain end due to the flow of current (I_D). Thus the reverse bias voltage between the P^+-gate and *N*-channel will vary at different points along the channel from a minimum at source end to a maximum at drain end. The depletion region width increases with increasing reverse bias voltage. Therefore the depletion layer width is also minimum at source end and maximum at drain end. This phenomenon makes the channel tapered (wedge shaped) since the effective channel width decreases from source end to drain end.

As V_{DS} is increased the cross-sectional area of the channel will decrease. At a certain value of V_{DS} called V_P, the channel width becomes nearly zero. The channel is said to be pinched off and V_P is so called pinch off voltage.

The channel acts as a resistor of resistance *R* given by

$$R = \frac{\rho l}{A} \qquad ...(5.1)$$

where ρ is the resistivity, *l* is the length and *A* is the area of the channel. The drain current is therefore given by

$$I_D = \frac{V_{DS}}{R} = \frac{AV_{DS}}{\rho l} \qquad ...(5.2)$$

Circuit Symbol, Diagram and Current-Voltage Characteristics of JFET

The circuit symbols of an *N*-channel and a *P*-channel JFET are shown in Figure 5.2 (*a*) and (*b*) respectively. The circuit diagram of an *N*-channel JFET is shown in Figure 5.3 which is used to experimentally determine the current-voltage characteristics of JFET.

NOTES

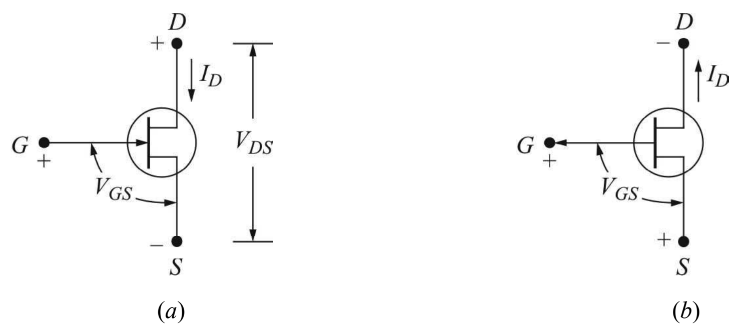


Fig. 5.2 *Circuit symbol for (a) N-channel JFET and (b) P-channel JFET*

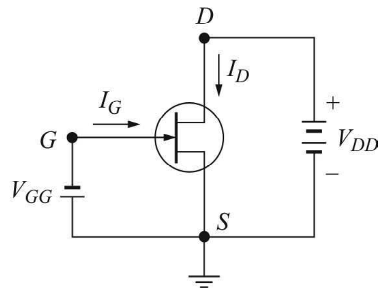


Fig. 5.3 *Circuit diagram of an N-channel JFET with gate and drain supply voltages*

The arrowheads on the gate terminals in the circuit symbols of Figure 5.2 (*a*) and (*b*) indicate the direction of gate current when the gate-source junction is forward biased. The gate-source junction is reverse biased for normal operation of JFET. The polarities of voltage sources V_{GG} and V_{DD} for *N*-channel JFET shown in Figure 5.3 indicate that V_{GS} is negative while V_{DS} is positive. The circuit diagram of *P*-channel JFET will be similar to that of *N*-channel JFET with the reversal of the polarities of V_{GG} and V_{DD}.

Electrons drift from source to drain and current flows from drain to source in an *N*-channel JFET. For a *P*-channel JFET holes drift from source to drain and current flows from source to drain.

JFET Characteristics

The current-voltage characteristics of JFET are of two types:

(*a*) **Drain Characteristics:** The plots of drain current (I_D) versus drain-to-source voltage (V_{DS}) for different values of gate-to-source voltage are known as drain characteristics. These characteristics can be experimentally determined.

(*b*) **Transfer Characteristics:** The plots of drain current (I_D) versus gate-to-source voltage (V_{GS}) for different values of drain-to-source voltage (V_{DS}) are known as transfer characteristics.

The drain and transfer characteristics of *N*-channel JFET in the common source configuration may be determined by using the circuit diagram shown in Figure 5.4.

NOTES

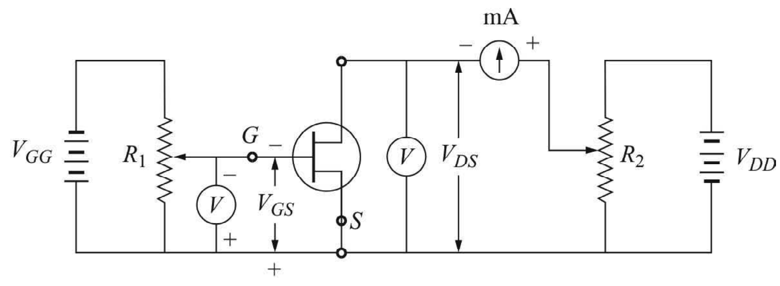


Fig. 5.4 *Circuit diagram to determine N-channel JFET characteristics*

The potentiometers R_1 and R_2 are used to vary the voltages V_{GS} and V_{DS} respectively. These voltages are measured by the voltmeters shown in Figure 5.5. The drain current I_D can be measured by the milliammeter connected in series with the JFET and the supply voltage V_{DD}.

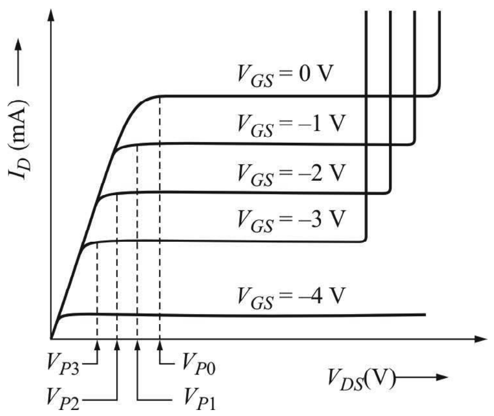


Fig. 5.5 *Drain characteristics of an N-channel JFET*

Drain Characteristics

The drain characteristics of an *N*-channel JFET shown in Figure 5.5 exhibit four different regions. These regions are separately shown in Figure 5.6 for V_{GS} = 0 V. The regions are

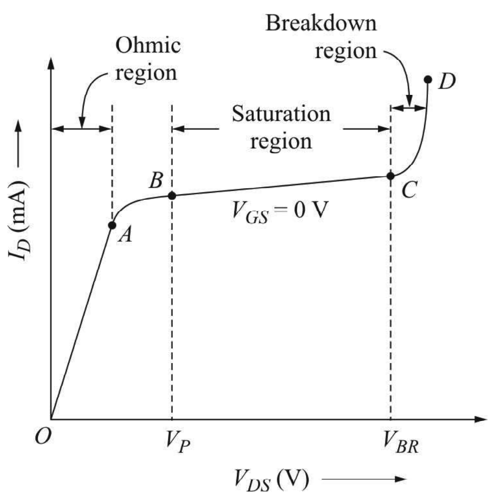


Fig. 5.6 *Drain characteristics of an N-channel JFET with V_{GS} = 0 V*

NOTES

1. **Ohmic Region:** In the ohmic region *OA* shown in Figure 5.6, the drain current increases linearly with the increase of V_{DS}, obeying ohm's law. The JFET acts like a voltage variable resistor in this region.
2. **The Region *AB*:** The curve *AB* is nonlinear and in this region the drain current increases at a slower rate than that in ohmic region. The increase of drain current with V_{DS} in this region follows an inverse square law rate. With the increase of V_{DS}, the reverse bias at the gate-source junction increases. Thus the depletion layer width increases and effective channel width decreases. The channel width is narrowest corresponding to the point *B* of the drain characteristics. This phenomenon is known as pinch-off and the value of V_{DS} at which this phenomenon occurs is known pinch-off voltage (V_P).
3. The region *BC* in the curve of shown in Figure 5.6 is known as pinch-off or saturation region where the drain current I_D remains constant with the increase of V_{DS}. This region is normally used for the operation of JFET as amplifier.
4. When V_{DS} is increased beyond *C*, I_D increases sharply. The region *CD* is called breakdown region. The rapid increase of I_D in this region is due to the breakdown of reverse biased gate source junction caused by impact ionization and avalanche multiplication of carriers. The drain to source voltage at which this breakdown takes places is denoted by V_{BR} called breakdown voltage.

The effect of increasing gate-to-source voltage (V_{GS}) on the drain characteristics is observed from Figure 5.5. With the increase of negative value of V_{GS}, the pinch-off voltage decreases ($V_{P3} < V_{P2} < V_{P1} < V_{P0}$) as well as the drain current (I_D) decreases. The decrease of V_P with the increase of negative value of V_{GS} can be explained in the following way. As negative V_{GS} increases the reverse bias across the gate-source junction increases and a smaller value of V_{DS} will widen the depletion region across the junction to pinch off the drain current.

Transfer Characteristics

The plot of drain current (I_D) versus gate-to-source voltage (V_{GS}) for constant drain-to-source voltage (V_{DS}) is known as transfer or transconductance characteristics shown in Figure 5.7. The circuit diagram of shown in Figure 5.4 may be used to determine experimentally the transfer characteristics. V_{DS} is here maintained constant at a suitable value greater than V_P. The gate-to-source voltage (V_{GS}) is decreased from zero till I_D decreases to zero. This value of V_{GS} is pinch off voltage, i.e., $V_{GS\,(\text{off})} = V_P$. The maximum drain current at $V_{GS} = 0$ V is I_{DSS}. The transfer characteristics is very nearly a parabola and can be analytically expressed by the following relation:

$$I_{DS} = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2 \qquad ..(5.3)$$

where I_{DSS} is the saturation drain current.

NOTES

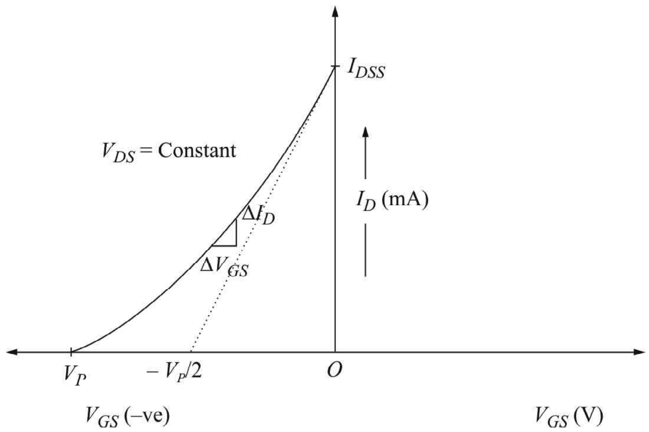


***Fig. 5.7** Transfer characteristics of N-channel JFET*

Characteristic Parameters of JFET

The drain current (I_D) is a function of both V_{DS} and V_{GS}. The following characteristic parameters describe the electrical behaviour of JFET. These are:

1. **Transconductance (g_m):** The ratio of small change in the drain current to the corresponding small change in gate-to-source voltage for constant drain-to-source voltage is called transconductance, i.e.,

$$g_m = \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}} = \left(\frac{\Delta I_D}{\Delta V_{GS}}\right)_{V_{DS} = \text{constant}} \quad ...(5.4)$$

The unit of g_m is mho. The value of g_m can be determined from the slope of transfer characteristics.

2. **Drain Resistance, r_d:** It is the ratio of small change in drain-to-source voltage to the corresponding change in drain current at constant gate-to-source voltage, i.e.,

$$r_d = \left(\frac{\partial V_{DS}}{\partial I_D}\right)_{V_{GS}} = \left(\frac{\Delta V_{DS}}{\Delta I_D}\right)_{V_{GS} = \text{constant}} \quad ...(5.5)$$

The value of r_d is obtained from the reciprocal of the slope of drain characteristics. It is also called AC or dynamic resistance. The unit of r_d is ohm.

3. **DC Drain Resistance (R_D):** The ratio of drain-to-source voltage to drain current at a fixed value of gate-to-source voltage is known as DC drain resistance.

This is given by

$$R_D = \left(\frac{V_{DS}}{I_D}\right)_{V_{GS} = \text{constant}} \quad ...(5.6)$$

It is static or ohmic resistance of the channel.

4. **Drain-Source on Resistance $(R_{DS})_{ON}$:** The drain resistance at $V_{GS} = 0$, i.e., when the depletion regions of the channel are absent is known as drain-source on resistance.

5. **Drain Conductance (g_d):** The reciprocal of r_d is called drain conductance.

NOTES

In other words, drain conductance (g_d) is defined to be the ratio of small change in drain current to the corresponding change in drain-to-source voltage at constant gate-to-source voltage, i.e.,

$$g_d = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} = \left(\frac{\Delta I_D}{\Delta V_{DS}}\right)_{V_{GS} = \text{constant}} \quad ...(5.7)$$

6. **Amplification Factor (μ):** It is the ratio of small change in the drain voltage to the corresponding change in the gate voltage at a constant drain current. It is mathematically expressed as

$$\mu = -\left(\frac{\partial V_{DS}}{\partial V_{GS}}\right)_{I_D} = -\left(\frac{\Delta V_{DS}}{\Delta V_{GS}}\right)_{I_D = \text{constant}} \quad ...(5.8)$$

The negative sign shows that when V_{GS} increases, V_{DS} must decrease to maintain I_D constant.

Relationship among μ, r_d and g_m: I_D depends on V_{DS} and V_{GS}. The functional relationship is given by

$$I_D = f(V_{DS}, V_{GS}) \quad ...(5.9)$$

If V_{DS} is changed by a small value from V_{DS} to ($V_{DS} + \Delta V_{DS}$) and V_{GS} is changed from V_{GS} to ($V_{GS} + \Delta V_{GS}$) then the corresponding small change in I_D is obtained from Equation (5.9) by applying Taylor's series expansion and neglecting higher order terms

$$\therefore \quad \Delta I_D = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} \Delta V_{DS} + \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{GS}} \Delta V_{GS}$$

Dividing both sides of the above equation by ΔV_{GS},

$$\frac{\Delta I_D}{\Delta V_{GS}} = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} \left(\frac{\partial V_{DS}}{\partial V_{GS}}\right) + \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}}$$

Taking I_D to be constant,

$$0 = \left(\frac{\partial I_D}{\partial V_{DS}}\right)_{V_{GS}} \left(\frac{\partial V_{DS}}{\partial V_{GS}}\right) + \left(\frac{\partial I_D}{\partial V_{GS}}\right)_{V_{DS}}$$

or,
$$0 = \frac{1}{r_d}(-\mu) + g_m$$

$$\therefore \quad \mu = r_d\, g_m \quad ...(5.10)$$

Thus the amplification factor is the product of the drain resistance and transconductance.

Expression for Saturation Drain Current

The transfer characteristic is represented approximately by a parabolic relation given in Equation (5.3) i.e.,

$$I_{DS} = I_{DSS}\left(1 - \frac{V_{GS}}{V_P}\right)^2$$

Differentiating the above equation with respect to V_{GS}, we obtain

$$\frac{\partial I_{DS}}{\partial V_{GS}} = I_{DSS} \times 2\left(1 - \frac{V_{GS}}{V_P}\right)\left(-\frac{1}{V_P}\right)$$

NOTES

$$\therefore \quad g_m = -\frac{2I_{DSS}}{V_P}\left(1 - \frac{V_{GS}}{V_P}\right) \quad \text{...(5.11)}$$

From Equation (5.3), $\left(1 - \frac{V_{GS}}{V_P}\right)$

$$= \sqrt{\frac{I_{DS}}{I_{DSS}}} \quad \text{...(5.12)}$$

Substituting in Equation (5.11), we get,

$$g_m = -\frac{2\sqrt{I_{DS}I_{DSS}}}{V_P} \quad \text{...(5.13)}$$

when $V_{GS} = 0$, let $g_m = g_{mo}$, then from equation (5.11)

$$g_{mo} = -\frac{2I_{DSS}}{V_P} \quad \text{...(5.14)}$$

Therefore from Equations (5.13) and (5.11)

$$g_m = g_{mo}\left(1 - \frac{V_{GS}}{V_P}\right) \quad \text{...(5.15)}$$

Thus the mutual conductance or transconductance varies as the square root of the saturation drain current. Also g_m decreases linearly with the increase of V_{GS}.

Slope of the Transfer Characteristics at I_{DSS}: From Equation (5.13),

$$g_m = \frac{-2\sqrt{I_{DS}I_{DSS}}}{V_P}$$

or,

$$\frac{\partial I_{DS}}{\partial V_{GS}} = \frac{-2\sqrt{I_{DS}I_{DSS}}}{V_P}$$

When $I_{DS} = I_{DSS}$,

$$\frac{\partial I_{DS}}{\partial V_{GS}} = -\frac{2I_{DSS}}{V_P} = \frac{I_{DSS}}{-V_P/2}$$

Thus the slope of transfer curve at $V_{GS} = 0$, i.e., $I_{DS} = I_{DSS}$ is given by the above equation. The tangent to the curve at $V_G = 0$ has an intercept of $-\frac{V_P}{2}$ on the V_{GS} axis as shown in Fig. 5.7.

The gate-to-source cut-off voltage, $V_{GS\,(\text{off})}$ is equal to the pinch-off voltage V_P in the drain characteristics, i.e., $V_P = |V_{GS\,(\text{off})}|$

$$\therefore \quad I_D = I_{DSS}\left[1 - \frac{V_{GS}}{V_{GS\,(\text{off})}}\right]^2 \quad \text{...(5.16)}$$

More about Pinch-off Region

It is observed from the drain characteristics of an *N*-channel JFET shown in Figure 5.5 that the drain current gets saturated when the drain-to-source voltage

(V_{DS}) exceeds the pinch-off voltage (V_P). The depletion layer is narrower at source end and becomes wider as one moves towards the drain end. The channel width is therefore constricted at the drain end as shown in Figure 5.8. The channel width is $2b$ and channel dimension, perpendicular to z-direction is W.

NOTES

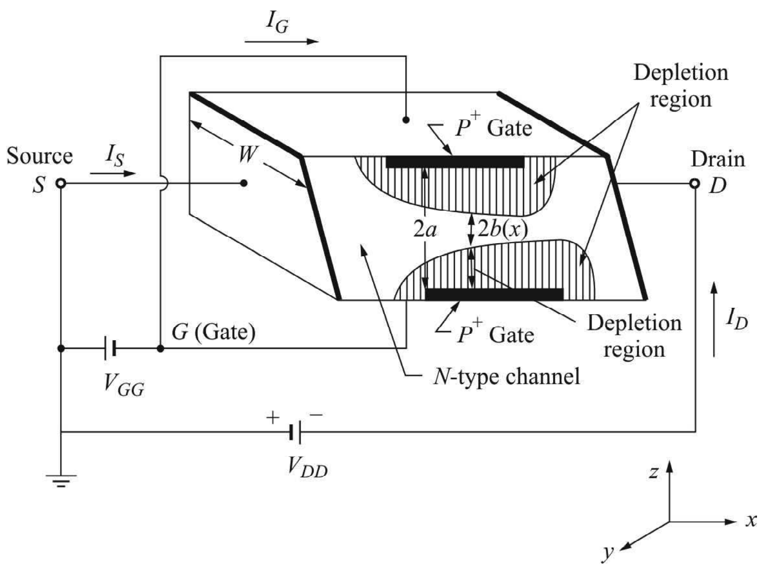


Fig. 5.8 *Three dimensional view of the basic structure of N-channel JFET*

Let us consider the situation where an electric field $E(x)$ is established along x-axis due to the application of V_{DS}. If a substantial drain current I_D flows, the drain end of the gate is more reverse biased than the source end. The depletion layer boundaries are not parallel to the centre of the channel but converge towards the drain end. It is assumed that the convergence of depletion region is gradual. Therefore one dimensional analysis will be valid. Considering the gradual channel approximation the drain current is given by

$$I_D = 2b(x)W\ eN_D\ \mu_n\ E(x) \qquad ...(5.17)$$

where the channel area at a distance x from the source is $A = 2b(x)W$, N_D is the donor concentration of N-type channel, and μ_n is the mobility of electrons.

As V_{DS} increases beyond pinch-off voltage, the constriction of the channel remains constant. The constriction of the channel is less towards the source allowing a constant current to flow in the channel. The saturation of drain current beyond V_P can be explained as follows by using Equation (5.17).

As V_{DS} increases, $E(x)$ and I_D increase whereas $b(x)$ decreases because the channel gets constricted. Hence the current density $J = \frac{I_D}{A} = \frac{I_D}{2b(x)W}$ increases. The complete pinch-off or total constriction of the channel is not possible. The channel width under complete pinch off is zero, i.e., $b = 0$ and so J would be infinite which is not physically possible. If J is allowed to increase without limit, $E(x)$ would follow the same provided μ_n is constant. Mobility is a function of electric field and remains constant for $E(x) < 10^3$ V/cm in N-type silicon. For moderate electric field in the range of 10^3 to 10^4 V/cm, μ_n is approximately inversely proportional to the square root of electric field. When the electric field at pinch off is more than 10^4 V/cm, μ_n is inversely proportional to $E(x)$. The drift velocity of the electrons $v_d = \mu_n E(x)$ remains constant and ohm's law breaks down. Therefore both I_D and b remain constant which explains the constant or saturation drain current at pinch off.

NOTES

5.2.1 MOSFET and MESFET: Structure, Working and Characteristics

In a junction field effect transistor, the gate is not isolated from the channel but forms a p^+n or n^+p junction with the channel. The insulated gates are observed in MOSFETs (Metal Oxide Semiconductor Field Effect Transistor) where the gate is isolated from the channel by an insulating layer. The different types of MOSFETs or Gate Isolation Type FETs will be described in this section.

Metal Oxide Semiconductor Field Effect Transistor (MOSFET)

Metal Oxide Semiconductor Field Effect Transistor (MOSFET) also known as Insulated Gate Field Effect Transistor (IGFET) is an important member of FET family. The most important advantage of MOSFETs over JFETs is their larger input impedance. The input impedance may be as high as hundreds of megohms which is due to the fact that the metal contact of the gate is separated from the semiconductor channel by silicon dioxide (SiO_2) which is an insulating layer. There are two different modes of MOSFET which are classified from the constructional and application points of view as (*i*) Enhancement MOSFETs and (*ii*) Depletion MOSFETs. Again each of these two types of MOSFETs may be classified as *N*-channel and *P*-channel.

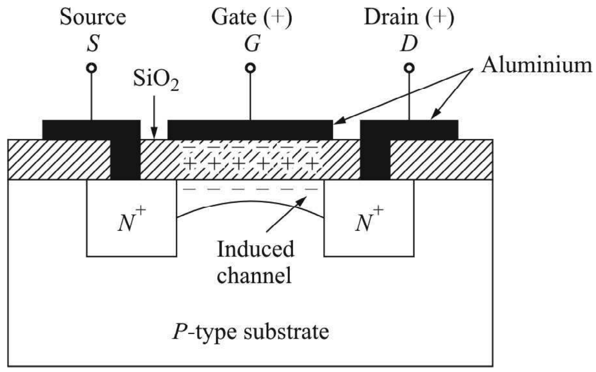
General Principle

Application of gate voltage produces a transverse electric field across the insulating silicon dioxide layer deposited on silicon substrate. The resistance and thickness of the conducting channel made of silicon can be controlled by varying the gate voltage.

The controlling transverse electric field reduces the concentration of majority carriers available for conduction in a depletion MOSFET whereas that electric field increases the majority carrier concentration in the channel of an enhancement MOSFET.

Enhancement MOSFET

The construction and structure of an *N*-channel enhancement MOSFET is shown in Figure 5.9 (*a*) and the circuit symbols of *N*-channel as well as *P*-channel enhancement MOSFET are shown in Figures 5.9 (*b*) and (*c*) respectively.



(*a*)

(*b*) (*c*)

Fig. 5.9 *(a) Structure of N-channel enhancement MOSFET,* (*b*) *and* (*c*) *Circuit symbols of N-channel and P-channel enhancement MOSFETs*

NOTES

The broken lines in the circuit symbols shown in Figure 6.9 (*b*) and (*c*) indicate continuous channel of enhancement MOSFETs. The structure and construction of *N*-channel enhancement MOSFET shown in Figure 1.9 (*a*) indicate that two highly doped N^+ regions are diffused in a lightly doped *P*-type silicon substrate. The terminals are metal (Al) contacts from these N^+ regions, called the Source (*S*) and Drain (*D*) terminals. They are separated by about 1 mil (25 μm). A thin insulating layer of SiO_2 of about 1000 Å thickness is grown over the surface of silicon substrate. Patterns are formed into the oxide layer by masking and photolithography. Finally source and drain ohmic contacts are made. The gate metal (Aluminium) is deposited on silicon dioxide layer to cover the whole channel region and a contact to this metal over the channel area is the gate terminal (*G*). The chip area of a MOSFET is 5 square mils or less which is only 5% of that of a BJT.

A MOS structure is formed by the gate metal along with insulating silicon dioxide layer and the semiconductor channel. The channel and gate act like two parallel plates of a capacitor separated by the dielectric (SiO_2) layer. This layer is responsible for high input impedance (10^{10} to 10^{15} Ω) for the MOSFET and also the reason for the name Insulated Gate Field Effect Transistor (IGFET) as SiO_2 is an insulating layer.

Principle of Operation of *N*-channel Enhancement MOSFET

If a positive voltage is applied between source (*S*) and drain (*D*), keeping the substrate at ground potential, no current will flow through the channel between the source and drain as the channel is not formed between two N^+ diffused islands. Now if a positive voltage is applied at the gate terminal (*G*) assuming the substrate at zero potential, the positive charges on the gate induce equal amount of negative charges on the substrate between source and drain. These negative charges of electrons which are minority carriers in the *P*-type substrate form an inversion layer between the two diffused N^+ layers. The positive gate voltage will establish a transverse electric field through the oxide which will end on induced negative charges (inversion layer) on the semiconductor surface. With the increase of positive gate voltage the amount of induced negative charge increases; thus extending the inversion layer between two N^+ diffused regions. The accumulation of electrons (negative charge) between the two N^+ diffused regions forms a conducting channel and current flows through the induced channel from the source to the drain. The magnitude of the drain current increases with the increase of positive gate voltage due to more and more accumulation of electrons in the top of *P*-type substrate as shown in Figure 5.10 (*a*). Since the drain current is enhanced by the positive gate voltage, the name 'enhancement' is appropriate for the *N*-channel MOSFET.

NOTES

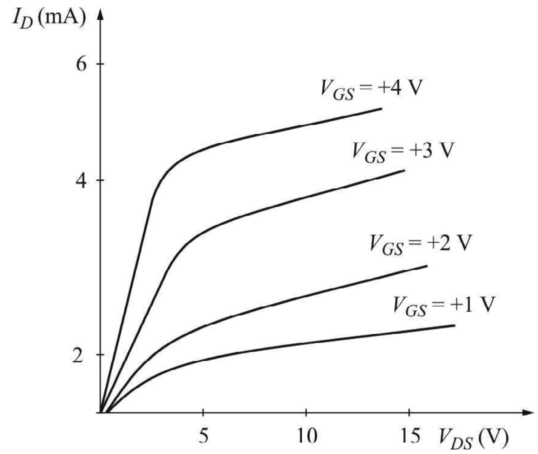
The typical drain and transfer characteristics of enhancement type *N*-channel MOSFET are shown in Figures 5.10 (*a*) and (*b*) respectively.

The transfer characteristic shown in Figures 5.10 (*b*) indicates that for $V_{GS} \leq 0$, a very low drain current of the order of a few nanoamperes flows which is designated as I_{DSS}. When V_{GS} is made positive and allowed to increase, the current I_D increases very slowly at first till V_{GS} reaches a threshold value called gate-source threshold voltage (V_{GST}) at which I_D attains a value say 10 μA. The manufacturer provides the value of V_{GST} in the specification sheet. As V_{GS} exceeds V_{GST} the drain current increases much more rapidly. The maximum value of drain current $I_{D\,(ON)}$ is shown in the transfer characteristics [Refer Figure 5.10 (*b*)] and the value of V_{GS} needed to obtain this current are given in the manufacturer's specification sheet.

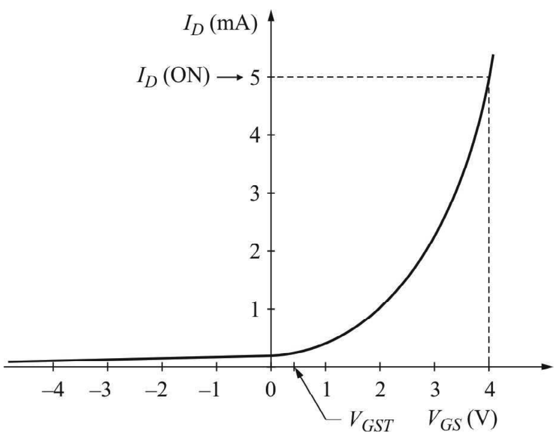
The drain current at any point in the transfer characteristics of shown in Figure 5.10 (*b*) is given by the relation

$$I_D = I_{DSS} + K[V_{GS} - V_{GST}]^2$$

where *K* is a constant whose value depends on the type of MOSFET. Its value can be determined from the specification sheet if $I_{D\,(ON)}$ and corresponding V_{GS} are substituted in the above equation.



(*a*)



(*b*)

Fig. 5.10 *(a) The drain characteristics, (b) The transfer characteristics of an N-channel enhancement MOSFET*

NOTES

The Depletion MOSFET

The structure and construction of N-channel depletion MOSFET is similar to that of an N-channel enhancement MOSFET except that a lightly doped N-channel is diffused between the source and drain regions as shown in Figure 5.11 (*a*).

An appreciable drain current I_{DSS} flows even when $V_{GS} = 0$ as the lightly doped N-layer between the source and drain acts as a channel. If the gate voltage is negative with respect to source, induced positive charges below the gate dielectric (SiO_2) deplete the doped N-layer in the channel and make the channel less conductive. The accumulation of positive charges in the channel restricts the flow of electrons between the source and drain and hence the drain current decreases as V_{GS} is made more and more negative. The drain characteristics shown in Figure 5.12 (*a*) shift downwards with increasing negative value of V_{GS}. An effective depletion of majority carriers, i.e., electrons in the channel due to negative gate voltage with respect to source takes place for which this type of MOSFET is designated as depletion MOSFET. The channel region near the drain is more depleted than that near the source due to the voltage drop caused by drain current. This situation is similar to that of pinch-off which occurs at the drain end of the channel of a JFET.

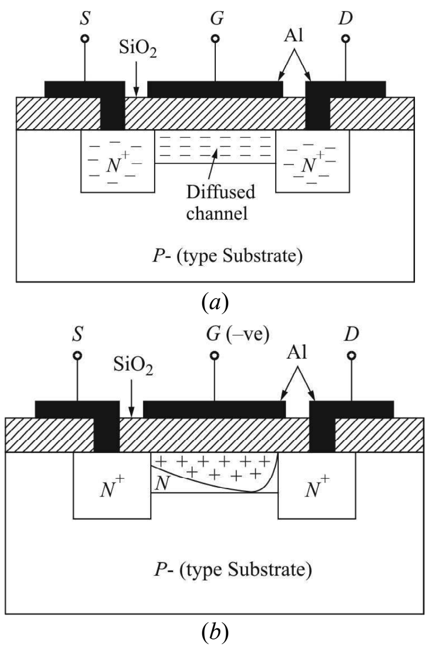


Fig. 5.11 *(a) An N-channel depletion MOSFET, (b) Channel depletion with negative gate voltage*

A depletion type MOSFET can exhibit enhancement mode if V_{GS} is made positive so that negative charges are induced in the channel through the SiO_2 of the gate capacitor. The drain current increases with the increase of positive gate voltage and the drain characteristics shown in Figure 5.12 (*a*) shift upwards. These characteristics are similar to the drain characteristics of N-channel enhancement MOSFET. Thus this type of MOSFET exhibits depletion mode for negative gate voltage and enhancement mode for positive gate voltage as indicated in shown in Figure 5.12 (*a*). Figure 5.12 (*b*) shows the transfer characteristics of both N-channel depletion and enhancement MOSFETs for negative and positive gate voltages respectively.

NOTES

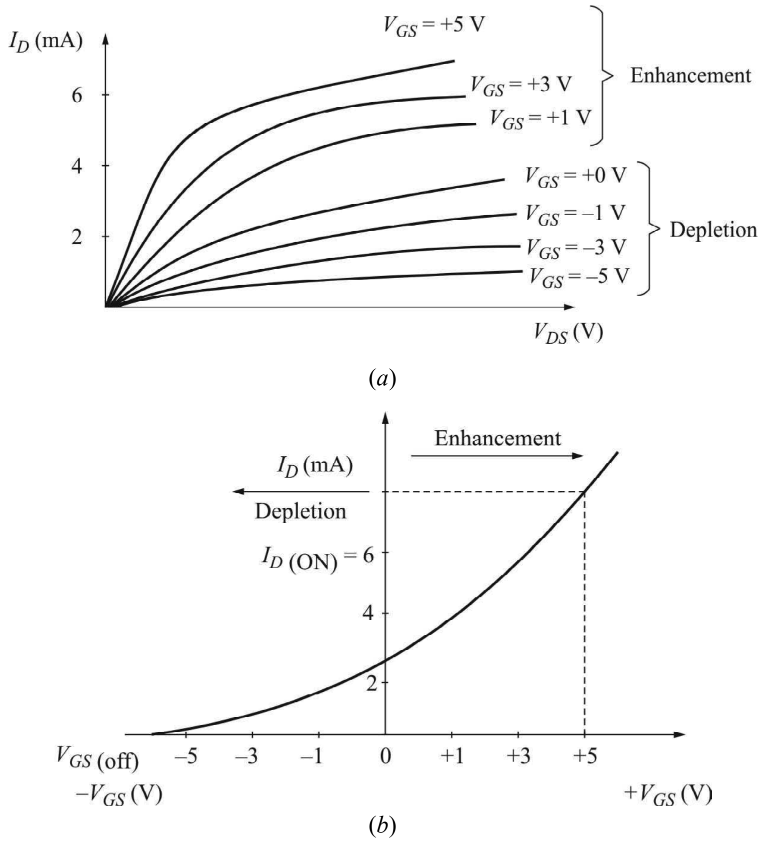


***Fig. 5.12** (a) The drain characteristics, (b) The transfer characteristics of N-channel MOSFET used in either enhancement or depletion made*

The manufacturer's specification sheet provides the gate-to-source cut off voltage $V_{GS\ (off)}$ at which I_D has some specified negligibly low value at a recommended value of V_{GS}. This value of $V_{GS\ (off)}$ corresponds to the pinch-off voltage (V_P) of a JFET.

P-Channel MOSFET

The structure of *P*-channel enhancement MOSFET consists of *N*-type silicon substrate in which two highly doped P^+ regions are diffused forming the source and drain regions. A negative gate voltage with respect to source induces positive charges to form the channel below the gate dielectric and make the channel conductive. To construct a *P*-channel depletion MOSFET, a lightly doped *P*-layer is diffused in the channel region between source and drain. Application of positive gate voltage induces negative charges below SiO_2 layer which deplete the positive charge carriers (holes) and makes the channel less conductive.

Figures 5.13 (*a*) and (*b*) show the circuit symbols for an *N*-channel depletion MOSFET while shown in Figure 5.13 (*c*) indicate that for a *P*-channel MOSFET. A thin vertical line just right to the gate (*G*) represents the channel. The drain (*D*) and source (*S*) terminals are connected to the top and bottom of the channel respectively. The arrow on the *P*-type substrate points towards the channel indicating that the channel is *N*-type. In some MOSFETs, the substrate terminal is accessible as shown in Figure 5.13 (*a*), i.e., such MOSFETs have four terminals. But in most MOSFETs, substrate is internally connected to the source. These MOSFETs have three terminals whose circuit symbols are shown in Figure

NOTES

5.13 (*b*). Figure 5.13 (*c*) shows the circuit symbol for a *P*-channel depletion MOSFET. This symbol is similar to that for an *N*-channel depletion MOSFET except that the direction of the arrow on the substrate is away from the channel.

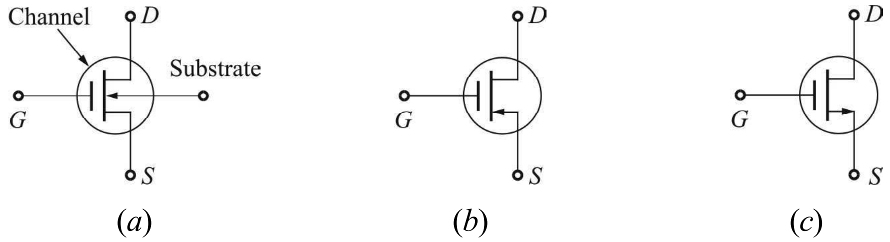


Fig. 5.13 *(a), (b) Circuit symbols for N-channel depletion MOSFETs,*
(c) Circuit symbol for P-channel depletion MOSFET

Comparison of MOSFET with JFET

1. The channel conductivity of both modes of MOSFETs is controlled by a transverse electric field applied from gate to channel through the gate dielectric (SiO_2) layer. In JFETs the conductivity of the channel is controlled by a transverse field across the reverse biased *P-N* junction.
2. The gate leakage current in a MOSFET is of the order of picoampere (10^{-12} A) while that in a JFET is of the order of nanoampere (10^{-9} A). The MOSFETs have much higher input impedance ($10^{10} - 10^{15}$ Ω) than JFETs whose input impedance is 10^8 Ω.
3. The drain characteristics of MOSFETs are less flat than those of JFETs. MOSFETs have lower drain resistance (1 to 50 kΩ) than JFETs (0.1 to 1 MΩ).
4. JFETs can operate only in the depletion mode while MOSFETs can operate in both depletion and enhancement modes.
5. Fabrication steps of MOSFETs are simpler than those of JFETs.
6. MOSFETs are sensitive to excess load voltage. These devices get damaged due to mishandling.
7. MOSFETs have zero offset voltage. The source and drain terminals can be interchanged because of the symmetrical nature of the device. Therefore MOSFETs are useful in analog signal switching.
8. MOSFETs find more useful and wider applications in digital VLSI circuits than JFETs. Special purpose digital CMOS circuits are available for low voltage and current requirement and low power dissipation.

Check Your Progress

1. Define source and drain.
2. How many terminals are there in JFET?
3. What is static characteristics of JFET?
4. Give the two different modes of MOSFET.
5. How is MOS diode constructed?
6. What do you mean by the P-channel MOSFET?

5.3 MICROWAVE DEVICES: TUNNEL DIODE, GUNN DIODE AND IMPATT DIODES

NOTES

Microwaves: Microwaves are incredibly short waves, as the name implies. In general, RF energy ranges from DC to infrared, and it is a type of electromagnetic energy. A quick check at the various frequency ranges reveals that the Microwave frequency range includes UHF (Ultra-High Frequency) and SHF (Super High Frequencies), with wave length (λ) ranging from 1 to 100 cm. Microwaves and low-frequency radio waves both work on the same underlying basis. The phenomena can be explained in terms of current flow in a closed electric circuit in this case. Low-frequency circuit elements, such as C, L, and R, are lumped together and may be easily identified and located in a circuit. Microwave circuitry, on the other hand, assumes that inductance and capacitance are spread throughout a transmission line. Microwaves are electromagnetic waves that range in frequency from 1 GHz to 1000 GHz (1 GHz = 109 Hz).

Frequency	**Band Designation**
3Hz—30 Hz	Ultra Low Frequency (ULF)
30 to 300 Hz	Extra Low Frequency (ELF)
300 to 3000 Hz (3 KHz)	Voice frequency, base band/telephony
3 KHz to 30 KHz	VLF
30 to 300 KHz	LF
300 to 3000 KHz (3 MHz)	MF
3 MHz to 30 MHz	HF
30 to 300 MHz	VHF
300 to 3000 MHz (3GHz)	Ultra High Frequency (UHF)
3 GHz to 30 GHz	SHF
30 to 300 GHz	EHF
300 to 3000 GHz (3 THz), (3 -30 THz, 30-3000 THz)	Infra-red frequencies

Advantages of Microwave

Microwaves have certain distinct benefits over low frequencies:

1. Increased Bandwidth Availability
2. Improved Directivity Properties
3. Fading Effect and Reliability
4. Power Requirements
5. Transparency Property

Fields of Application

- RADAR
- Surveillance (Air traffic control)
- Navigation (Direction finding)
- Meteorology
- Treatment of Diseases
- Microwave imaging (Surveying)
- Land Heating
- Industrial Quality Control
- Radio Astronomy
- Navigation via Global Positioning systems
- Remote Sensing
- Power Transmission

NOTES

The velocity-modulation hypothesis was used to achieve microwave production and amplification. Microwave solid-state devices such tunnel diodes, Gunn diodes, Transferred Electron Devices (TEDs), and avalanche transit-time devices have been created to execute these functionalities in the last two decades. TEDs and avalanche transit-time devices, as well as their design and subsequent development, were among the notable technical achievements. The negative resistance that can be employed for microwave oscillation and amplification is a common feature of all microwave solid state devices. TEDs and avalanche transit-time devices have made such rapid development that they are now widely recognized as one of the most important groups of microwave solid-state devices.

Microwave Devices

At microwave frequencies, solid-state microwave devices are becoming increasingly significant. These devices can be divided into four categories. Microwave Bipolar Junction Transistors (BJTs), Heterojunction Bipolar Transistors (HBTs), and tunnel diodes belong to the first group. Microwave Field-Effect Transistors (FETs) such as Junction Field-Effect Transistors (JFETs), Metal-Semiconductor Field-Effect Transistors (MESFETs), High Electron Mobility Transistors (HEMTs), Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs), metal-oxide-semiconductor transistors and memory devices, and Charge-Coupled Devices are included in the second group (CCDs). The transferred electron device is the third group, which is defined by the bulk effect of the semiconductor (TED). The Gunn diode, Limited Space charge Accumulation diode (LSA diode), Indium Phosphide diode (InP diode), and Cadmium Telluride diode are examples of these devices (CdTe diode). The Impact Ionisation Avalanche Transit-Time diodes (IMPATT diodes), Trapped Plasma Avalanche Triggered Transit-Time diodes (TRAPATT diodes), and barrier injected transit-time diodes are the devices of the fourth group, which are driven by the semiconductor’s avalanche effect (BARITT diodes). Table 2 lists all of those microwave solid-state devices.

NOTES

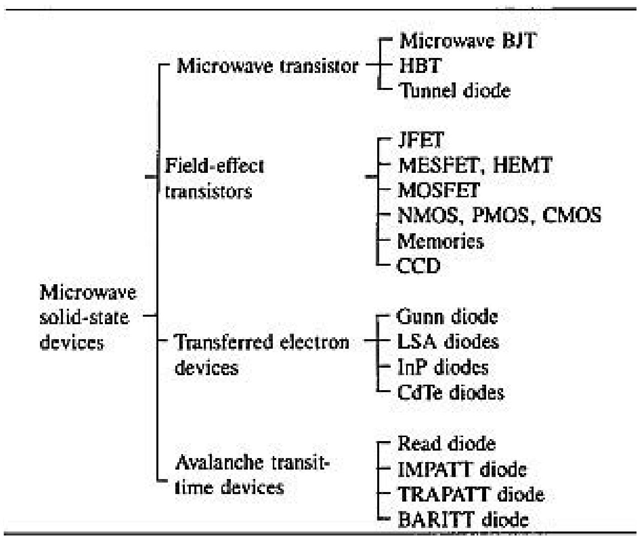


Fig. 5.14 *Microwave Solid State Devices*

Microwave Tunnel Diodes

Tunnel diodes' potential for microwave applications was soon proven after Esaki's famous work on tunnel diodes was published in 1958. Many scientists had noticed unusual properties of some p-n junctions prior to 1958, but the anomalies were quickly dismissed since they did not match the 'Classic' diode equation. Esaki, on the other hand, used quantum tunnelling theory to explain this strange phenomenon. The majority carrier effect is responsible for the tunnelling phenomenon. The quantum transition probability per unit time governs the tunnelling time of carriers over the potential energy barrier, rather than the standard transit-time idea, which states that the transit time is equal to the barrier width divided by the carrier velocity. Because of its low cost, light weight, high speed, low-power operation, low noise, and high peak-current to valley-current ratio, tunnel diodes are beneficial in numerous circuit applications such as microwave amplification, microwave oscillation, and binary memory.

Principles of Operation

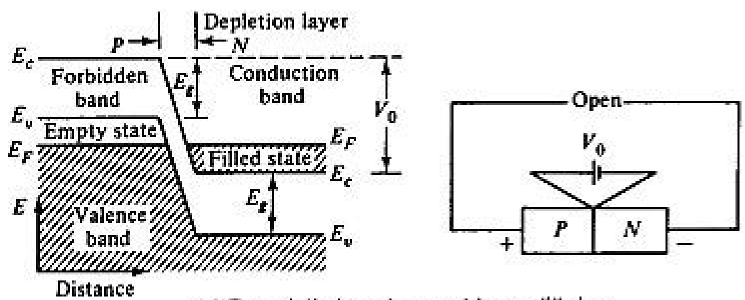
The tunnel diode is a p-n junction semiconductor diode with a negative resistance. The tunnel effect of electrons in the p-n junction causes the negative resistance. Both the p and n regions of the tunnel diode are heavily doped (impurity concentrations of 10^{19} to 10^{20} atoms/cm^3), and the depletion-layer barrier at the junction is extremely thin (on the order of 100 Å or 10^{-6} cm). Those particles can pass through the potential barrier if and only if their energy is equal to or greater than the height of the potential barrier. Quantum mechanically, however, if the barrier is smaller than 3 Å, particles have a good chance of tunnelling through the potential barrier even if they don't have enough kinetic energy to pass through the same barrier. There must be full energy levels on the side from which particles will tunnel and allowed empty states on the other side into which particles will enter at the same energy level, in addition to the barrier thinness. Let's look at the energy-band photos of a severely doped p-n diode to better comprehend tunnel effects. A tunnel diode's energy-band diagrams are shown in Figure 5.15 (1).

NOTES

The upper levels of electron energy of both the p type and the n type are lined up at the same Fermi level under open-circuit conditions or at zero-bias equilibrium, as shown in Figure 5.15-1(a). There is no charge movement in either direction across the junction because there are no filled states on one side of the junction that are at the same energy level as empty allowed states on the other side, as seen at point (a) of the volt-ampere characteristic curve of a tunnel diode shown in Figure 5.15 (2).

The Fermi level resides in the forbidden band in conventional diodes. The Fermi level exists in the valence band in p-type semiconductors and in the conduction band in n-type semiconductors because the tunnel diode is strongly doped. The energy curve in part (1) of as shown in Figure 5.15-l (b) is illustrated when the tunnel diode is forward biased by a voltage between zero and the value that would provide peak tunnelling current I_p ($0 < V < V_p$). As a result, the magnitude of the supplied forward-bias voltage lowers the potential barrier. On both sides, a disparity in Fermi levels is created. Because filled states in the n type's conduction band have the same energy level as allowed empty states in the p type's valence band, electrons tunnel through the barrier from the n type to the p type, causing a forward tunnelling current to flow from the p type to the then type, as shown in sector (1) of shown in Figure 5.15-2(a). The picture of the energy band changes as the forward bias is increased to V_p, as illustrated in part (2) of shown in Figure 5.15-l (b). The peak current Ip shown in Figure 5.15-2 is caused by a maximum number of electrons tunnelling across the barrier from the filled states in the n type to the empty states in the p type (a). The state indicated in component (3) of shown in Figure 5.15-l (b) is attained if the bias voltage is increased further. As illustrated in sector (3) of shown in Figure 5.15-2, the tunnelling current diminishes (a). Finally, the band structure of section (4) of shown in Figure 5.15-l (b) is attained at a very high bias voltage. No electrons can tunnel across the barrier now that there are no authorized vacant states in the p type at the same energy level as filled states in the n type, and the tunnelling current reduces to zero, as illustrated at point (4) of shown in Figure 5.15-2 (a).

The conventional injection current I at the p-n junction begins to flow when the forward-bias voltage V is increased over the valley voltage V_v. The forward voltage increases the injection current exponentially, as seen by the dashed curve shown in Figure 5.15-2 (a). The volt-ampere characteristic of the tunnel diode is determined by the total current, which is equal to the sum of the tunnelling and injection currents, as shown in Figure 5.15-2. (b). The total current achieves a minimal value I_v (or valley current) somewhere near the point where the tunnel diode characteristic meets the standard p-n diode characteristic, as shown in the diagram. Peak current to valley current (I_p/I_v) ratios can theoretically range from 50 to 100. However, in practice, this ratio is around 15.



(a) Tunnel diode under zero-bias equilibrium

NOTES

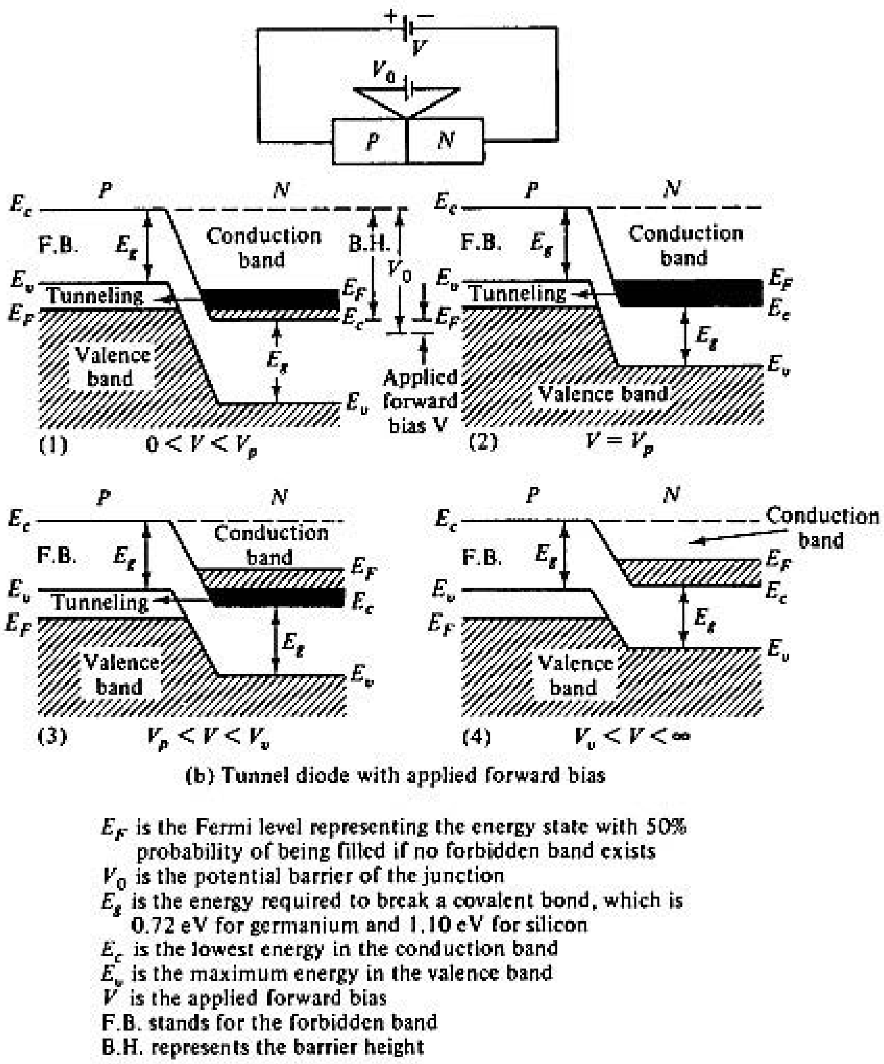


Fig. 5.15 *Energy band diagram of tunnel diode*

I-V Characteristics Under Different Condition

The tunnel diode is important in microwave oscillators and amplifiers because it exhibits a negative resistance characteristic in the region between peak current I_p and valley current I_v. The I-V characteristic of a tunnel diode with the load line is shown in Figure 5.16. The 'abc' load line crosses the characteristic curve three times. Points a and c are both secure, however point b is vulnerable. If the voltage and current change by approximately b, the final values of I and V will be found at point an or c, but not at b. Because the tunnel diode has two stable states for this load line and can be utilized as a binary device in switching circuits, the circuit is called bistable. The microwave oscillation or amplification caused by the tunnel diode is, however, our primary focus in this section. The second load line only intersects the I-V curve at point b.

The tunnel diode can now serve as a microwave amplifier or oscillator because this point is stable and has a dynamic negative conductance. Astable circuits have a load line crossing point in the negative resistance zone. A monostable circuit is shown by another load line crossing point a in the positive-resistance zone. As shown in Figure 5.16, the negative conductance is given by

NOTES

$$-g = \frac{\partial i}{\partial v}\bigg|_{V_b} = \frac{1}{-R_n}$$

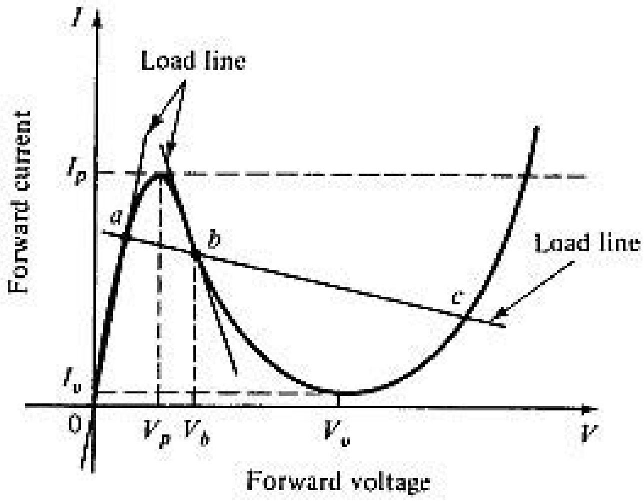


Fig. 5.16 *I-V characteristic of tunnel diode with load line.*

The negative resistance remains constant for a slight fluctuation of the forward voltage around V, and the diode circuit behavior is stable. Figure 5.17 shows a small-signal equivalent circuit for a tunnel diode operating in the negative-resistance region. The inductance and resistance of a tunnel diode's packaging circuit are denoted by R_s and L_s. The diode's junction capacitance C is normally measured at the valley point, while Rn is the diode's negative resistance. These parameters have typical values for a tunnel diode with a peak current I_p of I_O mA.

$-R_n = -30\ \Omega$ $R_s = 1\,0\Omega$ $L_s = 5$ nH C = 20 pF

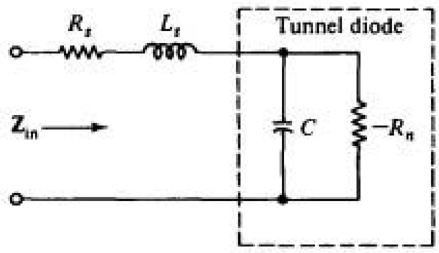


Fig. 5.1*7 Equivalent circuit of tunnel diode.*

The input impedance Z_{in} of the equivalent circuit shown in Figure 5.17 is given by:

$$\mathbf{Z}_{in} = R_s + j\omega L_s + \frac{R_n[\,j/(\omega C)]}{-R_n - j/(\omega C)}$$

$$\mathbf{Z}_{in} = R_s - \frac{R_n}{1 + (\omega R_n C)^2} + j\left[\omega L_s - \frac{\omega R_n^2 C}{1 + (\omega R_n C)^2}\right] \qquad (5.18)$$

For the resistive cutoff frequency, the real part of the input impedance Z_{in} must be zero. Consequently, from above Equation (5.18) the resistive cutoff frequency is given by:

$$f_c = \frac{1}{2\pi R_n C}\sqrt{\frac{R_n}{R_s} - 1}$$

NOTES

For the self-resonance frequency, the imaginary part of the input impedance must be zero. Thus,

$$f_r = \frac{1}{2\pi R_n C}\sqrt{\frac{R_n^2 C}{L_s} - 1}$$

The tunnel diode can be connected either in parallel or in series with a resistive load as an amplifier; its equivalent circuits are shown in Figure 5.18.

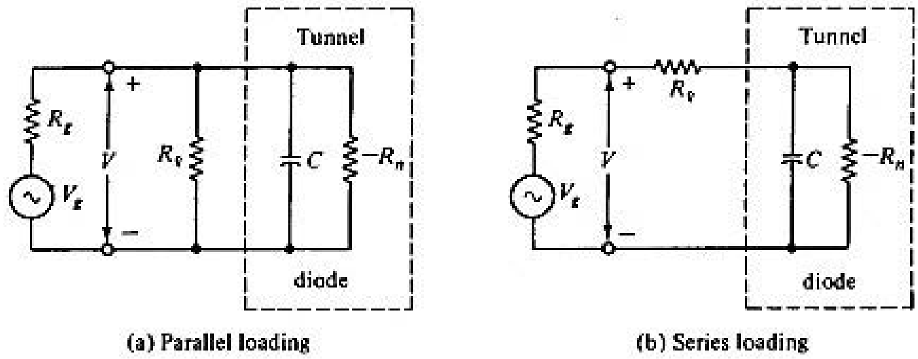


Fig. 5.18 *Equivalent circuits of tunnel diodes*

Parallel Loading

It can be shown in Figure 5.18 (a) that the output power in the load resistance is given by:

$$P_{out} = \frac{V^2}{R_\ell}$$

One part of this output power is generated by the small input power through the tunnel diode amplifier with a gain of A, and this part can be written

$$P_{in} = \frac{V^2}{AR_\ell}$$

Another part of the output power is generated by the negative resistance, and it is expressed as

$$P_n = \frac{V^2}{R_n}$$

Therefore

$$\frac{V^2}{AR_\ell} + \frac{V^2}{R_n} = \frac{V^2}{R_\ell} \qquad (5.19)$$

and the gain Equation (5.19) of a tunnel diode amplifier is given by:

$$A = \frac{R_n}{R_n - R_\ell}$$

When the negative resistance R_n of the tunnel diode approaches the load resistance R_l, the gain A approaches infinity and the system goes into oscillation.

Series Loading

NOTES

In the series circuit shown in Figure 5.18(b) the power gain A is given by

$$A = \frac{R_\ell}{R_\ell - R_n} = \frac{1}{1 - R_n/R_\ell}$$

The device remains stable in the negative-resistance region without switching if $R_l < R_n$.

As shown in Figure 5.19, a tunnel diode can be coupled to a microwave circulator to create a negative resistance amplifier. A microwave circulator is a multiport junction in which power can only travel in one direction: from port 1 to port 2, port 2 to port 3, and so on. Microwave circulators with four ports are the most frequent, despite the fact that the number of ports is not limited. If the circulator is perfect and has a positive real characteristic impedance R_o, a negative-resistance tunnel diode with a real component equal to - R_o and an imaginary part equal to zero can be used to build an infinite-gain amplifier. Figure 5.19 shows that the reflection coefficient is infinite. In general, the reflection coefficient is calculated as follows:

$$\Gamma = \frac{-R_n - R_0}{-R_n + R_0}$$

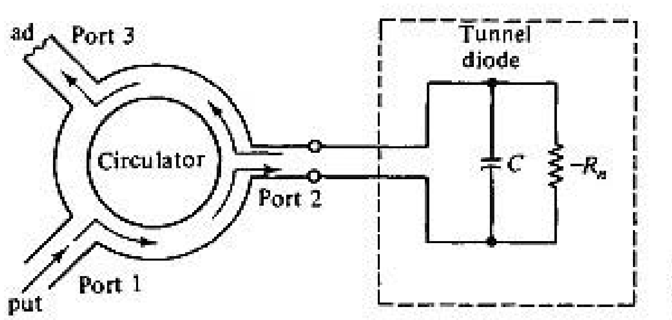


Fig. 5.19 *Tunnel diode connected to circulator.*

Introduction

Negative resistance is a typical feature of all active two-terminal solid-state devices. Over a wide range of frequencies, the true part of their impedance is negative. The current through the resistance and the voltage across it are in phase in a positive resistance. The voltage drop across a positive resistance is positive, and the resistance dissipates a power of (I^2R). However, in a negative resistance, the current and voltage are 180° out of phase. The voltage drop across a negative resistance is negative, and the power supply associated with the negative resistance generates a power of ($-I^2R$). In other words, positive resistances (passive devices) consume power, whereas negative resistances (active devices) generate power (active devices).

Gunn-Effect Diodes-GaAs Diode

J. B. Gunn, who found periodic fluctuations of current passing through then-type Gallium Arsenide (GaAs) specimens when the applied voltage surpassed a specific critical value in 1963, is the inventor of Gunn-effect diodes. B. C. DeLoach, R. C. Johnston, and B. G. Cohen developed the impact Ionization Avalanche Transit-Time (IMPATT) mechanism in silicon two years later, in 1965. This method uses the diode's avalanching and transit-time features to generate microwave frequencies. Later developments included the Limited Space Charge-Accumulation diode (LSA diode) and the Indium Phosphide diode (InP diode). In contrast to the tunnel diode, these are bulk devices in that microwave amplification and oscillation are generated from the bulk negative-resistance property of homogeneous semiconductors rather than the junction negative-resistance property of two distinct semiconductors.

The Gunn Affect

J. B. Gunn found a periodic fluctuation of current travelling through n-type gallium arsenide when the applied voltage surpassed a certain point in 1963, and the effect was named after him.

In 1954, Shockley proposed that two terminal negative resistance semiconductor devices have benefits over transistors at high frequencies. Ridley and Watkins published a paper in 1961 detailing a new approach for obtaining negative differential mobility in semiconductors. When carriers in a light mass, limited mobility, higher energy sub band have a high temperature, the principle is to heat them. Finally, Kroemer claimed that Ridley Watkins Hilsum's method of electron transport into valleys in conduction bands is the source of negative differential mobility.

A homogeneous n-type GaAs diode with ohmic contacts at the end surfaces is depicted in the schematic below. 'Above a critical voltage, corresponding to an electric field of 2000 to 4000 Volts/cm, the current in each specimen became a fluctuating function of time,' Gunn explained.

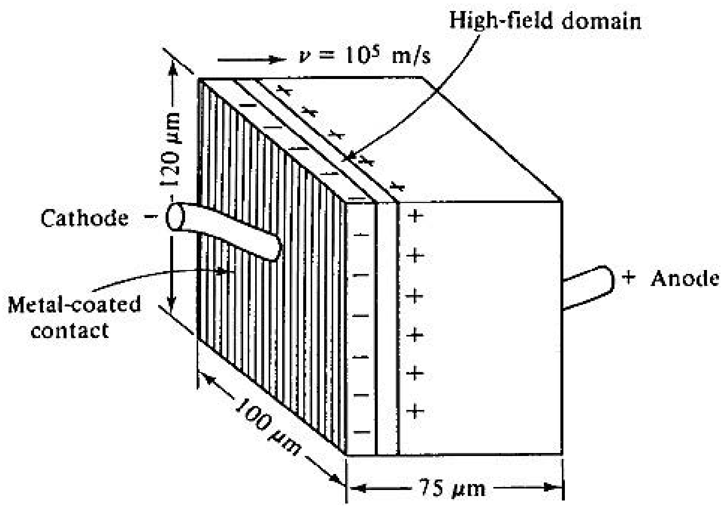


Fig. 5.20 *Schematic diagram for n-type GaAs diode.*

NOTES

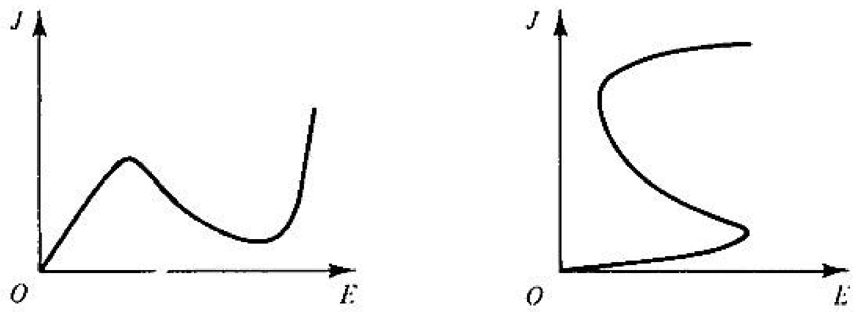
When the electric field is changed from zero to a threshold value, the carrier drift velocity increases linearly from zero to a maximum, according to Gunn's observations. The drift velocity is reduced and the diode exhibits negative resistance when the electric field exceeds the threshold value of 3000 V/cm for n-type GaAs. Figure 5.20 shows depicts the situation.

Hilsum and Ridley Watkins Hypothesis

The Gunn Effect has been explained in a variety of ways. In 1964, Kroemer stated that Gunn's data matched the Ridley-Watkins-Hilsum (RWH) theory perfectly.

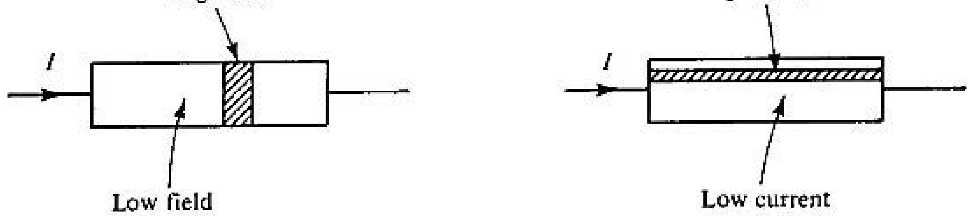
Differential Negative Resistance

The differential negative resistance formed in a bulk solid-state III-V compound when either a voltage (or electric field) or a current is applied to the sample's terminals is the basic premise of the Ridley-Watkins-Hilsum (RWH) theory. Negative-resistance devices can operate in two modes: voltage-controlled and current-controlled.



(a) Voltage-controlled mode (b) Current-controlled mode

The current density in the voltage-controlled mode can be multivalued, whereas the voltage in the current-controlled mode can be multivalued. The presence of a differential negative-resistance zone in the current density field curve has the primary impact of making the sample electrically unstable. As a result, in an attempt to achieve stability, the initially homogenous sample becomes electrically heterogeneous. High-field domains are created in the voltage-controlled negative-resistance mode, separating two low-field zones. The interfaces separating the low and high-field domains are parallel to the current direction and so reside in planes and perpendicular to it.



(a) High-field domain (b) High-current filament

NOTES

The current density J_o is formed when an electric field E_o (or voltage V_o) is supplied to the sample, for example. The current density decreases to J_2 as the applied field (or voltage) is increased to E_2 (or V_2). The current density increases to J_1 when the field (or voltage) is reduced to E_1 (or V_1). Figure 5.21 shows depicts these voltage-controlled negative resistance events (a). Similarly, the negative-resistance profile for the current regulated mode is depicted shown in Fig. 5.21. (b).

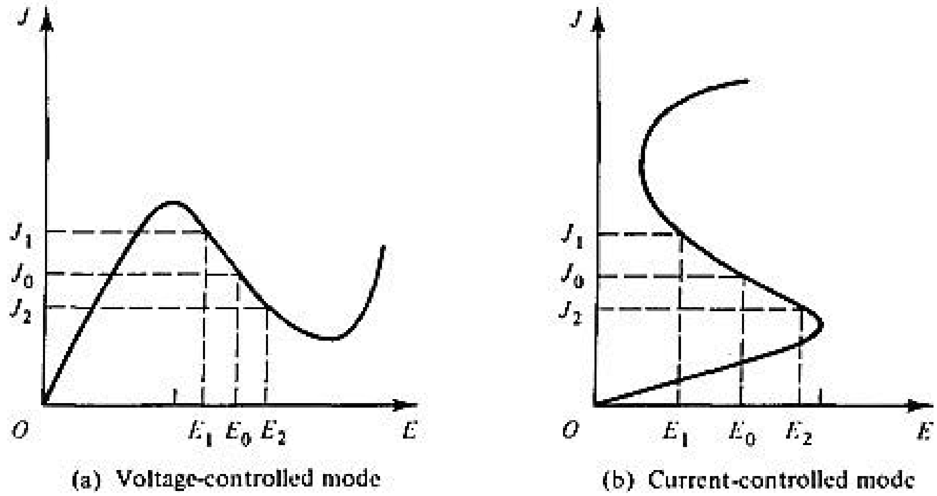


Fig. 5.21 *Multiple Values of Current Density for Negative Resistance*

Theory of the Two-Valley Model

Kroemer proposed a negative mass microwave amplifier in 1958 and 1959, a few years before the Gunn Effect was discovered. As shown in Figure 5.22, a high-mobility lower valley is separated from a low-mobility upper valley by an energy of 0.36 eV according to the energy band theory of then-type GaAs.

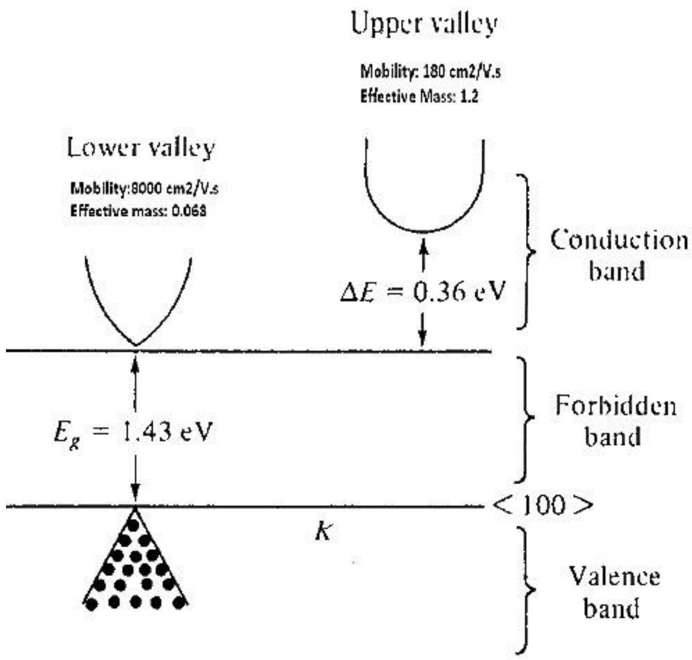


Fig. 5.22 *Two-Valley Model of Electron Energy Versus Wave Number for n-type GaAs*

NOTES

Table 5.1 presents data for two valleys in then-type GaAs, while Table 5.2 offers data for two-valley semiconductors.

Table 5.1 *Data for Two Valleys in GaAs*

Valley	Effective Mass M_e	Mobility μ	Separation ΔE
Lower	$M_{e\ell} = 0.068$	$\mu_\ell = 8000$ cm²/V-sec	$\Delta E = 0.36$ eV
Upper	$M_{eu} = 1.2$	$\mu_u = 180$ cm²/V-sec	$\Delta E = 0.36$ eV

Table 5.2 *Data for Two Valleys Semiconductors*

Semiconductor	Gap energy (at 300°K) E_g(eV)	Separation energy between two valleys ΔE(eV)	Threshold field E_{th}(KV/cm)	Peak velocity v_p(10^7 cm/s)
Ge	0.80	0.18	2.3	1.4
GaAs	1.43	0.36	3.2	2.2
InP	1.33	0.60	10.5	2.5
		0.80		
CdTe	1.44	0.51	13.0	1.5
InAs	0.33	1.28	1.60	3.6
InSb	0.16	0.41	0.6	5.0

Under an equilibrium condition, the electron densities in the lower and higher valleys are the same. No electrons will move to the higher valley when the applied electric field is lower than the electric field of the lower valley ($E < E_e$), as shown in Figure 5.23 (a). As shown in Figure 5.23, electrons will begin to transfer to the upper valley when the applied electric field is higher than that of the lower valley but lower than that of the upper valley ($E_e < E < E_u$) (b). When the applied electric field is greater than the upper valley's ($E_u < E$), all electrons flow to the upper valley, as shown in Figure 5.23 (c).

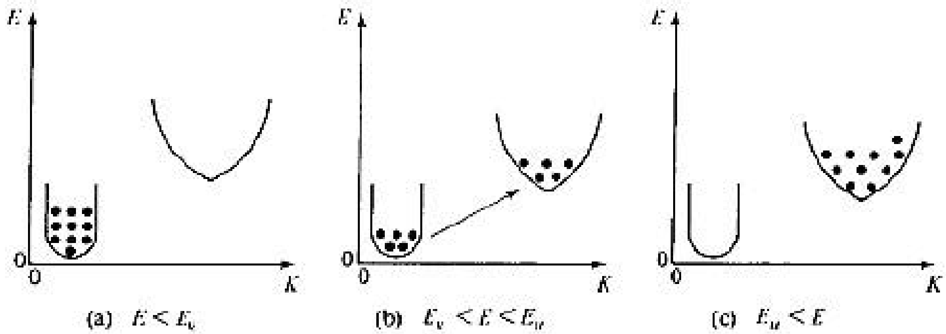


Fig. 5.23 *Transfer of electron densities*

The conductivity of n-type GaAs is if the electron concentrations in the bottom and upper valleys are n_c and n_u, respectively.

$$\sigma = e\,(\mu_e n_e + \mu_u n_u)$$

where e = the electron charge μ = the electron mobility $n = n_l + n_u$ is the electron density.

NOTES

Avalanche Transit-Time Devices

To generate a supply of holes and electrons, avalanche transit-time diode oscillators rely on the impact of voltage breakdown across a reverse-biased p-n junction. Scientists have wondered whether it is conceivable to create a two-terminal negative-resistance device since the development of modern semiconductor device theory. The tunnel diode was the first practical implementation of such a device. Its operation is dependent on the features of a forward-biased p-n junction with substantially doped p and n regions. The transferred electron devices and the avalanche transit-time devices are the other two devices.

The Gunn oscillators, also known as transferred electron devices, work by applying a de voltage to a bulk semiconductor. This gadget does not have any p-n junctions. Its frequency is determined by the load and the circuit's inherent frequency. At microwave frequencies, the avalanche diode oscillator produces a negative resistance by using carrier impact ionisation and drift in the high-field region of a semiconductor junction. Read first presented the device in a theoretical work where he looked at the negative-resistance features of an imagined n^+-p-i-p^+ diode. Avalanche oscillator has been reported in two separate modes. The IMPATT mode (impact ionization avalanche transit-time operation) is one of them. The normal de-to-RF conversion efficiency in this mode is 5 to 10%, and silicon diodes can operate at frequencies up to 100 GHz. The TRAPATT mode, which stands for trapped plasma avalanche triggered transit, is the alternative option. Its usual conversion efficiency ranges between 20% and 60%.

The BARITT (Barrier Injected Transit-Time) diode is another active microwave device. Long drift areas, similar to those of IMPATT diodes, are present. Rather than being taken from the plasma of an avalanche area, the carriers crossing the drift regions of BARITT diodes are created by minority carrier injection from forward-biased junctions. Several alternative topologies, including p-n-p, p-n-v-p, p-n-metal, and metal-n-metal, have been used as BARITT diodes. Although BARITT diodes have low noise figures of 15 dB, their bandwidth is restricted and their output power is low.

IMPATT Diodes

Physical Structures

An n^+-p-i-p^+ or p^+-n-i-n^+ structure Read diode was studied theoretically. The interplay of the impact ionization avalanche and the charge carrier transit time is its fundamental physical process. IMPATT diodes are thus named after Read-type diodes. Due to two effects, these diodes have a distinct negative resistance:

1. The impact ionization avalanche effect, which generates a 90° phase shift between the carrier current $I_o(t)$ and the ac voltage.
2. The transit-time effect, which causes the external current $I_e(t)$ to be 90° delayed in relation to the ac voltage.

NOTES

However, the first IMPATT functioning was produced from a basic p-n junction, as described by Johnston and his colleagues in 1965. Lee and his colleagues announced the world's first actual Read-type IMPATT diode. A negative resistance of the IMPATT diode may be generated from a junction diode with any doping profile, according to the small-signal theory proposed by Gilden. Many IMPATT diodes have a high doping avalanching area followed by a drift region with a low field that allows the carriers to pass through without avalanching. In the IMPATT diode family, the Read diode is the most basic. The one-sided abrupt p-n junction, the linearly graded p-n junction (or double-drift area), and the p-i-n diode, which are all depicted shown in Figure 5.24, are the others.

Resistant Negative

The actual component of the terminal impedance of a Read diode may be expressed as follows using small-signal analysis:

$$R = R_s + \frac{2L^2}{v_d \epsilon_s A} \frac{1}{1 - \omega^2/\omega_r^2} \frac{1 - \cos\theta}{\theta}$$

Where R_s = Passive resistance of the inactive region

V_D = Carrier drift velocity

L = Length of the drift space-charge region

A = Diode cross section

e_s = Semiconductor dielectric permittivity

Moreover, θ is the transit angle, given by:

$$\theta = \omega\tau = \omega\frac{L}{v_d}$$

and ω_r, is the avalanche resonant frequency, defined by

$$\omega_r \equiv \left(\frac{2\alpha' v_d I_0}{\epsilon_s A}\right)^{1/2} \qquad (5.20)$$

In the above Equation (5.20) the quantity α' is the derivative of the ionization coefficient with respect to the electric field. The number of ionizations per centimetre produced by a single carrier increases rapidly as the electric field increases. The variation of the negative resistance with the transit angle when $\omega > \omega_r$ is plotted shown in Figure 5.24. The peak value of the negative resistance occurs near θ = π. For transit angles larger than π and approaching 3π/2, the negative resistance of the diode decreases rapidly. For practical purposes, the Read-type IMPATT diodes work well only in a frequency range around the π transit angle.

i.e.,

$$f = \frac{1}{2\tau} = \frac{v_d}{2L}$$

NOTES

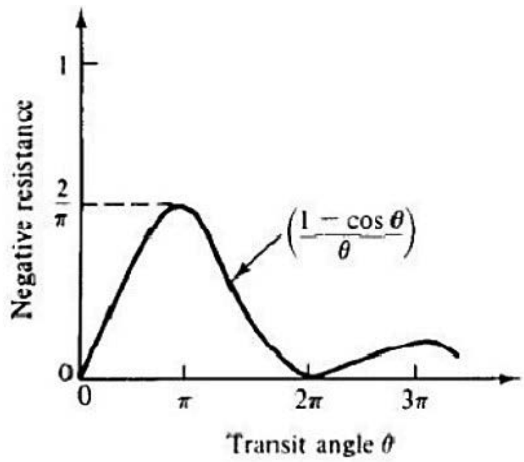


Fig. 5.24 *Negative resistance versus transit angle*

Power Output and Efficiency

The maximum output power of a single diode at a particular frequency is constrained by semiconductor materials and microwave circuitry impedance levels. The highest voltage that may be placed across the diode for a uniform avalanche is calculated as follows:

$$V_m = E_m L$$

The depletion length is L, and the maximum electric field is E_m. The breakdown voltage sets the upper limit for this maximum applied voltage. The avalanche breakdown process also limits the maximum current that the diode can carry, because current in the space-charge zone increases the electric field. The following formula calculates the maximum current:

$$I_m = J_m A = \sigma E_m A = \frac{\epsilon_s}{\tau} E_m A = \frac{v_d \epsilon_s E_m A}{L}$$

Therefore, the upper limit of the power input is given by:

$$P_m = I_m V_m = E_m^2 \epsilon_s v_d A$$

The capacitance across the space-charge region is defined as

$$C = \frac{\epsilon_s A}{L}$$

Substitution of above equations and using $2\pi f\tau = 1$ yield

$$P_m f^2 = \frac{E_m^2 v_d^2}{4\pi^2 X_c} \qquad (5.21)$$

This Equation (5.21) is nearly equivalent to the microwave power transistor's power frequency restriction. As $1/f^2$, the maximum power available to mobile carriers diminishes. At frequencies as high as 100 GHz, this electrical limit dominates silicon. The IMPATT diodes' efficiency is calculated as:

$$\eta = \frac{P_{ac}}{P_{dc}} = \left(\frac{V_a}{V_d}\right)\left(\frac{I_a}{I_d}\right)$$

The ratio of ac voltage to applied voltage is around 0.5 for an ideal Read-type IMPATT diode, and the ratio of ac current to de current is about $2/\pi$, resulting in an efficiency of about $1/\pi$ or more than 30%. The space-charge effect, the reverse-saturation-current effect, the high-frequency-skin effect, and the ionization-saturation effect all contribute to the efficiency of practical IMPATT diodes being less than 30%.

NOTES

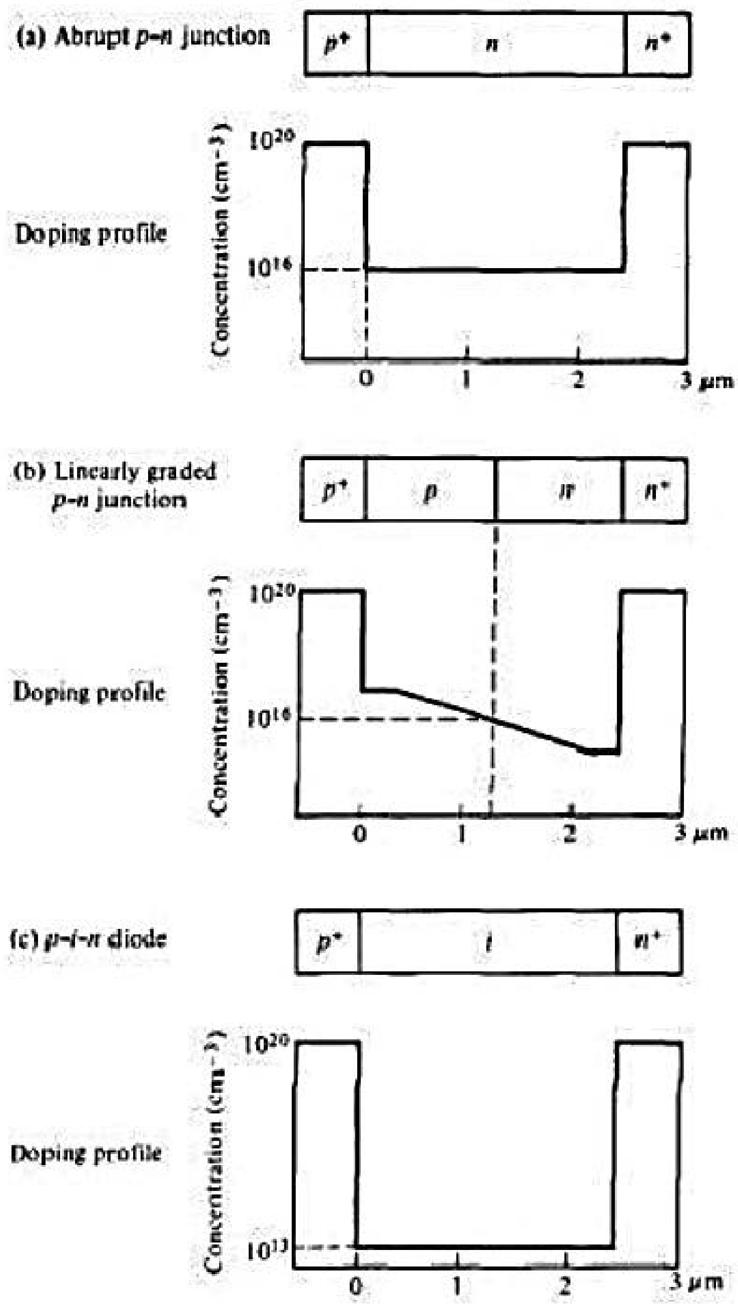


Fig. 5.25 *Three Typical Silicon IMPATT Diodes*

5.3.1 Parametric Devices

Physical Characteristics

A parametric device is one that employs a time-varying or nonlinear reactance (capacitance or inductance). Because capacitance or inductance, which is a reactive property, may be utilized to create capacitive or inductive excitation, the term parametric is derived from the phrase parametric excitation. Parametric amplification

NOTES

and oscillation are two types of parametric stimulation. Faraday defined several of the key features of nonlinear energy-storage devices in 1831, and Lord Rayleigh in 1883. Van der Ziel published the first nonlinear capacitance study in 1948. Because it was basically a reactive device with no thermal noise, van der Ziel argued that such a device would be beneficial as a low-noise amplifier in his study. Landon examined and published the experimental results of amplifiers, converters, and oscillators in 1949. Microwave electronics experts dreamed of a solid-state microwave device to replace the noisy electron beam amplifier during the solid-state era. Suhl presented a ferrite-based microwave solid-state amplifier in 1957. Weiss was the first to realize a microwave parametric amplifier after Suhl's idea in 1957. The parametric amplifier was finally identified after Suhl and Weiss' efforts. The solid-state varactor diode is now the most used parametric amplifier. The parametric diode, unlike microwave tubes, transistors, and lasers, is a reactive device that produces very little Johnson noise (thermal noise). One of the characteristics of a parametric amplifier is that it uses an ac power source rather than the dc power supply used by microwave tubes. The parametric amplifier is similar to a quantum amplifier laser or maser that uses an AC power source in this regard.

Nonlinear Reactance and Manley-Rowe Power Relations

The term reactance refers to a circuit element that both stores and releases electromagnetic energy, as opposed to the term resistance, which refers to a circuit part that just dissipates energy. Whenever the stored energy is concentrated primarily in the electric field, the reactance is referred to as capacitive; when the stored energy is concentrated primarily in the magnetic field, the reactance is referred to as inductive. Capacitive reactance is defined as follows: In microwave engineering, it is more convenient to discuss in terms of voltages and currents rather than electric and magnetic fields since voltages and currents are more readily understood. Therefore, a capacitive reactance is a type of circuit element where capacitance is defined as the relationship between charge on a capacitor and voltage across the capacitor. Then

$$C = \frac{Q}{V}$$

The capacitive reactance is considered to be nonlinear if the ratio is not linear. In this scenario, a nonlinear capacitance may be defined as the partial derivative of charge with respect to voltage. In other words,

$$C(v) = \frac{\partial Q}{\partial v}$$

The analogous definition of a nonlinear inductance is,

$$L(i) = \frac{\partial \Phi}{\partial i}$$

When a nonlinear reactance is impressed with voltages at two or more separate frequencies during the operation of parametric devices, mixing effects occur.

Small-Signal Method

It is assumed that the signal voltage V_s is much smaller than the pumping voltage V_p, and the total voltage across the nonlinear capacitance C(t) is given by:

$$v = v_s + v_p = V_s \cos(\omega_s t) + V_p \cos(\omega_p t)$$

where $V_s << V_p$. The charge on the capacitor can be expanded in a Taylor series about the point $V_s = 0$, and the first two terms are

NOTES

$$Q(v) = Q(v_1 + v_p) = Q(v_p) + \left.\frac{dQ(v_p)}{dv}\right|_{v_s=0} v_s$$

For convenience, it is assumed that

$$C(v_p) = \frac{dQ(v_p)}{dv} = C(t)$$

where C(Vp) is periodic with a fundamental frequency of Wp . If the capacitance C(Vp) is expanded in a Fourier series, the result is

$$C(v_p) = \sum_{n=0}^{\infty} C_n \cos(n\omega_p t)$$

Since Vp is a function of time, the capacitance C(Vp) is also a function of time. Then

$$C(t) = \sum_{n=0}^{\infty} C_n \cos(n\omega_p t) \tag{5.22}$$

The coefficients C_n are the magnitudes of the time-varying capacitance's harmonics. Coefficients C_n are not linear functions of the AC pumping voltage V_P in general. Due to the nonlinear nature of the junction capacitance C(t) of a parametric diode, the concept of superposition does not hold true for arbitrary AC signal amplitudes.

The current flowing through the capacitance C(t) is the time derivative of Equation (5.22) and it equals

$$i = \frac{dQ}{dt} = \frac{dQ(v_p)}{dt} + \frac{d}{dt}[C(t)v_s] \tag{5.23}$$

It is obvious that the nonlinear capacitance acts similarly to a time-varying linear capacitance for signals with considerably lower amplitudes than the pumping voltage's amplitude. The first term of Equation (5.23) yields a current at the pump frequency f_p and is not related to the signal frequency f_s.

Large-Signal Method

If the signal voltage is not too low in comparison to the pumping voltage, the Taylor series can be enlarged in a junction diode around a dc bias voltage V_o. The capacitance C of a junction diode is proportional to

$(\phi_o - V)^{-1/2} = V_o^{-1/2}$, where ϕ_o is the junction barrier potential and V is a negative voltage supply. Since

$$[V_0 + V_p \cos(\omega_p t)]^{-1/2} \approx V_0^{-1/2}\left(1 - \frac{V_p}{3V_0}\cos(\omega_p t)\right) \qquad \text{for } V_p \ll V_0$$

the capacitance C (t) can be expressed as

$$C(t) = C_0[1 + 2\gamma \cos(\omega_p t)]$$

The parameter is proportional to the pumping voltage V_P and represents the coupling effect between the voltages at the signal and output frequencies f_s and f_o.

Manley-Rowe Power Relations

NOTES

Manley and Rowe defined a set of general energy rules that govern the flow of energy into and out of an ideal nonlinear reactance. These connections play a critical role in deciding whether a parametric amplifier is capable of achieving power gain. Figure 5.26 shows the equivalent circuit for the Manley-Rowe derivation.

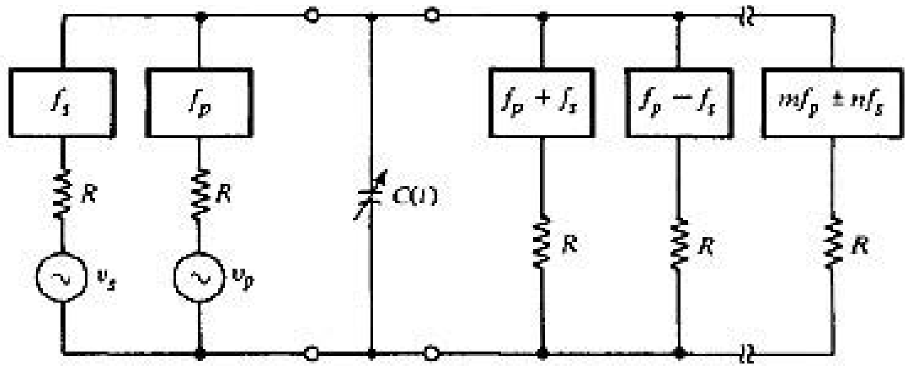


***Fig. 5.26** Equivalent Circuit for Manley-Rowe Derivation*

As shown in Figure 5.26, a signal generator and a pump generator operating at their respective frequencies f_s and f_p are applied to a nonlinear capacitance C(t), together with related series resistances and bandpass filters. Filters using resonant circuits are intended to reject power at all frequencies other than the signal frequency. When two frequencies f_s and f_p are applied, an unlimited number of resonant frequencies $mf_p \pm nf_p$ are created, where m and n are any integers between zero and infinity.

Each resonant circuit is presumptively perfect. The nonlinear susceptances have a low power loss. That is, the power supplied to the nonlinear capacitor at the pump frequency is equivalent to the power supplied to the capacitor at the other frequencies via the nonlinear interaction. Manley and Rowe defined the power relationships between the input power at the fs and f_p frequencies and the output power at the other frequencies $mf_p \pm nf_s$.

From above equation the voltage across the nonlinear capacitor C(t) can be expressed in exponential form as

$$v = v_p + v_s = \frac{V_p}{2}(e^{j\omega_p t} + e^{-j\omega_p t}) + \frac{V_s}{2}(e^{j\omega_s t} + e^{-j\omega_s t})$$

The general expression of the charge Q deposited on the capacitor is given by

$$Q = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} Q_{m,n} e^{j(m\omega_p t + n\omega_s t)}$$

In order for the charge Q to be real, it is necessary that

$$Q_{m,n} = Q^*_{-m,-n}$$

The total voltage v can be expressed as a function of the charge Q. A similar Taylor series expansion of V(Q) shows that

NOTES

$$v = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} V_{m,n} e^{j(m\omega_p t + n\omega_s t)}$$

In order for the voltage V to be real, it is required that

$$V_{m,n} = V^*_{-m,-n}$$

The current flowing through C(t) is the total derivative of above equation with respect to time. This is

$$i = \frac{dQ}{dt} = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} j(m\omega_p + n\omega_s) Q_{m,n} e^{j(m\omega_p t + n\omega_s t)}$$
$$= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} I_{m,n} e^{j(m\omega_p t + n\omega_s t)}$$

where $I_{m,n} = j(m\omega_p + n\omega_s)Q_{m,n}$ and $I_{m,n} = I^*_{-m,-n}$. Since the capacitance C(t) is assumed to be a pure reactance, the average power at the frequencies $mf_p + nf_s$ is

$$P_{m,n} = (V_{m,n} I^*_{m,n} + V^*_{m,n} I_{m,n})$$
$$= (V^*_{-m,-n} I_{-m,-n} + V_{-m,-n} I^*_{-m,-n}) = P_{-m,-n}$$

Then conservation of power can be written

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} P_{m,n} = 0 \qquad (5.24)$$

Multiplication of Equation (5.24) by a factor of $(m\omega_p + n\omega_s)/(m\omega_p + n\omega_s)$ and rearrangement of the resultant into two parts yield

$$\omega_p \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \frac{mP_{m,n}}{m\omega_p + n\omega_s} + \omega_s \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \frac{nP_{m,n}}{m\omega_p + n\omega_s} = 0$$

Since $\quad I_{m,n}/(m\omega_p + n\omega_s) = jQ_{m,n}$

$P_{m,n}/(m_p + n_s)$ thus becomes $- jV_{m,n}Q^*_{m,n} - jV_{-m,-n}Q^*_{m,-n}$ and is independent of p or s. For any combination of f_p and f_s, the resonating circuit external to the nonlinear capacitance C(t) may be set in such a way that the currents maintain all of the voltage amplitudes $V_{m,n}$ unaltered. Charges $Q_{m,n}$ remain unaltered as well, because they are functions of the voltages $V_{m,n}$. As a result, the frequencies f_p and f_s may be altered freely to meet the requirements.

$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \frac{mP_{m,n}}{m\omega_p + n\omega_s} = 0$$
$$\sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \frac{nP_{m,n}}{m\omega_p + n\omega_s} = 0 \qquad (5.25)$$

Equation (5.25) can be expressed as two terms:

$$\sum_{m=0}^{\infty} \sum_{n=-\infty}^{\infty} \frac{mP_{m,n}}{m\omega_p + n\omega_s} + \sum_{m=0}^{\infty} \sum_{n=-\infty}^{\infty} \frac{-mP_{m,n}}{-m\omega_p - n\omega_s} = 0$$

NOTES

Since $P_{m,n} = P_{-m,-n}$, then

$$\sum_{m=0}^{\infty} \sum_{n=-\infty}^{\infty} \frac{mP_{m,n}}{mf_p + nf_s} = 0 \quad (5.26)$$

Similarly,

$$\sum_{m=-\infty}^{\infty} \sum_{n=0}^{\infty} \frac{nP_{m,n}}{mf_p + nf_s} = 0 \quad (5.27)$$

where ω_p and ω_s have been replaced by f_p and f_s, respectively.

The typical formulations for the Manley-Rowe power relations are Equations (5.26) and (5.27). $P_{m,n}$ denotes the actual power flowing into or out of the nonlinear capacitor at $mf_p + nf_s$. The fundamental frequency of the pumping voltage oscillator is denoted by f_p, whereas the fundamental frequency of the signal voltage generator is denoted by f_s. The sign convention for the power term $P_{m,n}$ is as follows: power entering the nonlinear capacitance or power arriving from the two voltage generators is positive, whereas power leaving the nonlinear capacitance or power entering the load resistance is negative.

Consider the scenario in which the power output flow is permitted at a frequency of $f_p + f_s$, as shown in Figure 5.26. Each of the remaining harmonics is open circuited. Then only currents with the three frequencies f_p, f_s, and $f_p + f_s$ exist. Under these constraints, m and n range from -1 to +1, respectively. Then, Equations (5.26) and (5.27) become

$$\frac{P_{1,0}}{f_p} + \frac{P_{1,1}}{f_p + f_s} = 0 \quad (5.28)$$

$$\frac{P_{0,1}}{f_s} + \frac{P_{1,1}}{f_p + f_s} = 0 \quad (5.29)$$

where $P_{1,0}$ and $P_{0,1}$ denote the power provided by the two voltage generators at frequency f_p and f_s, respectively, and are regarded as positive. $P_{1,1}$ is deemed negative power when it flows from the reactance to the resistive load at a frequency of $f_p + f_s$.

According to Equation (5.28), the power gain is defined as the ratio of the power produced by the capacitor at a frequency of $f_p + f_s$ to the power absorbed by the capacitor at a frequency of f_s.

$$\text{Gain} = \frac{f_p + f_s}{f_s} = \frac{f_0}{f_s} \qquad \text{(for modulator)} \quad (5.30)$$

where $f_p + f_s = f_o$ and $(f_p + f_s) > f_p > f_s$. The ratio of the output frequency to the input frequency is the maximum power gain. The sum-frequency parametric amplifier or up-converter is a type of parametric device.

If the signal frequency is the sum of the pump frequency and the output frequency, Equation (5.29) predicts that the parametric device will have a gain of

$$\text{Gain} = \frac{f_s}{f_p + f_s} \qquad \text{(for demodulator)} \qquad (5.31)$$

where $f_s = f_p + f_o$ and $f_o = f_s - f_p$. The parametric down-converter is a sort of parametric device whose power gain is actually a loss.

NOTES

The power $P_{1,1}$ delivered at f_p is positive if the signal frequency is f_s, the pump frequency is f_p, and the output frequency is f_o, where $f_p = f_s + f_o$. $P_{1,0}$ and $P_{0,1}$ is also negative. In other words, rather than absorbing power, the capacitor provides it to the signal generator at f_s. The circuit may oscillate at both f_s and f_o, indicating that the power gain is unlimited, which is an unstable state. Another sort of parametric device is a negative-resistance parametric amplifier, which is also known as a negative-resistance parametric amplifier.

Parametric Amplifiers

To create the sum and difference frequencies in a superheterodyne receiver, a radio frequency signal is mixed with a signal from the local oscillator in a nonlinear circuit (the mixer). As shown in Figure 5.27, a parametric amplifier replaces the local oscillator with a pumping generator such as a reflex klystron and the nonlinear element with a time-varying capacitor such as a varactor diode (or inductor).

As shown in Figure 5.27, the nonlinear capacitor C is used to combine the signal frequency f_s and the pump frequency f_p. As a result, a voltage corresponding to the fundamental frequencies f_s and f_p emerges across C, as well as the sum and difference frequencies $mf_p \pm nf_s$. If a resistive load is placed across the idler circuit's terminals, an output voltage at the output frequency f_o can be created across the load. The idler circuit is the output circuit that does not require external stimulation. The idler circuit's output (or idler) frequency f_o is defined as the sum and difference of the signal and pump frequencies f_s and f_p.

i.e.,

$$f_o = mf_p \pm nf_s \qquad (5.32)$$

where m and n are positive integers from zero to infinity.

If fo > fs, the device is called a parametric up-converter. Conversely, if fo < fs, the device is known as a parametric down-converter.

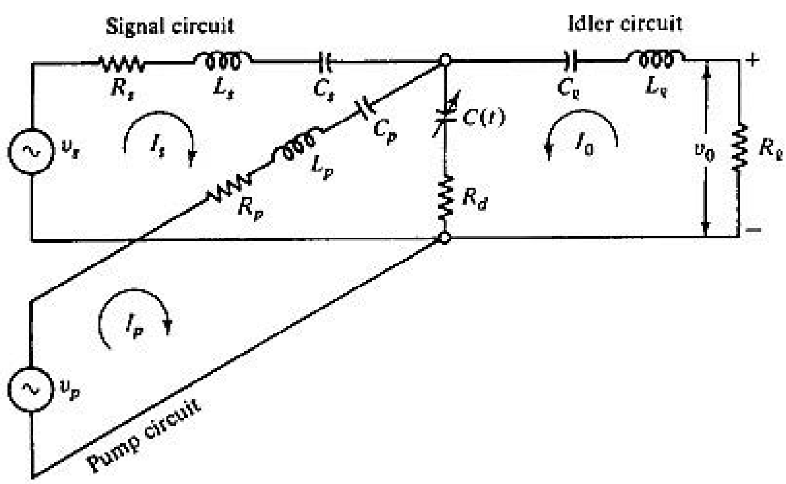


Fig. 5.27 *Equivalent Circuit for a Parametric Amplifier*

NOTES

Parametric Up-Converter

The following are the characteristics of a parametric up-converter:

1. The ouptut frequency equals the product of the signal and pump frequencies.
2. The parametric device has no power flow at frequencies other than those of the signal, pump, and output.

Power Gain

When these two requirements are satisfied, a parametric up-maximum converter's power increase is represented as

$$\text{Gain} = \frac{f_0}{f_s}\frac{x}{(1+\sqrt{1+x})^2}$$

$$\text{where } f_0 = f_p + f_s$$
$$x = \frac{f_s}{f_0}(\gamma Q)^2 \tag{5.33}$$
$$Q = \frac{1}{2\pi f_s C R_d}$$

Additionally, R_d is the series resistance of a p-n junction diode and γQ is the nonlinear capacitor's figure of merit. The value of can be thought of as a gain-degradation factor. As Rd decreases toward zero, the figure of merit γQ approaches infinity and the gain-degradation factor approaches unity. As a result, the power gain of a parametric up-converter is equal to f_o/f_s, as anticipated by the Manley-Rowe relations as given in Equation (5.30). γQ might be equivalent to 10 in a conventional microwave diode. If $f_o/f_s = 15$, Equation (5.33) yields a maximum gain of 7.3 dB.

Noise Figure

The parametric amplifier has an advantage over the transistor amplifier in terms of noise figure, as a pure reactance contributes no thermal noise to the circuit. For a parametric up-converter, the noise figure F is defined as:

$$F = 1 + \frac{2T_d}{T_0}\left[\frac{1}{\gamma Q} + \frac{1}{(\gamma Q)^2}\right] \tag{5.34}$$

Where T_d = Diode temperature in degrees Kelvin

To = 300 K is the ambient temperature in degrees Kelvin

γQ = Figure of merit for the nonlinear capacitor

γQ might be equivalent to 10 in a conventional microwave diode. When f_o/f_s equals 10 and $T_d = 300$ K, the minimum noise figure is 0.90 dB, as determined by Equation (5.34).

Bandwidth

The bandwidth of a parametric up-converter is proportional to the gain-degradation factor of the merit figure and the signal-to-output ratio. The bandwidth formula is as follows:

$$BW = 2\gamma\sqrt{\frac{f_o}{f_s}} \tag{5.35}$$

If $f_o/f_s = 10$ and $\gamma = 0.2$, the Bandwidth (BW) is equal to 1.264.

NOTES

Parametric Down-Converter

If a parametric amplifier's down conversion mode is desired, the signal frequency f_s must match the sum of the pump frequency f_s and the output frequency f_o. This means that the input power must be routed through the idler circuit, while the output power must exit the signal circuit, as shown in Figure 5.27. The down-conversion gain (in reality, a loss) is calculated as follows:

$$\text{Gain} = \frac{f_s}{f_o}\frac{x}{(1+\sqrt{1+x})^2} \tag{5.36}$$

Negative-Resistance Parametric Amplifier

If a sufficient percentage of power flows exclusively at the signal frequency f_s, the pump frequency f_p, and the idler frequency f_i, a regenerative state with the possibility of oscillation at both the signal and idler frequencies occurs. The idler frequency, $f_i = f_p - f_s$, is defined as the difference between the pump and signal frequencies. When the device is operated below the oscillation threshold, it functions as a bilateral negative-resistance parametric amplifier.

Power Gain

The output power is derived from the resistance R_i at a frequency f_i, and the conversion gain from f_s to f_i is calculated as follows:

$$\text{Gain} = \frac{4f_i}{f_s}\cdot\frac{R_g R_i}{R_{Ts}R_{Ti}}\cdot\frac{a}{(1-a)^2} \tag{5.37}$$

Where f_s = Signal frequency

f_p = Pump frequency

$f_i = f_p - f_s$ is the idler frequency

R_g = Output resistance of the signal generator

R_i = Output resistance of the idler generator

R_{Ts} = Total series resistance at s

R_{Ti} = Total series resistance at f;

$a = R/R_{Ts}$

$R = \gamma^2//(\omega_s\omega_i C^2 R_{Ti})$ is the equivalent negative resistance

Noise Figure

The optimum noise figure of a negative-resistance parametric amplifier is expressed as:

$$F = 1 + 2\frac{T_d}{T_0}\left[\frac{1}{\gamma Q} + \frac{1}{(\gamma Q)^2}\right] \tag{5.38}$$

Where γQ = Figure of merit for the nonlinear capacitor

T_o = 300°K is the ambient temperature in degrees Kelvin

T_d = Diode temperature in degrees Kelvin

It is interesting to note that the noise figure given by Equation (5.38) is identical to that for the parametric up-converter in Equation (5.34).

NOTES

Bandwidth

The maximum gain bandwidth of a negative-resistance parametric amplifier is given by

$$BW = \frac{\gamma}{2}\sqrt{\frac{f_i}{f_s \text{ gain}}} \quad (5.39)$$

If gain = 20 dB, $f_i = 4f_s$, and $\gamma = 0.30$, the maximum possible bandwidth for single-tuned circuits is about 0.03.

Degenerate Parametric Amplifier

A negative-resistance amplifier with a signal frequency equal to the idler frequency is referred to as a degenerate parametric amplifier or oscillator. Because the idler frequency f_i is equal to the difference between the pump frequency f_p and the signal frequency f_s, the signal frequency is only half that of the pump frequency.

Power Gain and Bandwidth

A degenerate parametric amplifier's power gain and bandwidth characteristics are identical to those of a parametric up converter. The power transmitted from the pump to the signal frequency is equal to the power transferred from the pump to the idler frequency when $f_s = f_i$ and $f_p = 2f_s$. The total power at the signal frequency is nearly equal to the total power at the idler frequency when the gain is large. As a result, the overall power in the pass band will gain 3 dB.

Noise Figure

In the following equations, the noise figures for a single-sideband and a double-sideband degenerate parametric amplifier are given:

$$F_{ssb} = 2 + \frac{2\bar{T}_d R_d}{T_0 R_g} \quad (5.40)$$

$$F_{dsb} = 1 + \frac{\bar{T}_d R_d}{T_0 R_g} \quad (5.41)$$

Where T_d = Average diode temperature in degrees Kelvin

T_o = 300°K is the ambient noise temperature in degrees Kelvin

R_d = Diode series resistance in ohms

R_g = External output resistance of the signal generator in ohms

The noise figure for double-sideband operation is 3 dB lower than that of single-sideband operation, as can be observed.

Applications

Which type of parametric amplifier to employ is determined by the needs of the microwave system? The upconverter is a single-ended device having a wide bandwidth and a low gain. Negative-resistance amplifiers are by definition bilateral and unstable devices with a limited bandwidth and a high gain. The degenerate parametric amplifier is the simplest form of parametric amplifier since it does not require a separate signal and idler circuit connected via the diode.

The up-converter, in comparison to the negative resistance parametric amplifier, offers the following advantages:

NOTES

1. A positive input impedance
2. Unconditionally stable and unilateral
3. Power gain independent of changes in its source impedance
4. No circulator required
5. A typical bandwidth on the order of 5%

At higher frequencies, when the up-converter becomes impractical, the negative resistance parametric amplifier in conjunction with a circulator becomes the best option. When a system requires a low noise figure, a degenerate parametric amplifier may be the obvious choice, since its double-sideband noise figure is smaller than that of an upconverter or a non-degenerate negative-resistance parametric amplifier. Additionally, the degenerate amplifier is a far simpler device to construct and operates at a very low pump frequency. In radar systems, a negative-resistance parametric amplifier may be preferable, as the system's needed frequency may exceed the X band. However, because the parametric amplifier is complex to fabricate and expensive to manufacture, there is a trend in microwave engineering to replace it in airborne radar systems with the GaAs Metal-Semiconductor Field-Effect Transistor (MESFET) amplifier.

5.4 RADIATIVE AND NON-RADIATIVE TRANSMITTER OR RADIATIVE AND NON-RADIATIVE RECOMBINATION

When electrons are given energy, they absorb it and make the upward transition to a high energy state. Excitations are the transitions of electrons from lower to higher energy states by absorbing external energy. Excitations are the transitions of electrons from lower to higher energy states by absorbing external energy. Electron hole recombination is the passage of electrons from higher energy levels to lower energy ones in semiconductors. Electrons release energy in any form during rest. There are two types of electron hole recombinations:

1. Radiative Recombination.
2. Non-radiative Recombination.

The radiative process is favoured in light emitting devices. We will never be able to eliminate nonradiative recombination, but we can diminish it. A shift from E_1 (of the Valence band) to E_2 is also feasible (of conduction band). Absorption is another name for this transformation. As a result, radiative translation requires the emission or absorption of a photon. These are permissible transitions in direct band gap semiconductors according to K-selection principles. When an electron transitions from one level to another, it is called a radiative transition. Photons are emitted or absorbed as a result of the energy difference between the levels.

$$E = E_2 - E_1 = h\upsilon$$

NOTES

In a semiconductor with a straight band gap. A photon is emitted or absorbed in a radiative transition. Radiative transition is used in every optical device. In contrast to non-radiative transitions, the electron can make the transition without the presence of a photon and the energy is transferred to another form, the most common of which is heat.

The transition energy is changed into forms other than light when phonons are emitted to crystal lattices or electrons are trapped in defects in non-radiative recombination. The transition energy is changed into forms other than light when phonons are emitted to crystal lattices or electrons are trapped in defects in non-radiative recombination. In the case of an indirect transition, the K-value is different if an electron is in the conduction band and a hole is in the valence band. In this case, the K selection criterion is not followed because ΔK is exceedingly large. As a result, non-radiative translation involves phonon emission or absorption.

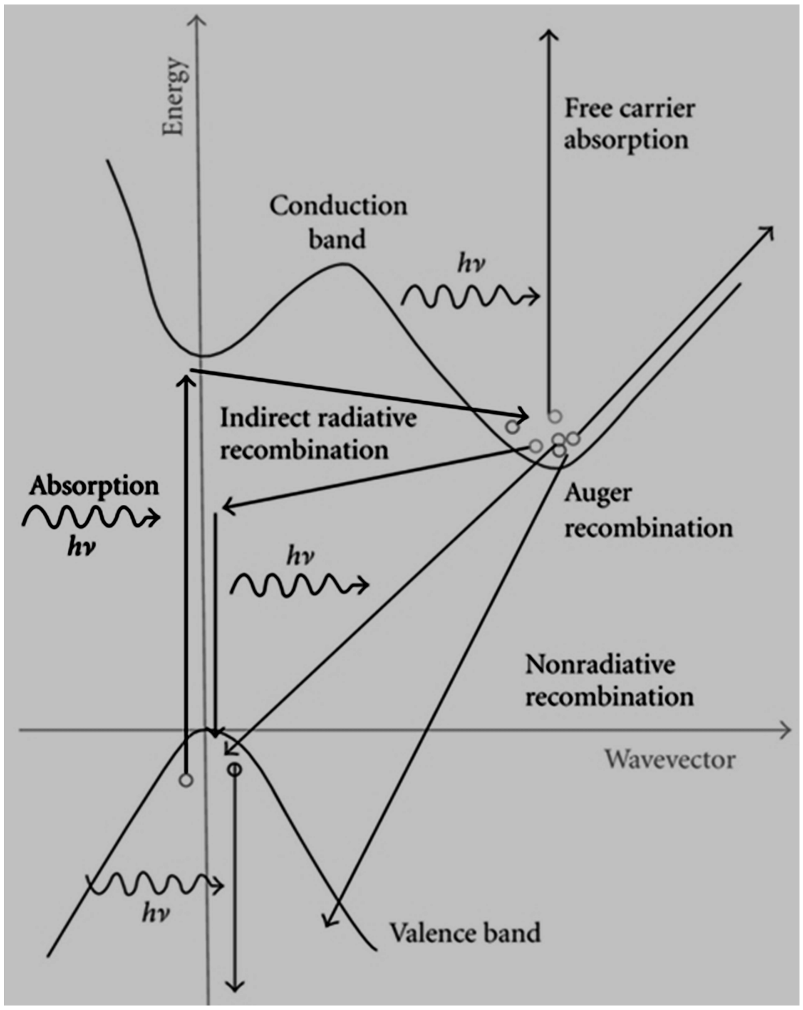


Fig. 5.28 *A Schematic Depiction of Silicon's Energy Bands*

As shown in Figure 5.28, a schematic depiction of silicon's energy bands. Indirect absorption is indicated by the black arrows. Red arrows indicate indirect radiative recombination aided by a phonon. Nonradiative recombination is seen in blue. Auger recombination is shown in green. Absorption of free carriers is shown by the colour orange.

NOTES

1. Radiative Electron-Hole Recombination

The law of mass action states that the product of electron and hole concentrations is constant at a given temperature when the system is in equilibrium.

$$n_0 p_0 = n_i^2 \tag{5.42}$$

Here,

n_0 - Equilibrium Electron Concentration.

p_0 - Equilibrium Hole Concentration.

n_i - Intrinsic Carrier Concentration.

The sum of equilibrium and excess carrier concentrations equals total carrier concentration.

$$n = n_0 + \Delta n \quad and \quad p = p_0 + \Delta p \tag{5.43}$$

R α np is the recombination rate, which is proportional to the product of electron and hole concentrations. The rate of recombination per unit of time per unit volume may therefore be calculated:

$$R = -\frac{dn}{dt} = -\frac{dp}{dt} = Bnp \tag{5.44}$$

This is the **bimolecular rate equation**, where **B** is the **bimolecular recombination coefficient**, which typically ranges from 10^{-11} to 10^{-9} cm^3/s for III-V semiconductors.

2. Radiative Recombination for Low-Level Exciton

Because electrons and holes are created and annihilated in pairs (via recombination), the steady-state hole and electron excess concentrations are identical, i.e., $\Delta n(t) = \Delta p(t)$, And we'll obtain this from Equation 5.44:

$$R = B[n_0 + \Delta n(t)][p_0 + \Delta p(t)] \tag{5.45}$$

The photo-generated carrier concentration is substantially smaller than the majority carrier concentration in the case of low-level excitation, $\Delta n << (n_0 + p_0)$, and we also get that $\Delta n(t) = \Delta p(t)$ from Equation 45.

$$R = Bn_i^2 + B(n_0 + p_0)\Delta n(t) = R_0 + R_{excess} \tag{5.46}$$

The rate equation yields a time-dependent carrier concentration of

$$\frac{dn(t)}{dt} = G - R = (G_0 + G_{excess}) - (R_0 + R_{exces}) \tag{5.47}$$

Excess carriers are created when a semiconductor is irradiated with light, $G_0 = R_0$

$$\frac{d}{dt}\Delta n(t) = -B(n_0 + p_0)\Delta n(t) \tag{5.48}$$

and the result will be,

$$\Delta n(t) = \Delta n_0 e^{-B(n_0+p_0)t} \tag{5.49}$$

NOTES

$$\Delta n(t) = \Delta n_0 e^{-t/\tau} \tag{5.50}$$

We obtain, after rewriting,

where $\tau = [B\,(n_0 + p_0)]^{-1}$ is the carrier lifetime.

The photo produced carrier concentration is substantially lower than the majority carrier concentration in the case of low-level excitation. The photogenerated carrier concentration, on the other hand, is significantly higher than the minority carrier concentration.

3. Radiative Recombination for High-Level Excitation

The photogenerated carrier concentration is greater than the equilibrium carrier concentration for high-level excitation, i.e., $\Delta n >> (n_0 + p_0)$. The bimolecular rate equation is thus

$$\frac{d}{dt}\Delta n(t) = -B\Delta n^2 \tag{5.51}$$

after that we get,

$$\Delta n(t) = \frac{1}{Bt + \Delta n_0^{-1}} \tag{5.52}$$

$$\tau(t) = t + \frac{1}{B\Delta n_0} \tag{5.53}$$

Here, the time constant will be,

With time, the lifespan of minority carriers rises. Low-level excitation conditions will be reached after a sufficiently long time and will approach the low-level value.

4. Non-Radiative Recombination in the Bulk

The electron energy is transformed to vibrational energy of lattice atoms, i.e., phonons, during non-radiative recombination. As a result, the electron energy is transformed into heat. Defects in the crystal structure are the most common cause of non-radiative recombination events. Unwanted foreign atoms, native flaws, and dislocations are examples of these consequences. All of these flaws have a different energy level structure than significant semiconductor atoms. And it's not uncommon for such flaws to produce one or more energy levels within the semiconductor's forbidden gap.

Energy levels within the semiconductor gap are effective recombination centres, especially when the energy level is near the centre of the gap. When an electron falls into a 'Trap,' which is an energy level within the bandgap induced by the presence of a foreign atom or a structural defect, trap-assisted recombination occurs. When the trap is full, it is unable to receive any more electrons. In a second step, the electron in the trap falls into an empty valence band state, completing the recombination process.

The rate of non-radiative recombination at a deep level with trap energy E_T and concentration N_T is as follows:

$$R_{SR} = \frac{p_0 \Delta n + n_0 \Delta p + \Delta n \Delta p}{(N_T v_p \sigma_p)^{-1}(n_0 + n_1 + \Delta n) + (N_T v_n \sigma_n)^{-1}(p_0 + p_1 + \Delta p)} \tag{5.54}$$

$\Delta n = \Delta p,$

v_n and v_p - are the electron and hole velocities, respectively. σ_n and σ_p - capture cross section of the traps n_1 and p_1 - electron and hole concentrations if the Fermi energy is located at the trap levels:

$$n_1 = n_i exp\left(\frac{E_T - E_{Fi}}{k_B T}\right) \quad p_1 = n_i exp\left(\frac{E_{Fi} - E_T}{k_B T}\right) \tag{5.55}$$

$$\frac{1}{\tau} = \frac{p_0 + n_0 p + \Delta n}{(N_T v_p \sigma_p)^{-1}(n_0 + n_1 + \Delta n) + (N_T v_n \sigma_n)^{-1}(p_0 + p_1 + \Delta p)} \tag{5.56}$$

E_{Fi} - Fermi level in the intrinsic semiconductors.

$R_{SR} = \Delta n / \tau$ may be used to calculate the non-radiative lifetime of surplus electrons, which gives

If we assume a p-type semiconductor, in which the holes are in the majority $p_0 >> n_0$ and $p_0 >> p_1$, and a slight deviation from equilibrium i.e., $\Delta n << p_0$ then the **minority carrier life**time is given by

$$\frac{1}{\tau} = \frac{1}{\tau_{n0}} = N_T v_{n\sigma n} \tag{5.57}$$

If the electrons were majority carriers, the lifetime would be

$$\frac{1}{\tau} = \frac{1}{\tau_{p0}} = N_T v_p \sigma_p \tag{5.58}$$

The rate of Shockley -Read recombination is restricted by the rate of minority carrier capture, because majority carrier capture is a far more probable event than minority carrier capture. We will get the following result from Equation 5.56:

$$\frac{1}{\tau} = \frac{p_0 + n_0 + \Delta n}{\tau_{p0}(n_0 + n_1 + \Delta n) + \tau_{n0}(p_0 + p_1 + \Delta p)} \tag{5.59}$$

For slight deviation from equilibrium, i.e $\Delta n << p_0$, will result in

$$\tau = \tau_{n0}\frac{p_0 + p_1}{p_0 + n_0} + \tau_{p0}\frac{n_0 + n_1 + \Delta n}{p_0 + n_0} \approx \tau_{n0}\frac{p_0 + p_1}{p_0 + n_0} \tag{5.60}$$

In an extrinsic semiconductor, tiny deviations from equilibrium have no effect on the lifetime.

$$\tau = \tau_{n0}\left(1 + \frac{p_0 + p_1}{p_0 + n_0}\right) \quad \textbf{(5.61)}$$

NOTES

We obtain Equation 5.60 by assuming that the trap catches electrons and holes at the same rate, , i.e., $v_n\sigma_n = v_p\sigma_p$ and $\tau_{n0} = \tau_{p0}$.

$$\tau_i = \tau_{n0}\left(1 + \frac{p_1 + n_1}{2n_i}\right) = \tau_{n0}\left[1 + cosh\left(\frac{E_T - E_{Fi}}{k_BT}\right)\right] \quad (5.62)$$

The equation becomes, i.e., $n_0 = p_0 = n_i$ for the exceptional case of intrinsic material.

E_{Fi} - the Fermi level, which is often towards the middle of the gap.

When the traplevel is at or near the midgap energy, the non-radiative life duration is minimized if if E_T -E_{Fi} is zero. The lifetime for such midgap levels is given by $\tau = 2\tau_{n0}$. Deep levels are effective recombination centers if they are near the centre of the gap, as evidenced by this.

The non-radiative life time is minimized, if E_T -E_{Fi} is zero, when the traplevel is at or close to the midgap energy. For such midgap levels the lifetime is given by $\tau=2\tau_{n0}$ This shows that deep levels are effective recombination centers if they are near in the middle of the gap.

Auger recombination is another non-radiative recombination in which kinetic energy is transferred to another electron.

Excitation of a free electron high into the conduction band, or a hole profoundly excited into the valence band, dissipates the energy that becomes accessible through electron-hole recombination (E_g). After that, the highly excited carrier will lose energy through numerous phonon emission until they reach the band edge.

Recombination rate due to the Auger process

$$R_{Auger} = C_p np^2 \quad \textbf{(5.63)}$$

and

$$R_{Auger} = C_n n^2 p \quad \textbf{(5.64)}$$

Only at very high excitation intensities or very high carrier injection currents does Auger recombination diminish the luminescence efficiency in semiconductors. The Auger recombination rate is relatively tiny at low carrier concentrations and can be ignored.

5. Non-Radiative Recombination at Surfaces

Due to the absence of nearby atoms, atoms at the surface cannot have the same bonding structure as atoms in the bulk. As a result, some valence orbitals fail to form a chemical connection. These partially filled electron orbitals, also known as dangling bonds, are electronic states that can be found in the semiconductor's forbidden gap and operate as recombination centers. Surface recombination causes a reduction in luminescence efficiency as well as surface heating due to non-radiative recombination. In electro luminescent devices, both effects are undesirable.

NOTES

Only when both types of carriers are present can surface recombination occur. The carrier-injected active zone, in which both types of carriers are present, must be far away from any surface in the design of LEDs. Carrier injection under a contact much smaller than the semiconductor chip can archive this.

6. Competition between Radiative and Non-Radiative Recombination

Non-radiative bulk recombination and Auger recombination, like surface recombination, can never be completely avoided. Native faults will exist in any semiconductor crystal. It's also challenging to make materials with impurity levels below a few Parts Per Billion (PPB). As a result, even the cleanest semiconductors have impurities in the $10^{12}cm^{-3}$ range.

The sum of the radiative and non-radiative probabilities determines the total likelihood of recombination.

$$\tau^{-1} = \tau_r^{-1} + \tau_{nr}^{-1} \tag{5.65}$$

where τ_r and τ_{nr} are the radiative and non-radiative lifetimes, respectively.

$$\eta_{int} = \frac{\tau_r^{-1}}{\tau_r^{-1} + \tau_{nr}^{-1}} \tag{5.66}$$

Radiative probability divided by total probability of recombination gives the relative probability of radiative recombination. As a result, the likelihood of radiative recombination or internal **quantum efficiency** can be calculated as follows:

The ratio of the number of light quanta emitted inside the semiconductor to the number of charge quanta undergoing recombination is known as the internal quantum efficiency. Due to the light escape problem, reabsorption in the substrate, or following reabsorption mechanism, not all photons emitted internally may escape from the semiconductor.

5.7 LIGHT DEPENDENT RESISTOR (LDR)

A Light Dependent Resistor (also called a photoresistor or LDR) is a resistive device whose resistance varies in response to incoming electromagnetic radiation. As a result, they are photosensitive devices. Additionally, they are referred to as photoconductors, photoconductive cells, or just photocells.

They are constructed using high-resistance semiconductor materials. There are several symbols used to denote a photoresistor or light-dependent resistor; one of the most often used is seen in the image below. The arrow denotes the direction of light falling on it.

NOTES

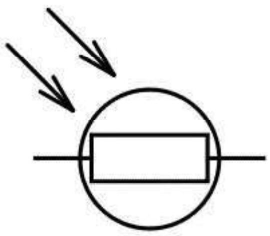


Fig. 5.29 Symbol of a Photoresistor (or LDR)

Principle of Operation

Photoresistors operate on the photoconductivity concept. Photoconductivity is an optical phenomenon in which the conductivity of a substance increases as it absorbs light.

When light strikes a device, i.e., when photons strike the device, the electrons in the semiconductor material's valence band are stimulated to the conduction band. To cause electrons to jump from the valence band to the conduction band, these photons in the incoming light must have an energy larger than the bandgap of the semiconductor material.

Thus, when sufficiently energetic light impacts the device, an increasing number of electrons are stimulated to the conduction band, resulting in a high number of charge carriers. As a result of this process, increasing current begins to flow through the device when the circuit is closed, implying that the device's resistance has lowered. This is the most often used LDR operating concept.

Characteristics of Photoresistor (LDR)

Photoresistor LDRs are light-dependent devices whose resistance decreases in the presence of light and increases in the absence of light. When a resistor that is light dependent is kept in the dark, its resistance is extremely high. This is referred to as dark resistance. It may be as high as 1012, and allowing the gadget to absorb light significantly reduces its resistance. When a steady voltage is supplied and the light intensity is raised, the current begins to grow. The resistance vs. illumination curve for a specific LDR is depicted in the image below.

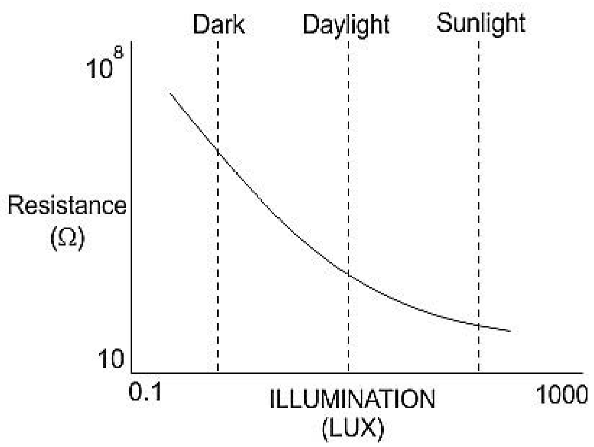


Fig. 5.30 A characteristic curve illustrating the fluctuation in resistance as a function of light.

NOTES

Photocells or light-dependent resistors are nonlinear devices. Their sensitivity changes according on the wavelength of incoming light. Certain photocells may be completely insensitive to a certain wavelength range. Different cells exhibit a range of spectral response curves depending on the material utilized.

When light is incident on a photocell, it typically takes between 8 and 12 milliseconds for the resistance to change, whereas it takes one or more seconds for the resistance to return to its initial value when the light is removed. This is referred to as the resistance recovery rate. Compressors of audio make use of this characteristic.

Additionally, LDRs have a lower sensing capability than photodiodes and phototransistors. (A photo diode and a photocell (LDR) are not synonymous; a photo diode is a p-n junction semiconductor device that converts light to electricity, whereas a photocell is a passive device that has no p-n junction and does not 'Convert' light to electricity.)

Types of Light Dependent Resistors

According to the materials used to create them, photoresistors (LDRs) can be classified into two classes. There are two types of photoresistors:

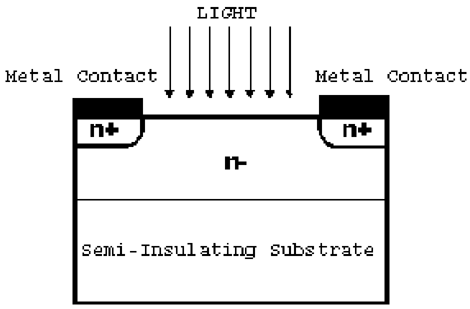
1. **Intrinsic photoresistors (Undoped semiconductor):** These are constructed entirely of semiconductor materials such as silicon or germanium. Electrons are stimulated from the valance band to the conduction band when photons with sufficient energy strike them, increasing the number of charge carriers.
2. **Extrinsic photoresistors:** Extrinsic photoresistors are semiconductor materials that have been doped with impurities known as dopants. These dopants form new energy bands above the electron-filled valence band. As a result, the band gap narrows and less energy is required to excite them. Generally, extrinsic photo resistors are employed for long wavelengths.

Construction of LDR

A light-dependent resistor is constructed by depositing a light-sensitive substance on an insulating substrate such as ceramic. To achieve the necessary resistance and power rating, the material is placed in a zigzag pattern. This zigzag pattern divides the metal deposition sites into two distinct sections.

Then, on both sides of the region, ohmic connections are established. The resistances of these contracts should be kept to a minimum to ensure that they vary mostly as a result of the action of light. Cadmium sulphide, cadmium selenide, indium antimonide, and cadmium sulphide are frequently used materials. Lead and cadmium are not used because they are toxic to the environment.

NOTES



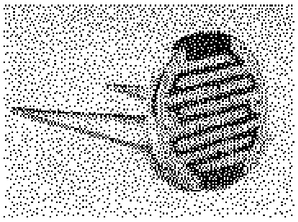


Fig. 5.31 *Schematic Representation for LDR.*

Applications of LDRs

Due to their inexpensive cost and simple construction, photoresistors or LDRs are frequently utilized as light sensors. Additional applications for photoresistors include the following:

1. Detect the absence or presence of light in the same way as a camera light meter does.
2. Used in the design of street lighting (can be combined with a good Arduino starter kit to act as a street light controller)
3. Alarm clocks
4. Burglar alarm circuits
5. Light intensity meters
6. Used as a component of a supervisory control and data acquisition system to conduct tasks like as counting the number of parcels on a moving conveyor belt.

5.4.2 Photodiode Detectors

A photodetector is a device which absorbs light and converts the optical energy to measurable electric current. Detectors are classified as

1. Thermal Detectors
2. Photon Detectors

Thermal Detectors

When sun shines on a device, it raises its temperature, which alters the device's electrical properties, such as its electrical conductivity. Examples of thermal detectors are thermopile (which is a series of thermocouples), pyroelectric detector etc.

Photon Detectors

Photon detectors follow the principle of photons to electrons conversion. Unlike the thermal detectors, such detectors are based on the rate of absorption of photons rather than on the rate of energy absorption. However, a device may absorb photons only if the energy of incident photons is above a certain minimum threshold. Photon detectors, in terms of the technology, could be based on:

i. Vacuum Tubes - e.g., Photomultipliers
ii. Semiconductors - e.g., Photodiodes

For optical fiber applications, semiconductor devices are preferred because of their small size, good responsivity and high speed.

Physical Processes in Light Detection

NOTES

Detection of radiation is essentially a process of its interaction with matter. Some of the prominent processes are

1. Photoconductivity
2. Photovoltaic Effect
3. Photoemissive Effect

Photoconductivity

Photoconductivity is an optical and electrical phenomenon in which electromagnetic radiation such as visible light, ultraviolet light, infrared light, or gamma radiation causes a substance to become more electrically conductive. The number of free electrons and holes rises when light is absorbed by a substance like a semiconductor, resulting in improved electrical conductivity. The light that reaches the semiconductor must have enough energy to raise electrons across the band gap or excite the impurities within the band gap in order to generate excitation.

Because semiconductors have a tiny band gap, it is possible to produce extra carriers by lighting a sample with light with a frequency greater than. This causes the sample's conductivity to rise, a phenomenon known as intrinsic photoconductivity. Except when the lighting is by an intense beam of light, the impact is not very evident at high temperatures. At low temperatures, illumination causes localized carriers to be excited to the conduction or valence band.

Even if an incident photon does not have enough energy to make an electron-hole pair, it can nevertheless excite impurity centers by forming a free electron-bound hole pair (for excitation at the donor level) or a free hole-bound electron pair (for excitation at the acceptor level) (for acceptor level). If E_i is the impurity ionization energy, the extrinsic photoconductivity radiation frequency should be at least E_i/h.

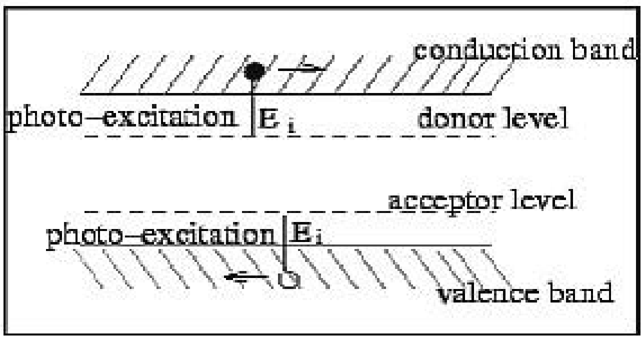


Fig. 5.32 *The Schematic Representation of Photoexcitation Process.*

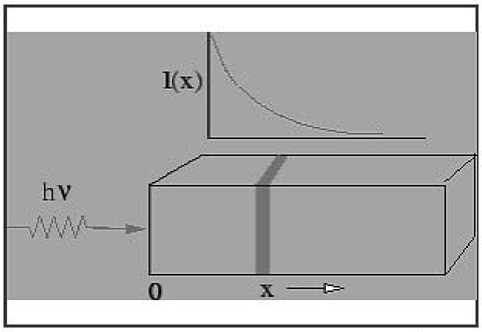
Consider a thin slab of semiconductor that is lit by a laser beam that propagates down its length (x-direction). Let I denote the intensity of the radiation (in watts/m^2) at a point x on the semiconductor. The power absorbed per unit length is αI if α = absorption coefficient per unit length. The intensity changes with distance along the sample length as follows:

NOTES

$$\frac{dI}{dx} = -\alpha I$$

which has solution

$$I = I_0 e^{-\alpha x}$$



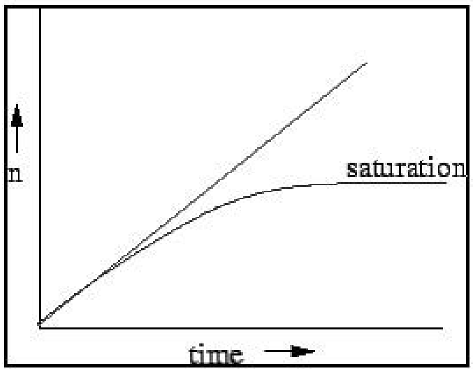


Fig. 5.33 *Schematic Representation of Laser Beam Propagation*

As shown in Figure 5.33, the schematic representation of laser beam propagation through a thin slab of semiconductor and corresponding graph of electron pair generated with time.

The number of pairs created per unit time is given by: If we define η as the quantum efficiency, which is the proportion of absorbed photons that produce electron-hole pairs, then the number of pairs produced per unit time is given by:

$$\Delta n = \Delta p = \frac{\eta \alpha I}{h\nu}$$

In theory, as the quantity of energy absorbed (and thus "n and "p) increases linearly with time, the illumination process will result in a continuing increase in the number of carriers. The excited couples, on the other hand, have a finite life duration (~10^{-7} to 10^{-2} s). Recombination of the pairs occurs as a result of this. As a pair is necessary in the procedure, the relevant life time is that of minority carriers. Recombination ensures that the number of extra carriers does not continue to grow endlessly, but instead reaches a saturation point.

NOTES

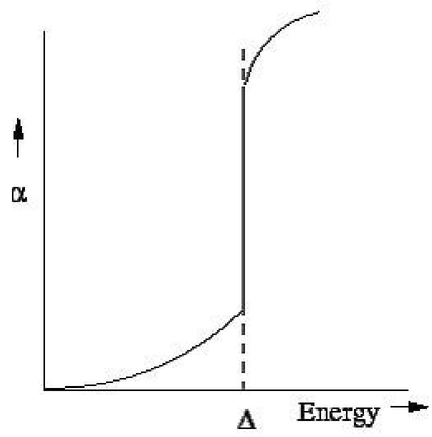


Fig. 5.34 *Variation of Absorption Coefficient (á) with Bandgap (Ä).*

Take the case of an n-type semiconductor. If the minority carriers' recombination life time is τ_p, the rate of change in carrier concentration is given by:

$$\frac{d}{dt}(\Delta p) = \frac{\eta \alpha I}{h\nu} - \frac{\Delta p}{\tau_p}$$

Under steady state condition dδp/dt = 0, which gives

$$\delta p = \frac{\eta \alpha I_p \tau_p}{h\nu} = \Delta n$$

This excess hole density leads to an additional conductivity

$$\Delta\sigma = q\delta n\mu_e + q\Delta p\mu_h$$

Photovoltaic Effect

In a material with a space charge layer, such as a p-n junction, the photovoltaic effect can occur. The detector material can absorb a photon with enough energy to excite an electron from the valence band to the conduction band. The contribution of the excited electron to the current can be seen. Without the use of a bias voltage, a photovoltaic detector can be used.

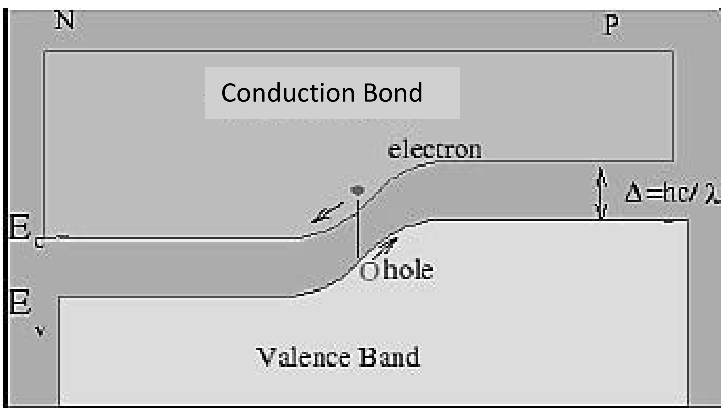


Fig. 5.35 *Schematic Illustration of Photoelectric Effect.*

NOTES

Photoemissive Process

Incident radiation generates electron emission from the photocathode, which is collected by an anode in a photoemissive process (also known as external photo effect). Photoemissive detectors have a competitive advantage over other detectors in terms of speed, gain, and noise. However, because the input photon must have enough energy to expel electrons from the photocathode, their spectrum range is limited. In the ultraviolet region, photoemissive detectors are thus a suitable choice.

Performance Parameters

A detector's performance is measured in terms of particular figures of merit.

Responsivity:

The ratio of the generated photocurrent (I) to the quantity of optical power (P_o) incident on the detector is the detector's sensitivity.

$$\mathcal{R} = \frac{I}{P_0}$$

Amperes/watt is the unit of responsivity.

Quantum Efficiency

A detector's ability to collect all photons and convert them to electron-hole pairs is limited. The quantum efficiency, which is commonly stated as a percentage, is defined as the number of electrons created per incident photon.

$$\eta = \frac{\text{No. of electrons produced}}{\text{No. of incident photons}} (\times 100\%)$$

If I = photocurrent in the external circuit and P_o = the incident optical power (dropping the percentage in the definition)

$$\eta = \frac{I/q}{P_0/h\nu}$$

Using this in the expression for the responsivity, we get

$$\mathcal{R} = \frac{I}{P_0} = \frac{q\eta}{h\nu} = \frac{q\eta}{hc}\lambda$$

The responsivity, therefore, depends on the wavelength λ. For an ideal photodetector η = 1, and R is linear with λ.

Spectral Response:

The spectral response of a detector is determined by how the output signal of the detector changes as the wavelength of incident radiation changes. Because quantum efficiency is proportional to wavelength, the response is not linear, as it would be if η = 1. The photon's energy must be high enough to stimulate an electron past the energy barrier ". The maximum wavelength that the detector will respond to if " is in eV is:

$$\lambda_{\max} \text{ (in nm)} = \frac{1240}{\Delta \text{ (ev)}}$$

However, the response does not fall off abruptly to zero for values of λ above the threshold. This is because, due to thermal energy of the molecules, the absorption coefficient α of the material of the device is found to be given by:

$$\alpha = \alpha_0 e^{E/\Delta}$$

NOTES

where E is the incident photon energy. For $\lambda > \lambda_{max}$, E < " so that the absorption of radiance becomes smaller.

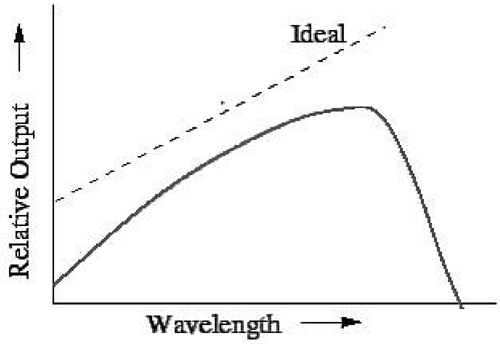


Fig. 5.36 *Variation of Relative Output vs Wavelength.*

Noise Equivalent Power:

Thermal fluctuation is a source of noise in a detector. Charged particles are constantly in motion. Even when no radiation is incident on a device, a background current is formed, which can be measured in nano-amperes or pico-amperes. This is referred to as dark current. The power of the signal must be greater than the noise signal in order for a detector to distinguish between such random noise and an incoming signal. The Signal to Noise Ratio (SNR) is defined in a detector design as:

$$\text{SNR} = \frac{Signal\ power}{Noise\ power}$$

The Noise Equivalent Power (NEP) of a detector is an important figure of merit. Due to noise effects, NEP is defined as the rms incident power that causes a current (or voltage) whose rms value is equal to the rms value of the current (voltage).

The NEP is usually provided for a detector at a specific wavelength and temperature. The bandwidth of incoming radiation used to measure NEP is typically set to 1 Hz. Within a bandwidth of "f, noise power is predicted to be proportional to Δf. Since the current (voltage) is proportional to the square root of the power, the noise current (voltage) is proportional to $\sqrt{\Delta f}$. The unit of NEP is, therefore, watts/$\sqrt{Hz}$.

Detectivity (D*)

Although some definitions distinguish between the two, they are frequently used interchangeably. D* is essentially the inverse of NEP normalized to the detector's unit area.

$$D^* = \frac{\sqrt{A}}{\text{NEP}}$$

The unit of D* is m-(Hz)$^{1/2}$ /w. (Detectivity is often defined as the inverse of NEP)

Photodiode

NOTES

The photo diode is a p-n junction semiconductor device that operates only in the reverse biased region. The photodiode sign is depicted in the diagram below.

Fig. 5.37 *Circuit symbol for photodiode.*

A photodiode is a p-n junction diode that absorbs photons and produces either a photo voltage or free carriers that cause photocurrent. They are used to detect optical signals as well as to convert optical power to electrical power. A p^+n junction diode with a strongly doped p^- side is shown in the figure. On the n side of the junction, the donor concentration is lower than on the p^+ side, where the acceptor concentration is higher. Thermal diffusion or ion implantation on n-type silicon form the very thin p^- layer on the device's front surface. The active area is coated with an antireflection coating (such as silicon nitride) to trap the majority of the light that falls on the device. The terminals are provided by metallized contacts.

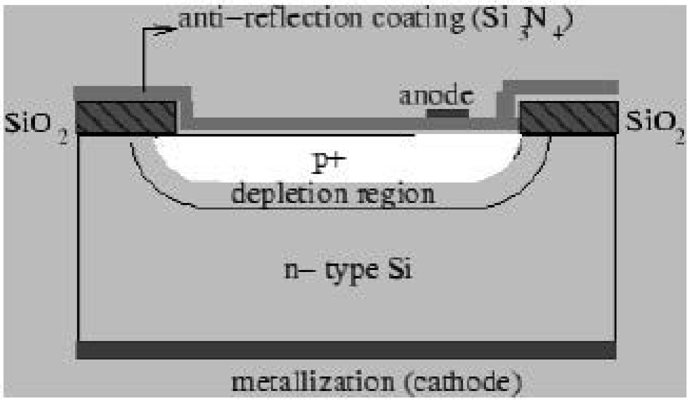


Fig. 5.38 *Schematic Illustration of Construction of Photodiode.*

Silicon is the most popular photodiode material. Its maximal sensitivity is in the I.R. between 800 and 950 nm, with a band gap of 1.1 eV. At shorter wavelengths, sensitivity decreases. Before reaching the junction, light for $\lambda < 700$ nm is absorbed in the p^- layer. As a result, the breadth of the p^- layer should be decreased to maximise sensitivity at shorter wavelengths. Because the p^- type region has more holes than the n^- type region, the holes diffuse to the n^- side and the electrons to the p^-side, resulting in an electric field gradient from the n-side to the p^- side. The built-in electric field is strong enough to prevent charges from moving farther through the depletion region. The depletion zone reaches all the way to the weakly doped n^- side. If the n^- side has a donor density of N_d per unit volume and the p^- side has an acceptor density of N_a per unit volume, the two sides lose an equal number of mobile carriers, leaving fixed charges on the p^+ and

NOTES

sides. As may be seen, the charge density distribution is as follows. The condition of charge neutrality requires

$$qN_a x_p = qN_d x_n$$

where q is the magnitude of electronic charge and x_p and x_n are respectively the widths of the depletion region in the p^- side and n-side.

One can determine the electric field on both sides by using Gauss's law of electrostatics, $\vec{\nabla}.\vec{E} = \rho/\varepsilon_o K$, where K is the dielectric constant and ρ is the charge density.

For $d < x < d + x_p$, $\rho = -N_a q$, so that

$$E(x) = \int_d^x \frac{\rho}{\epsilon_0 K} dx = -\frac{N_a q}{\epsilon_0 K}(x - d)(A)$$

$E(x = d) = 0$ $\quad d + x_p < x < d + x_p + x_n \quad \rho = N_d q$

$$E(x) = \frac{N_d q}{\epsilon_0 K}(x - d - x_p - x_n)(B)$$

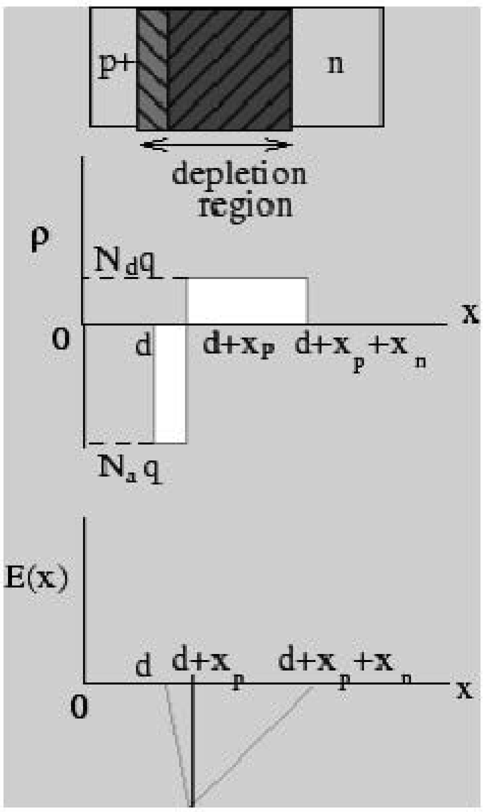


***Fig. 5.39** Variation of Charge Density (ñ) and X-Component of Electric Field in Depletion Region.*

NOTES

The maximum magnitude of the field occurs at x = d + x_p and is given by

$$E_0 = -\frac{N_a q x_p}{\epsilon_0 \kappa} = -\frac{N_d q x_n}{\epsilon_0 \kappa}$$

The requirement of charge neutrality makes the two formulations for E_o equal. The negative sign implies that the electric field is moving from the n to the p side.

Built in Potential

One can obtain an expression for the potential drop across the junction by integrating the electric field:

$$V = V(d + x_p + x_n) - V(d) = -\int_d^{d+x_p+x_n} E(x)dx$$

to obtain

$$V = \frac{q}{2\epsilon_0 \kappa}(N_a x_p^2 + N_d x_n^2)(C)$$

Sources of Noise in Detectors:

The presence of noise superimposed on the current or voltage output makes signal detection difficult. Some of these noises are caused by the photon field itself, while others are caused by the circuit. There are three main sources of noise in photon detectors.

Dark Noise

Optical power detection is made up of streams of photons that come at random intervals. A Poisson distribution is used to predict the number of photons that will arrive in a particular time interval. As a result, the signal fluctuates at random. There is no method to lessen or eliminate noise because it is statistical. Even in the absence of optical illumination, dark noise, or dark current shot noise, is caused by heat production of the e-h pair. An estimate of the black noise is given by for large quantities of reverse bias.

$$I_{n-d} = \sqrt{2qI_sB}$$

where I_s is the reverse bias saturation current in dark and B is the bandwidth of operation.

Shot Noise

Shot noise is also associated with the signal (photocurrent) itself because of statistical nature of generation of electron hole pairs due to photon absorption. The signal noise current is estimated by the formula

$$I_{n-s} = \sqrt{2qI_{ph}B}$$

where I_{ph} is the (signal) photocurrent.

NOTES

Johnson Noise

Johnson noise occurs when electrons in a resistive element move due to thermal motion, regardless of the supplied voltage. Noise is inherent in all resistors and is not the result of bad design or manufacturing. The formula for calculating the rms value of Johnson noise at temperature T is:

$$I_{n-J} = \sqrt{4kTB/R_{sh}}$$

where R is the shunt resistance.

Combining the above the noise the Signal to Noise Ratio (Power SNR) is given by:

$$SNR = \frac{I_{ph}^2}{I_{n-d}^2 + I_{n-s}^2 + I_{n-J}^2}$$

If the photocurrent is very high, the effect of dark noise and thermal noise may be neglected and the SNR is given by

$$SNR = \frac{I_{ph}^2}{I_{n-s}^2}$$

Principle of Operation

A photodiode is a type of photo detector that, depending on the mode of operation, can convert light into current or voltage. A large area photodiode is a common, classic solar cell used to generate electric solar power. A photodiode is made to work with a reverse bias. The deletion region has a considerable breadth. Due to a small number of charge carriers, it transports a modest amount of reverse current under normal conditions. When light strikes the p-n junction through a glass window, photons in the light attack the p-n junction, imparting energy to the valence electrons. As a result, valence electrons detach themselves from covalent bonds and become free electrons. As a result, more electron-hole pairs are formed. As a result, the overall number of minority charge carriers rises, and reverse current rises as well. This is the fundamental premise of photodiode operation.

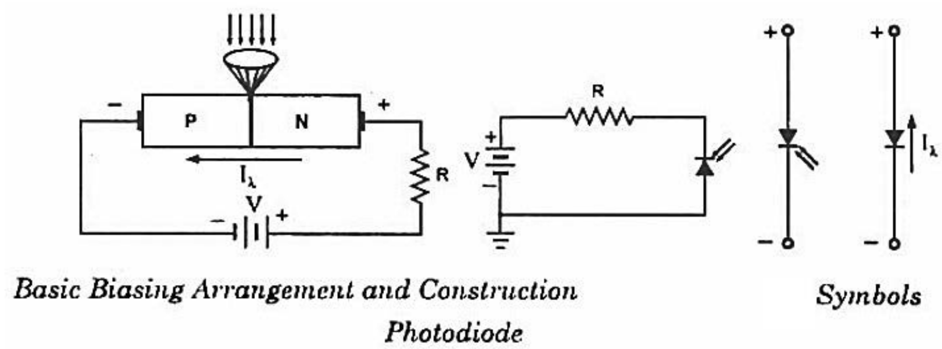


Fig. 5.40 *Biasing Arrangement, Construction and Circuit Symbol of Photodiode.*

Characteristics of Photodiode

When the p-n junction is reverse-biased, a reverse saturation current flows as the minority carriers are swept across the junction as thermally produced holes and

NOTES

electrons. As the temperature of the junction rises, more and more hole-electron pairs are formed, increasing the reverse saturation current I_0. Illuminating the junction produces the same effect. When light energy bombards a p-n junction, valence electrons are dislodged. The reverse current in the diode increases as more light strikes the junction. It's because when the degree of illumination rises, more and more charge carriers are generated. This is depicted in *figure* for various intensity levels. The current that exists when no light is incident is known as the dark current. It's worth noting that current becomes zero only when the applied bias is positive and equals V_Q. As illustrated in figure, the reverse saturation current I_o increases linearly with the luminous flux due to the approximately similar spacing between the curves for the same rise in luminous flux. Because all potential charge carriers have already been swept across the junction, an increase in reverse voltage has no effect on reverse current. It is essential to forward bias the junction by an amount equal to the barrier potential in order to reduce the reverse saturation current I_o to zero. As a result, the photodiode can function as a photoconductive device.

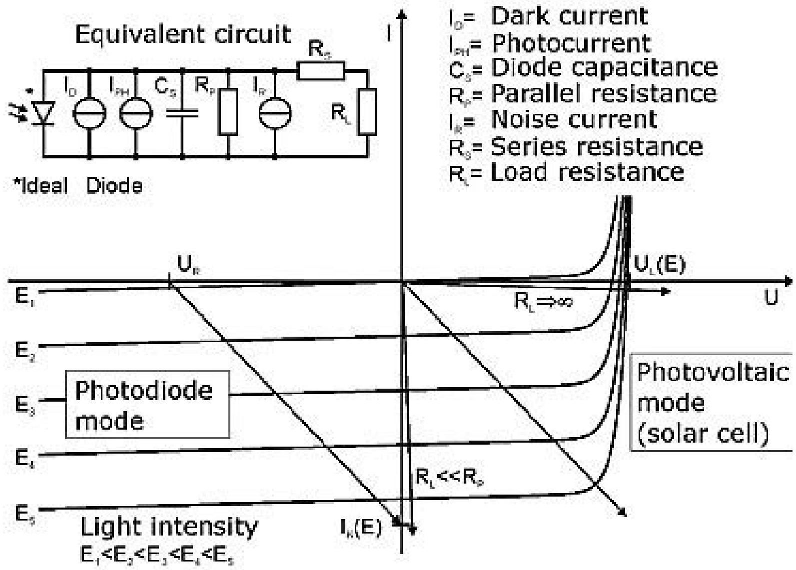


Fig. 5.41 *I-V Characteristics of Photodiode*

Minority charge carriers continue to be swept across the junction when the reverse bias applied across the photodiode is removed while the diode is lit. This raises the concentration of holes on the P-side and the concentration of electrons on the N-side. The barrier potential, on the other hand, is negative on the P side and positive on the N side, and was formed during junction construction by holes travelling from the P to the N side and electrons moving from the N to the P side. As a result, the flow of minority carriers tends to lower the potential barrier.

The minority carrier returns to the original side via the external circuit when an external circuit is connected across the diode terminals. The electrons that crossed the p-n junction now flow out of the N-terminal and back into the P-terminal. This indicates that the device is acting as a voltage cell, with the negative terminal on the N-side and the positive terminal on the P-side. As a result, the photodiode is both a photovoltaic and a photoconductive device.

Advantages

The following are some of the benefits of using a photodiode:

1. It can be utilized as a variable resistance device.
2. Extremely light-sensitive.
3. The rate of operation is really fast.

Disadvantages

The following are some of the disadvantages of using a photodiode:

1. Dark current that is temperature dependent.
2. Inconsistent temperature.
3. Current must be amplified in order to drive further circuits.

Applications

The basic application of photodiode is as:

1. An alarm system.
2. A tally system.

5.4.3 Solar Cells (Open Circuit Voltage, Short Circuit Element, and Fill Factor)

Solar cells are photovoltaic devices that turn light into electricity. Photovoltaic devices, such as solar cells, produce voltage when exposed to light. In 1839, Alexander-Edmond Becquerel discovered the photovoltaic effect in a junction created between a platinum electrode and an electrolyte (silver chloride). The photodiode's function is comparable to that of a solar cell (photodetector). An unbiased photodiode connected to a load is used (impedance). Solar cells and photodetectors have three fundamentally different characteristics:

1. For solar cells to be effective, they must be able to operate over a wide spectrum of wavelengths, as opposed to photodiodes (solar spectrum).
2. In order to optimize exposure, solar cells are typically large area devices.
3. Wh ile in photodiodes, Quantum Efficiency (QE) is used to measure Signal-to-Noise Ratio (SNR), in solar cells it is used to measure Power Conversion Efficiency (PCE). It is common practice for solar cells and the external load they are linked to be constructed to optimize the amount of power they produce.

The usual range of wavelengths in the solar spectrum is from 3 μm to 0.2 μm (i.e., the IR to the UV region). However, the level of intensity isn't the same everywhere.

Working Principle of Solar Cell

A p-n junction diode is a basic solar cell. As shown in Figure 5.42 depicts the device's schematic. The n-region is strongly doped and thin, allowing light to easily pass through it. Because the p-region is sparsely doped, the majority of the depletion region is on the p-side. The wavelength affects penetration, and the

NOTES

absorption coefficient increases as the wavelength lowers. Electron Hole Pairs (EHPs) are mostly formed in the depletion zone, where electrons flow to the n-region and holes to the p-region due to the built-in potential and electric field. When an external load is applied, the excess electrons are forced to travel through the load in order to recombine with the excess holes. Figure 5.42 shows that the p and n-regions also create electrons and holes. Shorter wavelengths (with a higher absorption coefficient) absorb in the n-region, while longer wavelengths absorb in the majority of the p-region. Some of the EHPs produced in these places may also contribute to the current situation. These are typically EHPs produced within the minority carrier diffusion length, L_e for electrons on the p-side and L_h for holes on the n-side. Carriers created in this region have the potential to disperse into the depletion region and contribute to the current. As a result, the total width of the zone contributing to solar cell current is $w_d + L_e + L_h$, where wd is the depletion width shown in Figure 5.43 depicts this metal electrodes on either side remove the carriers. To achieve the electrical contact, a finger electrode is employed on the top, leaving enough surface area for the light to enter. As shown in Figure 5.44 depicts the top electrode configuration. Consider a Si-based solar cell. Because the band gap, Eg, is 1.1 eV, wavelengths above 1.1 μm are not absorbed because the energy is less than the band gap. As a result, absorption is negligible at wavelengths higher than 1.1 μm. The absorption coefficient is quite high for wavelengths much lower than 1.1 μm, and EHPs are formed near the surface and can become stuck near surface imperfections. As shown in Figure 5.43, there is an ideal range of wavelengths where EHPs can contribute to photocurrent.

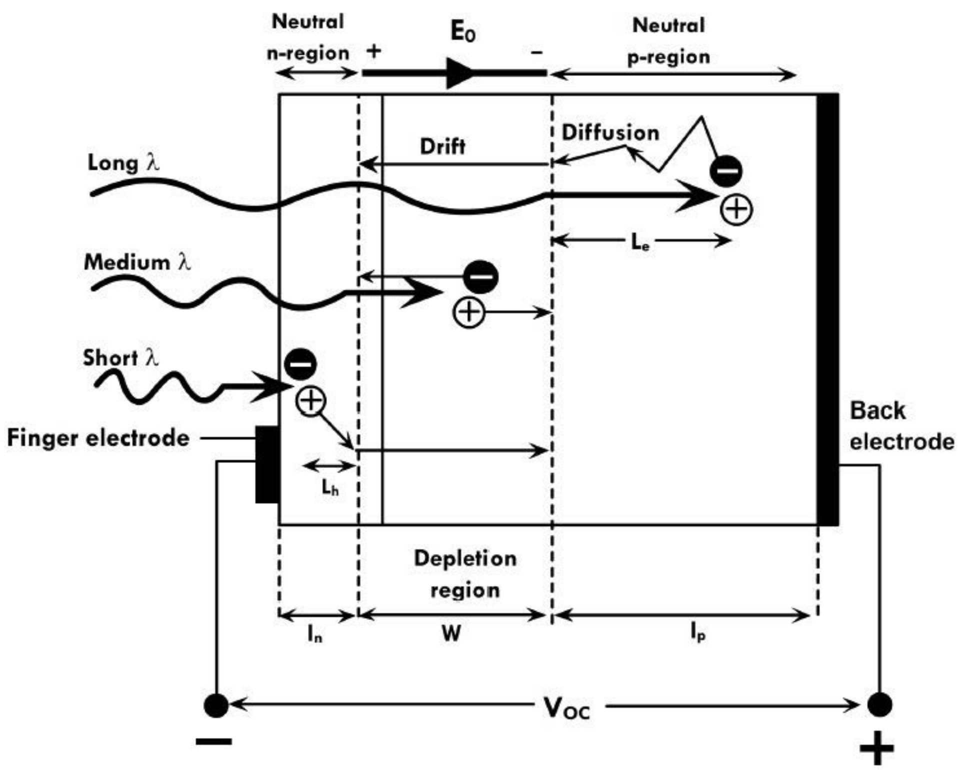


Fig. 5.42 *Principle of Operation of a p-n Junction Solar Cell.*

In the depletion area, radiation is absorbed and produces electrons and holes. The built-in potential separates them. Depending on the wavelength and thickness of the device, different sections of the spectrum can be absorbed.

NOTES

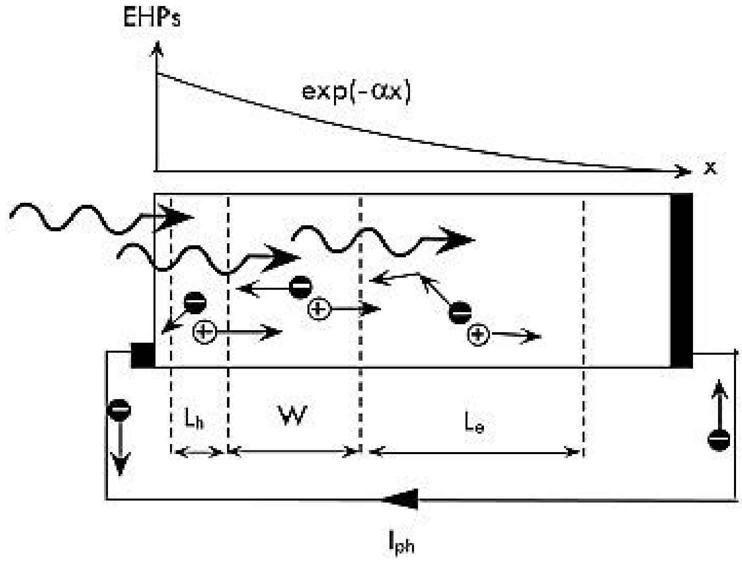


Fig. 5.43 *Photogenerated Carriers in a Solar Cell as a Result of Light Absorption.*

The width of the depletion area is given by 'W', whereas the minority carrier diffusion lengths in the n and p regions are given by L_h and L_e.

Because absorption decreases with depth, the depletion region must be close to the surface to maximize absorption. This is accomplished by using a thin n-region.

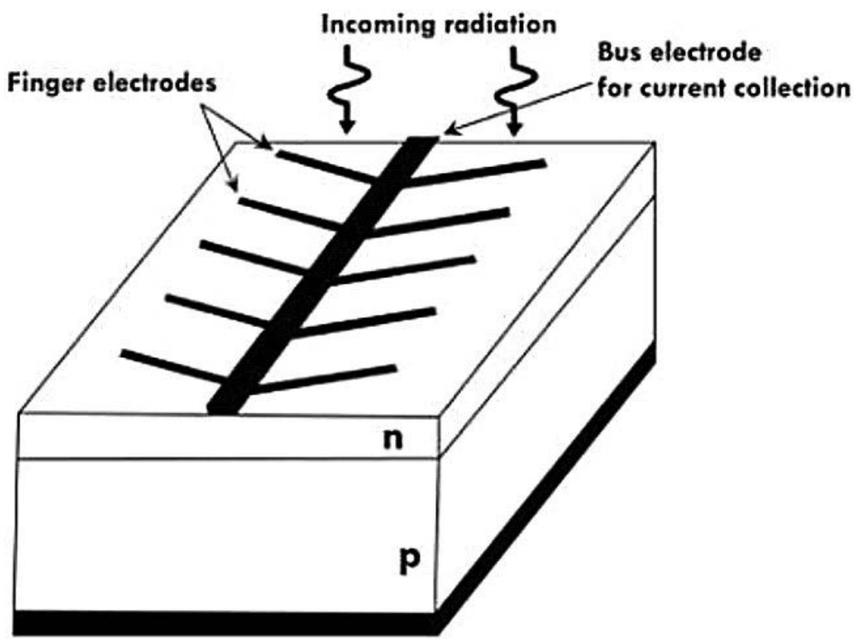


Fig. 5.44 *Finger Electrodes on a p-n Junction Solar Cell.*

The design comprises of a single bus electrode for transmitting current and finger electrodes that are thin enough to allow the solar cell to absorb adequate light.

NOTES

Solar Cell I-V Characteristics

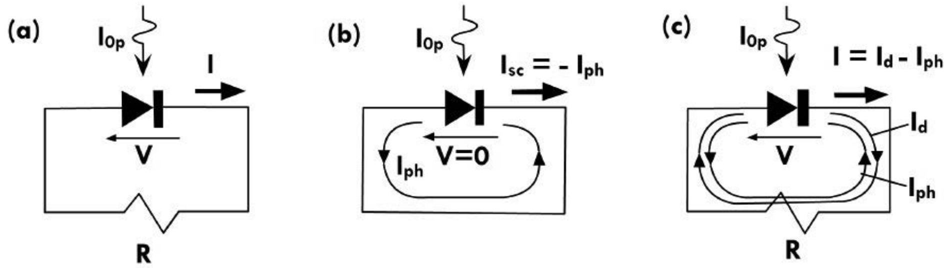


Fig. 5.45 *(a) p-n Junction Solar Cell with an External Load Illuminated. The Analogous Circuits (b) and (c) with and without an External Load.*

A photocurrent flows across the external circuit as a result of the illumination. When an external load is applied, the potential drop across it generates a forward bias current that is oblique to the photocurrent.

The I-V characteristics of a solar cell can be calculated by studying its equivalent circuit. The I-V characteristics are affected by the intensity of the incident radiation as well as the cell's operating point (external load). As shown in Figure 5.45 depicts a p-n junction solar cell under illumination. If the external circuit is a short circuit (the external load resistance is 0), the only current flowing is caused by the incident light's EHPs. This is referred to as the photocurrent, which is denoted by I_{ph}. This is also known as the short circuit current, I_{sc}. This is the inverse of the photo current and is proportional to the intensity of the incident radiation, I_{op}, according to the definition of current.

$$I_{sc} = \Delta I_{ph} = \Delta k I_{op} \tag{5.67}$$

where k is a constant that varies depending on the device k is an efficiency parameter that quantifies the conversion of light into EHPs. Consider the scenario of an external load R, as shown in Figure 5.45. As shown in Figure 5.46 depicts the analogous circuit for this scenario. V = IR denotes the voltage across the external load. This voltage works in opposition to the built-in potential and lowers the barrier to carrier injection across the junction. This is analogous to a p-n junction in forward bias, where the external bias produces minority carrier injection and higher current. This forward bias current works in opposition to the photocurrent created within the device as a result of solar radiation. This is because I_{ph} is formed by electrons travelling to the n side and holes travelling to the p side due to the electric field within the device, i.e., drift current, whereas forward bias current is generated by diffusion current caused by minority carrier injection. As a result, the net current can be written as:

$$\begin{aligned} I &= -I_{ph} + I_d \\ I_d &= I_{s0}\,[\exp(\frac{eV}{k_B T}) - 1] \\ I &= -I_{ph} + I_{s0}\,[\exp(\frac{eV}{k_B T}) - 1] \end{aligned} \tag{5.68}$$

where I_d is the forward bias current and can be expressed in terms of I_{s0}, the reverse saturation current, and V, the external voltage. Figure 5.46 depicts the overall I-V characteristic

NOTES

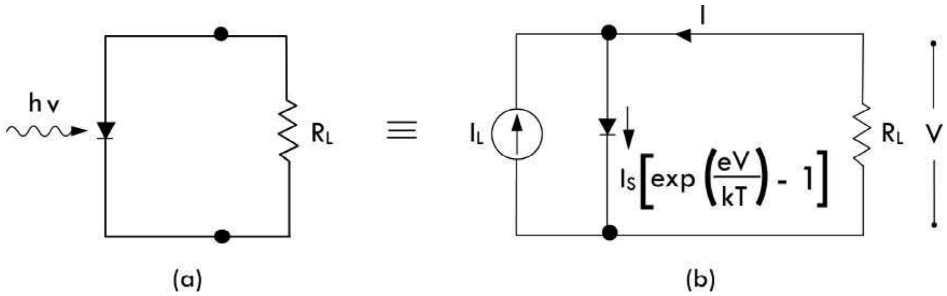


Fig. 5.46 *(a) A Solar Cell Connected to an External Load; (b) An Equivalent Circuit with a Constant Current Source, A Forward Biased p-n Junction, and the External Load.*

The constant current source is opposed by the current from the forward biased p-n junction.

The dark characteristic is similar to a p-n junction I-V curve in the absence of light. The presence of light (I_{ph}) causes the I-V curve to drop downward. Figure 5.47 shows a photo current I_{ph}, which is the current when the external voltage is zero, and an open circuit voltage, V_{oc}, which is the voltage when the net current in the circuit is zero. V_{oc} may be determined using equation 5.68 as:

$$\begin{aligned} I_{ph} &= I_{s0}\,[\exp(\frac{eV_{oc}}{k_BT}) - 1] \\ V_{oc} &\approx \frac{k_BT}{e}\,\ln[\frac{I_{ph}}{I_{s0}}] \end{aligned} \tag{5.69}$$

The greater the photon flux, the greater the value of I_{ph} (by Equation 5.67) and the greater the value of V_{oc}. Similarly, a lower I_{s0} can result in a greater V_{oc}. Because I_{s0} is the pn junction's reverse saturation current, it is given by

$$I_{s0} = n_i^2 e\,[\frac{D_e}{L_eN_A} + \frac{D_h}{L_hN_D}] \tag{5.70}$$

The reverse saturation current can be reduced by selecting a material with a higher band gap, E_g, resulting in a smaller n_i. However, this reduces the spectrum of wavelengths that can be absorbed by the material, resulting in a decrease in I_{ph}.

The overall power in the solar cell circuit can be calculated as follows:

$$P = IV = I_{s0}V\,[\exp(\frac{eV}{k_BT}) - 1] - I_{ph}V \tag{5.71}$$

Its derivative with respect to voltage should be zero for maximum power. This results in a cyclical relationship between current and voltage.

NOTES

$$\frac{dP}{dV} = 0$$

$$I_m \approx I_{ph}\,(1 - \frac{k_B T}{eV_m})$$

$$V_m \approx V_{oc} - \frac{k_B T}{e}\ln(1 + \frac{eV_m}{k_B T}) \tag{5.72}$$

$$P_m = I_m V_m \approx I_{ph}\,[V_{oc} - \frac{k_B T}{e}\ln(1 + \frac{eV_m}{k_B T}) - \frac{k_B T}{e}]$$

As shown in Figure 5.47 demonstrates this. The maximum power is given by the area under the curve, which corresponds to I_m and V_m. According to Equation 5.72, the maximum power is directly proportional to V_{oc} and can be raised by decreasing I_{s0}. This means that smaller n_i and larger E_g are preferable, but less radiation is absorbed as a result.

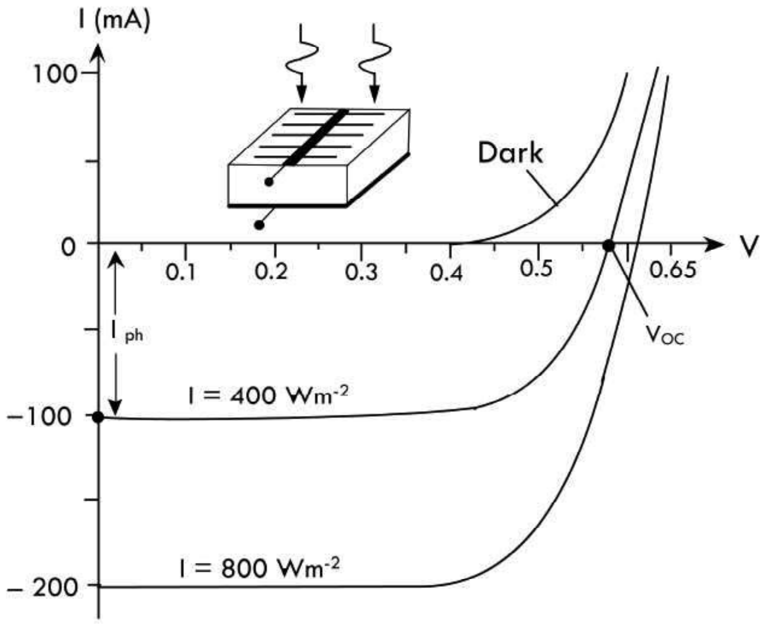


Fig. 5.47 *I-V Characteristics of a Si p-n Junction Solar Cell.*

As shown in Figure 5.47, I-V characteristics of a Si p-n junction solar cell in the dark and when illuminated with increasing intensity of light. With increasing illumination, both short circuit current and open circuit voltage rise.

Fill Factor

The maximum current and voltage from a solar cell are the short circuit current and open-circuit voltage, respectively. However, the power from the solar cell is zero at both of these operational locations. The 'Fill Factor,' abbreviated 'FF,' is a parameter that, along with V_{oc} and I_{sc}, determines the maximum power output of a solar cell. The FF is defined as the ratio of maximum power from a solar cell to the product of V_{oc} and I_{sc}, such that:

$$FF = \frac{P_{MP}}{V_{OC} \times I_{SC}}$$

$$FF = \frac{V_{MP}I_{MP}}{V_{OC}I_{SC}}$$

The FF is a graphical representation of the 'Squareness' of the solar cell, as well as the size of the largest rectangle that will fit in the IV curve (Refer Figure 5.48).

NOTES

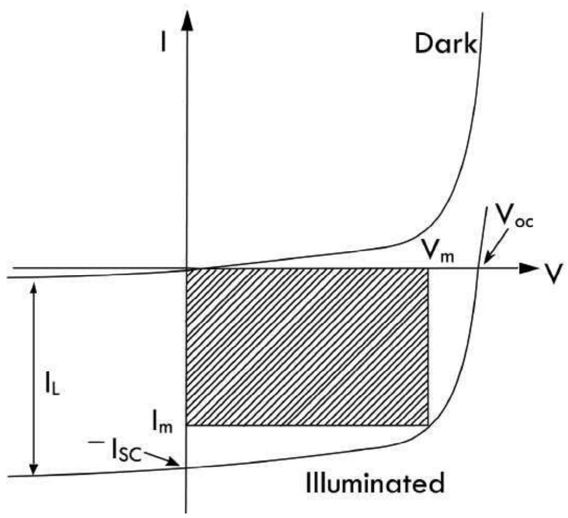


Fig. 5.48 *I-V Curve for a Solar Cell, with the Shaded Area Indicating Maximum Power.*

As shown in Figure 5.48, an I-V Curve for a Solar Cell, with the Shaded Area Indicating Maximum Power. V_m and I_m are the corresponding voltage and current. The value is determined by the applied external load.

5.5 LIGHT EMITTING DIODE (LED): LIGHT CONFINEMENT FACTOR

Light-emitting diodes are just forward-biased p-n junctions that spontaneously emit light. Radiative recombination of electron-hole pairs in the depletion area causes spontaneous emission (or electroluminescence). LEDs have a large angular bandwidth and a broad spectral bandwidth (20-150 nm). They are temporally and spatially incoherent sources.

Advantages of LED

1. **Simpler Fabrication:** There are no mirror facets and no striped geometry in some structures.
2. **Cost**: Because of the LED's simpler design, it has a substantially lower cost, which is likely to be maintained in the future.
3. **Reliability:** The LED does not degrade catastrophically and has proven to be significantly more resistant to slow decline than the infusion laser. It's also impervious to self-pulsation and modal noise.
4. **Generally Less Temperature Dependence:** Temperature has less of an impact on the light output against current characteristic than it does on the injection laser's equivalent characteristic. Furthermore, because the LED is not a threshold device, increasing the temperature does not cause the

threshold current to rise above the operating point, causing the device to shut down.

5. **Simpler Drive Circuitry:** This is owing to the lower driving currents and lower temperature dependence, which eliminates the need for temperature adjustment circuits.
6. **Linearity:** The LED should, in theory, have a linear light output versus current characteristic. When it comes to analogue modulation, this can be useful.

Drawbacks

(a) Optical Power Coupled into a Fibre is often Lower (Microwatts);

(b) Modulation Bandwidth is usually Smaller;

(c) Distortion of the Harmonic Spectrum.

LED Power and Efficiency

The surplus electrons and holes in the p- and n-type materials, respectively, can be used to calculate the power generated internally by an LED. When the device is forward biassed, carrier injection occurs at the contacts. Because the injected carriers are produced and recombined in pairs, the excess density of electrons Δn and holes Δp is equal, ensuring charge neutrality within the structure. In extrinsic materials, one carrier type will have a significantly higher concentration than the other, therefore the hole concentration will be much higher than the electron concentration in the p-type area, for example. The extra minority carrier density decays exponentially with time t in most cases, as shown by the relationship:

$$\boldsymbol{\Delta n = \Delta n(0)\exp(-t/\tau)} \tag{5.73}$$

n(0) is the initial excess electron density injected, and τ reflects the complete carrier recombination lifetime.

An equilibrium situation is reached when a steady current flows into the junction diode. The sum of the externally provided and thermal generation rates will be the overall rate at which carriers are generated in this situation. As a result, a rate equation for carrier recombination in the LED can be written as follows:

$$\frac{d(\Delta n)}{dt} = \frac{J}{ed} - \frac{\Delta n}{\tau} \quad (m^{-3}\,s^{-1}) \tag{5.74}$$

Setting the derivative in Equation (5.74) to zero yields the equilibrium condition. Hence:

$$\Delta n = \frac{J\tau}{ed} \quad (m^{-3}) \tag{5.75}$$

When a constant current is flowing into the junction region, Equation (5.75) gives the steady-state electron density.

In the steady state, Equation (5.74) shows that the total number of carrier recombination per second, or the recombination rate rt, will be:

$$r_t = \frac{J}{ed} \quad (m^{-3})$$

$$= r_r + r_{nr} \quad (m^{-3}) \tag{5.76}$$

Where r_r denotes the rate of radiative recombination per unit volume and r_{nr} denotes the rate of non-radiative recombination per unit volume. Furthermore, when the forward-biased current into the device is I, the total number of recombination per second R_t is calculated using Equation (5.77).

NOTES

$$R_t = \frac{i}{e} \tag{5.77}$$

The ratio of the radiative recombination rate to the overall recombination rate is known as the LED internal quantum efficiency η_{int}.

$$\eta_{int} = \frac{r_r}{r_t} = \frac{r_r}{r_r + r_{nr}}$$

$$= \frac{R_r}{R_t} \tag{5.78}$$

The total number of radiative recombination per second is given by Rr. When Equation (5.78) is rearranged and Equation (5.77) is substituted, the result is:

$$R_r = \eta_{int} \frac{i}{e} \tag{5.79}$$

Because R_r is likewise equal to the total number of photons generated per second, and each photon has an energy of hf joules, the optical power generated internally by the LED, P_{int}, is as follows:

$$P_{int} = \eta_{int} \frac{i}{e} hf \quad \text{(W)} \tag{5.80}$$

In terms of wavelength rather than frequency, the internally generated power is:

$$P_{int} = \eta_{int} \frac{hci}{e\lambda} \quad \text{(W)} \tag{5.81}$$

The radiative minority carrier lifetime is $\tau_r = \Delta n/r_r$ and the non-radiative minority carrier lifetime is $\tau_{nr} = \Delta n/r_{nr}$ for the exponential decay of excess carriers described by Equation (5.73). As a result of Equation (5.78), the internal quantum efficiency is as follows:

$$\eta_{int} = \frac{1}{1 + (r_{nr}/r_r)} = \frac{1}{1 + (\tau_r/\tau_{nr})} \tag{5.82}$$

Furthermore, if you write the overall recombination lifetime as $\tau = \Delta n/r_t$, you get:

$$\frac{1}{\tau} = \frac{1}{\tau_r} + \frac{1}{\tau_{nr}} \tag{5.83}$$

Hence,

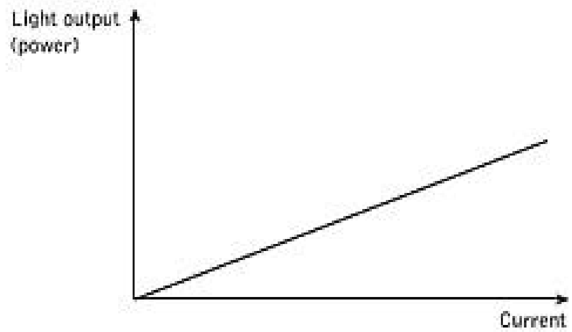
$$\eta_{int} = \frac{\tau}{\tau_r} \tag{5.84}$$

NOTES

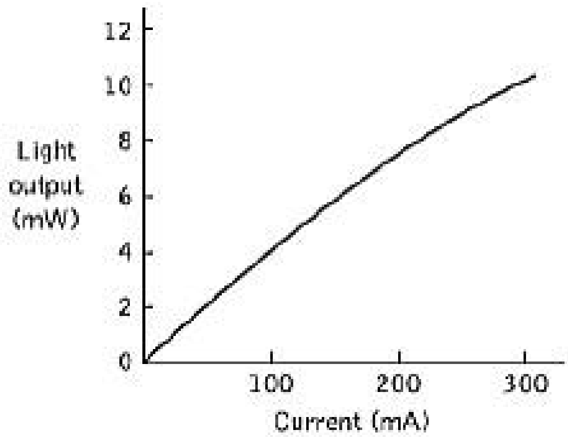
Characteristics of LEDs

Optical Output Power

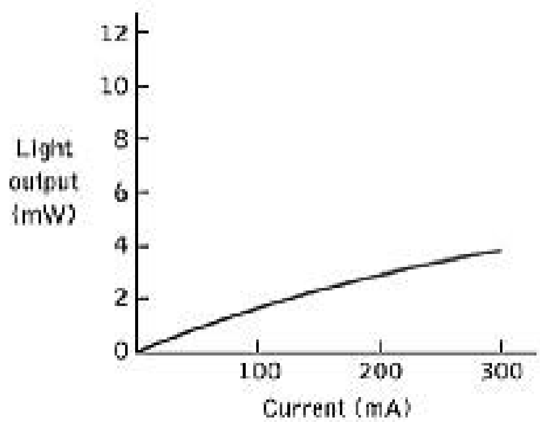
In compared to the majority of injection lasers, the LED is a very linear device, making it more ideal for analogue transmission where the linearity of the optical source is severely limited. However, LEDs do display considerable nonlinearities in practice, which are dependent on the design used. As a result, linear circuit techniques are frequently used to ensure the device's linear performance, allowing it to be used in high-quality analogue transmission systems.



(a)



(b)



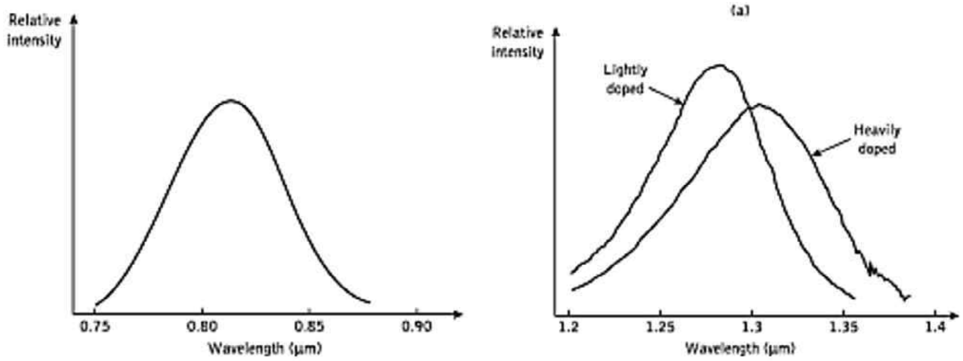
(c)

Fig. 5.49 *(a) Ideal LED Characteristics, (b) Surface Emitter with a 50 ìm Diameter Dot and (c) Edge Emitter with a 65 ìm Wide Stripe and 100 ìm Length Contact.*

Output Spectrum

At half maximum intensity, the spectral line width of an LED operating at room temperature in the 0.8 to 0.9 μm wavelength range is normally between 25 and 40 nm. The line width increases to roughly 50 to 160 nm for materials with reduced band gap energies operating in the 1.1 to 1.7 μm wavelength region. Figure shows examples of these two output spectra. Increased doping levels and the creation of band tail states cause line width to widen. This is seen in the output spectra of surface and edge-emitting LEDs, where the devices are often substantially doped and mildly doped, respectively.

NOTES



(a) (b)

Fig. 5.50 *(a) Output Spectrum of Spectral Linewidth of an LED (b) Comparative Output Spectrum of Lightly Doped and Heavily Doped LEDs.*

Diode Laser:

A **laser** is a device that produces light via an optical amplification process based on the stimulated emission of electromagnetic energy. **'Light Amplification by Stimulated Emission of Radiation'** is how the term 'Laser' came to be. A laser differs from other light sources in that it emits light in a coherent manner. Lasers have a wide range of applications. Optical disc drives, laser printers, and barcode scanners are examples of common consumer gadgets that utilise them. Both fiber-optic and free-space optical communication involve lasers.

Basic Concepts

Absorption and Emission of Radiation

Photons are discrete packets of energy or quanta that interact with matter. Furthermore, quantum theory says that atoms only exist in discrete energy states, and that light absorption and emission causes them to change from one discrete energy state to the next. The difference in energy E between the higher energy level E_2 and the lower energy state E_1 is related to the frequency of the absorbed or emitted radiation f by the expression:

$$E = E2 - E1 = hf$$

The Planck constant is $h = 6.626 \times 10^{-34}$ J s These atom's discrete energy states could be thought to correspond to electrons in certain energy levels relative to the nucleus. As a result, different atom energy states correlate to distinct electron configurations, and a single electron transition between two atom energy levels will result in an energy change suitable for photon absorption or emission.

NOTES

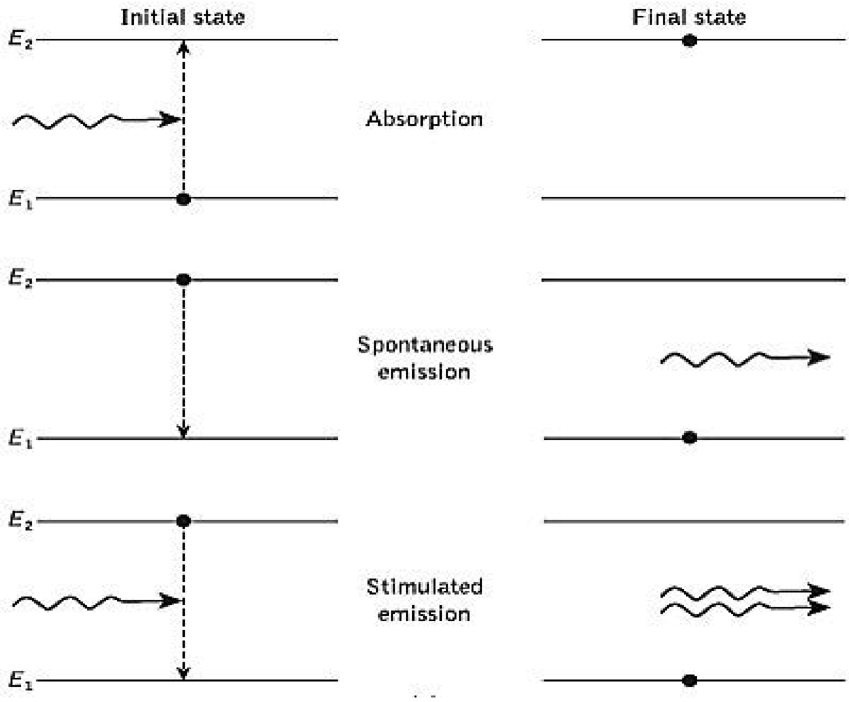


Fig. 5.51 Energy Spectrum of Absorption and Emission of Radiation.

This emission process can take place in one of two ways:

(a) By spontaneous emission, in which the atom spontaneously returns to a lower energy state;

(b) By stimulated emission, which occurs when a photon with an energy equal to the difference in energy between the two states ($E_2 - E_1$) interacts with an atom in the upper energy state, prompting it to return to the lower energy state and producing a second photon.

The laser's unique capabilities as an optical source are due to the stimulated emission process. Because the photon produced by stimulated emission has the same energy as the one that created it, the light associated with them has the same frequency. The light associated with the stimulating and stimulated photons is polarized similarly and is in phase. Coherent radiation is thus obtained, as opposed to spontaneous emission.

The Einstein Relations

Einstein established in 1917 that the rates of the three absorption, spontaneous emission, and stimulated emission transition processes were mathematically connected. He did so by assuming that the atomic system is in thermal equilibrium, which means that the rate of upward transitions must equal the rate of downward transitions. Boltzmann statistics, which describe the population of the two energy levels of such a system, are as follows:

$$\frac{N_1}{N_2} = \frac{g_1 \exp(-E_1/KT)}{g_2 \exp(-E_2/KT)} = \frac{g_1}{g_2} \exp(E_2 - E_1/KT)$$

$$= \frac{g_1}{g_2} \exp(hf/KT)$$

NOTES

Where N_1 and N_2 are the densities of atoms in energy levels E_1 and E_2, respectively, and g_1 and g_2 are the levels' respective degeneracy, K is Boltzmann's constant, and T is the absolute temperature.

The rate of upward transition or absorption is proportional to both N_1 and the spectral density ρ_f of the radiation energy at the transition frequency f, because the density of atoms in the lower or ground energy state E_1 is N_1. As a result, R_{12}, the upward transition rate, can be represented as:

$$R12 = N1\rho f B12$$

The spontaneous lifetime τ_{21} is the average duration an electron spends in the excited state before a transition occurs in spontaneous emission. If the density of atoms within the system with energy E_2 is N_2, then the spontaneous emission rate is given by the product of N_2 and $1/\tau_2$. This can be represented as N_2A_{21}, where A_{21} is the reciprocal of the spontaneous lifetime and is equal to the Einstein coefficient of spontaneous emission.

Similar to the rate of stimulated upward transition, the rate of stimulated downward transition of an electron from level 2 to level 1 can be calculated. As a result, the rate of stimulated emission is equal to

$$R_{21} = N_2A_{21} + N_2\rho_f B_{21}$$

The upward and downward transition rates must be identical in a system in thermal equilibrium, so $R_{12} = R_{21}$, or:

$$N_1\rho_f B_{12} = N_2A_{21} + N_2\rho_f B_{21}$$

It follows that:

$$\rho_f = \frac{N_2 A_{21}}{N_1 B_{12} - N_2 B_{21}}$$

and:

$$\rho_f = \frac{A_{21}/B_{21}}{(B_{12} N_1/B_{21} N_2) - 1}$$

Substituting values from equations

$$\rho_f = \frac{A_{21}/B_{21}}{[(g_1 B_{12}/g_2 B_{21})\ \exp(hf/KT)] - 1}$$

Planck demonstrated that the radiation spectral density of a black body emitting in the frequency range f to $f + df$ is determined by

$$\rho_f = \frac{8\pi h f^3}{c^3}\left[\frac{1}{\exp(hf/KT) - 1}\right]$$

After comparing equations

$$B_{12} = \left(\frac{g_2}{g_1}\right) B_{21}$$

NOTES

and

$$\frac{A_{21}}{B_{21}} = \frac{8\pi h f^3}{c^3}$$

The stimulated emission rate to the spontaneous emission rate is calculated as follows:

$$\frac{\text{Stimulated emission rate}}{\text{Spontaneous emission rate}} = \frac{B_{21}\rho_f}{A_{21}} = \frac{1}{\exp(hf/KT) - 1}$$

Population Inversion

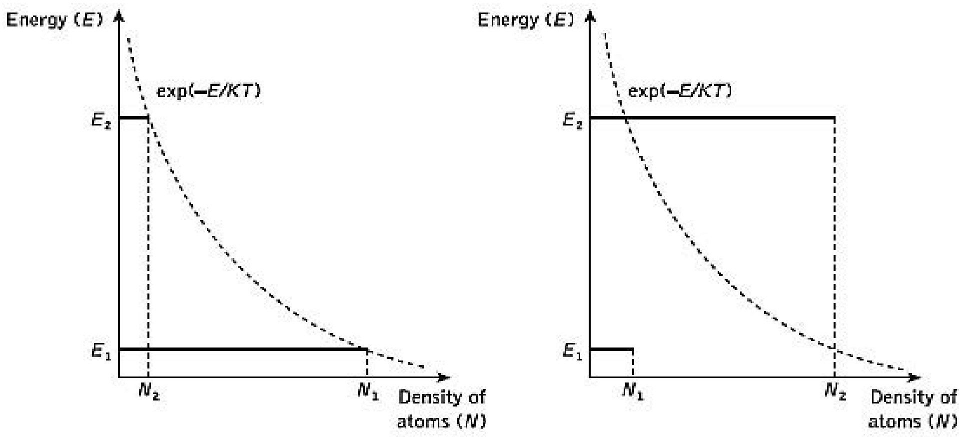


Fig. 5.52 *Energy (E) vs Atomic Density (N) for Population Inversion.*

The lower energy level E_1 of the two-level atomic system includes more atoms than the upper energy level E_2, which is common for structures at room temperature, under the criteria of thermal equilibrium described by the Boltzmann distribution. To achieve optical amplification, however, a non-equilibrium atom distribution must be created in which the population of the upper energy level is greater than that of the lower energy level (i.e., $N_2 > N_1$). Population inversion is the term for this situation.

It is important to excite atoms into the upper energy level E_2 and so obtain a non-equilibrium distribution in order to achieve population inversion. Pumping is a term used to describe a procedure that uses an external energy source. $B_{12} = B_{21}$ when the two levels are equally degenerate (or not degenerate). As a result, the chances of absorption and stimulated emission are equal, resulting in at least equal populations at both levels.

In systems with three or four energy levels, population inversion is possible. Both systems have a centre metastable state in which the atoms spend an extremely lengthy period to reach population inversion. The stimulated emission or lasing takes place from this metastable level.

NOTES

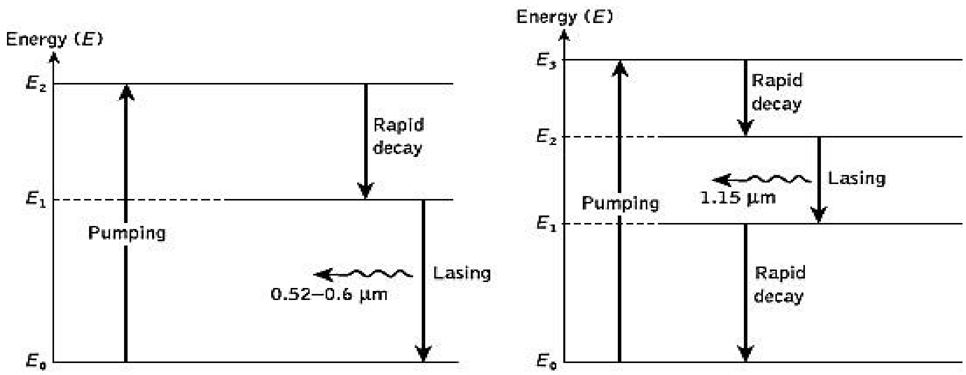


Fig. 5.53 *Energy Level Diagram of Population Inversion.*

Optical Feedback and Laser Oscillation

In a laser, light amplification happens when a photon collides with an atom in the excited energy state, causing the stimulated emission of a second photon, followed by the release of two more photons by both of these photons. When the electromagnetic waves associated with these photons are in phase, this process essentially causes avalanche multiplication, and amplified coherent emission is achieved. To produce this laser activity, photons must be contained within the laser medium and coherence conditions must be maintained. This is done by putting or creating planar or curved mirrors at either end of the amplifying medium. The optical cavity is more like an oscillator than an amplifier since it delivers positive feedback to the photons via reflection at the cavity's two ends. As a result, the optical signal is amplified and fed back numerous times as it passes through the medium.

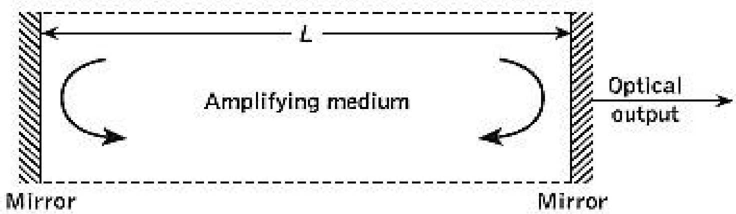


Fig. 5.54 *Optical Resonator for Light Amplification.*

Because the structure is a resonant cavity, when the amplifying medium has enough population inversion, the radiation builds up and establishes itself as standing waves between the mirrors. As a result, when the optical separation between the mirrors is L, the resonance condition along the cavity's axis is:

$$L = \frac{\lambda q}{2n}$$

where q is an integer, n is the refractive index of the amplifying medium, and λ is the emission wavelength. Discrete emission frequencies *f*, on the other hand, are defined as:

$$f = \frac{qc}{2nL}$$

The many integer values of q determine the distinct frequencies of oscillation within the laser cavity, and each represents a resonance or mode. A frequency interval δf separates these modes, where:

NOTES

$$\delta f = \frac{c}{2nL}$$

Assuming δf-f and $f = c/\lambda$, the mode separation in terms of the free space wavelength is:

$$\delta\lambda = \frac{\lambda\delta f}{f} = \frac{\lambda^2}{c}\delta f$$

Hence,

$$\delta\lambda = \frac{\lambda^2}{2nL}$$

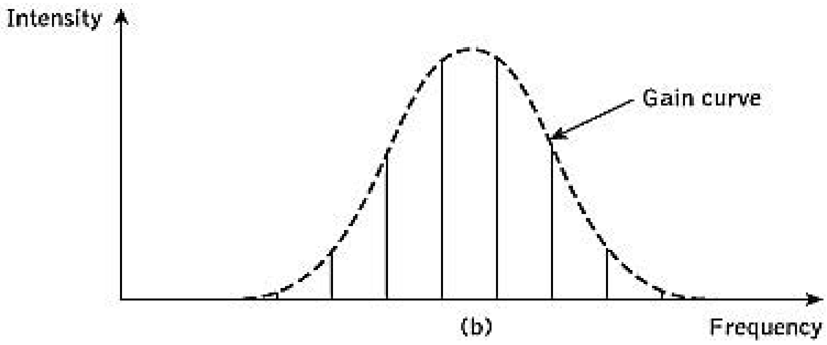


Fig. 5.55 *Intensity vs Frequency Curve (Gain Curve).*

Threshold Condition for Laser Oscillation

When the gain in the amplifying medium perfectly balances the total losses, steady-state conditions for laser oscillation are attained. As a result, whereas population inversion between the energy levels that provide the laser transition is required for oscillation to occur, it is not required for lasing to occur. We assume that the amplifying medium has a length L and fills the space between the two mirrors with reflectivity r_1 and r_2. The beam goes through the medium twice on each round trip. As a result, the light beam's fractional loss is:

Fractional loss = $r_1 r_2 \exp(-2AL)$

The rise in beam intensity caused by stimulated emission is shown to be exponential. As a result, if the stimulated emission gain coefficient per unit length is C cm^{-1}, the fractional round-trip gain is:

$$\text{Fractional Gain} = \exp(2CL)$$

Hence:

$$\exp(2CL) \times r_1 r_2 \exp(-2AL) = 1$$

And

$$r_1 r_2 \exp[2(C - A)L] = 1$$

NOTES

By rearranging the previous calculation, you may get the threshold gain per unit length:

$$\bar{g}_{th} = \bar{\alpha} + \frac{1}{2L} \ln \frac{1}{r_1 r_2}$$

The transmission loss through the mirrors is represented by the second term on the right-hand side.

Semiconductor Injection Laser

The presence of an optical cavity in the crystal structure to provide photon feedback encourages stimulated emission by recombination of the injected carriers in the semiconductor injection laser (also known as the Injection Laser Diode (ILD) or simply the injection laser). The injection laser thus has several significant benefits over conventional semiconductor sources (such as LEDs) that could be employed for optical communications.

The following are some of them:

1. The amplifying action of stimulated emission results in a high radiance. Injection lasers typically produce optical output power in the milliwatt range.
2. A narrow line width of 1 nm (10 Å) or less, which is effective for reducing the effects of material dispersion.
3. Modulation capabilities that currently extend into the gigahertz region and will undoubtedly be enhanced in the future.
4. Relative temporal coherence, which is thought to be necessary for heterodyne (coherent) detection in high-capacity systems but is only used in single-mode systems at the moment.
5. Good spatial coherence, allowing a lens to focus the output into a spot with a higher intensity than the distributed unfocused emission.

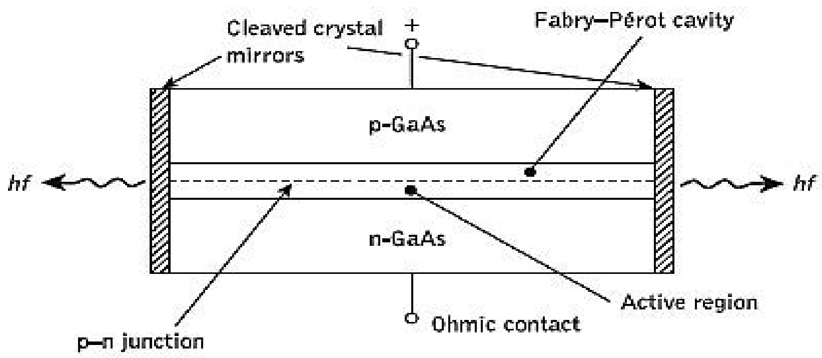


Fig. 5.56 *A GaAs Homojunction injection Laser with a Fabry–Pérot's Cavity is shown Schematically.*

The DH injection laser, which was made from lattice-matched III–V alloys, has carrier and optical confinement on both sides of the p–n junction, substantially improving its performance. This allowed these devices to function in a CW mode at 300 K with the appropriate heat sinking, which has obvious advantages for optical communications.

NOTES

Efficiency

The differential external quantum efficiency is the ratio of the increase in photon output rate to the increase in injected electrons. If 'P_e' represents the optical power output by the device, 'I' represents the current, 'e' represents the charge on an electron, and hf represents the photon energy, then:

$$\eta_D = \frac{dP_e/hf}{dI/e} \simeq \frac{dP_e}{dI(E_g)}$$

where E_g is the energy of the band gap in eV.

Efficiency, it should be mentioned, is a measure of the rate of change of optical output power with current, and so defines the slope of the output characteristic. The semiconductor laser's intrinsic quantum efficiency η_i,

$$\eta_i = \frac{\text{number of photons produced in the laser cavity}}{\text{number of injected electrons}}$$

The expression connects it to the differential external quantum efficiency

$$\eta_D = \eta_i \left[\frac{1}{1 + (2\bar{\alpha}L/\ln(1/r_1r_2))} \right]$$

Where A is the laser cavity's loss coefficient, L is its length, and r_1 and r_2 are the cleaved mirror reflectivity.

Another important measure is overall efficiency (external quantum efficiency), η_T which is defined as:

$$\eta_T = \frac{\text{total number of output photons}}{\text{total number of injected electrons}}$$

$$= \frac{P_e/hf}{I/e} \simeq \frac{P_e}{IE_g}$$

When the injection current I exceeds the threshold current I_{th}, the power emitted P_e changes linearly:

$$\eta_T \simeq \eta_D \left(1 - \frac{I_{th}}{I}\right)$$

When the injection current is large (e.g., I = 5I_{th}), then η_T H" η_D, the overall efficiency is high (about 15 to 25%), but when the current is low (I H"2I_{th}), the total efficiency is low (around 15 to 25%).

The device's external power efficiency (or device efficiency) η_{ep} in converting electrical input to optical output is calculated as follows:

$$\eta_{ep} = \frac{P_e}{P} \times 100 = \frac{P_e}{IV} \times 100\%$$

For the total efficiency we find:

$$\eta_{ep} = \eta_T \left(\frac{E_g}{V}\right) \times 100\%$$

Stripe Geometry

The refractive index step at the heterojunction interfaces in the DH laser structure enables optical confinement in the vertical direction, while lasing occurs across the entire width of the device.

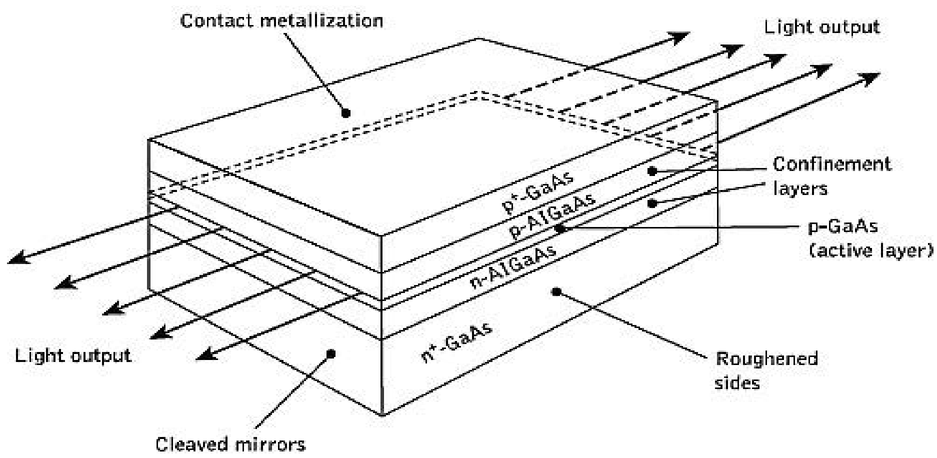


Fig. 5.57 Schematic representation of DH laser.

The sides of the cavity are simply made by roughening the device's edges in order to prevent unwanted emission in these directions and restrict the number of horizontal transverse modes, as shown in Figure. However, the large emission area causes a number of issues, including severe heat sinking, numerous filaments lasing in the comparatively small active region, and an inadequate light output shape for efficient coupling to cylindrical fibers.

Laser structures with an active region that does not extend to the device's edges were created to solve these concerns while also lowering the needed threshold current. The use of stripe geometry in the structure to create optical containment in the horizontal plane was a typical technique.

Laser modes

The laser cavity generates a vast number of modes, which are stored in the LASER. As a result, the laser output will only include longitudinal modes that fall inside the gain curve's spectral width, as shown in figure.

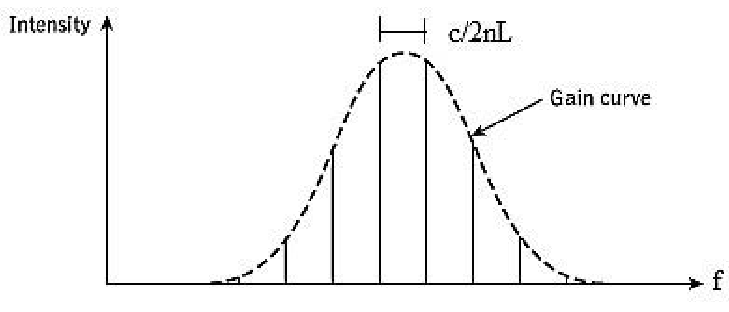


Fig. 5.58 Intensity vs frequency graph of gain curve.

This results in resonant modes that are transverse to the propagation direction. These transverse electromagnetic modes are denoted by TEM_{lm} in the same way

NOTES

that waveguide transverse modes are denoted by TEM_{lm}, where *l* and m denote the number of transverse modes. All components of the propagating wave front are in phase in the TEM_{00} mode. However, with higher order modes (TEM_{10}, TEM_{11}, and so on), when phase reversals produce the various mode patterns, this is not the case. Thus, a laser that exclusively works in theTEM_{00} mode can achieve the highest degree of coherence and spectral purity. Higher order transverse modes only exist when the cavity width is large enough for them to oscillate.

By limiting the width of the optical cavity, the precise stripe geometry prevents the production of higher order lateral modes, leaving only one lateral mode that produces the output spectrum.

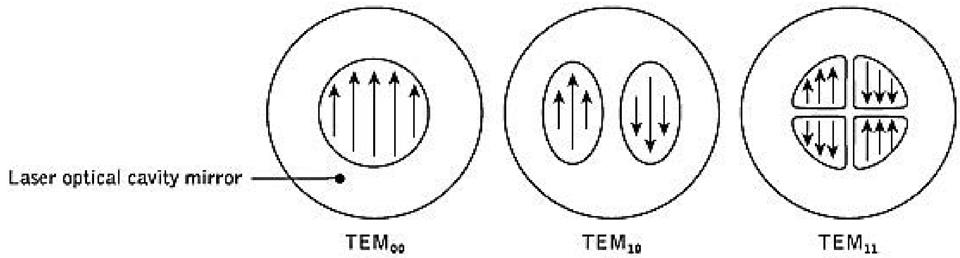


Fig. 5.59 *Schematic illustration of TEM_{00}, TEM_{10} and TEM_{11} modes of operations.*

Single-Mode Operation

A laser's optical output must only comprise a single longitudinal and transverse mode for single-mode operation. As a result, the single-mode device's emission spectral width is much lower than the enlarged transition line width. However, if the aperture of the resonant cavity is reduced to the point where only the TEM_{00} mode is supported, single transverse mode functioning can be achieved. To achieve single-mode operation, all but one of the longitudinal modes must be eliminated. Reduce the length of the cavity until the frequency separation of the adjacent modes, given by $\delta f = c/2nL$, is greater than the laser transition line width or gain curve, as one means of attaining single longitudinal mode operation. Within the laser cavity, just the single mode that falls within the transition line width can oscillate.

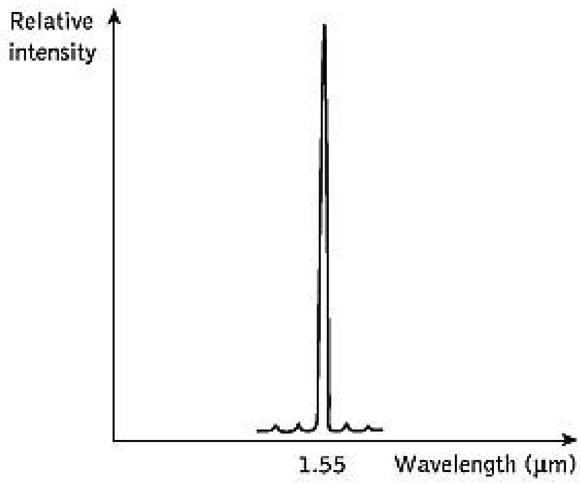


Fig. 5.60 *Relative Intensity vs Wavelength Plot.*

NOTES

External Quantum Efficiency (g_{th})

The quantity of photons emitted per radiative electron–hole pair recombination above threshold is denoted as the external quantum efficiency η_{ext}.

η_{ext} is computed experimentally using the straight-line section of the curve for emitted optical power *P* vs drive current *I*, which yields

$$\eta_{ext} = 0.806\ \lambda$$

Laser Diode Rate Equation

The rate equations that control the interaction of photons and electrons in the active region can be used to calculate the relationship between optical output power and diode drive current. For a p-n junction with a carrier confinement region of depth d, the rate of equation are given by:

$$= C_{n\varphi} + R_{sp} \tag{5.85}$$

= stimulated emission+ spontaneous emission+ photon loss; which governs the number of photons φ and

$$= -C_{n\varphi} \tag{5.86}$$

= injection + spontaneous recombination+ stimulated emission; which governs the number of electrons n

whereC = Coefficient describing the strength of the optical absorption

R_{sp} = Rate of spontaneous emission into lasing mode

τ_{ph}= Photon life time

τ_s= Spontaneous recombination lifetime

For a steady state scenario, calculating Equations (5.85) and (5.86) yields an expression for output power.

We have; in the first equation, assuming R_{sp} is insignificant and nothing that dφ/dt must be positive is small, we have; in the second equation, assuming R_{sp} is negligible and nothing that dφ/dt must be positive is small, we have;

$$C_{n-1}/\ \tau_{ph}e–0 \tag{5.87}$$

This demonstrates that in order for φ to increase, n must reach a threshold value of n_{th}. So, from Equation (5.85), when no. of photons φ = 0, this threshold value can be represented in terms of the threshold current J_{th} required to maintain an inversion level n = n_{th} in steady state.

When spontaneous emission is the lone decay mechanism, this expression defines the current requirement to maintain an excess electron density in the laser.

Now evaluate the photon and electron rate equations at the lasing threshold in steady state.

$$0=C_{nth}\varphi_s+R_{sp}-\varphi_s/\tau_{ph} \tag{5.88}$$

$$0 = -C_{nth}\varphi_s \tag{5.89}$$

When these two equations are combined, the number of photons per unit volume is calculated.

$$\Phi_s = (J-J_{th})+\ \tau_{ph}R_{sp}$$

NOTES

5.5.1 Optical Gain and Threshold Current for Lasing

Optical Gain: Gain or amplification is a process in laser physics in which the medium transmits a portion of its energy to the produced electromagnetic radiation, resulting in an increase in optical power. All lasers operate on the same basis. Gain is a measure of a laser medium's capacity to increase optical power in a quantitative sense.

As it passes through the medium, the gain is defined as the derivative of the logarithm of power P. The gain, indicated by G, is the factor by which an input beam is amplified by a medium.

$$G = \frac{\mathrm{d}}{\mathrm{d}z}\ln(P) = \frac{\mathrm{d}P/\mathrm{d}z}{P}$$

where z denotes the propagation direction The transverse profile of the beam is ignored in this equation.

The gain can be stated in terms of reduced and excited state populations N_1 and N_2 in the simple quasi two-level system:

$$G = \sigma_e N_2 - \sigma_a N_1$$

The effective emission and absorption cross-sections are denoted by σ_e and σ_a, respectively. The gain is negative in the case of non-pumped media.

Threshold for Lasing: The lasing threshold is the lowest excitation level at which stimulated emission takes precedence over spontaneous emission in a laser's output. With increasing excitation below the threshold, the laser's output power gradually grows. The laser is said to be lasing above the threshold.

When the sum of all the losses encountered by light in one round trip of the laser's optical cavity equals the optical gain of the laser medium, the lasing threshold is attained. Assuming steady-state operation, this can be represented as:

$$R_1 R_2 \exp(2g_{\text{threshold}}\, l)\exp(-2\alpha l) = 1.$$

The mirror (power) reflectivities are R_1 and R_2, the length of the gain medium is l, the round-trip threshold power gain is $\exp(2g_{threshold}l)$, and the round-trip power loss is $\exp(-2\alpha l)$. This equation distinguishes between localized losses owing to the mirrors, which the researcher has control over, and dispersed losses such as absorption and scattering, which the experimenter has no control over. The distributed losses are often out of the experimenter's control.

For any particular laser ($\alpha = \alpha_o$), the optical loss is nearly constant, especially around the threshold. The threshold condition can be rearranged as follows if this assumption is true:

$$g_{\text{threshold}} = \alpha_0 - \frac{1}{2l}\ln(R_1 R_2).$$

Because $R_1 R_2 < 1$, both terms on the right side are positive, increasing the needed threshold gain parameter. Low distributed losses and high reflectivity mirrors are required to keep the gain parameter $g_{threshold}$ as low as possible. The presence of l in the denominator suggests that increasing the length of the gain medium will reduce the needed threshold gain, however this is not always the case. Because α_o often increases with l due to diffraction losses, the dependency on l is more convoluted.

NOTES

Check Your Progress

7. Define the term microwaves.
8. Name the advantages of microwaves.
9. Why we use Gunn-Effect Diodes-GaAs Diode?
10. What is meant by Gunn Effect?
11. State the parametric devices.
12. What are different types of recombination?
13. What is Light Dependent Resistor (LDR)?
14. How will you classify the detectors?
15. What are the fundamentally different characteristics of solar cells and photodetectors?
16. How will you define the Light Emitting Diode (LED)?
17. Define the term optical gain.

5.6 ANSWERS TO 'CHECK YOUR PROGRESS'

1. Metallic contacts at the two ends of the bar are ohmic contacts (terminals) called source and drain.
2. The JFET has three terminals which are known as Source (S), Drain (D) and Gate (G).
3. The current-voltage characteristics of JFET are of two types:
   (*a*)Drain Characteristics: The plots of drain current (I_D) versus drain-to-source voltage (V_{DS}) for different values of gate-to-source voltage are known as drain characteristics. These characteristics can be experimentally determined.
   (*b*)Transfer Characteristics: The plots of drain current (I_D) versus gate-to-source voltage (V_{GS}) for different values of drain-to-source voltage (V_{DS}) are known as transfer characteristics.
4. Two different modes of MOSFET which are classified from the constructional and application points of view as (*i*) Enhancement MOSFETs and (*ii*) Depletion MOSFETs. Again each of these two types of MOSFETs may be classified as *N*-channel and *P*-channel.
5. A MOS structure is formed by the gate metal along with insulating silicon dioxide layer and the semiconductor channel. The channel and gate act like two parallel plates of a capacitor separated by the dielectric (SiO_2) layer. This layer is responsible for high input impedance (10^{10} to 10^{15} Ω) for the MOSFET and also the reason for the name Insulated Gate Field Effect Transistor (IGFET) as SiO_2 is an insulating layer.
6. The structure of *P*-channel enhancement MOSFET consists of *N*-type silicon substrate in which two highly doped P^+ regions are diffused forming the source and drain regions. A negative gate voltage with respect to source

NOTES

induces positive charges to form the channel below the gate dielectric and make the channel conductive.

7. Microwaves are incredibly short waves, as the name implies. In general, RF energy ranges from DC to infrared, and it is a type of electromagnetic energy. A quick check at the various frequency ranges reveals that the Microwave frequency range includes UHF (Ultra-High Frequency) and SHF (Super High Frequencies), with wave length (ë) ranging from 1 to 100 cm.
8. Microwaves have certain distinct benefits over low frequencies:
   - Increased Bandwidth Availability
   - Improved Directivity Properties
   - Fading Effect and Reliability
   - Power Requirements
   - Transparency Property
9. This method uses the diode's avalanching and transit-time features to generate microwave frequencies.
10. When the electric field is changed from zero to a threshold value, the carrier drift velocity increases linearly from zero to a maximum, according to Gunn's observations. The drift velocity is reduced and the diode exhibits negative resistance when the electric field exceeds the threshold value of 3000 V/cm for n-type GaAs.
11. A parametric device is one that employs a time-varying or nonlinear reactance (capacitance or inductance). Because capacitance or inductance, which is a reactive property, may be utilized to create capacitive or inductive excitation, the term parametric is derived from the phrase parametric excitation.
12. There are two types of electron hole recombinations:
    - Radiative Recombination.
    - Non-radiative Recombination.
13. A Light Dependent Resistor (also called a photoresistor or LDR) is a resistive device whose resistance varies in response to incoming electromagnetic radiation. As a result, they are photosensitive devices. Additionally, they are referred to as photoconductors, photoconductive cells, or just photocells.
14. Detectors are classified as
    - Thermal Detectors
    - Photon Detectors
15. Solar cells and photodetectors have three fundamentally different characteristics:
    - For solar cells to be effective, they must be able to operate over a wide spectrum of wavelengths, as opposed to photodiodes (solar spectrum).
    - In order to optimize exposure, solar cells are typically large area devices.

- While in photodiodes, Quantum Efficiency (QE) is used to measure Signal-to-Noise Ratio (SNR), in solar cells it is used to measure Power Conversion Efficiency (PCE). It is common practice for solar cells and the external load they are linked to be constructed to optimize the amount of power they produce.

16. Light-emitting diodes are just forward-biased p-n junctions that spontaneously emit light. Radiative recombination of electron-hole pairs in the depletion area causes spontaneous emission (or electroluminescence). LEDs have a large angular bandwidth and a broad spectral bandwidth (20-150 nm). They are temporally and spatially incoherent sources.
17. Gain or amplification is a process in laser physics in which the medium transmits a portion of its energy to the produced electromagnetic radiation, resulting in an increase in optical power. All lasers operate on the same basis. Gain is a measure of a laser medium's capacity to increase optical power in a quantitative sense.

5.7 SUMMARY

- Metallic contacts at the two ends of the bar are ohmic contacts (terminals) called Source and Drain.
- Source is an *N*-channel JFET the terminal which is connected to the negative pole of the DC source providing drain voltage (V_{DD}) and through which the majority carries, i.e., electrons enter into the semiconductor bar is called Source.
- Drain is the terminal of an *N*-channel JFET connected to the positive pole of the drain voltage source (V_{DD}) and through which the majority carriers, i.e., electrons leave the bar is known as Drain.
- Gate is the heavily doped P^+ regions on the two sides of the *N*-type silicon bar forming *P-N* junctions which are joined together are called the gate. A voltage V_{GS} is applied between the gate and source to reverse bias the *P-N* junctions.
- The plots of drain current (I_D) versus drain-to-source voltage (V_{DS}) for different values of gate-to-source voltage are known as drain characteristics. These characteristics can be experimentally determined.
- The plots of drain current (I_D) versus gate-to-source voltage (V_{GS}) for different values of drain-to-source voltage (V_{DS}) are known as transfer characteristics.
- Metal Oxide Semiconductor Field Effect Transistor (MOSFET) also known as Insulated Gate Field Effect Transistor (IGFET) is an important member of FET family. The most important advantage of MOSFETs over JFETs is their larger input impedance.
- Application of gate voltage produces a transverse electric field across the insulating silicon dioxide layer deposited on silicon substrate. The resistance and thickness of the conducting channel made of silicon can be controlled by varying the gate voltage.

NOTES

- A MOS structure is formed by the gate metal along with insulating silicon dioxide layer and the semiconductor channel. The channel and gate act like two parallel plates of a capacitor separated by the dielectric (SiO_2) layer. This layer is responsible for high input impedance (10^{10} to 10^{15} Ω) for the MOSFET and also the reason for the name Insulated Gate Field Effect Transistor (IGFET) as SiO_2 is an insulating layer.
- The structure of *P*-channel enhancement MOSFET consists of *N*-type silicon substrate in which two highly doped P^+ regions are diffused forming the source and drain regions. A negative gate voltage with respect to source induces positive charges to form the channel below the gate dielectric and make the channel conductive.
- The channel conductivity of both modes of MOSFETs is controlled by a transverse electric field applied from gate to channel through the gate dielectric (SiO_2) layer. In JFETs the conductivity of the channel is controlled by a transverse field across the reverse biased *P-N* junction.
- Microwaves are incredibly short waves, as the name implies. In general, RF energy ranges from DC to infrared, and it is a type of electromagnetic energy. A quick check at the various frequency ranges reveals that the Microwave frequency range includes UHF (Ultra-High Frequency) and SHF (Super High Frequencies), with wave length (ë) ranging from 1 to 100 cm.
- The velocity-modulation hypothesis was used to achieve microwave production and amplification.
- Tunnel diodes' potential for microwave applications was soon proven after Esaki's famous work on tunnel diodes was published in 1958.
- The tunnel diode is important in microwave oscillators and amplifiers because it exhibits a negative resistance characteristic in the region between peak current Ip and valley current Iv.
- Gunn-effect diodes-GaAs diode method uses the diode's avalanching and transit-time features to generate microwave frequencies.
- When the electric field is changed from zero to a threshold value, the carrier drift velocity increases linearly from zero to a maximum, according to Gunn's observations. The drift velocity is reduced and the diode exhibits negative resistance when the electric field exceeds the threshold value of 3000 V/cm for n-type GaAs.
- The Gunn Effect has been explained in a variety of ways. In 1964, Kroemer stated that Gunn's data matched the Ridley-Watkins-Hilsum (RWH) theory perfectly.
- Kroemer proposed a negative mass microwave amplifier in 1958 and 1959, a few years before the Gunn Effect was discovered.
- A parametric device is one that employs a time-varying or nonlinear reactance (capacitance or inductance). Because capacitance or inductance, which is a reactive property, may be utilized to create capacitive or inductive excitation, the term parametric is derived from the phrase parametric excitation.

NOTES

- Excitations are the transitions of electrons from lower to higher energy states by absorbing external energy.
- A Light Dependent Resistor (also called a photoresistor or LDR) is a resistive device whose resistance varies in response to incoming electromagnetic radiation. As a result, they are photosensitive devices. Additionally, they are referred to as photoconductors, photoconductive cells, or just photocells.
- Light-emitting diodes are just forward-biased p-n junctions that spontaneously emit light. Radiative recombination of electron-hole pairs in the depletion area causes spontaneous emission (or electroluminescence). LEDs have a large angular bandwidth and a broad spectral bandwidth (20-150 nm). They are temporally and spatially incoherent sources.
- Gain or amplification is a process in laser physics in which the medium transmits a portion of its energy to the produced electromagnetic radiation, resulting in an increase in optical power. All lasers operate on the same basis. Gain is a measure of a laser medium's capacity to increase optical power in a quantitative sense.

5.8 KEY TERMS

- **Drain (*D*):** The terminal of an *N*-channel JFET connected to the positive pole of the drain voltage source (V_{DD}) and through which the majority carriers, i.e., electrons leave the bar is known as Drain.
- **Parametric devices:** A parametric device is one that employs a time-varying or nonlinear reactance (capacitance or inductance). Because capacitance or inductance, which is a reactive property, may be utilized to create capacitive or inductive excitation, the term parametric is derived from the phrase parametric excitation.
- **LDR:** A Light Dependent Resistor (also called a photoresistor or LDR) is a resistive device whose resistance varies in response to incoming electromagnetic radiation. As a result, they are photosensitive devices. Additionally, they are referred to as photoconductors, photoconductive cells, or just photocells.
- **Optical gain:** Gain or amplification is a process in laser physics in which the medium transmits a portion of its energy to the produced electromagnetic radiation, resulting in an increase in optical power. All lasers operate on the same basis. Gain is a measure of a laser medium's capacity to increase optical power in a quantitative sense.

5.9 SELF-ASSESSMENT QUESTIONS AND EXERCISES

Short-Answer Questions

1. Define the term JFET.
2. Differentiate between MOSFET and MESFET.

NOTES

3. Write in brief about tunnel diode, Gunn diode and impatt diodes.
4. What are parametric devices?
5. State the radiative and non-radiative transmitter.
6. Define the term LDR.
7. What is photodiode detectors?
8. How will you define the solar cells?
9. What do you mean by the LED?
10. State the optical gain and threshold current for lasing.

Long-Answer Questions

1. Explain the structure, working and characteristics of JFET with the help of giving examples.
2. Discuss the structure, working and characteristics of MOSFET and MESFET. Give appropriate examples.
3. What are microwave tunnel diodes? Explain with the help of giving examples.
4. Discuss the physical characteristics of parametric devices with the help of examples.
5. Discuss the radiative and non-radiative transmitter. Give appropriate examples.
6. Explain the Light Dependent Resistor (LDR) with the help of examples.
7. Describe the physical processes in light detection with the help of relevant examples.
8. Explain the fundamentally different characteristics of solar cells and photodetectors. Give appropriate examples.
9. What do you understand by the Light Emitting Diode (LED)? Discuss the advantages and disadvantages of light emitting diode.
10. Analyse the optical gain and threshold current for lasing with the help of examples.

5.10 FURTHER READING

Basavaraj, B. and H.N. Shivashankar. 2004. *Basic Electronics*. New Delhi: Vikas Publishing House Pvt. Ltd.

Black, Uyless D. 1993. *Computer Networks.* New Jersey: Prentice-Hall.

Kumar, A. Anand. 2003. *Fundamentals of Digital Circuits*. New Delhi: Prentice-Hall of India.

Mano, Morris. 1979. *Digital Logic and Computer Design*. New Delhi: Prentice-Hall of India.

Roth, Charles. 2001. *Fundamentals of Logic Design*. Thomson Learning.

Shepherd, Robert D. 1998. *Introduction to Computers and Technology*. EMC/ Paradigm Publishing.

Stallings, William. 1996. *Data and Computer Communications.* New Jersey: Prentice-Hall.

Tanenbaum, Andrew. 2002. *Computer Networks.* New Jersey: Prentice-Hall PTR.

Yarbarough, John M. 1996. *Digital Logic Applications and Design*. Thomson Learning.

Carlson Bruce and Paul B. Crilly. 2010. Communication System. Noida: Tata McGraw-Hill Education.

Lathi B. P. and Zhi Ding. 2010. Modern Digital and Analog Communication Systems. Oxford: Oxford University Press.

Kennedy George and Bernard Davis. 1993. Electronic Communication Systems. London: Glencoe/ McGraw-Hill.

NOTES

NOTES

NOTES

NOTES