

Course Material

On

High Frequency Engineering

Subject Code: PET6I102



Course: - B. Tech

Discipline: - Electronics and Telecommunication Engineering

Semester: - 6th

Syllabus

HIGH FREQUENCY ENGINEERING (PET6I102)

MODULE-I

Microwave Tubes- Limitations of conventional tubes, construction, operation; Properties of Klystron Amplifier, reflex Klystron, Magnetron, Travelling Wave Tube (TWT); Backward Wave Oscillator (BWO); Crossed field amplifiers.

MODULE-II

Microwave Solid State Devices- Limitation of conventional solid state devices at Microwaves; Transistors (Bipolar, FET); Diodes (Tunnel, Varactor, PIN), Transferred Electron Devices (Gunn diode); Avalanche transit time effect (IMPATT, TRAPATT, SBD); Microwave Amplification by Stimulated Emission of Radiation (MASER).

MODULE-III

Microwave Components- Analysis of Microwave components using s-parameters, Junctions (E, H, Hybrid), Directional coupler; Bends and Corners; Microwave posts, S.S. tuners, Attenuators, Phase shifter, Ferrite devices (Isolator, Circulator, Gyrator); Cavity resonator.

MODULE-IV

Introduction to Radar Systems- Basic Principle-Block diagram and operation of Radar; Radar range Equation; Pulse Repetition Frequency (PRF) and Range Ambiguities.

Doppler Radars- Doppler determination of velocity, Continuous Wave (CW) radar and its limitations, Frequency Modulated Continuous Wave (FMCW) radar, Basic principle and operation of Moving Target Indicator (MTI) radar, Delay line cancellers, Blind speeds and staggered PRFs.

Scanning and Tracking Techniques- Various scanning techniques (Horizontal, vertical, spiral, palmer, raster, nodding); Angle tracking systems (Lobe switching, conical scan, mono pulse).

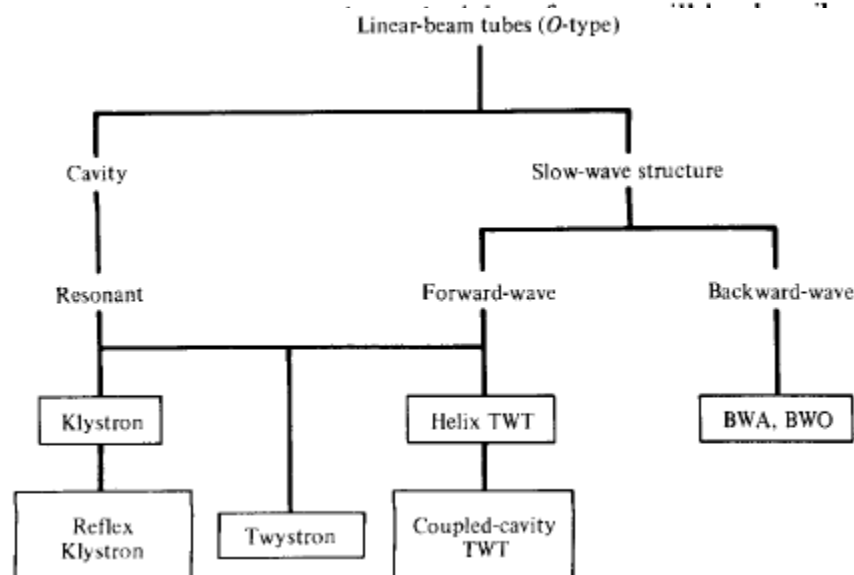
- COPYRIGHT IS NOT RESERVED BY AUTHORS.
- AUTHOR IS NOT RESPONSIBLE FOR ANY LEGAL ISSUES ARISING OUT OF ANY COPYRIGHT DEMANDS
- AND/OR REPRINT ISSUES CONTAINED IN THIS MATERIALS.
- THIS IS NOT MEANT FOR ANY COMMERCIAL PURPOSE & ONLY MEANT FOR PERSONAL USE OF STUDENTS FOLLOWING SYLLABUS PRINTED IN PREVIOUS PAGE.
- READERS ARE REQUESTED TO SEND ANY TYPING ERRORS CONTAINED, HEREIN.

Module-I

Introduction:

We turn now to a quantitative and qualitative analysis of several conventional vacuum tubes and microwave tubes in common use. The conventional vacuum tubes, such as triodes, tetrodes, and pentodes, are still used as signal sources of low output power at low microwave frequencies. The most important microwave tubes at present are the linear-beam tubes (*O* type)

The paramount *O*-type tube is the two-cavity klystron, and it is followed by the reflex klystron. The helix traveling-wave tube (TWT), the coupled-cavity TWT, the forward-wave amplifier (FWA), and the backward-wave amplifier and oscillator (BWA and BWO) are also *O*-type tubes, but they have nonresonant periodic structures for electron interactions. The Twystron is a hybrid amplifier that uses combinations of klystron and TWT components. The switching tubes such as krytron, thyatron, and planar triode are very useful in laser modulation.



In a linear-beam tube a magnetic field whose axis coincides with that of the electron beam is used to hold the beam together as it travels the length of the tube. *O*-type tubes derive their name from the French *TPO* (*tubes à propagation des ondes*) or from the word *original* (meaning the original type of tube). In these tubes electrons receive potential energy from the dc beam voltage before they arrive in the microwave interaction region, and this energy is converted into their kinetic energy. In the microwave interaction region the electrons are either accelerated or decelerated by the microwave field and then bunched as they drift down the tube. The bunched electrons, in turn, induce current in the output structure. The electrons then give up their kinetic energy to the microwave fields and are collected by the collector.

O-type traveling-wave tubes are suitable for amplification. At present, klystron and TWT amplifiers can deliver a peak power output up to 30 MW (megawatts) with a beam voltage on the order of 100 kV at the frequency of 10 GHz. The average power outputs are up to 700 kW. The gain of these tubes is on the order of 30 to 70 dB, and the efficiency is from 15 to 60%. The bandwidth is from 1 to 8% for klystrons and 10 to 15% for TWTs.

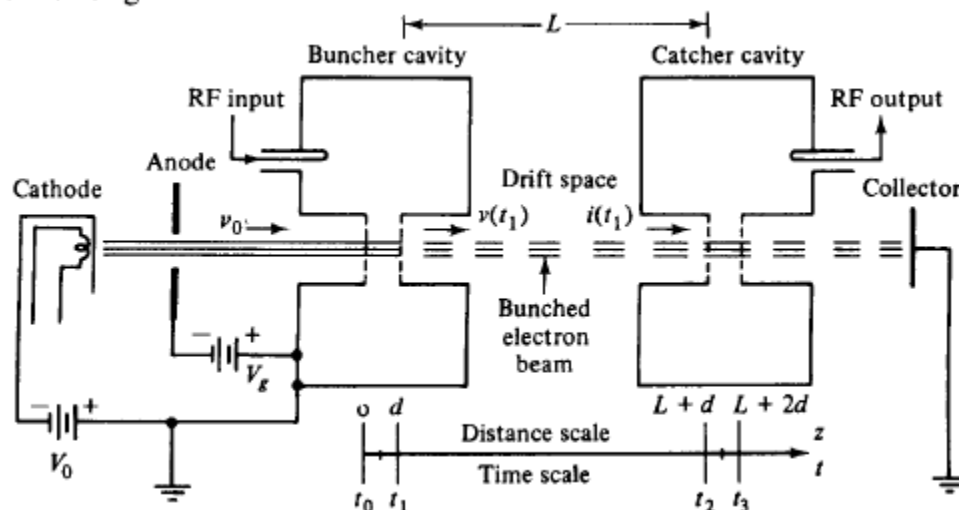
Limitations of conventional tubes

Conventional vacuum triodes, tetrodes, and pentodes are less useful signal sources at frequencies above 1 GHz because of lead-inductance and interelectrode-capacitance effects, transit-angle effects, and gain-bandwidth product limitations.

Klystron:

The two-cavity klystron is a widely used microwave amplifier operated by the principles of velocity and current modulation. All electrons injected from the cathode arrive at the first cavity with uniform velocity. Those electrons passing the first cavity gap at zeros of the gap voltage (or signal voltage) pass through with unchanged velocity; those passing through the positive half cycles of the gap voltage undergo an increase in velocity; those passing through the negative swings of the gap voltage undergo a decrease in velocity. As a result of these actions, the electrons gradually bunch together as they travel down the drift space. The variation in electron velocity in the drift space is known as *velocity modulation*. The density of the electrons in the second cavity gap varies cyclically with time. The electron beam contains an ac component and is said to be current-modulated. The maximum bunching should occur approximately midway between the second cavity grids during its retarding phase; thus the kinetic energy is transferred from the electrons to the field of the second cavity. The electrons then emerge from the second cavity with reduced velocity and finally terminate at the collector. The characteristics of a two-cavity klystron amplifier are as follows:

1. Efficiency: about 40%.
2. Power output: average power (CW power) is up to 500 kW and pulsed power is up to 30 MW at 10 GHz.
3. Power gain: about 30 dB.



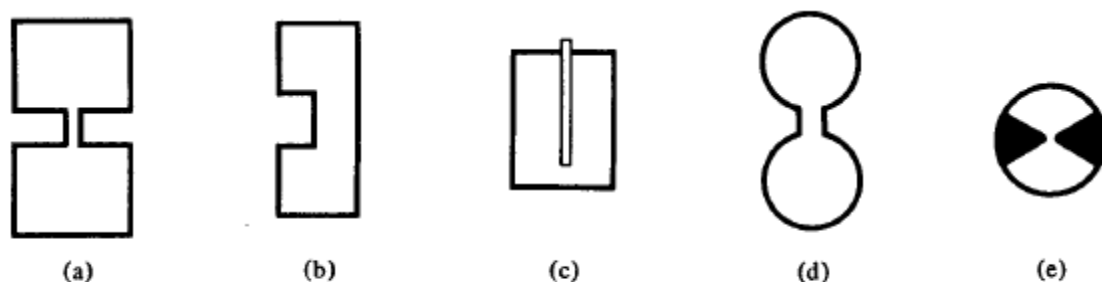
Two cavity Klystron

Figure above shows the schematic diagram of a two-cavity klystron amplifier. The cavity close to the cathode is known as the *buncher cavity* or input cavity, which velocity-modulates the electron beam. The other cavity is called the *catcher cavity* or output cavity; it catches energy from the bunched electron beam. The beam then passes through the catcher cavity and is terminated at the collector. The quantitative analysis of a two-cavity klystron can be described in four parts under the following assumptions:

1. The electron beam is assumed to have a uniform density in the cross section of the beam.
2. Space-charge effects are negligible.
3. The magnitude of the microwave signal input is assumed to be much smaller than the dc accelerating voltage.

Reentrant Cavities:

At a frequency well below the microwave range, the cavity resonator can be represented by a lumped-constant resonant circuit. When the operating frequency is increased to several tens of megahertz, both the inductance and the capacitance must be reduced to a minimum in order to maintain resonance at the operating frequency. Ultimately the inductance is reduced to a minimum by short wire. Therefore the reentrant cavities are designed for use in klystrons and microwave triodes. A reentrant cavity is one in which the metallic boundaries extend into the interior of the cavity. Several types of reentrant cavities are shown in Fig. below.



Reentrant cavities. (a) Coaxial cavity. (b) Radial cavity. (c) Tunable cavity. (d) Toroidal cavity. (e) Butterfly cavity.

Bunching Process:

The effect of velocity modulation produces bunching of the electron beam—or current modulation. The electrons that pass the buncher at $V_s = 0$ travel through with unchanged velocity v_0 and become the bunching center. Those electrons that pass the buncher cavity during the positive half cycles of the microwave input voltage V_s travel faster than the electrons that passed the gap when $V_s = 0$. Those electrons that pass the buncher cavity during the negative half cycles of the voltage V_s travel slower than the electrons that passed the gap when $V_s = 0$. At a distance of ΔL along the beam from the buncher cavity, the beam electrons have drifted into dense clusters. Figure 9-2-8 shows the trajectories of minimum, zero, and maximum electron acceleration.

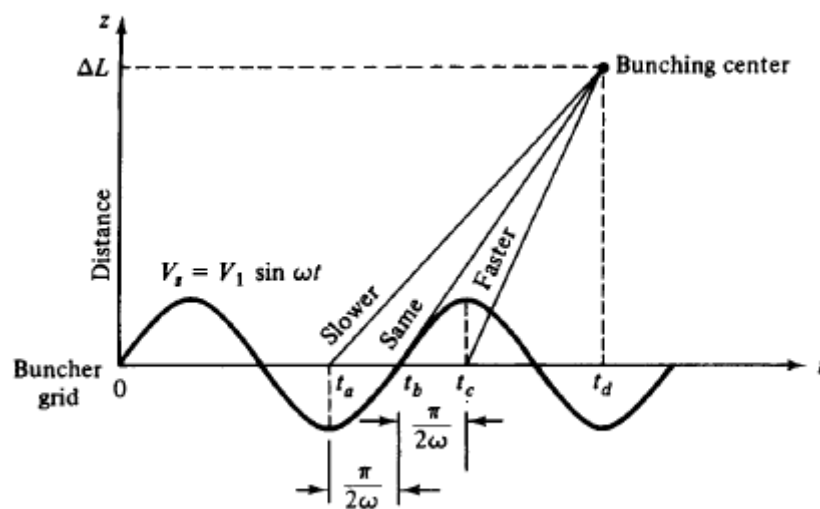
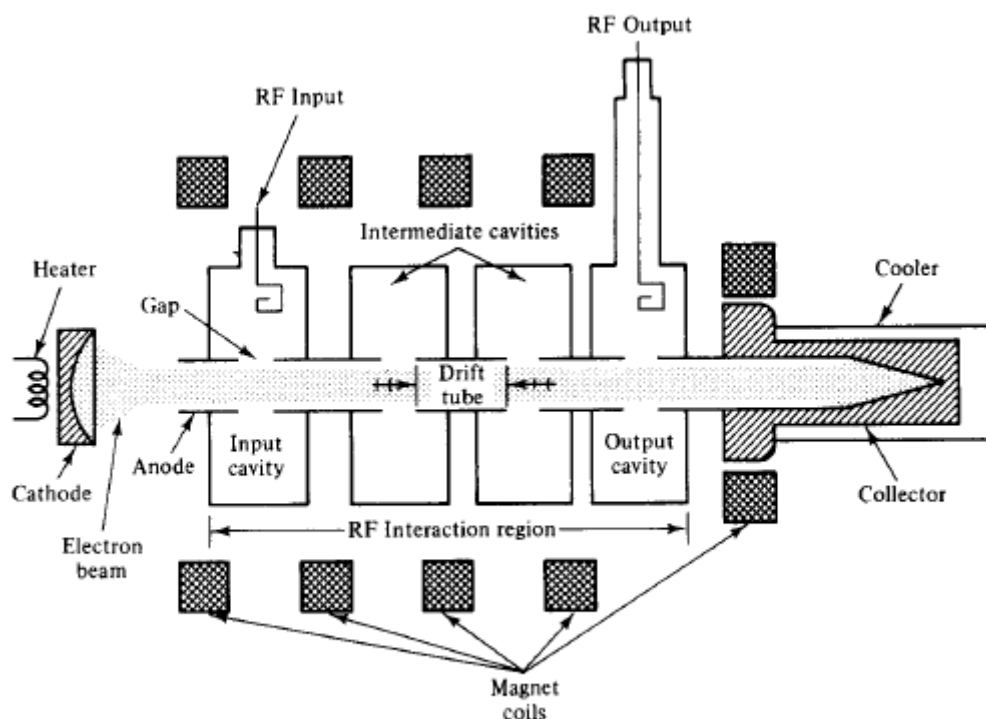


Figure Bunching distance.

Multicavity Klystron Amplifier:

The typical power gain of a two-cavity klystron amplifier is about 30 dB. In order to achieve higher overall gain, one way is to connect several two-cavity tubes in cascade, feeding the output of each of the tubes to the input of the following one. Besides using the multistage techniques, the tube manufacturers have designed and produced multicavity klystron to serve the high-gain requirement as shown in Fig.



Schematic diagram of a four-cavity klystron amplifier.

In a multicavity klystron each of the intermediate cavities, placed at a distance of the bunching parameter X of 1.841 away from the previous cavity, acts as a buncher with the passing electron beam inducing a more enhanced RF voltage than the previous cavity, which in turn sets up an increased velocity modulation. The spacing between the consecutive cavities would therefore diminish because of the requirement of X being 1.841 and an increasing velocity modulating RF voltage as the beam progresses through the various cavities. Keeping the intercavity distance constant, an increasing beam voltage V_0 could be used in the subsequent cavities.

When the two-cavity klystron amplifier was discussed, it was assumed that the space-charge effect was negligible, and it was not considered because of the assumed small density of electrons in the beam for low-power amplifier. However, when high-power klystron tubes are analyzed the electron density of the beam is large and the forces of mutual repulsion of the electrons must be considered. When the electrons perturbate in the electron beam, the electron density consists of a dc part plus an RF perturbation caused by the electron bunches. The space-charge forces within electron bunches vary with the size and shape of an electron beam. In an infinitely wide beam, the electric fields (and, thus, the space-charge forces) are constrained to act only in the axial direction. In a finite beam, the electric fields are radial as well as axial with the result that the axial component is reduced in comparison to the infinite beam. With reduced axial space-charge force, the plasma frequency is reduced and the plasma wavelength is increased.

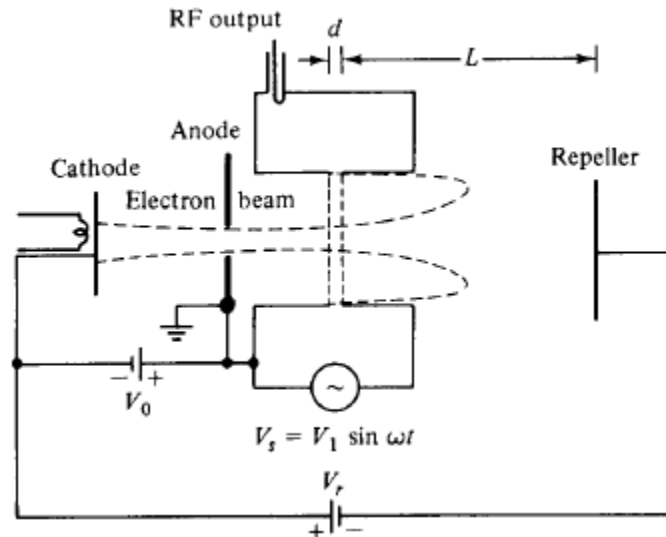
Reflex Klystron:

The reflex klystron is a single cavity variable frequency time-base generator of low power and load efficiency

APPLICATION:

- It is widely used as in radar receiver
- Local oscillators in microwave receiver
- Portable microwave rings
- Pump oscillator in parametric amplifier

If a fraction of the output power is fed back to the input cavity and if the loop gain has a magnitude of unity with a phase shift of multiple 2π , the klystron will oscillate. However, a two-cavity klystron oscillator is usually not constructed because, when the oscillation frequency is varied, the resonant frequency of each cavity and the feedback path phase shift must be readjusted for a positive feedback. The reflex klystron is a single-cavity klystron that overcomes the disadvantages of the two-cavity klystron oscillator. It is a low-power generator of 10 to 500-mW output at a frequency range of 1 to 25 GHz. The efficiency is about 20 to 30%. This type is widely used in the laboratory for microwave measurements and in microwave receivers as local oscillators in commercial, military, and airborne Doppler radars as well as missiles. The theory of the two-cavity klystron can be applied to the analysis of the reflex klystron with slight modification. A schematic diagram of the reflex klystron is shown in Fig.



The electron beam injected from the cathode is first velocity-modulated by the cavity-gap voltage. Some electrons accelerated by the accelerating field enter the repeller space with greater velocity than those with unchanged velocity. Some electrons decelerated by the retarding field enter the repeller region with less velocity. All electrons turned around by the repeller voltage then pass through the cavity gap in bunches that occur once per cycle. On their return journey the bunched electrons pass through the gap during the retarding phase of the alternating field and give up their kinetic energy to the electromagnetic energy of the field in the cavity. Oscillator output energy is then taken from the cavity. The electrons are finally collected by the walls of the cavity or other grounded metal parts of the tube.

Or

- Reflex cavity klystron consists of an electron gun, filament surrounded by cathode and a floating electron at cathode potential
- Electron gun emits electron with constant velocity

$$\frac{1}{2}mv^2 = qVa$$

$$V = \sqrt{\frac{2qVa}{m}} \text{ m/s}$$

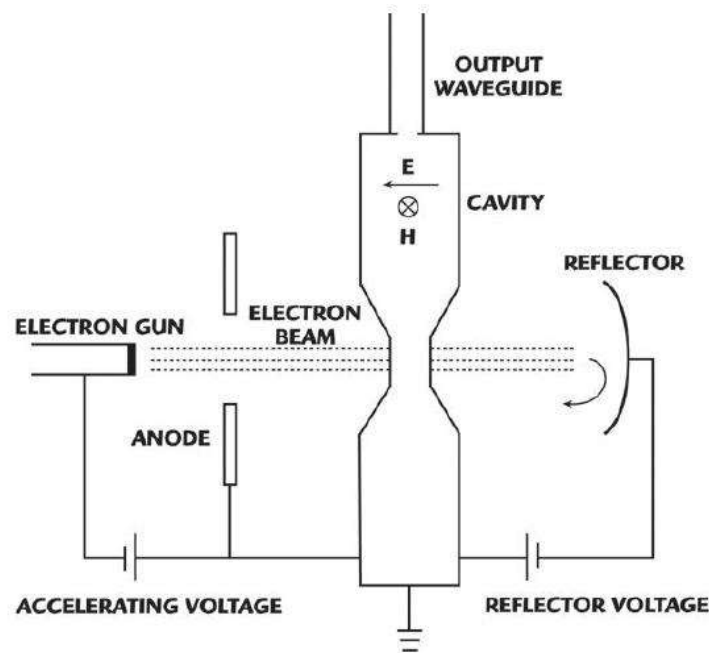
- The electron that are emitted from cathode with constant velocity enter the cavity where the velocity of electrons is changed or modified depending upon the cavity voltage.
- The oscillations is started by the device due to high quality factor and to make it sustained we have to apply the feedback.

Hence there are the electrons which will bunch together to deliver the energy act a time to the RF signal.

- Inside the cavity velocity modulation takes place. Velocity modulation is the process in which the velocity of the emitted electrons are modified or change with respect to cavity voltage. The exit velocity or velocity of the electrons after the cavity is given as

$$V' = \sqrt{\frac{2q(Va + Vi \sin(\omega t))}{m}}$$

- In the cavity gap the electrons beams get velocity modulated and get bunched to the drift space existing between cavity and repellar.
- Bunching is a process by which the electrons take the energy from the cavity at a different time and deliver to the cavity at the same time.



Bunching continuously takes place for every negative going half cycle and the most appropriate time for the electrons to return back to the cavity, when the cavity has positive peak. So that it can give maximum retardation force to electron.

- It is found that when the electrons return to the cavity in the second positive peak that is 1 whole $\frac{3}{4}$ cycle. ($n=1\pi$). It is obtained max power and hence it is called dominant mode.

The electrons are emitted from cathode with constant anode voltage V_a , hence the initial entrance velocity of electrons is

$$V = \sqrt{\frac{2qVa}{m}} \text{ m/s}$$

Inside the cavity the velocity is modulated by the cavity voltage $V_i \sin(\omega t)$ as,

$$V = \sqrt{\frac{2q(V_a + V_i \sin(\omega t))}{m}}$$

$$V = \sqrt{\frac{2qV_a}{m} + \frac{2qV_a K_a V_i \sin \omega t}{V_{am}}}$$

$$V = \sqrt{V_o^2 + \frac{V_o^2 K_i V_i \sin \omega t}{V_a}}$$

$$V = V_o \sqrt{1 + \frac{K_i V_i \sin \omega t}{V_a}}$$

$$V = \sqrt{1 + \frac{V_o^2 K_i V_i \sin \omega t}{V_a}}$$

When, $\frac{K_i V_i}{V_a}$ = Depth of velocity

$$V = V_o \left(1 + \frac{V_o^2 K_i V_i \sin \omega t}{2V_a} \right)$$

Magnetron:

DIFFERENCE BETWEEN REFLEX KLYSTRON AND MAGNETRON:

REFLEX KLYSTRON	MAGNETRON
❖ It is a linear tube in which the magnetic field is applied to focus the electron and electric field is applied to drift the electron. ❖ In klystron the bunching takes place only inside the cavity which is very small, hence generate low power and low frequency.	❖ In magnetron the magnetic field and electric field are perpendicular to each other hence it is called as cross field device. ❖ In magnetron the interacting or bunching space is extended so the efficiency can be increased.

APPLICATION:

- Used as oscillator.
- Used in radar communication.
- Used in missiles.

➤ Used in microwave oven (in the range of frequency of 2.5Ghz). Types of magnetron:

Microwave cross field Tubes:

In the **last section** several commonly used linear-beam tubes were described in detail. In these tubes, the dc magnetic field that is in parallel with the dc electric field is used merely to focus the electron beam. In crossed-field devices, however, the dc magnetic field and the dc electric field are perpendicular to each other. In all crossed-field tubes, the dc magnetic field plays a direct role in the RF interaction process.

Crossed-field tubes derive their name from the fact that the dc electric field and the dc magnetic field are perpendicular to each other. They are also called *M*-type tubes after the French TPOM (*tubes à propagation des ondes à champs magnétique*: tubes for propagation of waves in a magnetic field). In a crossed-field tube, the electrons emitted by the cathode are accelerated by the electric field and gain velocity, but the greater their velocity, the more their path is bent by the magnetic field. If an RF field is applied to the anode circuit, those electrons entering the circuit during the retarding field are decelerated and give up some of their energy to the RF field. Consequently, their velocity is decreased, and these slower electrons will then travel the dc electric field far enough to regain essentially the same velocity as before. Because of the crossed-field interactions, only those electrons that have given up sufficient energy to the RF field can travel all the way to the anode. This phenomenon would make the *M*-type devices relatively efficient. Those electrons entering the circuit during the accelerating field are accelerated by means of receiving enough energy from the RF field and are returned back toward the cathode. This back-bombardment of the cathode produces heat in the cathode and decreases the operational efficiency.

Here, several commonly used crossed-field tubes such as magnetrons, forward-wave crossed-field amplifiers (FWCFAs), backward-wave crossed-field amplifiers (BWCFA or amplitrons), and backward-wave crossed-field oscillators (BWCFO or carcinotrons) are studied.

Cylindrical magnetron: The cylindrical magnetron was developed by Boot and Randall in early 1940.

Coaxial magnetron: The coaxial magnetron introduced the principle of integrating a stabilizing cavity into the magnetron geometry.

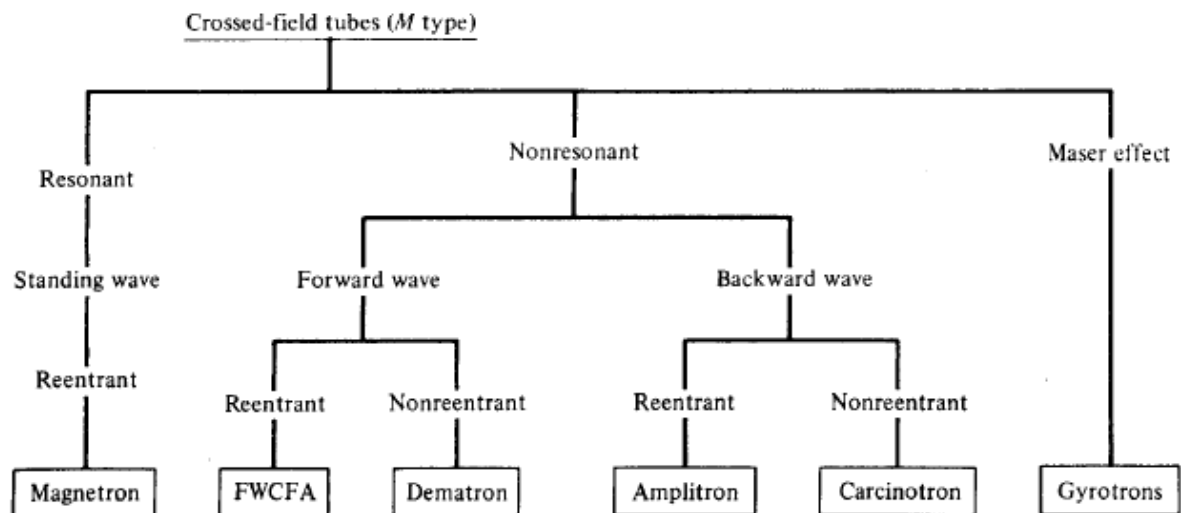
Voltage-tunable magnetron: The voltage-tunable magnetron has the cathode-anode geometry of the conventional magnetron, but its anode can be tuned easily.

Inverted magnetron: The inverted magnetron has the inverted geometry of the conventional magnetron with the cathode placed on the outside surrounding the anode and microwave circuit.

Forward-wave crossed-field amplifier (FWCFA): The forward-wave crossed-field amplifier is also called the *M*-type forward-wave amplifier.

Backward-wave crossed-field amplifier (BWCFA): The backward-wave crossed-field amplifier was developed by Raytheon Company in 1960 and its trade name is Amplitron. It is a broadband, high-power, high-gain, and high-efficiency microwave tube, and it has many applications such as in airborne radar systems and spaceborne communications systems.

Backward-wave crossed-field oscillator (BWCFO): In 1960, the French Company developed a new type of crossed-field device in which an injection gun replaced the conventional cylindrical cathode of the magnetron. Its trade name is Carcinotron, and it is also called the *M*-type backward-wave oscillator.



Magnetron Oscillator:

All magnetrons consist of some form of anode and cathode operated in a dc magnetic field normal to a dc electric field between the cathode and anode. Because of the crossed field between the cathode and anode, the electrons emitted from the cathode are influenced by the crossed field to move in curved paths. If the dc magnetic field is strong enough, the electrons will not arrive in the anode but return instead to the cathode. Consequently, the anode current is cut off. Magnetrons can be classified into three types:

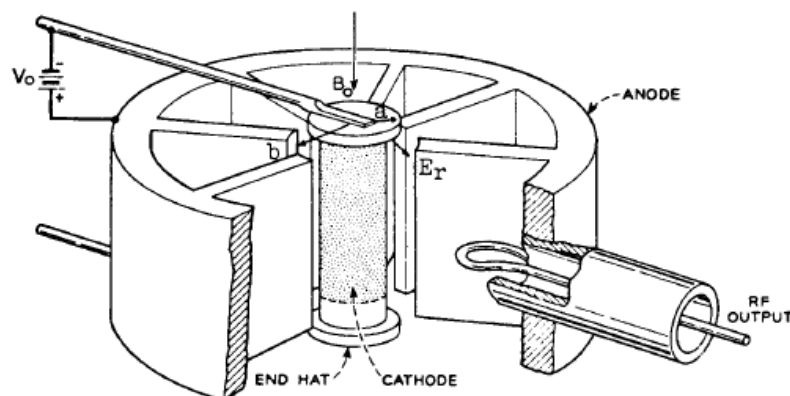
1. *Split-anode magnetron*: This type of magnetron uses a static negative resistance between two anode segments.
2. *Cyclotron-frequency magnetrons*: This type operates under the influence of synchronism between an alternating component of electric field and a periodic oscillation of electrons in a direction parallel to the field.
3. *Traveling-wave magnetrons*: This type depends on the interaction of electrons with a traveling electromagnetic field of linear velocity. They are customarily referred to simply as *magnetrons*.

Negative-resistance magnetrons ordinarily operate at frequencies below the microwave region. Although cyclotron-frequency magnetrons operate at frequencies in microwave range, their power output is very small (about 1 W at 3 GHz), and their efficiency is very low (about 10% in the split-anode type and 1% in the single-anode type). Thus, the first two types of magnetrons are not considered in this text. In this section, only the traveling-wave magnetrons such as the cylindrical magnetron, linear (or planar) magnetron, coaxial magnetron, voltage-tunable magnetron, inverted coaxial magnetron, and the frequency-agile magnetron will be discussed.

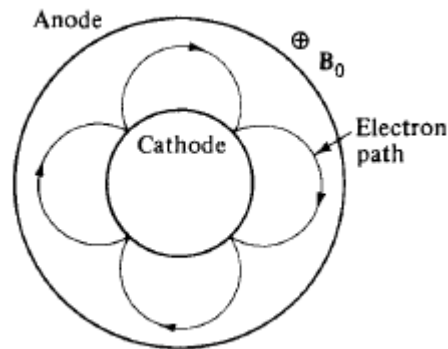
Cylindrical Magnetron:

This type of magnetron is also called a *conventional magnetron*.

In a cylindrical magnetron, several reentrant cavities are connected to the gaps. The dc voltage V_0 is applied between the cathode and the anode. The magnetic flux density B_0 is in the positive z direction. When the dc voltage and the magnetic flux are adjusted properly, the electrons will follow cycloidal paths in the cathode-anode space under the combined force of both electric and magnetic fields as shown in Fig.



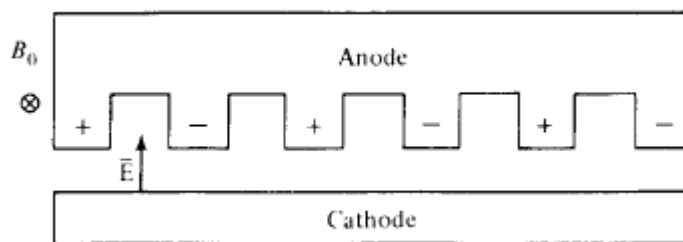
Schematic diagram of a cylindrical magnetron.



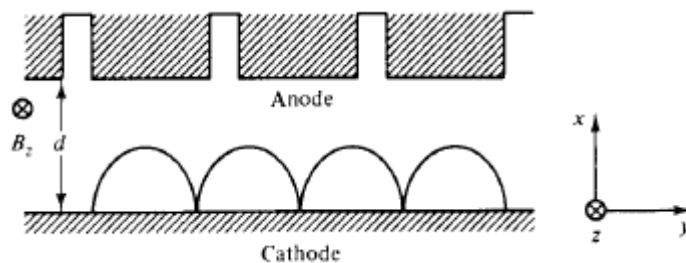
Electron path in a cylindrical magnetron.

Linear Magnetron:

The schematic diagram of a linear magnetron is shown in Fig. . In the linear magnetron as shown in Fig. , the electric field E_x is assumed in the positive x direction and the magnetic flux density B_z in the positive z direction. The differential equations of motion of electrons in the crossed-electric and magnetic fields can be written



Schematic diagram of a linear magnetron.

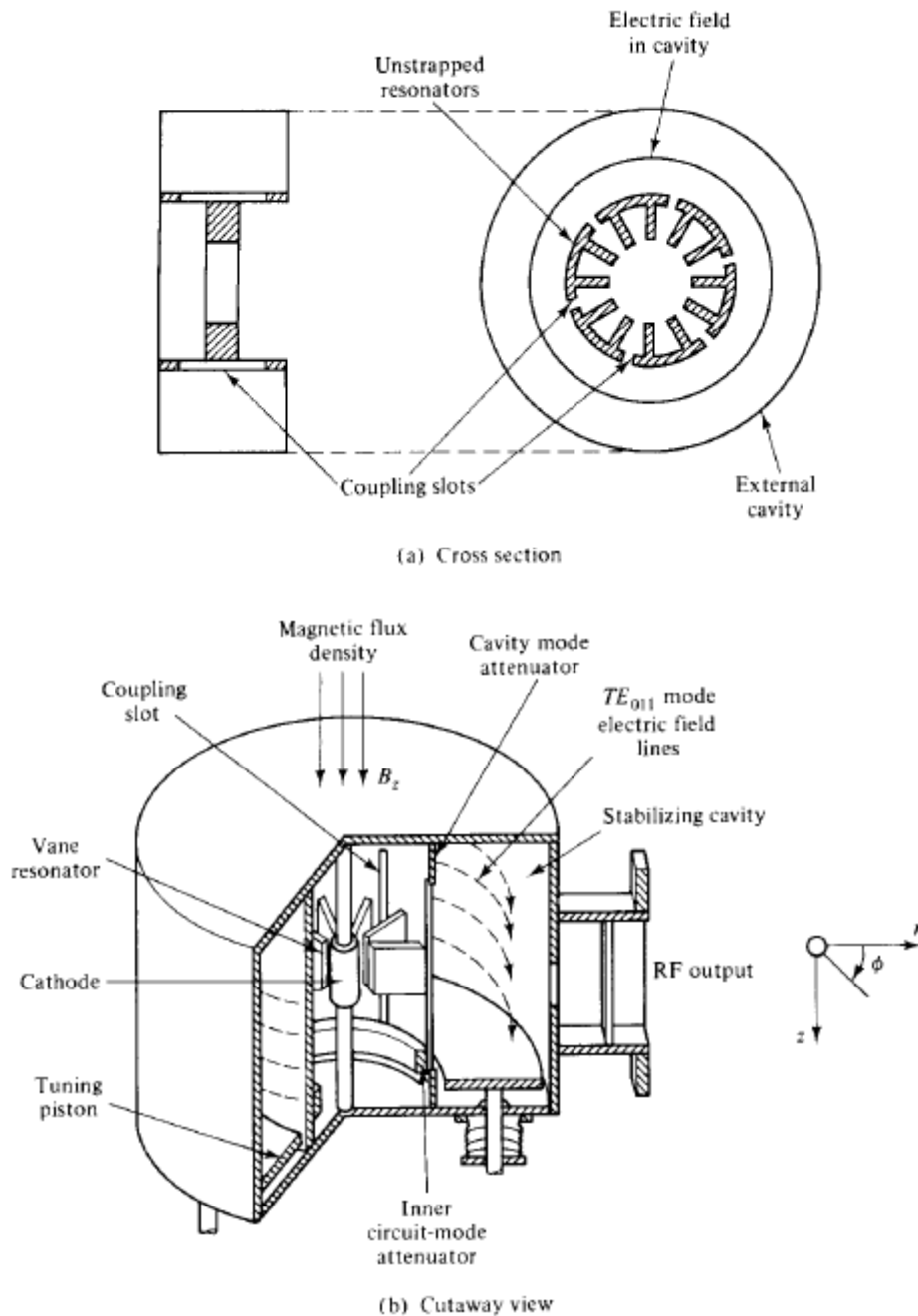


Electron path in a linear magnetron.

Coaxial Magnetron:

The coaxial magnetron is composed of an anode resonator structure surrounded by an inner-single, high- Q cavity operating in the TE_{011} mode as shown in Fig.

The slots in the back walls of alternate cavities of the anode resonator structure tightly couple the electric fields in these resonators to the surrounding cavity. In the π -mode operation, the electric fields in every other cavity are in phase, and so they couple in the same direction into the surrounding cavity. As a result, the surrounding coaxial cavity stabilizes the magnetron in the desired π -mode operation.

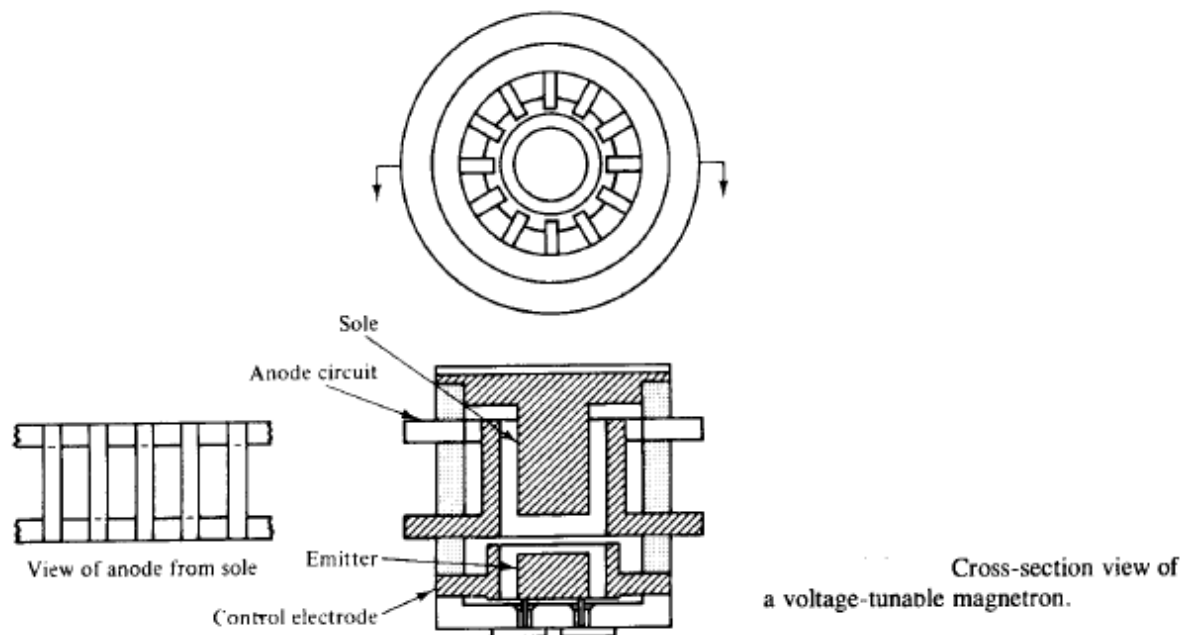


In the desired TE_{011} mode, the electric fields follow a circular path within the cavity and reduce to zero at the walls of the cavity. Current flow in the TE_{011} mode is in the walls of the cavity in circular paths about the axis of the tube. The undesired modes are damped out by the attenuator within the inner slotted cylinder near the ends of the coupling slots. The tuning mechanism is simple and reliable. As the straps are not required, the anode resonator for the coaxial magnetron can be larger and less complex than for the conventional strapped magnetron. Thus cathode loading is lower, and voltage gradients are reduced.

Voltage Tunable Magnetron:

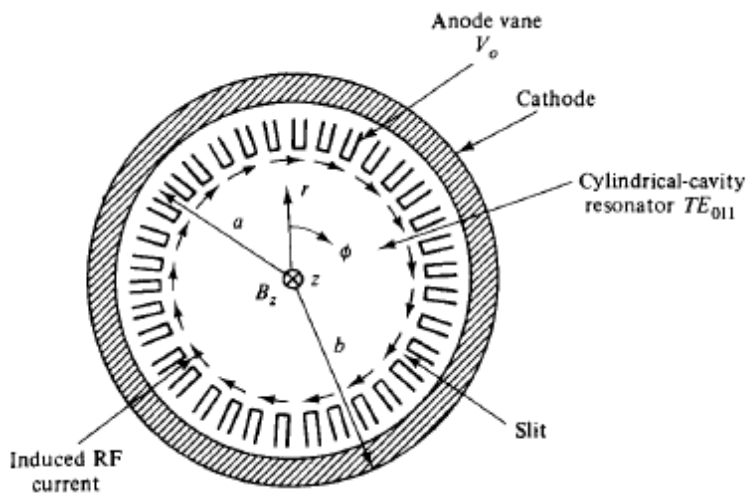
The voltage-tunable magnetron is a broadband oscillator with frequency changed by varying the applied voltage between the anode and sole. As shown in Fig. the electric beam is emitted from a short cylindrical cathode at one end of the device. Electrons are formed into a hollow beam by the electric and magnetic forces near the cathode and then are accelerated radially outward from the cathode. The electron beam is then injected into the region between the sole and the anode. The beam rotates about the sole at the rate controlled by the axial magnetic field and the dc voltage applied between the anode and the sole.

The voltage-tunable magnetron uses a low- Q resonator, and its bandwidth may exceed 50% at low-power levels. In the π -mode operation, the bunch process of the hollow beam occurs in the resonator, and the frequency of oscillation is determined by the rotational velocity of the electron beam. In other words, the oscillation frequency can be controlled by varying the applied dc voltage between the anode and sole. Power output can be adjusted to some extent through the use of the control electrode in the electron gun. At high-power levels and high frequencies, the bandwidth percentage is limited. However, at low-power levels and low frequencies, the bandwidth may approach 70%.



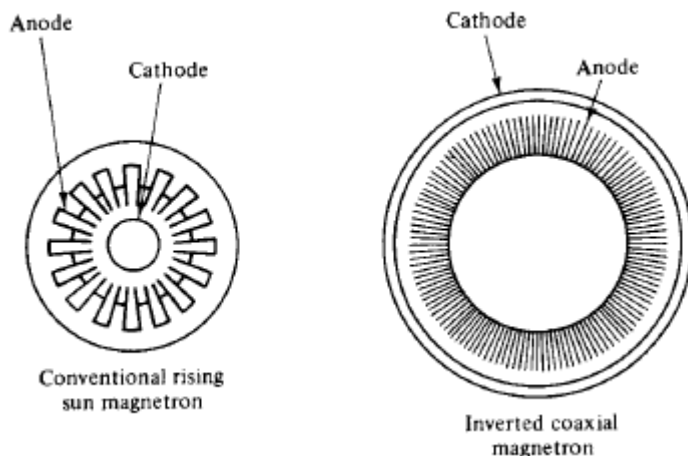
Inverted Coaxial Magnetron:

A magnetron can be built with the anode and cathode inverted—that is, with the cathode surrounding the anode. For some time, the basic problem of mode suppression prevented its use. In an inverted coaxial magnetron, the cavity is located inside a slotted cylinder, and a resonator vane array is arranged on the outside. The cathode is built as a ring around the anode. Figure shows the schematic diagram of an inverted coaxial magnetron.



Schematic diagram of an inverted coaxial magnetron.

The advantage of an inverted coaxial magnetron design is that the cathode current density can be reduced to one-tenth of that used in cathode-centered magnetrons. Thus the millimeter magnetron is a practical and long-life device. The output waveguide can be in the circular electric mode that has extremely low transmission loss. Figure below compares the inverted coaxial magnetron with a conventional magnetron designed for the same frequency. It should be noted that the cathode sizes are quite different.



Comparison of a conventional magnetron with an inverted coaxial magnetron.

Frequency Agile Coaxial Magnetron:

A frequency-agile coaxial magnetron differs from a standard tunable magnetron. The frequency agility (FA) of a coaxial magnetron is defined as the capability to tune the output frequency of the radar with sufficiently high speed to produce a pulse-to-pulse frequency change greater than the amount required effectively to obtain decorrelation of adjacent radar echoes. The frequency-agile magnetron, together with appropriate receiver integration circuits, can reduce target scintillation, increase the detectability of target in a clutter environment, and improve resistance to electronic countermeasures (ECM). The increase of the pulse-to-pulse frequency separation will improve the radar system performance. Furthermore, the greater the pulse-to-pulse frequency separation, the more difficult it will be to center a jamming transmitter on the radar frequency for effective interference with system operation.

The frequency-agile coaxial magnetrons are classified into three types:

1. *Dither magnetrons*: The output RF frequency varies periodically with a constant excursion, constant rate, and a fixed center frequency.
2. *Tunable/dither magnetrons*: The output RF frequency varies periodically with a constant excursion and constant rate, but the center frequency can be manually tuned by hand or mechanically tuned by a servomotor.
3. *Accutune magnetrons*: The output RF frequency variations are determined by the waveforms of an externally generated, low-level voltage signal. With proper selection of a tuning waveform, the accutune magnetron combines with features of dither and tunable/dither magnetrons, together with a capability for varying the excursion, rate, and tuning waveform. Figure 10-1-15 shows an accutune magnetron.

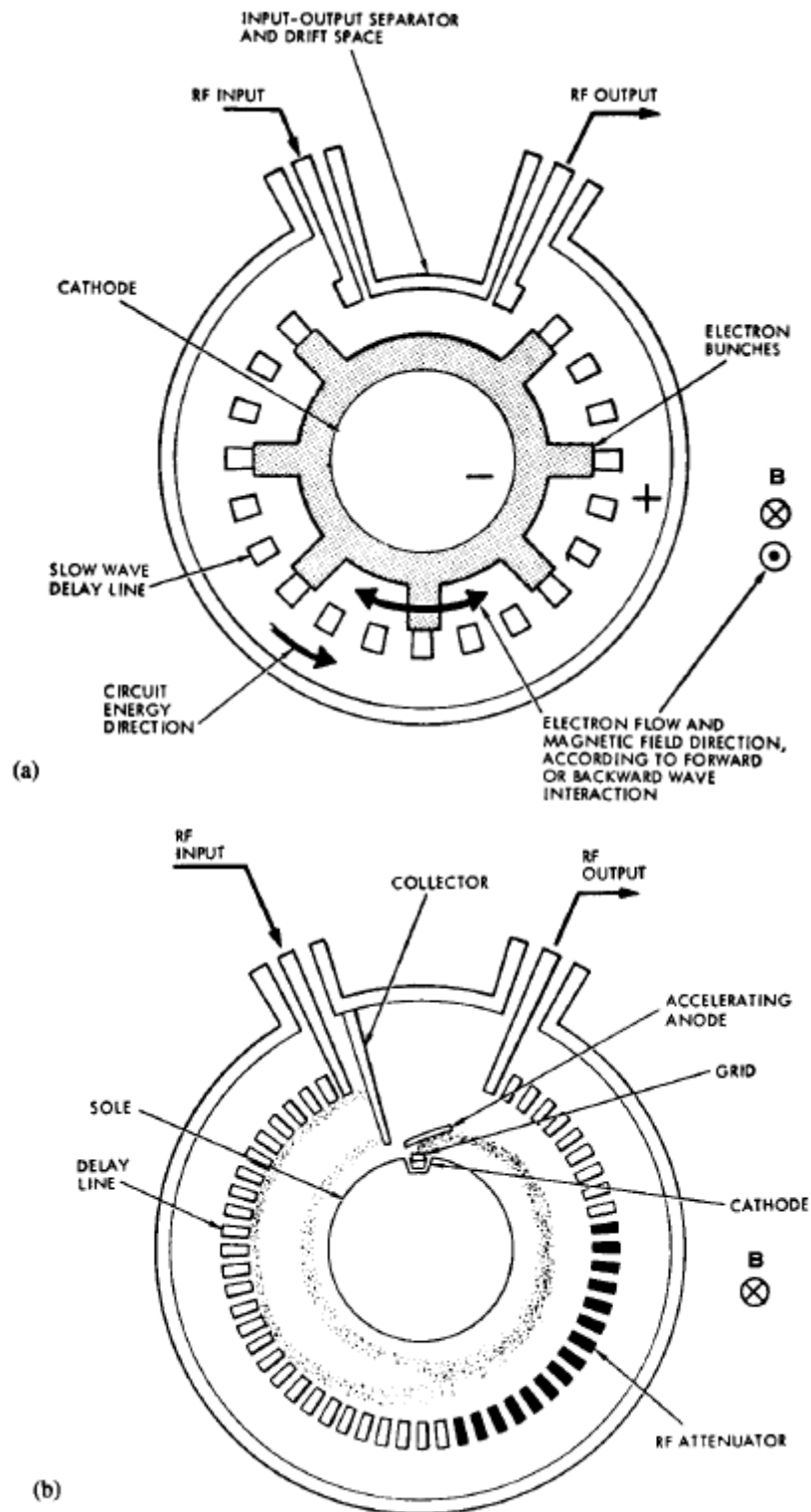
Forward Wave Cross field Amplifier:

The crossed-field amplifier (CFA) is an outgrowth of the magnetron. CFAs can be grouped by their mode of operation as forward-wave or backward-wave types and by their electron stream source as emitting-sole or injected-beam types. The first group concerns the direction of the phase and group velocity of the energy on the microwave circuit.

Since the electron stream reacts to the RF electric field forces, the behavior of the phase velocity with frequency is of prime concern. The second group emphasizes the method by which electrons reach the interaction region and how they are controlled. This can be seen in the schematic diagrams of Fig.

In the forward-wave mode, the helix-type slow-wave structure is often selected as the microwave circuit for the crossed-field amplifier; in the backward-wave mode, the strapped bar line represents a satisfactory choice. A structure of strapped crossed-field amplifier is shown in Fig.

In the emitting-sole tube, the current emanated from the cathode is in response to the electric field forces in the space between the cathode and anode. The amount of current is a function of the dimension, the applied voltage, and the emission properties of the cathode. The perveance of the interaction geometry tends to be quite high, about $5 \text{ to } 10 \times 10^{-10}$, which results in a high-current and high-power capability at relative low voltage. In the injected-beam tube the electron beam is produced in a separate gun assembly and is injected into the interaction region.



Schematic diagrams of CFAs.

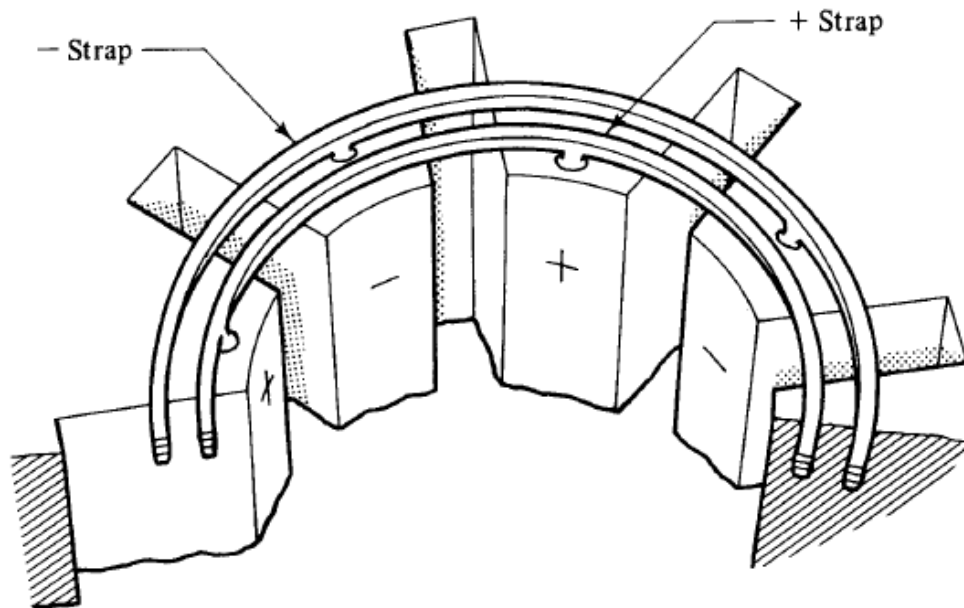
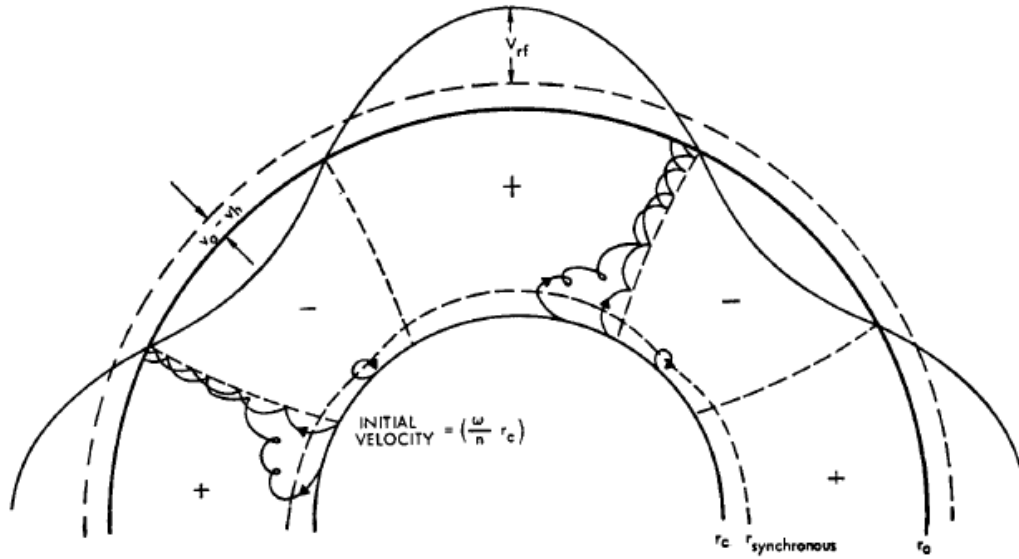


Diagram of a strapped CFA.

The beam-circuit interaction features are similar in both the emitting-sole and the injected-beam tubes. Favorably phased electrons continue toward the positively polarized anode and are ultimately collected, whereas unfavorably phased electrons are directed toward the negative polarized electrode.

In linear-beam interaction, as discussed for traveling-wave tubes, the electron stream is first accelerated by an electric gun to the full dc velocity; the dc velocity is approximately equal to the axial phase velocity of the RF field on the slow-wave structure. After interaction occurs, the spent electron beam leaves the interaction region with a low-average velocity. The difference in velocity is accounted for by the RF energy created on the microwave circuit. In the CFA, the electron is exposed to the dc electric field force, magnetic field force, and the electric field force of the RF field, and even to the space-charge force from other electrons. The last force is normally not considered in analytic approaches because of its complexity. Under the influence of the three forces, the electrons travel in spiral trajectories in a direction tending along equipotentials. The exact motion has been subject to much analysis by means of a computer. Figure below shows the pattern of the electron flow in the CFA by computerized techniques. It can be seen that when the spoke is positively polarized or the RF field is in the positive half cycle, the electron speeds up toward the anode; while the spoke is negatively polarized or the RF field is in the negative half cycle, the electrons are returned toward the cathode. Consequently, the electron beam moves in a spiral path in the interaction region.



Motion of electrons in CFA.

The total power generated in a given CFA is independent of the RF input power, as long as the input power exceeds the threshold value for spoke stability at the input. The power generated can be increased only by increasing the anode voltage and current. Neglecting circuit attenuation, the output power of the CFA is equal to the sum of the input power and the power generated in the interaction region. That is, the power gain of a CFA is given by

$$g = \frac{P_{out}}{P_{in}} = \frac{P_{in} + P_{gen}}{P_{in}} = 1 + \frac{P_{gen}}{P_{in}}$$

where $P_{out} = P_{in} + P_{gen}$

P_{in} = RF input power

P_{gen} = RF power induced into the anode circuit by electrons

Therefore, the CFA is not a linear amplifier but rather is termed a *saturated amplifier*.

The efficiency of a CFA is defined as the product of the electronic efficiency η_e and the circuit efficiency η_c .

The overall efficiency is then expressed as

$$\eta = \eta_c \eta_e = \frac{P_{out} - P_{in}}{V_{a0} I_{a0}}$$

where P_{out} = RF output power

P_{in} = RF input power

$\eta_e = P_{gen}/P_{dc}$

$P_{dc} = V_{a0} I_{a0}$ is dc power

V_{a0} = anode dc voltage

I_{a0} = anode dc current

The circuit efficiency is defined as

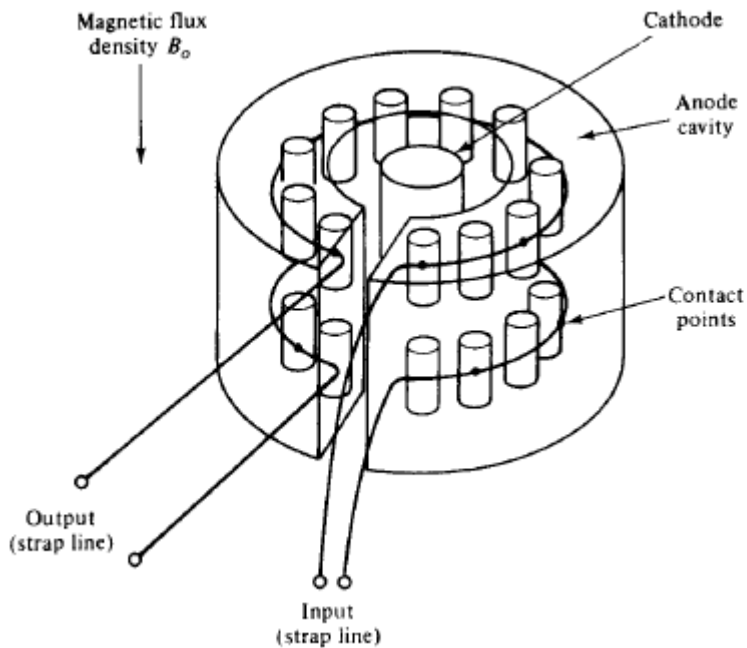
$$\eta_c = \frac{\eta}{\eta_e} = \frac{P_{out} - P_{in}}{P_{gen}}$$

where $P_{gen} = \eta_e V_{a0} I_{a0}$

Backward Wave Crossed field Amplifier (Amplitron):

The trade name of the backward-wave crossed-field amplifier (BWCFA) is Amplitron, and its schematic diagram is shown in Fig.

The anode cavity and pins comprise the resonator circuits. A pair of pins and the cavity are excited in opposite phase by the strap line. The electron beam and the electromagnetic waves interact in the resonant circuits. The BWCFA can deliver 3-MW pulse with 10- μ s duration at S-band, and the tube gain reaches 8 dB.



Schematic diagram of Amplitron.

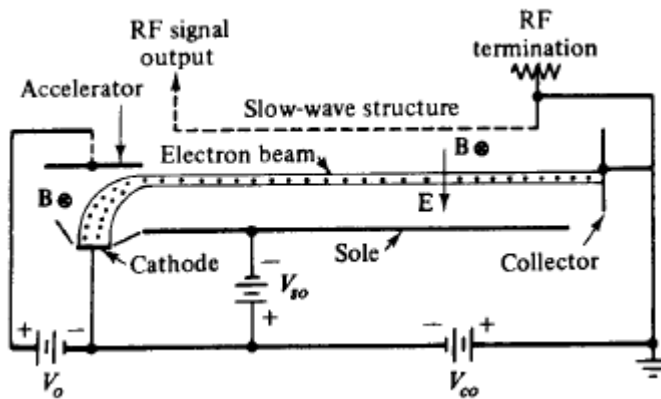
The highly successful QK434 Amplitron produced by Raytheon exhibited stable gain as high as 16 dB, power output levels ranging from a few hundred kilowatts to 3 MW, and efficiencies ranging from 60% for normal power levels to 76% for high-power, low-gain operation. The tube is commonly used in air surveillance radar and military pulsed radar.

The two-stage superpower Amplitron, also manufactured by Raytheon, generated 425 kW of CW power at an efficiency of 76%. The gain was 9 dB and the bandwidth was 5% at a mean frequency of 3 GHz. The tube was used for all high-data-rate transmission in the Apollo program.

Backward Wave Crossed field Oscillator (Carcinotron):

The backward-wave crossed-field oscillator of *M*-Carcinotron has two configurations: linear *M*-carcinotron and circular *M*-carcinotron.

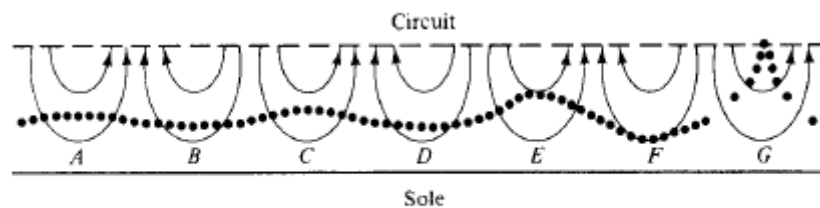
Linear *M*-Carcinotron. The *M*-Carcinotron oscillator is an *M*-type backward-wave oscillator. The interaction between the electrons and the slow-wave structure takes place in a space of crossed field. A linear model of the *M*-Carcinotron oscillator is shown in Fig.



Linear model of an *M*-Carcinotron oscillator.

The slow-wave structure is in parallel with an electrode known as the sole. A dc electric field is maintained between the grounded slow-wave structure and the negative sole. A dc magnetic field is directed into the page. The electrons emitted from the cathode are bent through a 90° angle by the magnetic field. The electrons interact with a backward-wave space harmonic of the circuit, and the energy in the circuit flows opposite to the direction of the electron motion. The slow-wave structure is terminated at the collector end, and the RF signal output is removed at the electron-gun end. Since the *M*-Carcinotron is a crossed-field device, its efficiency is very high, ranging from 30 to 60%.

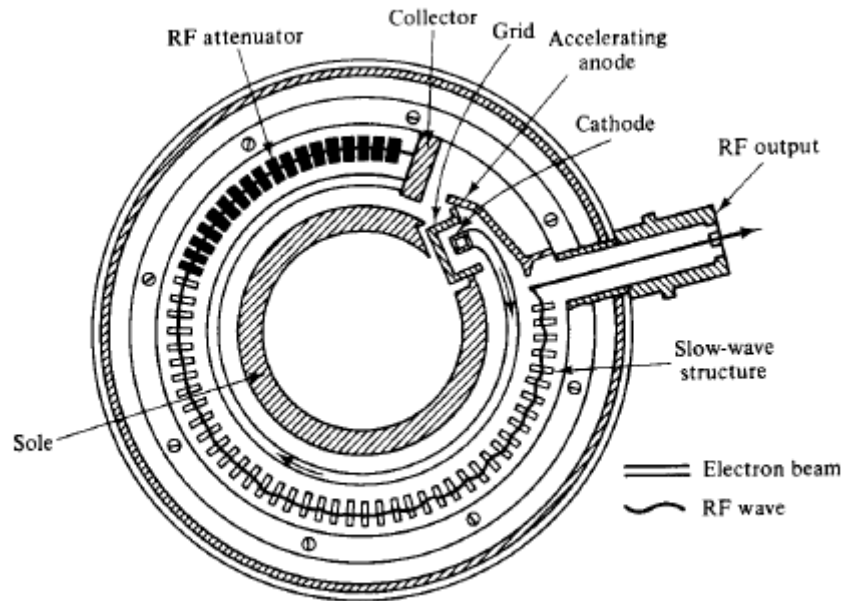
The perturbed electrons moving in synchronism with the wave in a linear *M*-Carcinotron are shown in Fig. below.



Beam electrons and electric field lines in an *M*-Carcinotron.

Electrons at position *A* near the beginning of the circuit are moving toward the circuit, whereas electrons at position *B* are moving toward the sole. Farther down the circuit, electrons at position *C* are closer to the circuit, and electrons at position *D* are closer to the sole. However, electrons at position *C* have departed a greater distance from the unperturbed path than have electrons at position *D*. Thus, the electrons have lost a net amount of potential energy, this energy having been transferred to the RF field. The reason for the greater displacement of the electrons moving toward the circuit is that these electrons are in stronger RF fields, since they are closer to the circuit. Electrons at position *G* have moved so far from the unperturbed position that some of them are being intercepted on the circuit. The length from position *A* through position *G* is a half cycle of the electron motion.

Circular *M*-Carcinotron. The *M*-Carcinotrons are generally constructed in the circular reentrant form as shown in Fig. . The slow-wave structure and sole are circular and nearly reentrant to conserve magnet weight. The sole has the appearance of the cathode in a magnetron.



Schematic diagram of a circular M-Carcinotron.

In the circular configurations, the delay line is terminated at the collector end by spraying attenuating material on the surfaces of the conductors. The output is taken from the gun end of the delay line which is an interdigital line. Clearly, in this case, the electron drift velocity has to be in synchronism with a backward-space harmonic.

Helix Traveling Wave Tubes (TWTs):

Since Kompfner invented the helix traveling-wave tube (TWT) in 1944 [11], its basic circuit has changed little. For broadband applications, the helix TWTs are almost exclusively used, whereas for high-average-power purposes, such as radar transmitters, the coupled-cavity TWTs are commonly used.

In previous sections klystrons and reflex klystrons were analyzed in some detail. Before starting to describe the TWT, it seems appropriate to compare the basic operating principles of both the TWT and the klystron. In the case of the TWT, the microwave circuit is nonresonant and the wave propagates with the same speed as the electrons in the beam. The initial effect on the beam is a small amount of velocity modulation caused by the weak electric fields associated with the traveling wave. Just as in the klystron, this velocity modulation later translates to current modulation, which then induces an RF current in the circuit, causing amplification. However, there are some major differences between the TWT and the klystron:

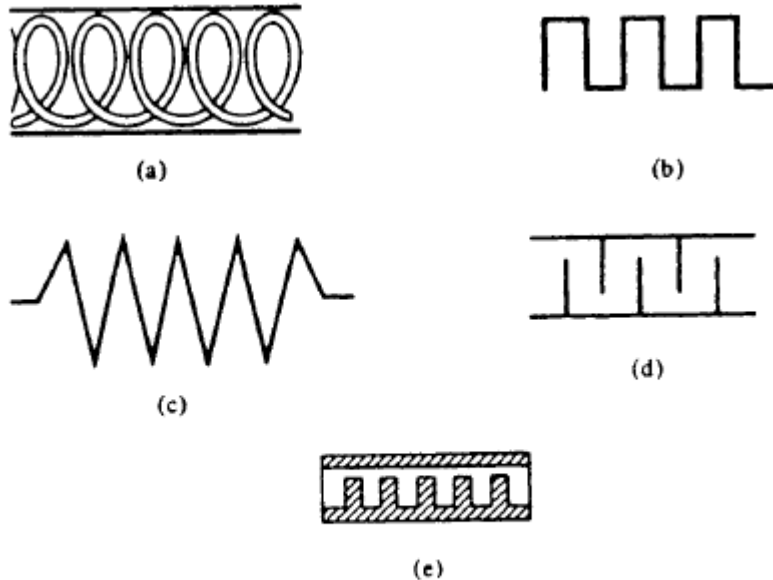
1. The interaction of electron beam and RF field in the TWT is continuous over the entire length of the circuit, but the interaction in the klystron occurs only at the gaps of a few resonant cavities.
2. The wave in the TWT is a propagating wave; the wave in the klystron is not.
3. In the coupled-cavity TWT there is a coupling effect between the cavities, whereas each cavity in the klystron operates independently.

A helix traveling-wave tube consists of an electron beam and a slow-wave structure. The electron beam is focused by a constant magnetic field along the electron beam and the slow-wave structure. This is termed an *O*-type traveling-wave tube. The slow-wave structure is either the helical type or folded-back line. The applied signal propagates around the turns of the helix and produces an electric field at the center of the helix, directed along the helix axis. The axial electric field progresses with a velocity that is very close to the velocity of light multiplied by the ratio of helix pitch to helix circumference. When the electrons enter the helix tube, an interaction takes place between the moving axial electric field and the moving electrons. On the average, the electrons transfer energy to the wave on the helix. This interaction causes the signal wave on the helix to become larger. The electrons entering the helix at zero field are not affected by the signal wave; those electrons entering the helix at the accelerating field are accelerated, and those at the retarding field are decelerated. As the electrons travel further along the helix, they bunch at the collector end. The bunching shifts the phase by $\pi/2$. Each electron in the bunch encounters a stronger retarding field. Then the microwave energy of the electrons is delivered by the electron bunch to the wave on the helix. The amplification of the signal wave is accomplished. The characteristics of the traveling-wave tube are:

Frequency range:	3 GHz and higher
Bandwidth:	about 0.8 GHz
Efficiency:	20 to 40%
Power output:	up to 10 kW average
Power gain:	up to 60 dB

Slow Wave Structures:

As the operating frequency is increased, both the inductance and capacitance of the resonant circuit must be decreased in order to maintain resonance at the operating frequency. Because the gain-bandwidth product is limited by the resonant circuit, the ordinary resonator cannot generate a large output. Several nonresonant periodic circuits or slow-wave structures are designed for producing large gain over a wide bandwidth.



Slow-wave structures. (a) Helical line. (b) Folded-back line. (c) Zigzag line. (d) Interdigital line. (e) Corrugated waveguide.

Slow-wave structures are special circuits that are used in microwave tubes to reduce the wave velocity in a certain direction so that the electron beam and the signal wave can interact. The phase velocity of a wave in ordinary waveguides is greater than the velocity of light in a vacuum. In the operation of traveling-wave and magnetron-type devices, the electron beam must keep in step with the microwave signal. Since the electron beam can be accelerated only to velocities that are about a fraction of the velocity of light, a slow-wave structure must be incorporated in the microwave devices so that the phase velocity of the microwave signal can keep pace with that of the electron beam for effective interactions. Several types of slow-wave structures are shown in Fig.

The commonly used slow-wave structure is a helical coil with a concentric conducting cylinder

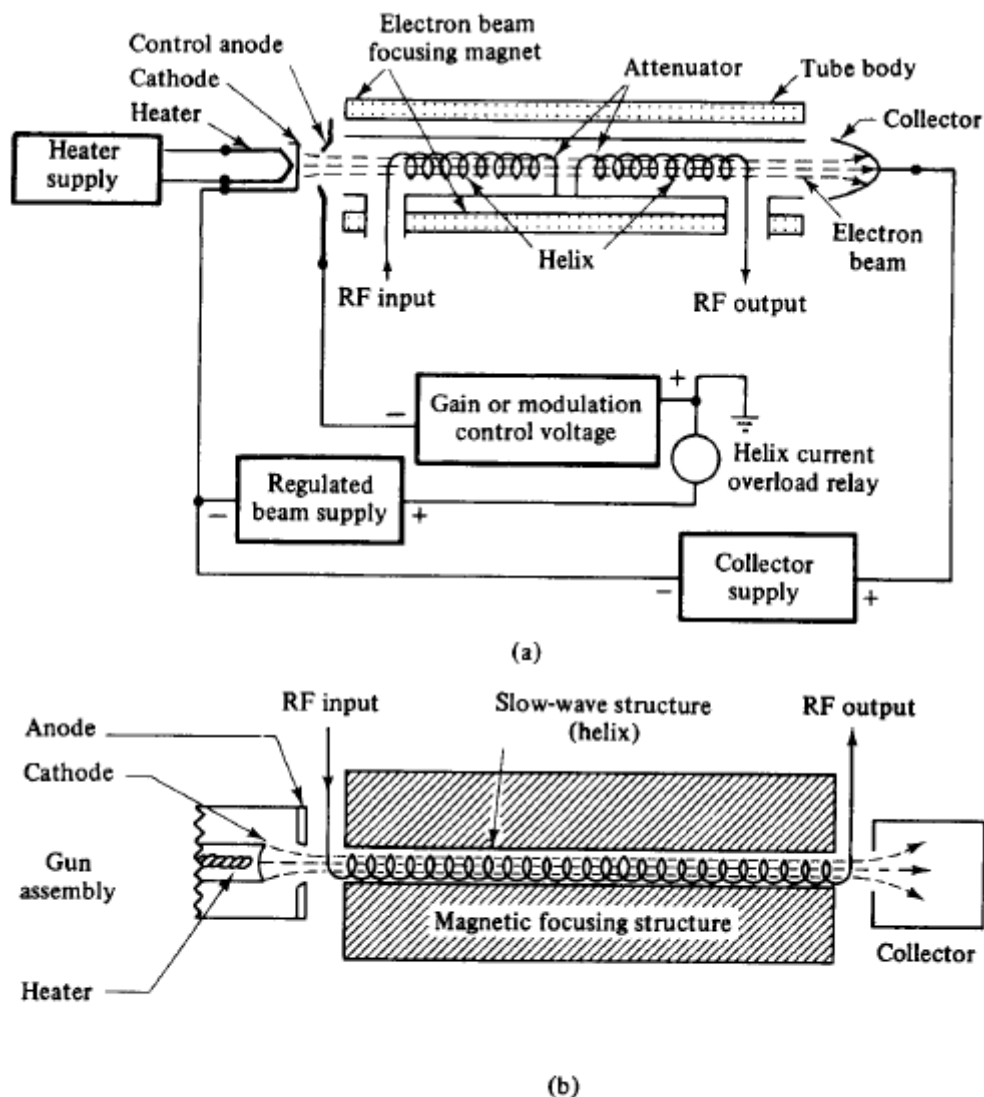
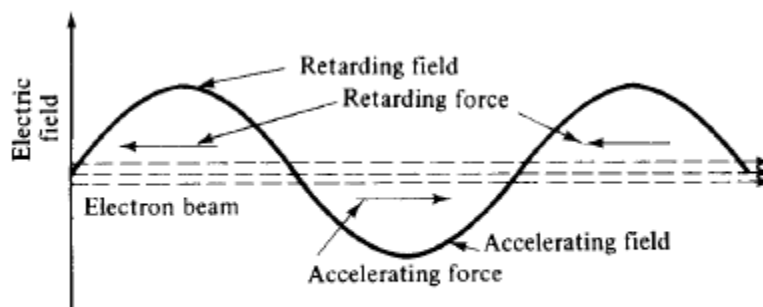
Amplification process:

Diagram of helix traveling-wave tube: (a) schematic diagram of helix traveling-wave tube; (b) simplified circuit.

The electrons entering the retarding field are decelerated and those in the accelerating field are accelerated. They begin forming a bunch centered about those electrons that enter the helix during the zero field. This process is shown in Fig. below.



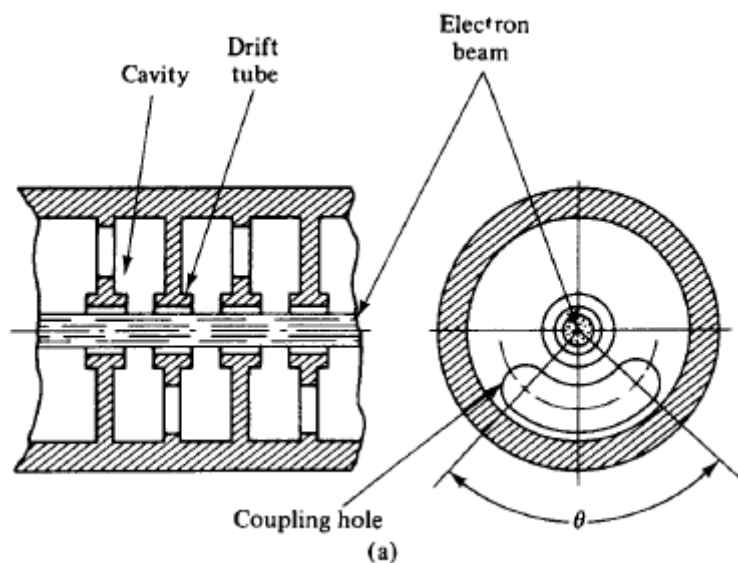
Interactions between electron beam and electric field.

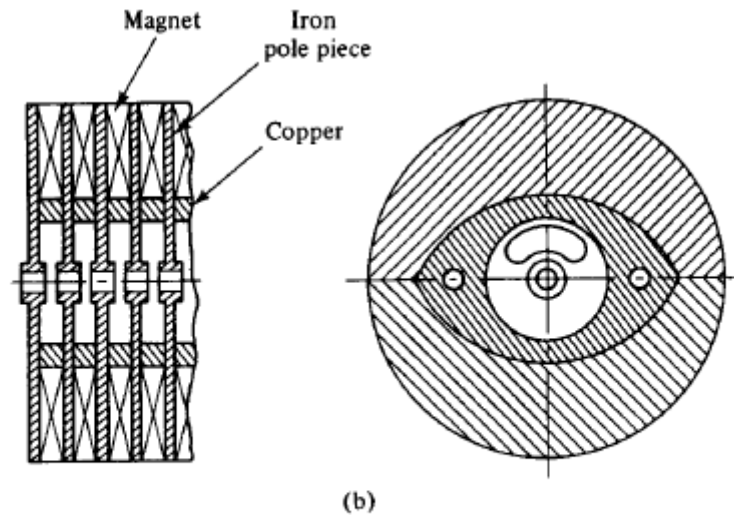
Since the dc velocity of the electrons is slightly greater than the axial wave velocity, more electrons are in the retarding field than in the accelerating field, and a great amount of energy is transferred from the beam to the electromagnetic field. The microwave signal voltage is, in turn, amplified by the amplified field. The bunch continues to become more compact, and a larger amplification of the signal voltage occurs at the end of the helix. The magnet produces an axial magnetic field to prevent spreading of the electron beam as it travels down the tube. An attenuator placed near the center of the helix reduces all the waves traveling along the helix to nearly zero so that the reflected waves from the mismatched loads can be prevented from reaching the input and causing oscillation. The bunched electrons emerging from the attenuator induce a new electric field with the same frequency. This field, in turn, induces a new amplified microwave signal on the helix.

Coupled Cavity Traveling Wave Tube:

Helix traveling-wave tubes (TWTs) produce at most up to several kilowatts of average power output. Here we describe coupled-cavity traveling-wave tubes, which are used for high-power applications.

The term *coupled cavity* means that a coupling is provided by a long slot that strongly couples the magnetic component of the field in adjacent cavities in such a manner that the passband of the circuit is mainly a function of this one variable. Figure below shows two coupled-cavity circuits that are principally used in traveling-wave tubes.





Coupled-cavity circuits in the TWTs. (a) Basic coupled-cavity circuit. (b) Coupled-cavity circuit with integral periodic-permanent-magnet (PPM) focusing.

As far as the coupling effect is concerned, there are two types of coupled-cavity circuits in traveling-wave tubes. The first type consists of the fundamentally forward-wave circuits that are normally used for pulse applications requiring at least half a megawatt of peak power. These coupled-cavity circuits exhibit negative mutual inductive coupling between the cavities and operate with the fundamental space harmonic. The cloverleaf and centipede circuits belong to this type. The second type is the first space-harmonic circuit, which has positive mutual coupling between the cavities. These circuits operate with the first spatial harmonic and are commonly used for pulse or continuous wave (CW) applications from one to several hundred kilowatts of power output. In addition, the long-slot circuit of the positive mutual coupling-cavity circuit operates at the fundamental spatial

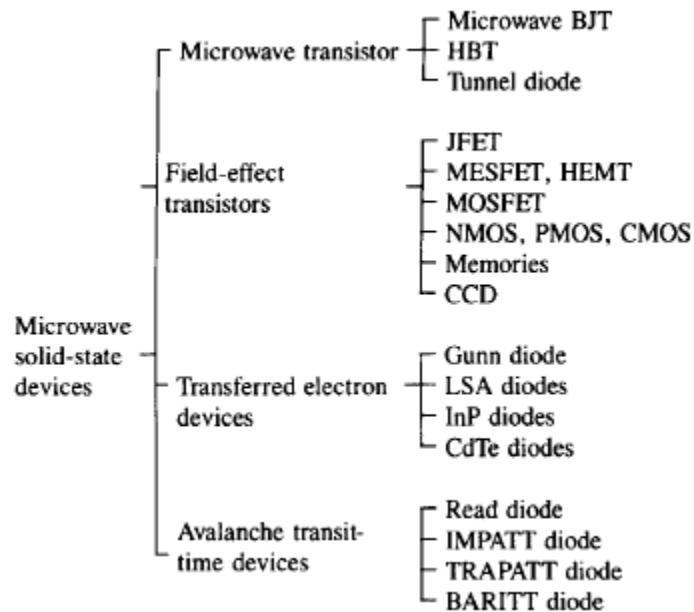
Any repetitive series of lumped LC elements constitute a propagating filter-type circuit. The coupled cavities in the traveling-wave tube are usually highly over-coupled, resulting in a bandpass-filter-type characteristic. When the slot angle (θ) as shown in Fig. is larger than 180° , the passband is close to its practical limits. The drift tube is formed by the reentrant part of the cavity, just as in the case of a klystron. During the interaction of the RF field and the electron beam in the traveling-wave tube a phase change occurs between the cavities as a function of frequency. A decreasing phase characteristic is reached if the mutual inductance of the coupling slot is positive, whereas an increasing phase characteristic is obtained if the mutual inductive coupling of the slot is negative.

The amplification of the traveling-wave tube interaction requires that the electron beam interact with a component of the circuit field that has an increasing phase characteristic with frequency. The circuit periodicity can give rise to field components that have phase characteristics as shown in Fig. In Fig. the angular frequency (ω) is plotted as a function of the phase shift ($\beta \ell$) per cavity. The ratio of ω to β is equal to the phase velocity. For a circuit having positive mutual in-

ductive coupling between the cavities, the electron beam velocity is adjusted to be approximately equal to the phase velocity of the first forward-wave spatial harmonic. For the circuits with negative mutual inductive coupling, the fundamental branch component of the circuit wave is suitable for synchronism with the electron beam and is normally used by the traveling-wave tube. The coupled-cavity equivalent circuit has been developed by Curnow as shown in Fig. .

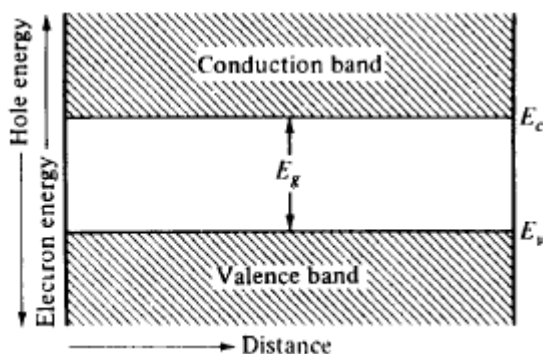
Module-II

Microwave Solid State Devices



The energy bands of a semiconductor play a major role in their electrical behavior. For any semiconductor, there is a forbidden energy region in which no allowable states can exist. The energy band above the forbidden region is called the *conduction band*, and the bottom of the conduction band is designated by E_c . The energy band below the forbidden region is called the *valence band*, and the top of the valence band is designated by E_v . The separation between the energy of the lowest conduction band and that of the highest valence band is called the *bandgap energy* E_g , which is the most important parameter in semiconductors.

Electron energy is conventionally defined as positive when measured upward whereas the hole energy is positive when measured downward. A simplified band diagram is shown in Fig. below.



Energy-band diagram.

APPLICATIONS OF MICROWAVE SOLID-STATE DEVICES

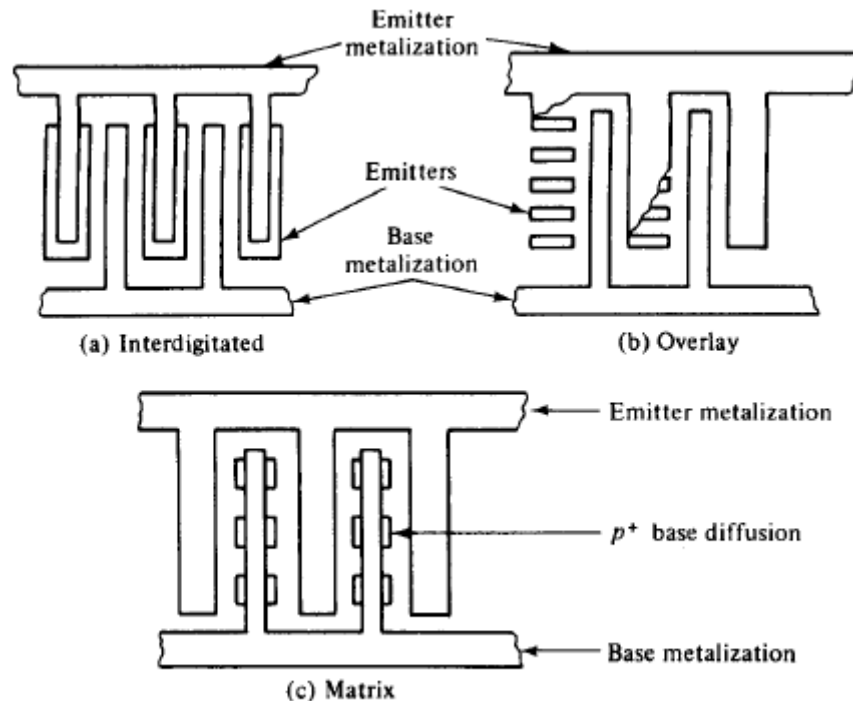
Devices	Applications	Advantages
Transistor	L-band transmitters for telemetry systems and phased array radar systems L- and S-band transmitters for communications systems	Low cost, low power supply, reliable, high CW power output, light weight
TED	C-, X-, and Ku-band ECM amplifiers for wideband systems X- and Ku-band transmitters for radar systems, such as traffic control	Low power supply (12 V), low cost, light weight, reliable, low noise, high gain
IMPATT	Transmitters for millimeter-wave communications systems	Low power supply, low cost, reliable, high CW power output, light weight
TRAPATT	S-band pulsed transmitters for phased array radar systems	High peak and average power, reliable, low power supply, low cost
BARITT	Local oscillators in communications and radar receivers	Low cost, low power supply, reliable, low noise

Microwave Bipolar Transistors:

The invention of the transistor (contraction for transfer resistor) by William Shockley and his coworkers at Bell Laboratory in 1948 had a revolutionary impact on electronic technology in general and on solid-state devices in particular. Since then transistors and related semiconductor devices have replaced vacuum tubes for lower-power sources. Microwave power transistor technology has advanced significantly during the past three decades. The microwave transistor is a nonlinear device, and its principle of operation is similar to that of the low-frequency device, but requirements for dimensions, process control, heat sinking, and packaging are much more severe.

For microwave applications, the silicon (Si) bipolar transistors dominate for frequency range from UHF to about S band (about 3 GHz). As the technology improves, the upper frequency limit for these devices is continuously being extended, and at the present time the devices are capable of producing useful power up to 22 GHz. The majority of bipolar transistors of current interest are fabricated from silicon, although GaAs devices offer prospects for improvements in operating frequency, in high temperatures, and in radiation hardness. The Si bipolar transistor is inexpensive, durable, integrative, and offers gain much higher than available with competing field-effect devices. It has moderate noise figure in RF amplifiers and $1/f$ noise characteristics that are about 10–20 dB superior to GaAs MESFETs. For these reasons, the Si bipolar transistors dominate in amplifier applications for the lower microwave frequencies and are often the devices of choice for local oscillators.

All microwave transistors are now planar in form and almost all are of the silicon $n-p-n$ type. The geometry can be characterized as follows: (a) interdigitated, (b) overlay, and (c) matrix (also called mesh or emitter grid) as shown in Fig. The interdigitated type is for a small signal and power, but the overlay type and matrix type are for small power only.



Surface geometries of microwave power transistor.

For high-frequency applications, the n - p - n structure is preferred because the electron mobility ($\mu_n = 1500 \text{ cm}^2/\text{V} \cdot \text{s}$) is much higher than the hole mobility ($\mu_p = 450 \text{ cm}^2/\text{V} \cdot \text{s}$).

Although there are many ways of fabricating a transistor, diffusion and ion implantation are generally used. For example, the structure would typically start with a lightly doped n -type epitaxial layer as the collector. The base region would be formed by counter-doping the base region p -type by diffusion. The emitter would be formed by a shallow heavily doped n -type diffusion or by ion implantation. The emitter and base contacts are generally located on the semiconductor surface in an interdigital, planar arrangement. The interdigital geometry always provides for $n + 1$ base fingers, where n is the number of emitter fingers. The number of fingers varies with the application, with more fingers required as the output power capability of the transistor increases. Additional fingers, however, increase the device parasitics and degrade the noise and upper frequency capability of the devices.

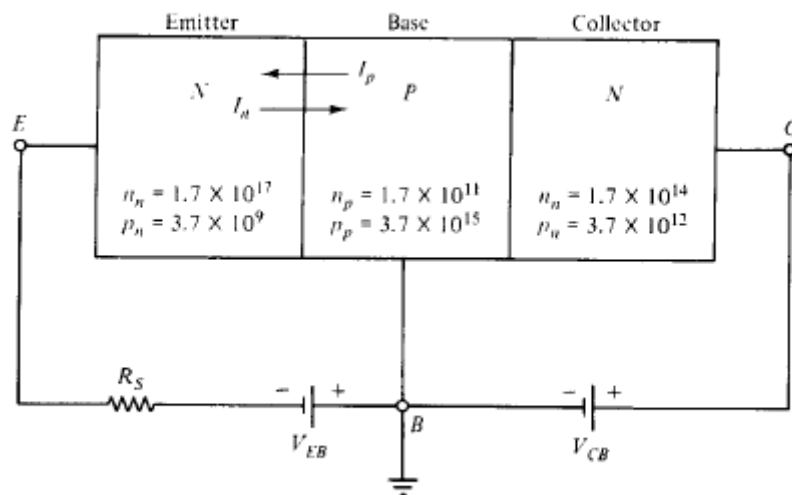
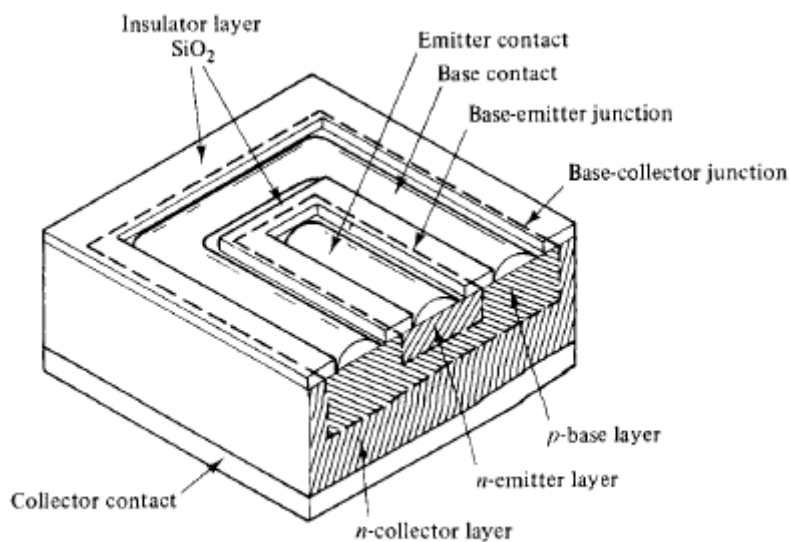
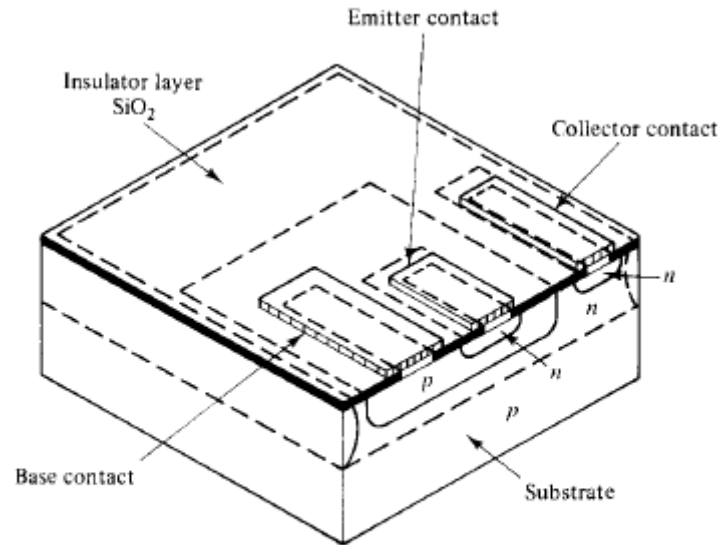
Carrier densities of an *n-p-n* transistor.

Figure below shows two schematic diagrams for a bipolar junction transistor (BJT): (a) the cross section of a discrete *n-p-n* planar BJT and (b) the cross section of a chip-type *n-p-n* integrated BJT.

The *p-n-p* bipolar junction transistor is a complementary structure of the *n-p-n* BJT by interchanging *p* for *n* and *n* for *p*. The *p-n-p* BJT is basically fabricated by first forming an *n*-type layer in the *p*-type substrate; then a *p*⁺-type region is developed in the *n* layer. Finally, metallic contacts are introduced to the *p*⁺ region and *p* layer through the windows opened in the oxide layer and to the *p* region at the bottom.

(a) Discrete *n-p-n* planar BJT

(b) Integrated chip-type n - p - n BJT

Schematic diagrams of bipolar junction transistors.

Hetero Junction Bipolar Transistor (HBTs):

Bipolar transistors can be constructed as homojunction or heterojunction types of transistors. When the transistor junction is jointed by two similar materials such as silicon to silicon or germanium to germanium, it is a *homojunction transistor*. The transistor junction formed by two different materials, such as Ge to GaAs, is called a *heterojunction transistor*. So far only the ordinary homojunction transistors have been discussed. In this section we study the heterojunction transistors.

When the lattice constants of two semiconductor materials are matched, they can be formed together as a heterojunction transistor. This lattice condition is very important because the lattice mismatch could introduce a large number of interface states and degrade the heterojunction operation. Currently, Ge and GaAs are the two materials commonly used for heterojunction structures because their lattice constants ($a = 5.646 \text{ \AA}$ for Ge and $a = 5.653 \text{ \AA}$ for GaAs) are matched to within 1%. Since each material may be either p type or n type, there are four possible heterojunction combinations:

1. p -Ge to p -GaAs junction
2. p -Ge to n -GaAs junction
3. n -Ge to p -GaAs junction
4. n -Ge to n -GaAs junction

Fig. below shows the model diagram of a heterojunction transistor formed by n -Ge, p -GaAs, and n -GaAs materials.

When an n -Ge and a p -GaAs are isolated, their Fermi energy levels are not aligned, as shown in Fig. below

In Fig. below, the vacuum level is used as reference, the work function is denoted by ϕ , n -Ge is designated as 1, and p -GaAs is referred to as 2. The different energies of the conduction-band edge and the valence-band edge are given by

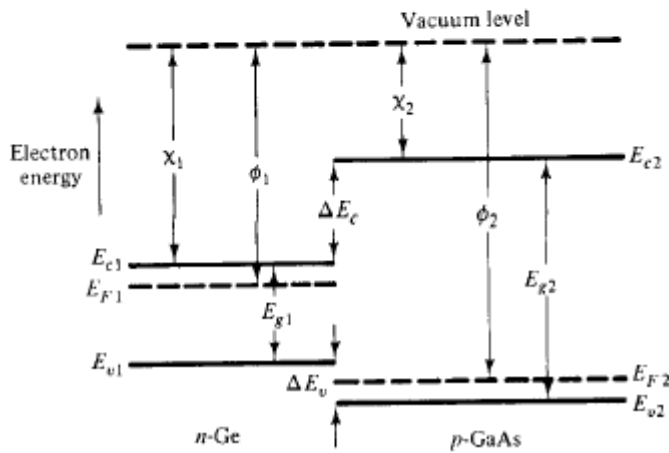
$$\Delta E_c = \chi_1 - \chi_2$$

and

$$\Delta E_v = E_{g2} - E_{g1} - \Delta E_c$$

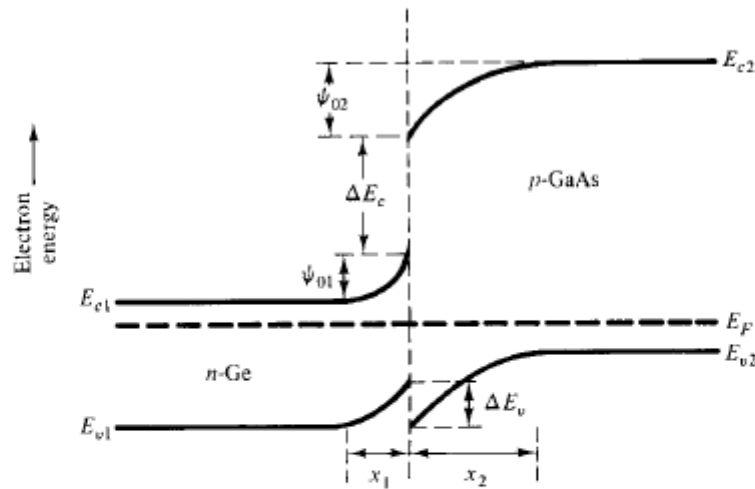
where χ = electron affinity in eV

E_g = bandgap energy in eV



Energy-band diagram for isolated n -Ge and p -GaAs.

When the two semiconductor materials are joined together, their Fermi energy levels are aligned and their energy bands are depleted at the junction, as shown in Fig.



Energy-band diagram for an n -Ge- p -GaAs junction.

When the two materials are joined together, the electrons in the n -Ge are injected into the p -GaAs, and the holes in the p -GaAs are transferred to the n -Ge until their Fermi energy levels are aligned. As a result, the energy bands at the junction are depleted or bent. The bending energy then creates a built-in voltage in both sides of the junction. The total built-in voltage is expressed by

$$\psi_0 = \psi_{01} + \psi_{02}$$

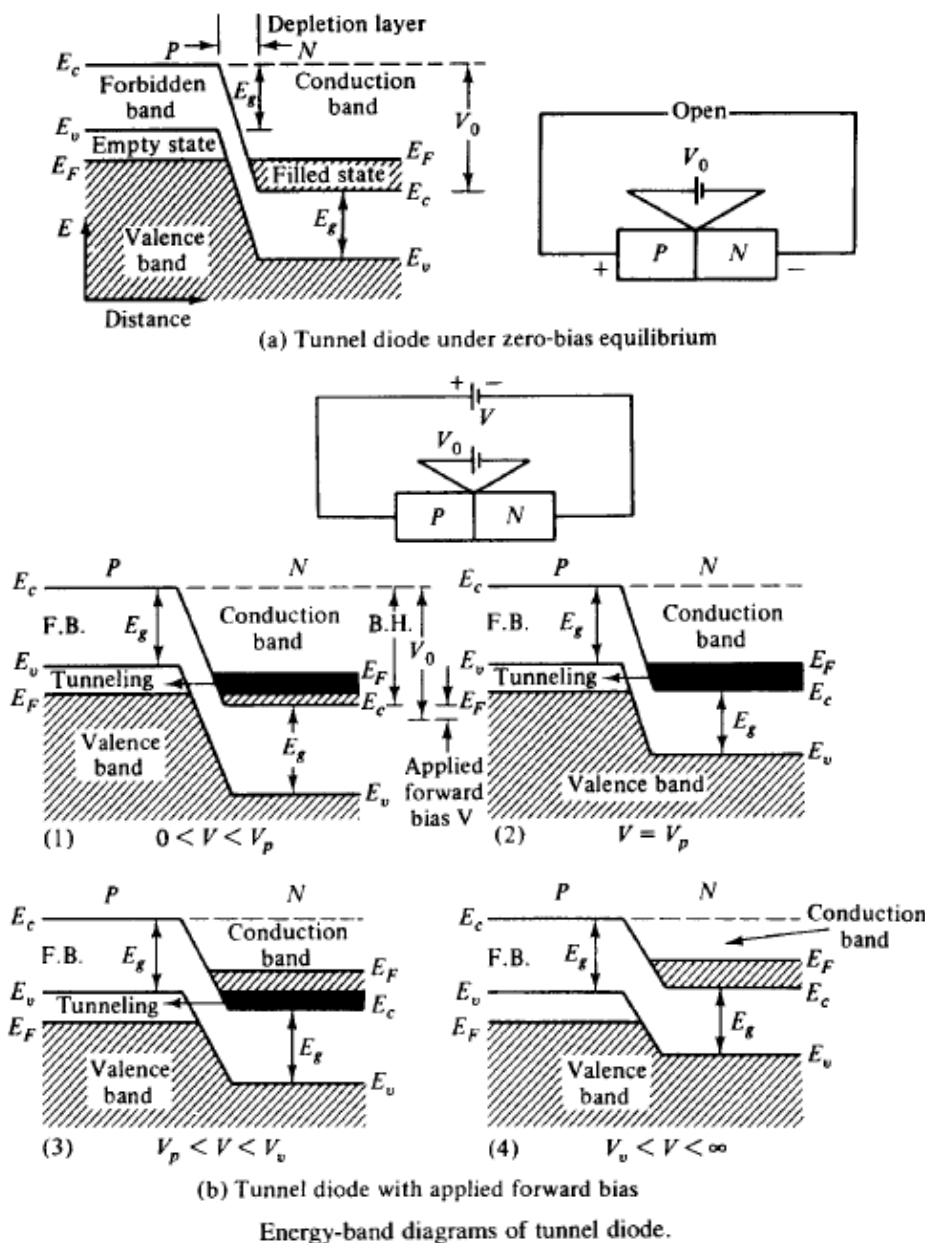
where ψ_{01} = barrier potential or portion of built-in voltage in n -Ge

ψ_{02} = barrier potential or portion of built-in voltage in p -GaAs

Microwave Tunnel Diodes:

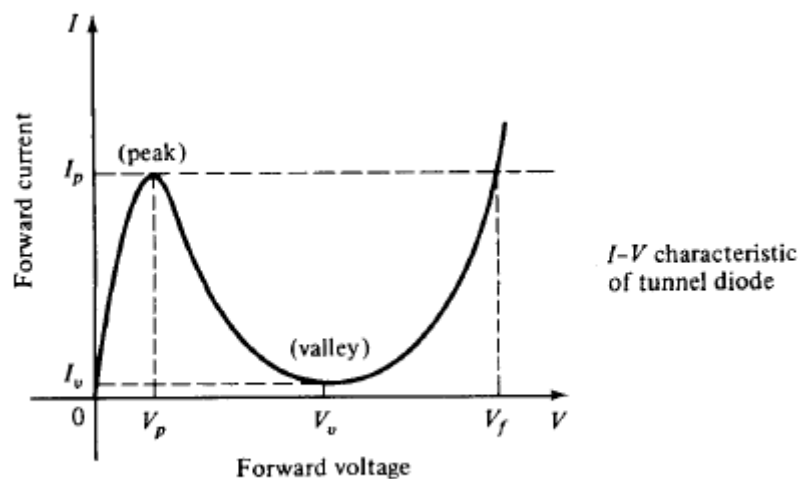
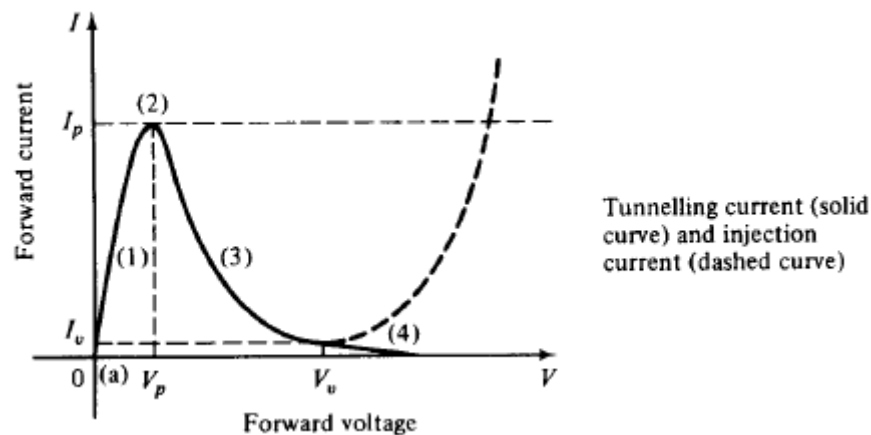
After the publication of Esaki's classic paper on tunnel diodes in 1958, the potential of tunnel diodes for microwave applications was quickly established. Prior to 1958 the anomalous characteristics of some p - n junctions were observed by many scientists, but the irregularities were rejected immediately because they did not follow the "classic" diode equation. Esaki, however, described this anomalous phenomenon by applying a quantum tunneling theory. The tunneling phenomenon is a majority carrier effect. The tunneling time of carriers through the potential energy barrier is not governed by the classic transit-time concept—that the transit time is equal to the barrier width divided by the carrier velocity—but rather by the quantum transition probability per unit time. Tunnel diodes are useful in many circuit applications in microwave amplification, microwave oscillation, and binary memory because of their low cost, light weight, high speed, low-power operation, low noise, and high peak-current to valley-current ratio.

The tunnel diode is a negative-resistance semiconductor p - n junction diode. The negative resistance is created by the tunnel effect of electrons in the p - n junction. The doping of both the p and n regions of the tunnel diode is very high—impurity concentrations of 10^{19} to 10^{20} atoms/cm³ are used—and the depletion-layer barrier at the junction is very thin, on the order of $100 \text{ \AA}$ or 10^{-6} cm . Classically, it is possible for those particles to pass over the barrier if and only if they have an energy equal to or greater than the height of the potential barrier. Quantum mechanically, however, if the barrier is less than $3 \text{ \AA}$ there is an appreciable probability that particles will tunnel through the potential barrier even though they do not have enough kinetic energy to pass over the same barrier. In addition to the barrier thinness, there must also be filled energy states on the side from which particles will tunnel and allowed empty states on the other side into which particles penetrate through at the same energy level. In order to understand the tunnel effects fully, let us analyze the energy-band pictures of a heavily doped p - n diode. Figure below shows energy-band diagrams of a tunnel diode.



E_F is the Fermi level representing the energy state with 50% probability of being filled if no forbidden band exists
 V_0 is the potential barrier of the junction
 E_g is the energy required to break a covalent bond, which is 0.72 eV for germanium and 1.10 eV for silicon
 E_c is the lowest energy in the conduction band
 E_v is the maximum energy in the valence band
 V is the applied forward bias
 F.B. stands for the forbidden band
 B.H. represents the barrier height

Under open-circuit conditions or at zero-bias equilibrium, the upper levels of electron energy of both the p type and n type are lined up at the same Fermi level as shown in Fig. (a). Since 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, there is no flow of charge in either direction across the junction and the current is zero, as shown at point (a) of the volt-ampere characteristic curve of a tunnel diode in Fig. below.



Ampere-voltage characteristics of tunnel diode.

In ordinary diodes the Fermi level exists in the forbidden band. Since the tunnel diode is heavily doped, the Fermi level exists in the valence band in p -type and in the conduction band in n -type semiconductors. When the tunnel diode is forward-biased by a voltage between zero and the value that would produce peak tunneling current I_p ($0 < V < V_p$), the energy diagram is shown in part (1) of Fig. (b). Accordingly, the potential barrier is decreased by the magnitude of the applied forward-bias voltage. A difference in Fermi levels in both sides is created. Since there are filled states in the conduction band of the n type at the same energy level as allowed empty states in the valence band of the p type, the electrons tunnel through the barrier from the n type to the p type, giving rise to a forward tunneling current

from the p type to the n type as shown in sector (1) of Fig. above. As the forward bias is increased to V_p , the picture of the energy band is as shown in part (2) of Fig.

A maximum number of electrons can tunnel through the barrier from the filled states in the n type to the empty states in the p type, giving rise to the peak current I_p in Fig. . If the bias voltage is further increased, the condition shown in part (3) of Fig. (b) is reached. The tunneling current decreases as shown in sector (3) of Fig. above. Finally, at a very large bias voltage, the band structure of part (4) of Fig. (b) is obtained. Since there are now no allowed empty states in the p type at the same energy level as filled states in the n type, no electrons can tunnel through the barrier and the tunneling current drops to zero as shown at point (4) of Fig.

When the forward-bias voltage V is increased above the valley voltage V_v , the ordinary injection current I at the p - n junction starts to flow. This injection current is increased exponentially with the forward voltage as indicated by the dashed curve of Fig. above. The total current, given by the sum of the tunneling current and the injection current, results in the volt-ampere characteristic of the tunnel diode as shown in Fig. above. It can be seen from the figure that the total current reaches a minimum value I_v (or valley current) somewhere in the region where the tunnel-diode characteristic meets the ordinary p - n diode characteristic. The ratio of peak current to valley current (I_p/I_v) can theoretically reach 50 to 100. In practice, however, this ratio is about 15.

Microwave Field Effect Transistors:

After Shockley and his coworkers invented the transistor in 1948, he proposed in 1952 a new type of field-effect transistor (FET) in which the conductivity of a layer of a semiconductor is modulated by a transverse electric field [1]. In a conventional transistor both the majority and the minority carriers are involved; hence this type of transistor is customarily referred to as a *bipolar* transistor. In a field-effect transistor the current flow is carried by majority carriers only; this type is referred to as a *unipolar* transistor. In addition, the field-effect transistors are controlled by a voltage at the third terminal rather than by a current as with bipolar transistors. Our purpose here is to describe the physical structures, principles of operation, microwave characteristics, and power-frequency limitations of unipolar field-effect transistors. Since the microwave field-effect transistor has the capability of amplifying small signals up to the frequency range of X band with low-noise figures, it has lately replaced the parametric amplifier in airborne radar systems because the latter is complicated in fabrication and expensive in production.

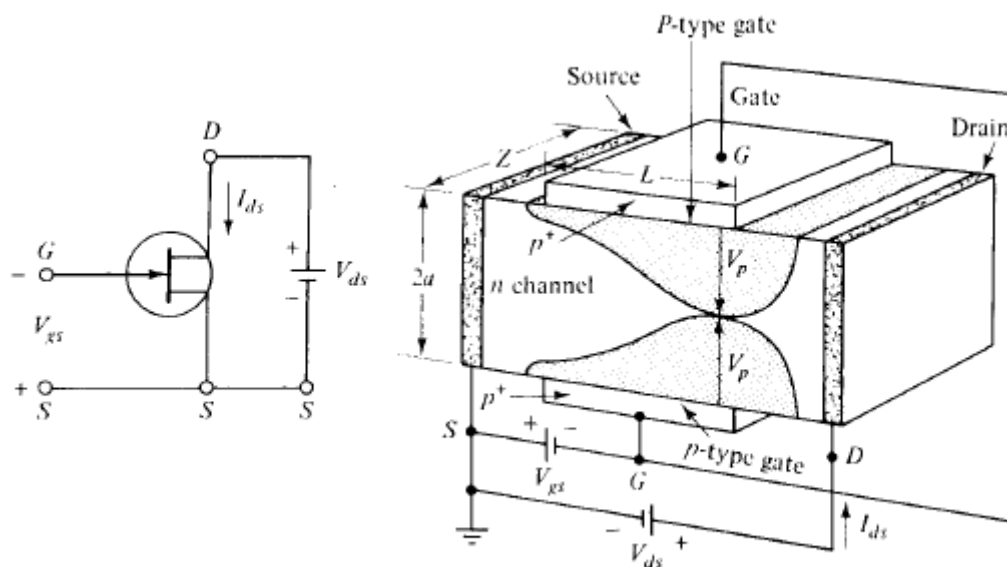
The unipolar field-effect transistor has several advantages over the bipolar junction transistor:

1. It may have voltage gain in addition to current gain.
2. Its efficiency is higher than that of a bipolar transistor.
3. Its noise figure is low.
4. Its operating frequency is up to X band.
5. Its input resistance is very high, up to several megohms.

Junction Field Effect Transistor:

Unipolar field-effect transistors may be in the form of either a p - n junction gate or a Schottky-barrier gate. The former is called a *junction field-effect transistor* (JFET), and the latter is referred to as a MESFET or *metal-semiconductor field-effect transistor*. The JFET was originally proposed by Shockley. Figure below shows the schematic diagram and circuit symbol for an n -channel JFET.

The n -type material is sandwiched between two highly doped layers of p -type material that is designated P^+ . This type of device is called an n -channel JFET. If the middle part is a p -type semiconductor, the device is called a p -channel JFET. The two p -type regions in the n -channel JFET shown in Fig. are referred to as the *gates*. Each end of the n channel is joined by a metallic contact. In accordance with the directions of the biasing voltages shown in Fig. , the left-hand contact which supplies the source of the flowing electrons is referred to as the *source*, whereas the right-hand contact which drains the electrons out of the material is called the *drain*. The circuit symbol for an n -channel JFET is also shown in Fig. . The direction of the drain current I_d is flowing from the drain to the device.



Schematic diagram and circuit symbol for an n -channel JFET.

For a p -channel JFET, the polarities of the two biasing voltages V_g and V_d are interchanged, the head of the arrow points away from the device, and the drain current I_d flows away from the device. Since electrons have higher mobility than holes, the n -channel JFET provides higher conductivity and higher speed and is preferred in most applications to the p -channel JFET.

Under normal operating conditions, when the gate voltage V_g is zero, the drain current I_d is also zero. The channel between the gate junctions is entirely open. When a small drain voltage V_d is applied between the drain and source, the n -type semiconductor bar acts as a simple resistor, and the current I_d increases linearly with V_d . If a reverse gate voltage V_g is applied across the p - n gate junctions, the majority of free electrons are depleted from the channel, and the space-charge regions are extended into the channel. As the drain voltage V_d is further increased, the space-charge regions expand and join together, so that all free electron carriers are completely depleted in the joined region. This condition is called *pinch off*. When the channel is pinched off, the drain current I_d remains almost constant, but the drain voltage V_d is continuously increased.

The JFET has a conducting channel between the source and the drain electrodes when the gate bias voltage is zero. This is the ON state, and the transistor is called a *normally ON JFET*. In order to reach the OFF state, a gate voltage must be applied to deplete all carriers in the channel. As a result, this device is referred to as the *depletion-mode JFET* or *D-JFET*.

Metal Semiconductor Field Effect Transistors (MESFETs):

In 1938, Schottky suggested that the potential barrier could arise from stable space charges in the semiconductor alone without the presence of a chemical layer. The model derived from his theory is known as the *Schottky barrier*.

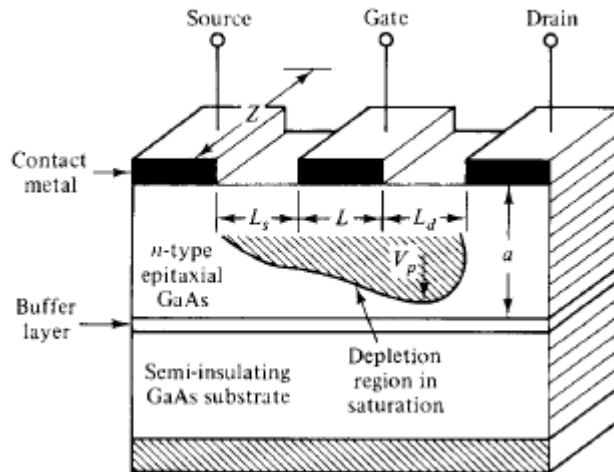
If the field-effect transistor is constructed with a metal-semiconductor Schottky-barrier diode, the device is called a *metal-semiconductor field-effect transistor* (MESFET). The material may be either silicon or gallium arsenide (GaAs), and the channel type may be either n channel or p channel.

Since GaAs MESFETs have the capability of amplifying small signals up to the frequency range of X band with low-noise figure, they have lately replaced the parametric amplifiers in airborne radar systems because the latter are complicated to fabricate and expensive to produce.

The GaAs MESFET has higher electron mobility, higher electric field, and higher electron saturation drift velocity than silicon devices, so its output power is also greater. Another special feature is its lower noise figure, accounted for by its higher electron mobility. Therefore the GaAs MESFETs are very commonly used in microwave integrated circuits for high-power, low-noise, and broadband amplifier applications.

The unipolar transistor such as a GaAs MESFET can be developed by using either the epitaxial process or the ion implantation method.

Figure below shows schematically a simple MESFET in GaAs.

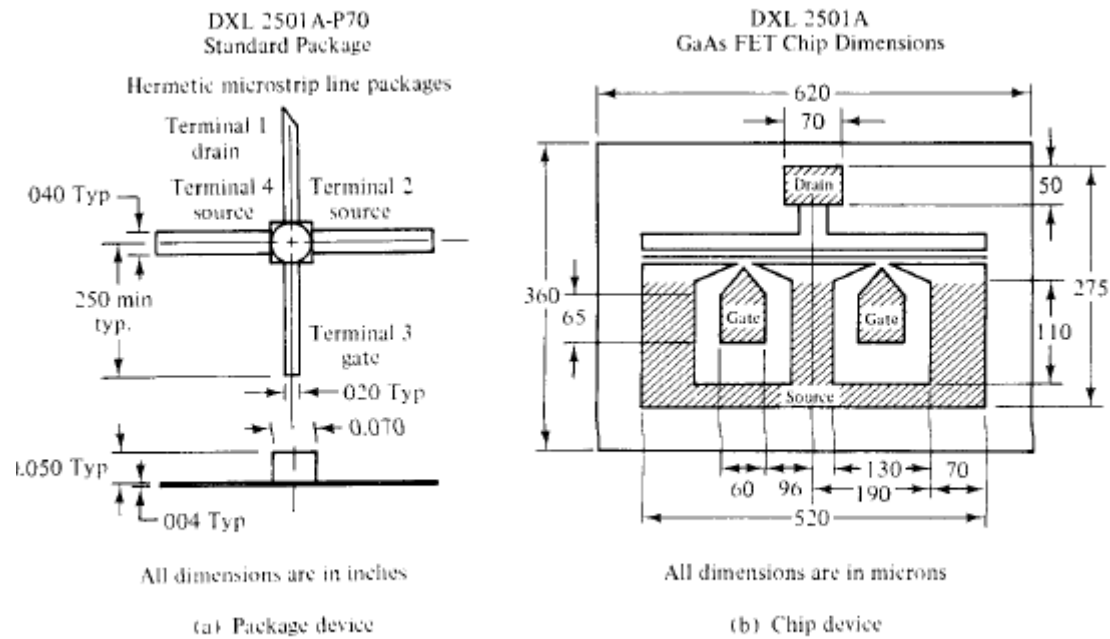


Schematic diagram of a GaAs MESFET.

In GaAs MESFETs the substrate is doped with chromium (Cr), which has an energy level near the center of the GaAs bandgap. As Cr is the dominant impurity, the Fermi level is pinned near the center of the bandgap. Thus, a very high-resistivity substrate (near 10^8 ohm-cm) generally results, and it is commonly called the *semi-insulator GaAs substrate*. On this nonconducting substrate a thin layer of lightly doped *n*-type GaAs is grown epitaxially to form the channel region of the field-effect transistor. In many cases a high resistivity GaAs epitaxial layer, called the *buffer layer*, is grown between the *n*-type GaAs layer and the substrate. The photolithographic process may be used to define the patterns in the metal layers such as Au-Ge for source and drain ohmic contacts and in the Al layer for the Schottky barrier-gate contact. The reason for using GaAs instead of Si is that GaAs has higher electron mobility and can operate at higher temperature and higher power.

The GaAs MESFET can also be grown by using ion implantation. A thin *n*-type layer can be formed at the surface of the substrate by implanting Si or a donor impurity Se from column VI of the element periodic table. However, the ion implantation process requires an anneal to remove the radiation damage. In either the fully implanted device or the epitaxial device, the source and drain contacts may be improved by further n^+ implantation in these regions.

After the fabrication processes are completed, the individual transistor is separated from the wafer, and this discrete transistor is called a *chip device*. The chip device is then alloyed to a header to provide a contact to the collector region, and Au or Al wires are bonded to the metallized regions to serve as leads to the emitter and base. This bonded chip device is then named a *package transistor*. Figure below shows both the package and chip GaAs MESFET devices.

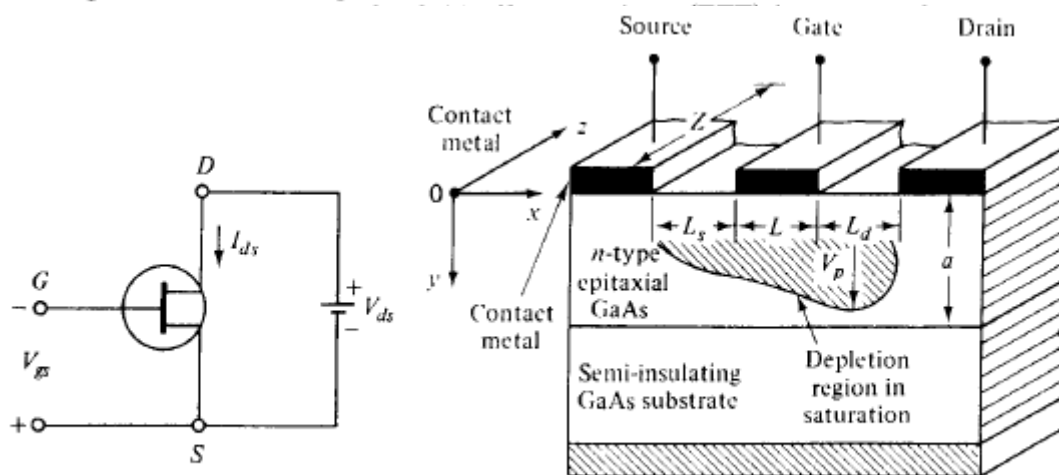


Dimensions of package and chip GaAs MESFETs.

The MESFET device is of an interdigitated structure, fabricated by using an n -type GaAs epitaxial film about $0.15\text{--}0.35\text{-}\mu\text{m}$ thick on a semi-insulating substrate about $100\text{ }\mu\text{m}$. The n -channel layer is doped with either sulphur or tin in a doping concentration N between $8 \times 10^{16}/\text{cm}^3$ and $2 \times 10^{17}/\text{cm}^3$. The electron mobility in the layer is in the range of 3000 to $4500\text{ cm}^2/\text{V}\cdot\text{s}$. The Schottky barrier-gate is evaporated aluminum. The source and drain contacts are Au-Ge, Au-Te, or Au-Te-Ge alloys. A contact metallization pattern of gold is used to bring the source, drain, and gate contacts out to bonding pads over the semi-insulating substrate. A buffer layer of about $3\text{ }\mu\text{m}$ with a doping concentration of 10^{15} to 10^{16} cm^{-3} is often fabricated between the active n -type epitaxial layer and the semi-insulating substrate.

Below figure for JFET, a voltage is applied in the direction to reverse-bias the p - n junction between the source and the gate, while the source and the drain electrodes are forward-biased. Under this bias condition, the majority carriers (electrons) flow in the n -type epitaxial layer from the source electrode, through the channel beneath the gate, to the drain electrode. The current in the channel

causes a voltage drop along its length so that the Schottky barrier-gate electrode becomes progressively more reverse-biased toward the drain electrode. As a result, a charge-depletion region is set up in the channel and gradually pinches off the channel against the semi-insulating substrate toward the drain end. As the reverse bias between the source and the gate increases, so does the height of the charge-depletion region. The decrease of the channel height in the nonpinched-off region will increase the channel resistance. Consequently, the drain current I_d will be modulated by the gate voltage V_g . This phenomenon is analogous to the characteristics of the collector current I_c versus the collector voltage V_c with the base current I_b as a parameter in a bipolar transistor. In other words, a family of curves of the drain current I_d versus the voltage V_{ds} between the source and drain with the gate voltage V_g as a parameter will be generated in an unipolar GaAs MESFET,



Schematic diagram and circuit symbol of a GaAs MESFET.

For a fixed drain-to-source voltage V_{ds} , the drain current I_d is a function of the reverse-biasing gate voltage V_g . Because the drain current I_d is controlled by the field effect of the gate voltage V_g , this device is referred to as the field-effect transistor. When the drain current I_d is continuously increasing, the ohmic voltage drop between the source and the channel reverse-biases the p - n junction further. As a result, the channel is eventually pinched off. When the channel is pinched off, the drain current I_d will remain almost constant even though the drain-to-source voltage V_d is continuously increased.

High Electron Mobility Transistors (HEMTs):

The evolution of high-speed GaAs integrated circuits (ICs) is the result of continuous technological progress utilizing the superior electronic properties of gallium arsenide compared with those of silicon. Electron mobility in the MESFET channel with typical donor concentrations of about 10^{17} cm^{-3} ranges from 4000 to 5000 $\text{cm}^2/\text{V}\cdot\text{s}$ at room temperature. The mobility in the channel at 77° K is not too much higher than at room temperature because of ionized impurity scattering. In undoped GaAs, however, electron mobility of 2 to $3 \times 10^5 \text{ cm}^2/\text{V}\cdot\text{s}$ has been achieved at 77° K. The mobility of GaAs with feasibly high electron concentrations for facilitating the fabrication of devices was found to increase through modulation-doping technique: demonstrated in GaAs–AlGaAs superlattices. A high electron-mobility transistor (HEMT), based on a modulation-doped GaAs–AlGaAs single heterojunction structure, was developed. HEMTs have exhibited lower noise figure and higher gain at microwave frequencies up to 70 GHz, and it is possible to construct HEMT amplifiers at even higher frequencies of operation. The major improvements over MESFETs include shorter gate lengths, reduced gate and source contact resistances, and optimized doping profiles.

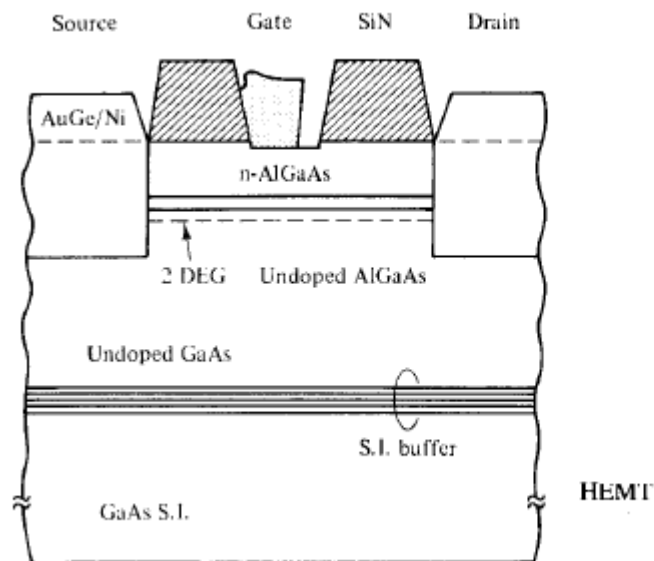
The basic structure of a HEMT is a selectively doped GaAs–AlGaAs heterojunction structure as shown in Fig. below.

An undoped GaAs layer and an Si-doped n -type AlGaAs layer are successively grown on a semi-insulating GaAs substrate. A two-dimensional electron gas (2-DEG) is created between the undoped and n -type layers. A buffer layer is sandwiched between the undoped GaAs layer and the semi-insulator substrate.

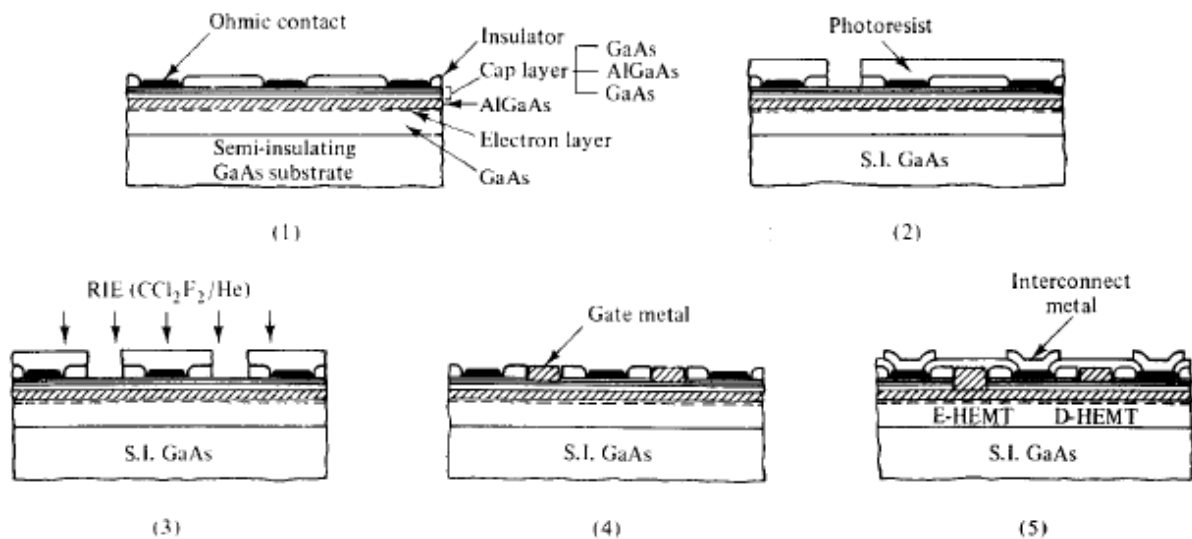
The HEMT can be fabricated by using the integrated-circuit techniques. Fig. 6-3-2 indicates the sequence for the self-aligned gate procedures in the fabrication of large-scale integration HEMTs, including the E-(enhancement-mode) and D-(depletion-mode) HEMTs.

The processing steps include the following:

1. Ohmic contact formation: The active region is isolated by a shallow mesa step (180 nm), which is almost achieved in a single process, and can be made nearly planar. The source and drain for E- and D-HEMTs are metallized with an AuGe eutectic alloy and an Au overlay alloy to produce ohmic contacts with the electron layer.
2. Opening gate windows: The fine gate patterns are formed for E-HEMTs as the top GaAs layer and thin $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ stopper are etched off by nonselective chemical etching.
3. Selective dry etching: With the same photoresist process for formation of gate patterns in D-HEMTs, selective dry etching is performed to remove the top GaAs layer for D-HEMTs and also to remove the GaAs layer under the thin $\text{Al}_{0.3}\text{Ga}_{0.7}\text{As}$ stopper for E-HEMTs.
4. Gate metallization: Schottky contacts for the E- and D-HEMT gate are provided by depositing Al. The Schottky gate contacts and the GaAs top layer for ohmic contacts are then self-aligned to achieve high-speed performance.
5. Interconnect metallization: Finally, electrical connections from the interconnecting metal, composed of Ti, Pt, and Au, to the device terminals are provided through contact holes etched in a crossover insulator film.

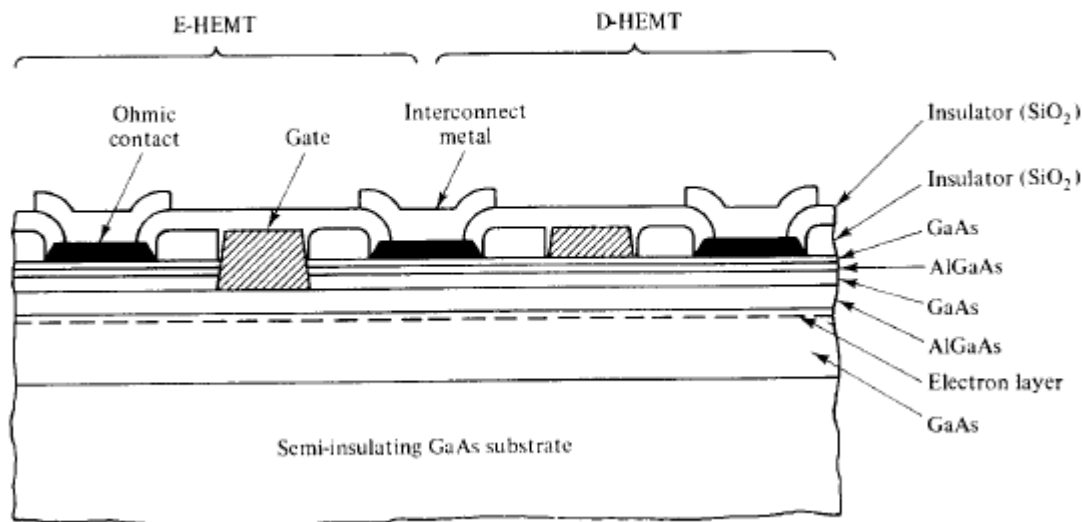


A cross section of a



Processing steps for HEMT direct-coupled FET logic (DCFL) circuits.

Figure below shows the cross-sectional view of a typical self-aligned structure of E- and D-HEMTs forming an inverter for the direct-coupled FET logic (DCFL) circuit configuration .



Cross-sectional view of a DCFL HEMT.

Since GaAs has higher electron affinity, free electrons in the AlGaAs layer are transferred to the undoped GaAs layer where they form a two-dimensional high-mobility electron gas within 100 Å of the interface. The *n*-type AlGaAs layer of the HEMT is depleted completely through two depletion mechanisms

1. Trapping of free electrons by surface states causes the surface depletion.
2. Transfer of electrons into the undoped GaAs layer brings about the interface depletion.

The Fermi energy level of the gate metal is matched to the pinning point, which is 1.2 eV below the conduction band. With the reduced AlGaAs layer thickness, the electrons supplied by donors in the AlGaAs layer are insufficient to pin the surface Fermi level, and the space-charge region is extended into the undoped GaAs layer. As a result, band bending is moving upward and the two-dimensional electron gas (2-DEG) does not appear. When a positive voltage higher than the threshold voltage is applied to the gate, electrons accumulate at the interface and form a two-dimensional electron gas.

The electron concentration can control D-(depletion-mode) and E-(enhancement-mode) HEMT operations. As temperature decreases, electron mobility, which is about 8000 cm²/V·s at 300° K, increases dramatically to 2×10^5 cm²/V·s at 77° because of reduced phonon scattering. When the temperatures decrease further, the electron mobility of 1.5×10^6 cm²/V·s at 50°K and 2.5×10^6 cm²/V·s at 4.5°K have been demonstrated.

Transferred Electron Devices (TEDs)

Difference between microwave transistors and TEDs

- Transistors operate with either junctions or gates but TEDs are bulk devices having no gates and junctions.
- Transistors are made up of elemental semiconductors i.e silicon/germanium but TEDs are fabricated from compound semiconductors such as Gallium Arsenide (GaAs), Indium Phosphide (InP) and Cadmium Telluride (CdTe).
- Transistors operate with warm electrons where as TEDs operate with hot electrons

Warm Electron: electrons whose energy is not much greater than that of thermal energy of electrons in semiconductor.

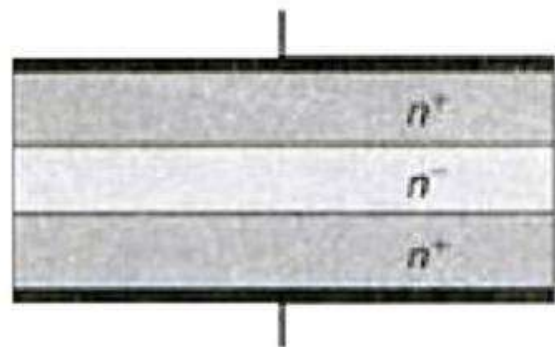
Hot Electron: electron whose energy is very much greater than that of thermal energy of electrons in semiconductor.

Gunn Diode:

- Two terminal negative resistance microwave device having more advantages over microwave transistors operating at higher frequencies.
- Generate microwave signals from around 1 GHz up to frequencies of possibly 100 GHz. May also be used to work as an amplifier.
- It is also known as transferred electron device, TED. Although is referred to as a diode, the devices does not possess a junction. Instead the device uses an effect known as the Gunn effect.
- Such type of semiconductor device which have only n-type doped semiconductor material.



Gunn Diode Symbol



Gunn diode physical structure

Gunn diodes are fabricated from a single piece of n-type semiconductor. The most common materials are gallium Arsenide (GaAs) and Indium Phosphide (InP). However other materials including Ge, CdTe, InAs, InSb, ZnSe and others have been used. The device is simply an n-type bar with n⁺ contacts. It is necessary to use n-type material because the transferred electron effect is only applicable to electrons and not holes found in a p-type material.

Within the device there are three main areas, which can be roughly termed the top, middle and bottom areas. The most common method of manufacturing a Gunn diode is to grow and epitaxial layer on a degenerate n⁺ substrate. The active region is between a few microns and a few hundred micron thick. This active layer has a doping level between 10^{14}cm^{-3} and 10^{16}cm^{-3} - this is considerably less than that used for the top and bottom areas of the device. The thickness will vary according to the frequency required.

The top n⁺ layer can be deposited epitaxially or doped using ion implantation. Both top and bottom areas of the device are heavily doped to give n⁺ material. This provides the required high conductivity areas that are needed for the connections to the device.

Devices are normally mounted on a conducting base to which a wire connection is made. The base also acts as a heat sink which is critical for the removal of heat. The connection to the other terminal of the diode is made via a gold connection deposited onto the top surface. Gold is required because of its relative stability and high conductivity.

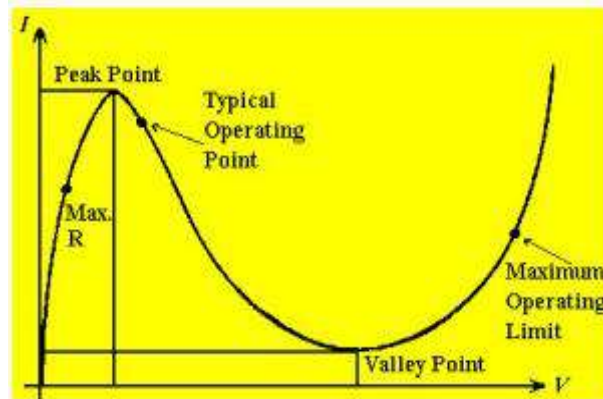
During manufacture there are a number of mandatory requirements for the devices to be successful - the material must be defect free and it must also have a very uniform level of doping.

GaAs and some other semiconductor materials have one extra-energy band in their electronic band structure instead of having only two energy bands, viz. valence band and conduction band like normal semiconductor materials. These GaAs and some other semiconductor materials consist of three energy bands, and this extra third band is empty at initial stage.

Because of this, an increase in the forward voltage increases the field strength (for field strengths where applied voltage is greater than the threshold voltage value), then the number of electrons reaching the state at which the effective mass increases by decreasing their velocity, and thus, the current will decrease.

Thus, if the field strength is increased, then the drift velocity will decrease; this creates a negative incremental resistance region in V-I relationship. Thus, increase in the voltage will

increase the resistance by creating a slice at the cathode and reaches the anode. But, to maintain a constant voltage, a new slice is created at the cathode. Similarly, if the voltage decreases, then the resistance will decrease by extinguishing any existing slice.



As shown in the above graph, initially the current starts increasing in this diode, but after reaching a certain voltage level (at a specified voltage value called as threshold voltage value), the current decreases before increasing again. The region where the current falls is termed as a negative resistance region, and due to this it oscillates. In this negative resistance region, this diode acts as both oscillator and amplifier, as in this region, the diode is enabled to amplify signals.

From Gunn's observation the carrier drift velocity is linearly increased from zero to a maximum when the electric field is varied from zero to a threshold value. When the electric field is beyond the threshold value of 3000 V/cm for the *n*-type GaAs, the drift velocity is decreased and the diode exhibits negative resistance. This situation is shown in Fig. 7-1-2.

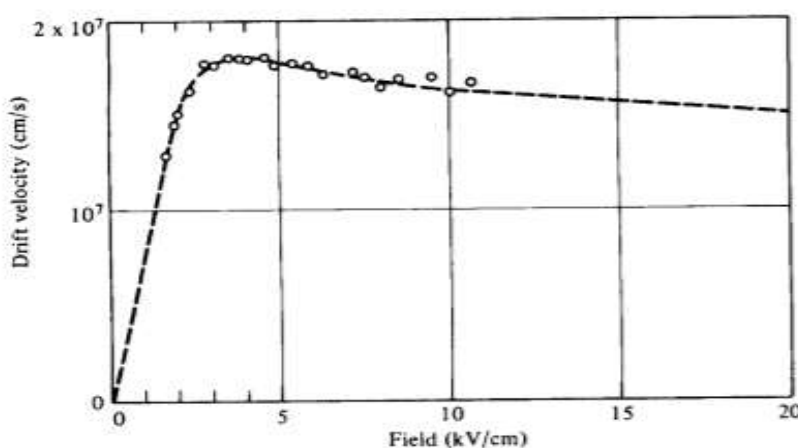


Figure 7-1-2 Drift velocity of electrons in *n*-type GaAs versus electric field. (After J. B. Gunn [8]; reprinted by permission of IBM, Inc.)

MODES OF OPERATION

Since Gunn first announced his observation of microwave oscillation in the n -type GaAs and n -type InP diodes in 1963, various modes of operation have been developed, depending on the material parameters and operating conditions. As noted, the formation of a strong space-charge instability depends on the conditions that enough charge is available in the crystal and that the specimen is long enough so that the necessary amount of space charge can be built up within the transit time of the electrons. This requirement sets up a criterion for the various modes of operation of bulk negative-differential-resistance devices. Copeland proposed four basic modes of operation of uniformly doped bulk diodes with low-resistance contacts [17] as shown in Fig. 7-3-1.

1. Gunn oscillation mode: This mode is defined in the region where the product of frequency multiplied by length is about 10^7 cm/s and the product of doping multiplied by length is greater than 10^{12} /cm². In this region the device is unstable because of the cyclic formation of either the accumulation layer or the high-field domain. In a circuit with relatively low impedance the device operates in the high-field domain mode and the frequency of oscillation is near the intrinsic frequency. When the device is operated in a relatively high- Q cavity and coupled properly to the load, the domain is quenched or delayed (or both) before nucleating. In this case, the oscillation frequency is almost entirely determined by the resonant frequency of the cavity and has a value of several times the intrinsic frequency.

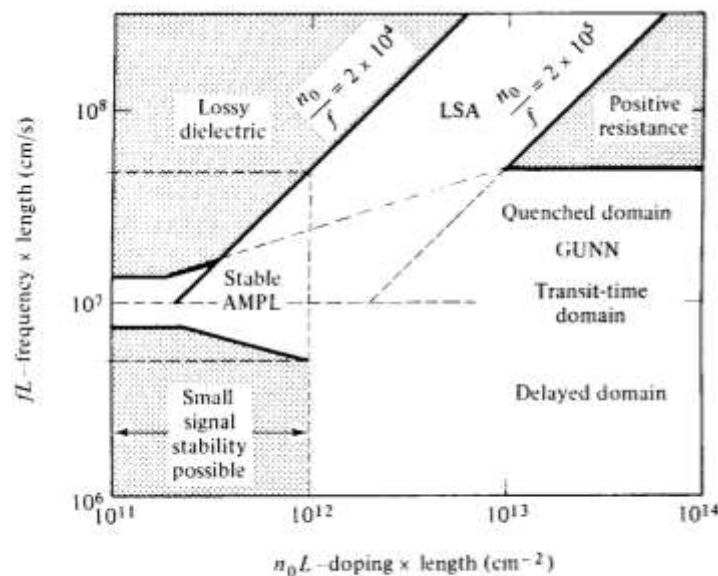


Figure 7-3-1 Modes of operation for Gunn diodes. (After J. A. Copeland [17]; reprinted by permission of IEEE, Inc.)

2. Stable amplification mode: This mode is defined in the region where the product of frequency times length is about 10^7 cm/s and the product of doping times length is between 10^{11} and $10^{12}/\text{cm}^2$.
3. LSA oscillation mode: This mode is defined in the region where the product of frequency times length is above 10^7 cm/s and the quotient of doping divided by frequency is between 2×10^4 and 2×10^5 .
4. Bias-circuit oscillation mode: This mode occurs only when there is either Gunn or LSA oscillation, and it is usually at the region where the product of frequency times length is too small to appear in the figure. When a bulk diode is biased to threshold, the average current suddenly drops as Gunn oscillation begins. The drop in current at the threshold can lead to oscillations in the bias circuit that are typically 1 kHz to 100 MHz [18].

The first three modes are discussed in detail in this section. Before doing so, however, let us consider the criterion for classifying the modes of operation.

Gunn diode as an oscillator

Whilst the Gunn diode has a negative resistance region, it is interesting to see a little more about how this happens and how it acts as an oscillator.

At microwave frequencies, it is found that the dynamic action of the diode incorporates elements resulting from the thickness of the active region.

When the voltage across the active region reaches a certain point a current is initiated that travels across the active region. During the time when the current pulse is moving across the active region the potential gradient falls preventing any further pulses from forming. Only when the pulse has reached the far side of the active region will the potential gradient rise, allowing the next pulse to be created.

It can be seen that the time taken for the current pulse to traverse the active region largely determines the rate at which current pulses are generated. It is this that determines the frequency of operation.

To see how this occurs, it is necessary to look at the electron concentration across the active region. Under normal conditions the concentration of free electrons would be the same regardless of the distance across the active diode region. However a small perturbation may occur resulting from noise from the current flow, or even external noise - this form of noise will always be present and acts as the seed for the oscillation. This grows as it passes across the active region of the Gunn diode.

The increase in free electrons in one area cause the free electrons in another area to decrease forming a form of wave.

The peak will traverse across the diode under the action of the potential across the diode, and growing as it traverses the diode as a result of the negative resistance.

A clue to the reason for this unusual action can be seen if the voltage and current curves are plotted for a normal diode and a Gunn diode. For a normal diode the current increases with voltage, although the relationship is not linear. On the other hand the current for a Gunn diode starts to increase, and once a certain voltage has been reached, it starts to fall before rising again. The region where it falls is known as a negative resistance region, and this is the reason why it oscillates.

Gunn diode advantages & disadvantages

Gunn diode advantages

- High bandwidth
- High reliability
- Low manufacturing cost
- Fair noise performance (does not use avalanche principle).
- Relatively low operating voltage

Gunn diode disadvantages

- Low efficiency below about 10 GHz
- Poor stability – frequency varies with bias and temperature
- FM noise high for some applications
- Small tuning range

Avalanche Transit Time Devices

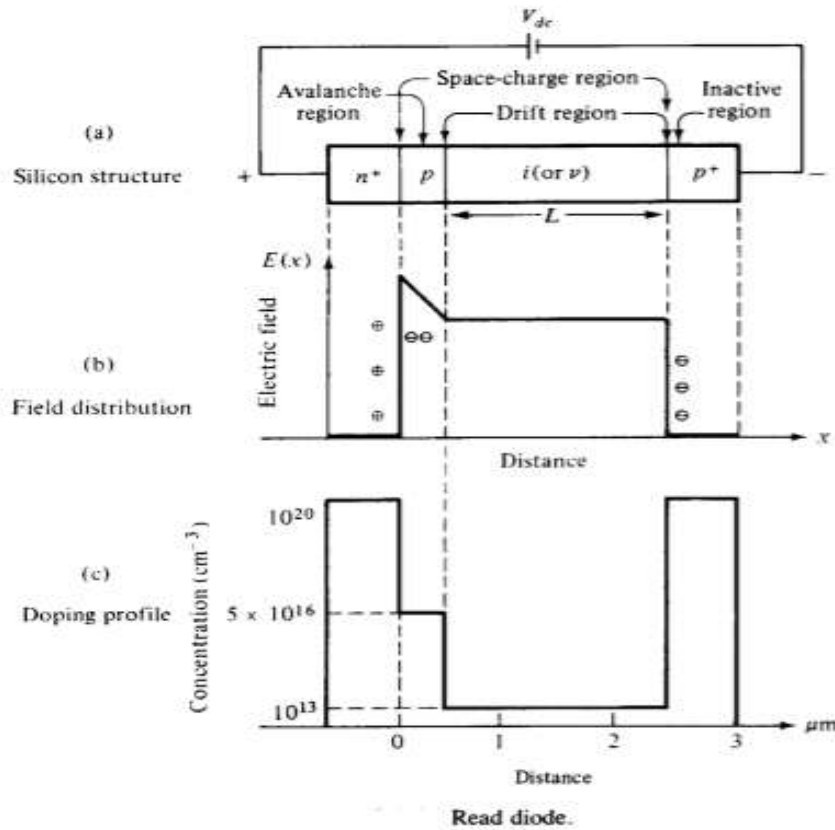
- The process of having a delay between voltage and current, in avalanche together with transit time, through the material is said to be Negative resistance. The devices that help to make a diode exhibit this property are called as Avalanche transit time devices.
- Avalanche transit time diode oscillators rely on the effect of voltage breakdown across a reverse biased p-n junction to produce a supply of holes and electrons.
- Two terminal negative resistance devices.
- The avalanche diode oscillator uses carrier impact ionization and drift in the high field region of a semiconductor junction to produce a negative resistance at microwave frequencies.
- Two distinct modes of avalanche oscillator have been observed as follows
 - (1) IMPATT (Impact Ionization Avalanche Transit Time) mode, where dc-to-RF conversion efficiency is 5 to 10% and frequencies are as high as 100 GHz with silicon diode.
 - (2) TRAPATT (Trapped Plasma Avalanche triggered transit) mode, where dc-to-RF conversion efficiency is 20 to 60%
- Another type of active microwave device is BARITT (Barrier Injected Transit time) diode.

IMPATT/ READ Diode:

- It is a high-power semiconductor diode used in high-frequency microwave electronics devices. It has negative resistance property and can be used as an oscillators and amplifiers at microwave frequencies. It operate at frequencies of about 3 and 100 GHz, or higher.
- The basic operating principle of IMPATT diode can be well explained by refereeing the first avalanche diode, READ diode.

The Read diode is an n^+p-i-p^+ structure, where the superscript plus sign denotes very high doping and the i or v refers to intrinsic material. The device consists essentially of two regions. One is the thin p region at which avalanche multiplication occurs. This region is also called the high-field region or the avalanche region. The other is the i or v region through which the generated holes must drift in moving to the p^+ contact. This region is also called the intrinsic region or the drift region. The p region is very thin. The space between the n^+p junction and the $i-p^+$ junction is called the space-charge region. Similar devices can be built in the $p^+-n-i-n^+$ structure, in which electrons generated from avalanche multiplication drift through the i region.

The Read diode oscillator consists of an n^+p-i-p^+ diode biased in reverse and mounted in a microwave cavity. The impedance of the cavity is mainly inductive and is matched to the mainly capacitive impedance of the diode to form a resonant circuit. The device can produce a negative ac resistance that, in turn, delivers power from the dc bias to the oscillation.



Avalanche Multiplication

When the reverse-biased voltage is well above the punchthrough or breakdown voltage, the space-charge region always extends from the n^+ - p junction through the p and i regions to the i - p^+ junction. The fixed charges in the various regions are shown in Fig. 8-1-1(b). A positive charge gives a rising field in moving from left to right. The maximum field, which occurs at the n^+ - p junction, is about several hundred kilovolts per centimeter. Carriers (holes) moving in the high field near the n^+ - p junction acquire energy to knock valence electrons into the conduction band, thus producing hole-electron pairs. The rate of pair production, or avalanche multiplication, is a sensitive nonlinear function of the field. By proper doping, the field can be given a relatively sharp peak so that avalanche multiplication is confined to a very narrow region at the n^+ - p junction. The electrons move into the n^+ region and the holes drift through the space-charge region to the p^+ region with a constant velocity v_d of about 10^7 cm/s for silicon. The field throughout the space-charge region is above about 5 kV/cm. The transit time of a hole across the drift i -region L is given by

$$\tau = \frac{L}{v_d} \quad (8-1-1)$$

and the avalanche multiplication factor is

$$M = \frac{1}{1 - (V/V_b)^n} \quad (8-1-1a)$$

where V = applied voltage

V_b = avalanche breakdown voltage

n = 3–6 for silicon is a numerical factor depending on the doping of p^+ - n or n^+ - p junction

The breakdown voltage for a silicon p^+-n junction can be expressed as

$$|V_b| = \frac{\rho_n \mu_n \epsilon_s |E_{max}|^2}{2} \quad (8-1-1b)$$

where ρ_n = resistivity

μ_n = electron mobility

ϵ_s = semiconductor permittivity

E_{max} = maximum breakdown of the electric field

A theoretical Read diode made of an $n^+-p-i-p^+$ or $p^+-n-i-n^+$ structure has been analyzed. Its basic physical mechanism is the interaction of the impact ionization avalanche and the transit time of charge carriers. Hence the Read-type diodes are called IMPATT diodes. These diodes exhibit a differential negative resistance by two effects:

1. The impact ionization avalanche effect, which causes the carrier current $I_o(t)$ and the ac voltage to be out of phase by 90°
2. The transit-time effect, which further delays the external current $I_e(t)$ relative to the ac voltage by 90°

it has been confirmed that a negative resistance of the IMPATT diode can be obtained from a junction diode with any doping profile. Many IMPATT diodes consist of a high doping avalanching region followed by a drift region where the field is low enough that the carriers can traverse through it without avalanching. The Read diode is the basic type in the IMPATT diode family. The others are the one-sided abrupt $p-n$ junction, the linearly graded $p-n$ junction (or double-drift region), and the $p-i-n$ diode, all of which are shown in Fig. 8-2-1. The principle of operation of these devices, however, is essentially similar to the mechanism described for the Read diode.

TRAPATT Diode:

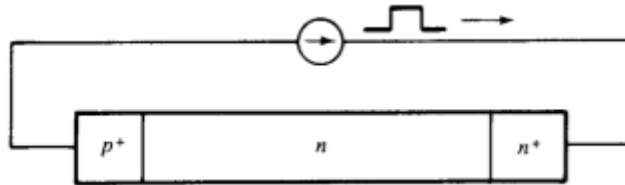
- The TRAPATT diode is normally used as a microwave oscillator. It has the advantage of a greater level of efficiency when compared to an IMPATT microwave diode. Typically the DC to RF signal conversion efficiency may be in the region of 20% to 60% which is particularly high.
- The TRAPATT diode is based around the initial concept of the IMPATT but it has been enhanced by increasing the doping level between the junction and the anode.

The abbreviation TRAPATT stands for *trapped plasma avalanche triggered transit mode*. It is a high-efficiency microwave generator capable of operating from several hundred megahertz to several gigahertz. The basic operation of the oscillator is a semiconductor $p-n$ junction diode reverse-biased to current densities well in excess of those encountered in normal avalanche operation. High-peak-power diodes are typically silicon n^+-p-p^+ (or p^+-n-n^+) structures with the n -type depletion region width varying from 2.5 to 12.5 μm . The doping of the depletion region is generally such that the diodes are well "punched through" at breakdown; that is, the dc electric field in the depletion region just prior to breakdown is well above the saturated drift-velocity level. The device's p^+ region is kept as thin as possible at 2.5 to 7.5 μm . The TRAPATT diode's diameter ranges from as small as 50 μm for CW operation to 750 μm at lower frequency for high-peak-power devices.

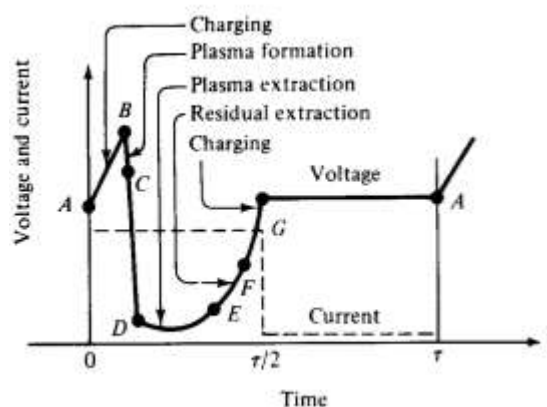
- Diode comprises of two layers of heavily doped P+ and N+ region and a N doped third layer is used to separate the heavily doped layers as shown in figure. The doping concentration of N region is such that the depletion in this region is just at

breakdown. P+ region is kept thin and is of 2.5 to 2.7 μm . Typically the construction of the device consists of a P+, N, N+, although where for higher power levels an N+, P, P+ structure is better. Silicon is also typically used in the fabrication of these devices.

- The TRAPATT is excited using a current pulse. This causes the electric field to increase to a critical value where avalanche multiplication occurs. At this point the field collapses locally due to the generated plasma.



- Diode is operated in reverse biased. This reverse bias causes increase in the electric field between P+ and N region and the minority carriers generated attains a very large velocity. Working of the diode can be explained with the help of following diagram (Time Vs Voltage and current plot).
- At the instant A, the diode current is on. Since current is thermal and diode is reverse biased, it charges like capacitor due to reverse biased condition. This charging of P+ and N region increases the electric field above the breakdown voltage. A heavy current is generated due to breakdown and particle current inside diode increases above the external current due to this electric field in the depletion region decreases. This drop in field is shown by curve from B to C. During this period the E-field is so large that the avalanche continuous and a dense plasma of electron and holes is created. As some of the electrons and holes drift out of the ends of the depletion layer, the field is further depressed and traps the remaining plasma. The voltage decreases to point D. Till the total plasma charge is removed the voltage increase to E and once the residue electrons and holes are removed the voltage is further improved to F. From F to G diode charges up again like a fixed capacitor. Since there is no internally generated charge voltage remains to G only and current to external circuit is zero. After half the period the cycle repeats. This entire action of generation of current through avalanche effect produces a pulse in the external circuit.



- Although the TRAPATT diode provides a much higher level of efficiency than the IMPATT, its major disadvantage is that the noise levels on the signal are even higher than they are when using an IMPATT. It also has very high levels of harmonics as a result of the short current pulses that are used. A balance needs to be made between the different options according to the particular application.

** BARITT Diode:

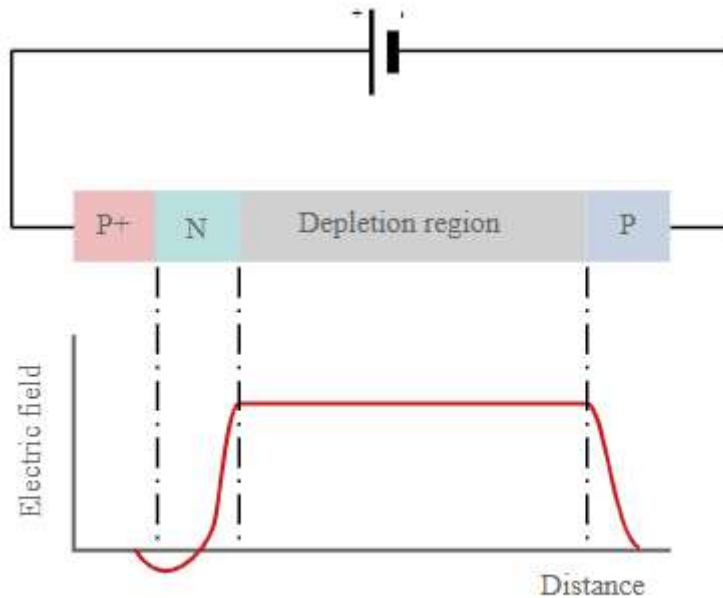
- Commonly referred to as **Barrier Injection Transit-Time Diode** has many similarities to the more widely used IMPATT diode.
- The main difference between BARITT Diode and other Diode is that BARITT uses thermionic emission whereas other diode works on avalanche multiplication.
- One of the advantages of using this form of emission is that the process is far less noisy and as a result the BARITT does not suffer from the same noise levels as does the IMPATT.
- It usually used for Microwave Signal Generations of frequencies up to 25 Ghz for Silicon (Si) Material and 90 GHz for Gallium-Arsenide (GaAs).

BARITT diodes, meaning *barrier injected transit-time* diodes, are the latest addition to the family of active microwave diodes. They have long drift regions similar to those of IMPATT diodes. The carriers traversing the drift regions of BARITT diodes, however, are generated by minority carrier injection from forward-biased junctions instead of being extracted from the plasma of an avalanche region.

Several different structures have been operated as BARITT diodes, including *p-n-p*, *p-n-v-p*, *p-n-metal*, and *metal-n-metal*. For a *p-n-v-p* BARITT diode, the forward-biased *p-n* junction emits holes into the *v* region. These holes drift with saturation velocity through the *v* region and are collected at the *p* contact. The diode exhibits a negative resistance for transit angles between π and 2π . The optimum transit angle is approximately 1.6π .

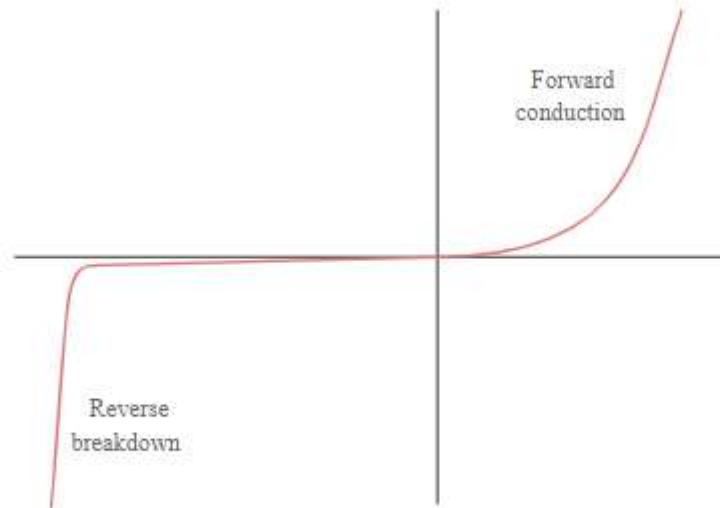
Such diodes are much less noisy than IMPATT diodes. Noise figures are as low as 15 dB at C-band frequencies with silicon BARITT amplifiers. The major disadvantages of BARITT diodes are relatively narrow bandwidth and power outputs limited to a few milliwatts.

- Essentially the BARITT diode consists of two back to back diodes. When a potential is applied across the device, most of the potential drop occurs across the reverse biased diode.
- The device has areas often referred to as the emitter, base, intermediate or drift area and the collector.
- If the voltage is then increased until the edges of the depletion region meet, then a condition known as punch through occurs.
- In terms of the operation of the device, the depletion or drift region needs to be completely free of carriers and this means that punch through occurs to the base-emitter region without there being avalanche breakdown of the base collector junction.
- It can be seen within the diagram that the punch through voltages, V_{pt} are different for the two directions. This difference results from asymmetry in the two junctions and can be controlled during the manufacture stages of the diode. They can be made to be different or almost the same.

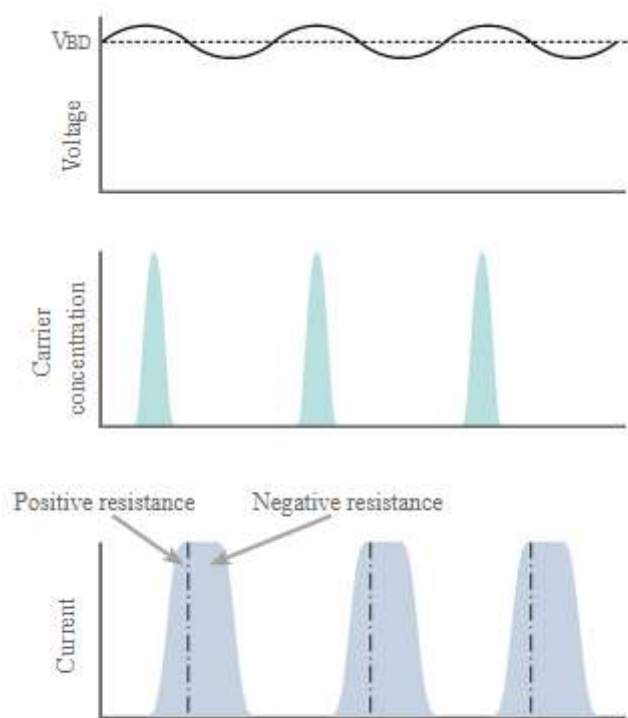


BARITT diode construction and operation

- After a charge is injected, it travels to the substrate with the saturation velocity.
- As seen from the diagram, it can be seen that the injection current is in phase with the RF voltage waveform. This results in a non-ideal current waveform situation which flows in the positive resistance region and therefore losses are higher in the BARITT than in an IMPATT.
- The terminal current pulse width is determined by the transit time which is L/v_{sat} (Where the electrodes are spaced L apart and v_{sat} is the saturation velocity). This constitutes around three quarters of the cycle.
- In view of the physical restraints of the BARITT diode, the power capability decreases approximately as the square of the frequency because higher frequencies require a smaller separation between the electrodes and this in turn limits the voltages that can be used. Also the efficiency falls away with increasing frequency. For low frequency operation it may be around 5% or a little more.



BARITT diode I-V curve or characteristic



BARITT diode current & voltage waveforms

Microwave Amplification by Stimulated Emission of Radiation (MASER)

- A maser is a device that produces coherent electromagnetic waves through amplification by stimulated emission OR the device which produces and amplifies electromagnetic radiation mainly in the microwave region of the spectrum.
- Masers produce extremely sharp radiation with low internal noise, and they serve as high-precision frequency references.
- The maser operates according to the same basic principle as the laser and shares many of its characteristics.

Physical structure:

- A maser oscillator requires a source of excited atoms or molecules and a resonator to store their radiation. The excitation must force more atoms or molecules into the upper energy level than in the lower, in order for amplification by stimulated emission to predominate over absorption.
- For wavelengths of a few millimetres or longer, the resonator can be a metal box whose dimensions are chosen so that only one of its modes of oscillation coincides with the frequency emitted by the atoms; that is, the box is resonant at the particular frequency, much as a kettle drum is resonant at some particular audio frequency.
- The losses of such a resonator can be made quite small, so that radiation can be stored long enough to stimulate emission from successive atoms as they are excited. Thus, all the atoms are forced to emit in such a way as to augment this stored wave. Output is obtained by allowing some radiation to escape through a small hole in the resonator.

General Principle of operation

- When atoms or molecules of an appropriate substance (called a medium) are bombarded with photons of a particular frequency, they go into an excited (higher) energy state and emit photons of the same frequency. In this sense, the maser involves stimulated emission of radiation. By putting the amplifying medium in a resonant cavity (or cavity resonator), feedback is created that can produce radiation that is coherent. Electromagnetic waves are said to be coherent when they are propagated at the same frequency in the same phase, and they move in the same direction. By contrast, electromagnetic waves from most other

sources have a range of different frequencies, they are in different phases (relative to one another), and they are propagated in practically all directions.

- Radio waves emitted by a maser have nearly the same frequency and their transmission over long distances is highly efficient. In the first maser to be developed, the medium in the resonant cavity was ammonia gas. In this case, the molecules of ammonia oscillated at a particular frequency between two energy states. More recently, a ruby maser has been developed, in which a ruby crystal is placed in the resonant cavity. The dual noble gas maser is an example of a nonpolar medium in a maser.

Some common types of masers are noted below. The names indicate the medium present in the resonant cavity.

Atomic beam masers

- Ammonia maser
- Hydrogen maser

Gas masers

- Rubidium maser

Solid State masers

- Ruby maser.

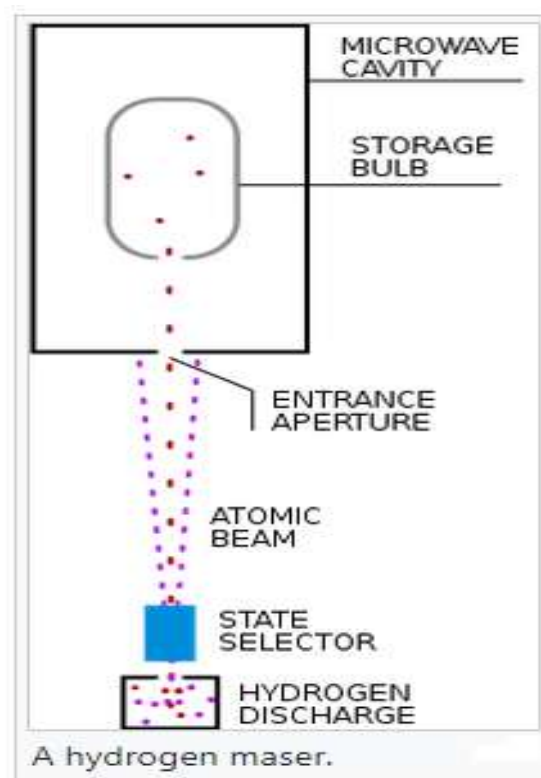
Hydrogen maser

- Today, the most important type of maser is the hydrogen maser, which provides a sharp and constant oscillating signal. It is based on transitions in atomic hydrogen that occur at a frequency of 1421 megahertz. This maser is used as an atomic frequency standard. Together with other types of atomic clocks, they constitute the "Temps Atomique International" or TAI. This is the international time scale, which is coordinated by the Bureau International des Poids et Mesures, or BIPM

Principle of Operation:

- First, a beam of atomic hydrogen is produced by exposing hydrogen gas at low pressure to a radio-frequency discharge. (See the box on the bottom of the diagram on the right.)

- The next step is known as "state selection." To get some stimulated emission, it is necessary to create a population inversion of the atoms—that is, most of the atoms need to be in the excited energy state (rather than in a lower energy state). This is done in a manner similar to the famous Stern-Gerlach experiment. After passing through an aperture and a magnetic field, many of the atoms in the beam are left in the upper energy level of the lasing transition. From this state, the atoms can decay to the lower energy state and emit some microwave radiation.
- A high quality factor microwave cavity confines the microwaves and reinjects them repeatedly into the atom beam. The stimulated emission amplifies the microwaves on each pass through the beam. This combination of amplification and feedback defines all oscillators. The resonant frequency of the microwave cavity is $1420\,405\,751.768\text{ Hz}$, which is exactly tuned to the hyperfine structure of hydrogen.
- A small fraction of the signal in the microwave cavity is coupled into a coaxial cable and sent to a coherent receiver.
- The microwave signal coming out of the maser is very weak in power (a few picowatts (pW)). The frequency of the signal is fixed but extremely stable. The coherent receiver is used to amplify the signal and change the frequency. This is done using a series of phase-locked loops and a high-performance quartz oscillator.



Astrophysical Maser

- Maser-like stimulated emission has also been observed in nature from interstellar space, and it is frequently called "superradiant emission" to distinguish it from laboratory masers. Such emission is observed from molecules such as water (H_2O), hydroxyl radicals ($\bullet\text{OH}$), methanol (CH_3OH), formaldehyde (HCHO), and silicon monoxide (SiO). Water molecules in star-forming regions can undergo a population inversion and emit radiation at about 22.0 GHz, creating the brightest spectral line in the radio universe. Some water masers also emit radiation from a rotational transition at a frequency of 96 GHz.

Microwave Cavities

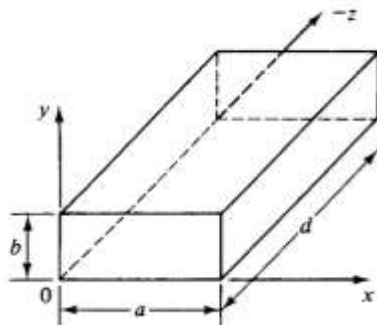
- Resonators can be constructed from closed sections of waveguide, which should not be surprising since waveguides are a type of transmission Line. Because of radiation Loss from open-ended waveguide, waveguide resonators are usually short circuited at both ends, thus forming a closed box or cavity. Electric and magnetic energy is stored within the cavity and power can be dissipated in the metallic walls of the cavity as well as in the dielectric filling the cavity. Coupling to the resonator can be done through a small aperture or a small probe or a loop.

OR

In general, a cavity resonator is a metallic enclosure that confines the electromagnetic energy. The stored electric and magnetic energies inside the cavity determine its equivalent inductance and capacitance. The energy dissipated by the finite conductivity of the cavity walls determines its equivalent resistance. In practice, the rectangular-cavity resonator, circular-cavity resonator, and reentrant-cavity resonator are commonly used in many microwave applications.

Theoretically a given resonator has an infinite number of resonant modes, and each mode corresponds to a definite resonant frequency. When the frequency of an impressed signal is equal to a resonant frequency, a maximum amplitude of the standing wave occurs, and the peak energies stored in the electric and magnetic fields are equal. The mode having the lowest resonant frequency is known as the *dominant mode*.

Rectangular waveguide cavity resonator



The geometry of rectangular cavity resonator spreads as

$$0 \leq x \leq a; \quad 0 \leq y \leq b; \quad 0 \leq z \leq d$$

The wave equations in the rectangular resonator should satisfy the boundary condition of the zero tangential $\mathbf{E}$ at four of the walls. It is merely necessary to choose the harmonic functions in z to satisfy this condition at the remaining two end

walls. These functions can be found if

$$H_z = H_{0z} \cos\left(\frac{m\pi x}{a}\right) \cos\left(\frac{n\pi y}{b}\right) \sin\left(\frac{p\pi z}{d}\right) \quad (\text{TE}_{mnp})$$

where $m = 0, 1, 2, 3, \dots$ represents the number of the half-wave periodicity in the x direction

$n = 0, 1, 2, 3, \dots$ represents the number of the half-wave periodicity in the y direction

$p = 1, 2, 3, 4, \dots$ represents the number of the half-wave periodicity in the z direction

and

$$E_z = E_{0z} \sin\left(\frac{m\pi x}{a}\right) \sin\left(\frac{n\pi y}{b}\right) \cos\left(\frac{p\pi z}{d}\right) \quad (\text{TM}_{mnp})$$

where $m = 1, 2, 3, 4, \dots$

$n = 1, 2, 3, 4, \dots$

$p = 0, 1, 2, 3, \dots$

The separation equation for both TE and TM modes is given by

$$k^2 = \left(\frac{m\pi}{a}\right)^2 + \left(\frac{n\pi}{b}\right)^2 + \left(\frac{p\pi}{d}\right)^2$$

For a lossless dielectric, $k^2 = \omega^2 \mu \epsilon$; therefore, the resonant frequency is expressed by

$$f_r = \frac{1}{2\sqrt{\mu\epsilon}} \sqrt{\left(\frac{m}{a}\right)^2 + \left(\frac{n}{b}\right)^2 + \left(\frac{p}{d}\right)^2} \quad (\text{TE}_{mnp}, \text{TM}_{mnp})$$

For $a > b < d$, the dominant mode is the TE_{101} mode.

In general, a straight-wire probe inserted at the position of maximum electric intensity is used to excite a desired mode, and the loop coupling placed at the position of maximum magnetic intensity is utilized to launch a specific mode.

The maximum amplitude of the standing wave occurs when the frequency of the impressed signal is equal to the resonant frequency.

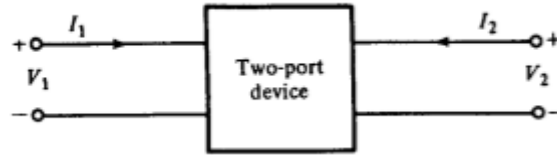
The quality factor Q is a measure of the frequency selectivity of a resonant or antiresonant circuit, and it is defined as

$$Q = 2\pi \frac{\text{maximum energy stored}}{\text{energy dissipated per cycle}} = \frac{\omega W}{P}$$

where W is the maximum stored energy and P is the average power loss.

Microwave Hybrid Circuits:

A microwave circuit ordinarily consists of several microwave devices connected in some way to achieve the desired transmission of a microwave signal. The interconnection of two or more microwave devices may be regarded as a microwave junction. Commonly used microwave junctions include such waveguide tees as the E -plane tee, H -plane tee, magic tee, hybrid ring (rat-race circuit), directional coupler, and the circulator.



A two-port network is shown in above figure. From network theory a two-port device can be described by a number of parameter sets, such as the H , Y , Z , and $ABCD$ parameters.

$$H \text{ parameters: } \begin{bmatrix} V_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ V_2 \end{bmatrix} \quad \begin{aligned} V_1 &= h_{11}I_1 + h_{12}V_2 \\ I_2 &= h_{21}I_1 + h_{22}V_2 \end{aligned}$$

$$Y \text{ parameters: } \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad \begin{aligned} I_1 &= y_{11}V_1 + y_{12}V_2 \\ I_2 &= y_{21}V_1 + y_{22}V_2 \end{aligned}$$

$$Z \text{ parameters: } \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \quad \begin{aligned} V_1 &= z_{11}I_1 + z_{12}I_2 \\ V_2 &= z_{21}I_1 + z_{22}I_2 \end{aligned}$$

$$ABCD \text{ parameters: } \begin{bmatrix} V_1 \\ I_1 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} V_2 \\ -I_2 \end{bmatrix} \quad \begin{aligned} V_1 &= AV_2 - BI_2 \\ I_1 &= CV_2 - DI_2 \end{aligned}$$

All these network parameters relate total voltages and total currents at each of the two ports. For instance,

$$h_{11} = \left. \frac{V_1}{I_1} \right|_{V_2=0} \quad (\text{short circuit})$$

$$h_{12} = \left. \frac{V_1}{V_2} \right|_{I_1=0} \quad (\text{open circuit})$$

If the frequencies are in the microwave range, however, the H , Y , and Z parameters cannot be measured for the following reasons:

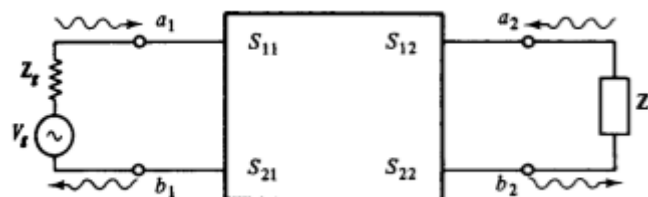
1. Equipment is not readily available to measure total voltage and total current at the ports of the network.
2. Short and open circuits are difficult to achieve over a broad band of frequencies.
3. Active devices, such as power transistors and tunnel diodes, frequently will not have stability for a short or open circuit.

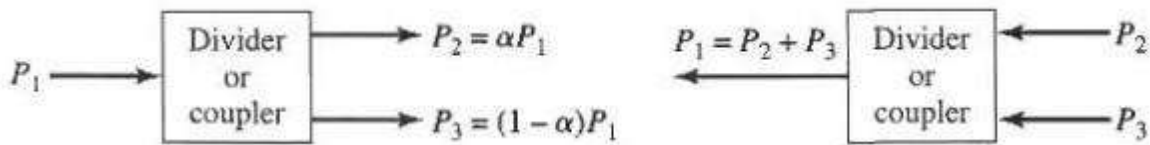
Consequently, some new method of characterization is needed to overcome these problems. The logical variables to use at the microwave frequencies are traveling waves rather than total voltages and total currents. These are the S parameters, which are expressed as

$$b_1 = S_{11}a_1 + S_{12}a_2$$

$$b_2 = S_{21}a_1 + S_{22}a_2$$

the S parameters of a two-port network is shown in below Fig.



POWER DIVIDERS:

- Power dividers and combiners are the passive microwave components that are used for power division and combination in microwave frequency range
- In this case, the input power is divided into two or more signals of lesser power.
- The power divider has certain basic parameters like isolation, coupling factor and directivity.

T-JUNCTION POWER DIVIDER USING WAVEGUIDE:

The T-junction power divider is a 3-port network that can be constructed either from a transmission line or from the waveguide depending upon the frequency of operation.

In microwave circuits a waveguide or coaxial-line junction with three independent ports is commonly referred to as a *tee junction*. From the *S*-parameter theory of a microwave junction it is evident that a tee junction should be characterized by a matrix of third order containing nine elements, six of which should be independent. The characteristics of a three-port junction can be explained by three theorems of the tee junction. These theorems are derived from the equivalent-circuit representation of the tee junction. Their statements follow

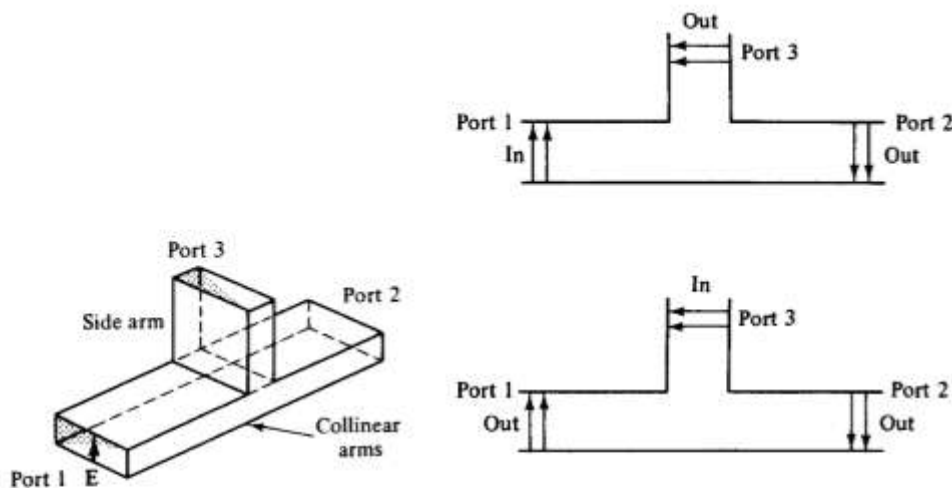
1. A short circuit may always be placed in one of the arms of a three-port junction in such a way that no power can be transferred through the other two arms.
2. If the junction is symmetric about one of its arms, a short circuit can always be placed in that arm so that no reflections occur in power transmission between the other two arms. (That is, the arms present matched impedances.)
3. It is impossible for a general three-port junction of arbitrary symmetry to present matched impedances at all three arms.

For very high frequency, power divider using waveguide is of 4 types

- E-Plane Tee
- H-Plane Tee
- E-H Plane Tee/Magic Tee
- Rat Race Tee

E-Plane/ Series Tee

- It can be constructed by making a rectangular slot along the wide dimension of the main waveguide and inserting another auxiliary waveguide along the direction so that it becomes a 3-port network.
- Port-1 and Port-2 are called collinear ports and Port-3 is called the E-arm.
- E-arm is parallel to the electric field of the main waveguide.
- If the wave is entering into the junction from E-arm it splits or gets divided into Port-1 and Port-2 with equal magnitude but opposite in phase
- If the wave is entering through Port-1 and Port-2 then the resulting field through Port-3 is proportional to the difference between the instantaneous field from Port-1 and Port-2



if the E -plane tee is perfectly matched with the aid of screw tuners or inductive or capacitive windows at the junction, the diagonal components of the scattering matrix, S_{11} , S_{22} , and S_{33} , are zero because there will be no reflection. When the waves are fed into the side arm (port 3), the waves appearing at port 1 and port 2 of the collinear arm will be in opposite phase and in the same magnitude. Therefore

$$S_{13} = -S_{23}$$

It should be noted that it does not mean that S_{13} is always positive and S_{23}

is always negative. The negative sign merely means that S_{13} and S_{23} have opposite signs. For a matched junction, the S matrix is given by

$$S = \begin{bmatrix} 0 & S_{12} & S_{13} \\ S_{21} & 0 & S_{23} \\ S_{31} & S_{32} & 0 \end{bmatrix} \quad (4-4-13)$$

From the symmetry property of S matrix, the symmetric terms in Eq. (4-4-13) are equal and they are

$$S_{12} = S_{21} \quad S_{13} = S_{31} \quad S_{23} = S_{32} \quad (4-4-14)$$

From the zero property of S matrix, the sum of the products of each term of any column (or row) multiplied by the complex conjugate of the corresponding terms of any other column (or row) is zero and it is

$$S_{11}S_{12}^* + S_{21}S_{22}^* + S_{31}S_{32}^* = 0 \quad (4-4-15)$$

Hence

$$S_{13}S_{23}^* = 0 \quad (4-4-16)$$

This means that either S_{13} or S_{23}^* , or both, should be zero. However, from the unity property of S matrix, the sum of the products of each term of any one row (or column) multiplied by its complex conjugate is unity; that is,

$$S_{21}S_{21}^* + S_{31}S_{31}^* = 1 \quad (4-4-17)$$

$$S_{12}S_{12}^* + S_{32}S_{32}^* = 1 \quad (4-4-18)$$

$$S_{13}S_{13}^* + S_{23}S_{23}^* = 1 \quad (4-4-19)$$

Substitution of Eq. (4-4-14) in (4-4-17) results in

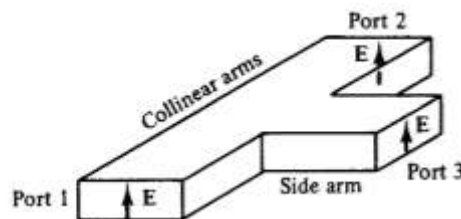
$$|S_{12}|^2 = 1 - |S_{13}|^2 = 1 - |S_{23}|^2 \quad (4-4-20)$$

Equations (4-4-19) and (4-4-20) are contradictory, for if $S_{13} = 0$, then S_{23} is also zero and thus Eq. (4-4-19) is false. In a similar fashion, if $S_{23} = 0$, then S_{13} becomes zero and therefore Eq. (4-4-20) is not true. This inconsistency proves the statement that the tee junction cannot be matched to the three arms. In other words, the diagonal elements of the S matrix of a tee junction are not all zeros.

In general, when an E -plane tee is constructed of an empty waveguide, it is poorly matched at the tee junction. Hence $S_{ii} \neq 0$ if $i = j$. However, since the collinear arm is usually symmetric about the side arm, $|S_{13}| = |S_{23}|$ and $S_{11} = S_{22}$. Then the S matrix can be simplified to

$$S = \begin{bmatrix} S_{11} & S_{12} & S_{13} \\ S_{12} & S_{11} & -S_{13} \\ S_{13} & -S_{13} & S_{33} \end{bmatrix} \quad (4-4-21)$$

H-Plane/ Shunt Tee



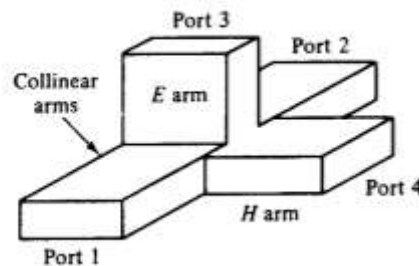
An H -plane tee is a waveguide tee in which the axis of its side arm is “shunting” the E field or parallel to the H field of the main guide

It can be seen that if two input waves are fed into port 1 and port 2 of the collinear arm, the output wave at port 3 will be in phase and additive. On the other hand, if the input is fed into port 3, the wave will split equally into port 1 and port 2 in phase and in the same magnitude. Therefore the S matrix of the *H*-plane tee is similar to Eqs. (4-4-13) and (4-4-21) except that

$$S_{13} = S_{23} \quad (4-4-22)$$

- An H-plane tee is formed by making a rectangular slot along the width of the main waveguide and inserting an auxiliary waveguide along this direction.
- In this case, the axis of the H-arm is parallel to the plane of the main waveguide.
- The wave entering through H-arm splits up through Port-1 and Port-2 with equal magnitude and same phase
- If the wave enters through Port-1 and Port-2 then the power through Port-3 is the phasor sum of those at Port-1 and Port-2.
- E-Plane tee is called PHASE DELAY and H-Plane tee is called PHASE ADVANCE.

E-H Plane Tee/Magic Tee



- It is a combination of E-Plane tee and H-Plane tee.
- If two waves of equal magnitude and the same phase are fed into Port-1 and Port-2, the output will be zero at Port-3 and additive at Port-4.
- If a wave is fed into Port-4 (H-arm) then it will be divided equally between Port-1 and Port-2 of collinear arms (same in phase) and will not appear at Port-3 or E-arm.
- If a wave is fed in Port-3 then it will produce an output of equal magnitude and opposite phase at Port-1 and Port-2 and the output at Port-4 will be zero.
- If a wave is fed in any one of the collinear arms at Port-1 or Port-2, it will not appear in the other collinear arm because the E-arm causes a phase delay and the H-arm causes phase advance.

The magic tee is commonly used for mixing, duplexing, and impedance measurements. Suppose, for example, there are two identical radar transmitters in equipment stock. A particular application requires twice more input power to an antenna than either transmitter can deliver. A magic tee may be used to couple the two transmitters to the antenna in such a way that the transmitters do not load each other. The two transmitters should be connected to ports 3 and 4, respectively, as shown in Fig. 4-4-8. Transmitter 1, connected to port 3, causes a wave to emanate from port 1 and another to emanate from port 2; these waves are equal in magnitude but opposite in phase. Similarly, transmitter 2, connected to port 4, gives rise to a wave at port 1 and another at port 2, both equal in magnitude and in phase. At port 1 the two opposite waves cancel each other. At port 2 the two in-phase waves add together; so double output power at port 2 is obtained for the antenna as shown in Fig. 4-4-8.

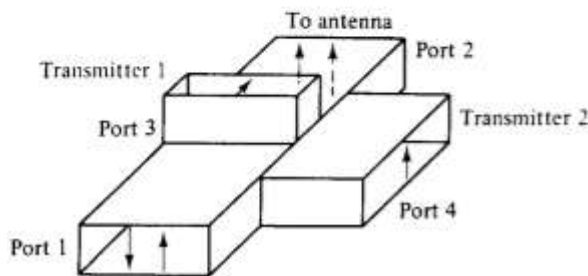
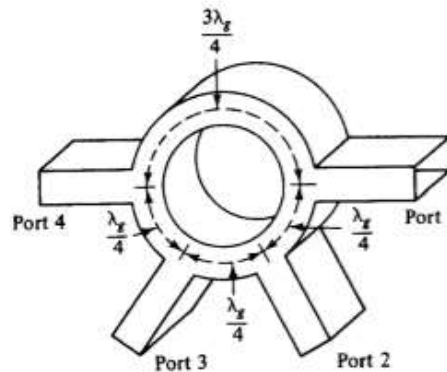


Figure 4-4-8 Magic tee-coupled transmitters to antenna.

Rat Race Tee / Hybrid Ring



A hybrid ring consists of an annular line of proper electrical length to sustain standing waves, to which four arms are connected at proper intervals by means of series or parallel junctions.

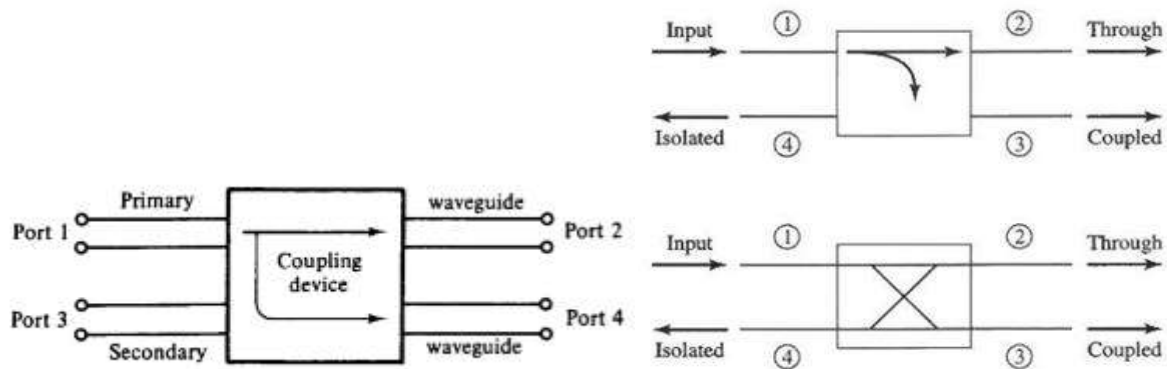
The hybrid ring has characteristics similar to those of the hybrid tee. When a wave is fed into port 1, it will not appear at port 3 because the difference of phase shifts for the waves traveling in the clockwise and counterclockwise directions is 180° . Thus the waves are canceled at port 3. For the same reason, the waves fed into port 2 will not emerge at port 4 and so on.

The S matrix for an ideal hybrid ring can be expressed as

$$\mathbf{S} = \begin{bmatrix} 0 & S_{12} & 0 & S_{14} \\ S_{21} & 0 & S_{23} & 0 \\ 0 & S_{32} & 0 & S_{34} \\ S_{41} & 0 & S_{43} & 0 \end{bmatrix} \quad (4-4-24)$$

It should be noted that the phase cancellation occurs only at a designated frequency for an ideal hybrid ring. In actual hybrid rings there are small leakage couplings, and therefore the zero elements in the matrix of Eq. (4-4-24) are not quite equal to zero.

Directional Coupler



- It is a 4- port waveguide junction consisting of a primary waveguide 1-2 and a secondary waveguide 3-4.
- When all the ports are terminated in their characteristic impedance there is free transmission of power without reflection between port-1 and port-2 and no power transmission takes place between port-1 and port-3 or port-2 and port-4 as no coupling exists.
- The characteristic of a directional coupler is expressed in terms of its coupling factor and directivity.
- The coupling factor is the measure of ratio of power levels in primary and secondary lines.
- Directivity is the measure of how well the forward travelling wave in the primary waveguide couples only to a specific port of the secondary waveguide.
- In ideal case, directivity is infinite i.e. power at port-3 =0 because port-2 and port-4 are perfectly matched.
- Let wave propagates from port-1 to port-2 in primary line then:

$$\text{Coupling factor (dB)} = 10 \log_{10} (P_1/P_4)$$

$$\text{Directivity (dB)} = 10 \log_{10} (P_4/P_3)$$

Where P_1 =power input to port-1

P_3 =power output from port-3

P_4 =power output from port-4

It should be noted that port 2, port 3, and port 4 are terminated in their characteristic impedances. The coupling factor is a measure of the ratio of power levels in the primary and secondary lines. Hence if the coupling factor is known, a fraction of power measured at port 4 may be used to determine the power input at port 1. This significance is desirable for microwave power measurements because no disturbance, which may be caused by the power measurements, occurs in the primary line. The directivity is a measure of how well the forward traveling wave in the primary waveguide couples only to a specific port of the secondary waveguide. An ideal directional coupler should have infinite directivity. In other words, the power at port 3 must be zero because port 2 and port 4 are perfectly matched. Actually, well-designed directional couplers have a directivity of only 30 to 35 dB.

Several types of directional couplers exist, such as a two-hole directional coupler, four-hole directional coupler, reverse-coupling directional coupler (Schwinger coupler), and Bethe-hole directional coupler (refer to Fig. 4-5-2). Only the very commonly used two-hole directional coupler is described here.

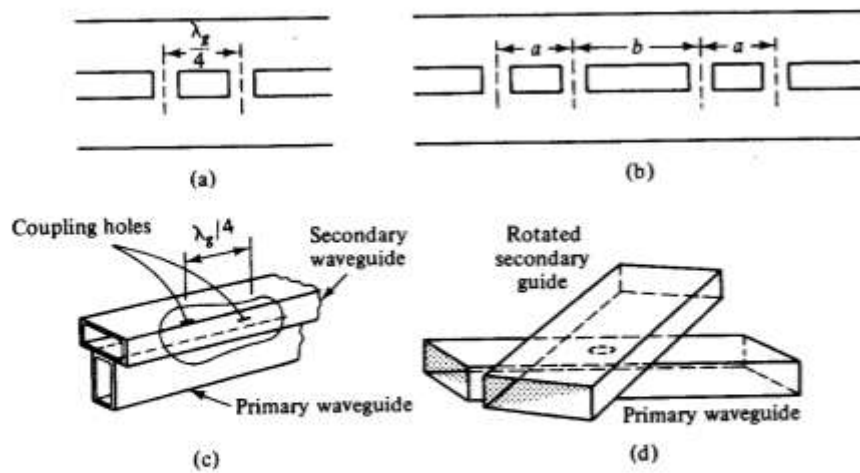


Figure 4-5-2 Different directional couplers. (a) Two-hole directional coupler. (b) Four-hole directional coupler. (c) Schwinger coupler. (d) Bethe-hole directional coupler.

In a directional coupler all four ports are completely matched. Thus the diagonal elements of the S matrix are zeros and

$$S_{11} = S_{22} = S_{33} = S_{44} = 0 \quad (4-5-4)$$

As noted, there is no coupling between port 1 and port 3 and between port 2 and port 4. Thus

$$S_{13} = S_{31} = S_{24} = S_{42} = 0 \quad (4-5-5)$$

Consequently, the S matrix of a directional coupler becomes

$$\mathbf{S} = \begin{bmatrix} 0 & S_{12} & 0 & S_{14} \\ S_{21} & 0 & S_{23} & 0 \\ 0 & S_{32} & 0 & S_{34} \\ S_{41} & 0 & S_{43} & 0 \end{bmatrix} \quad (4-5-6)$$

Equation (4-5-6) can be further reduced by means of the zero property of the S matrix, so we have

$$S_{12}S_{14}^* + S_{32}S_{34}^* = 0 \quad (4-5-7)$$

$$S_{21}S_{23}^* + S_{41}S_{43}^* = 0 \quad (4-5-8)$$

Also from the unity property of the S matrix, we can write

$$S_{12}S_{12}^* + S_{14}S_{14}^* = 1 \quad (4-5-9)$$

Equations (4-5-7) and (4-5-8) can also be written

$$|S_{12}||S_{14}| = |S_{32}||S_{34}| \quad (4-5-10)$$

$$|S_{21}||S_{23}| = |S_{41}||S_{43}| \quad (4-5-11)$$

Since $S_{12} = S_{21}$, $S_{14} = S_{41}$, $S_{23} = S_{32}$, and $S_{34} = S_{43}$, then

$$|S_{12}| = |S_{34}| \quad (4-5-12)$$

$$|S_{14}| = |S_{23}| \quad (4-5-13)$$

Let

$$S_{12} = S_{34} = p \quad (4-5-14)$$

where p is positive and real. Then from Eq. (4-5-8)

$$p(S_{23}^* + S_{41}) = 0 \quad (4-5-15)$$

Let

$$S_{23} = S_{41} = jq \quad (4-5-16)$$

where q is positive and real. Then from Eq. (4-5-9)

$$p^2 + q^2 = 1 \quad (4-5-17)$$

The S matrix of a directional coupler is reduced to

$$S = \begin{bmatrix} 0 & p & 0 & jq \\ p & 0 & jq & 0 \\ 0 & jq & 0 & p \\ jq & 0 & p & 0 \end{bmatrix} \quad (4-5-18)$$

Two-Hole Directional Couplers

A two-hole directional coupler with traveling waves propagating in it is illustrated in Fig. 4-5-3. The spacing between the centers of two holes must be

$$L = (2n + 1) \frac{\lambda_g}{4} \quad (4-5-3)$$

where n is any positive integer.

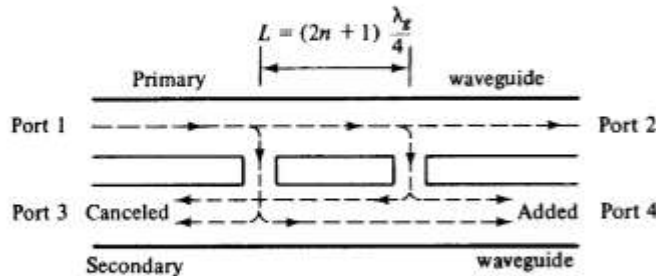


Figure 4-5-3 Two-hole directional coupler.

A fraction of the wave energy entered into port 1 passes through the holes and is radiated into the secondary guide as the holes act as slot antennas. The forward waves in the secondary guide are in the same phase, regardless of the hole space, and are added at port 4. The backward waves in the secondary guide (waves are progressing from right to left) are out of phase by $(2L/\lambda_g)2\pi$ rad and are canceled at port 3.

Waveguide corners, Bends and Twists

The waveguide corner, bend, and twist are shown in Fig. 4-4-10. These waveguide components are normally used to change the direction of the guide through an arbitrary angle.

In order to minimize reflections from the discontinuities, it is desirable to have the mean length L between continuities equal to an odd number of quarter-wave-lengths. That is,

$$L = (2n + 1) \frac{\lambda_g}{4} \quad (4-4-25)$$

where $n = 0, 1, 2, 3, \dots$, and λ_g is the wavelength in the waveguide. If the mean length L is an odd number of quarter wavelengths, the reflected waves from both ends of the waveguide section are completely canceled. For the waveguide bend, the minimum radius of curvature for a small reflection is given by Southworth [2] as

$$R = 1.5b \quad \text{for an } E \text{ bend} \quad (4-4-26)$$

$$R = 1.5a \quad \text{for an } H \text{ bend} \quad (4-4-27)$$

where a and b are the dimensions of the waveguide bend as illustrated in Fig. 4-4-10.

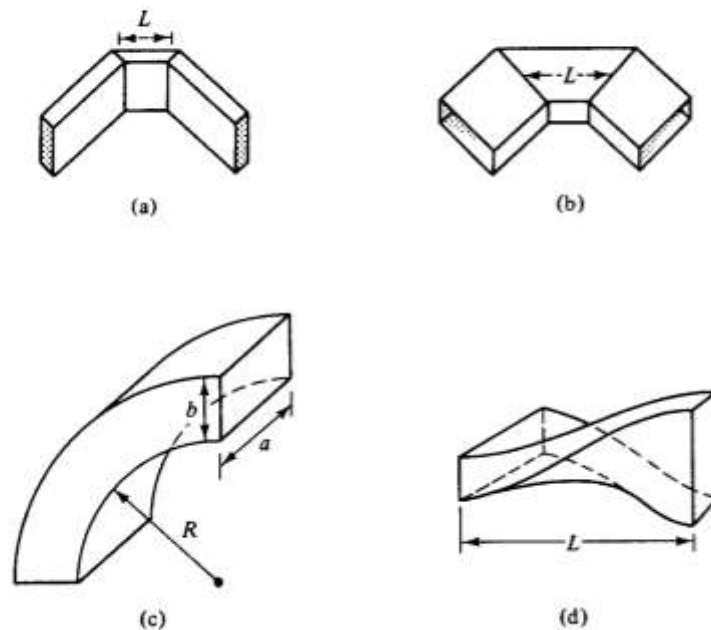


Figure 4-4-10 Waveguide corner, bend, and twist. (a) *E*-plane corner. (b) *H*-plane corner. (c) Bend. (d) Continuous twist.

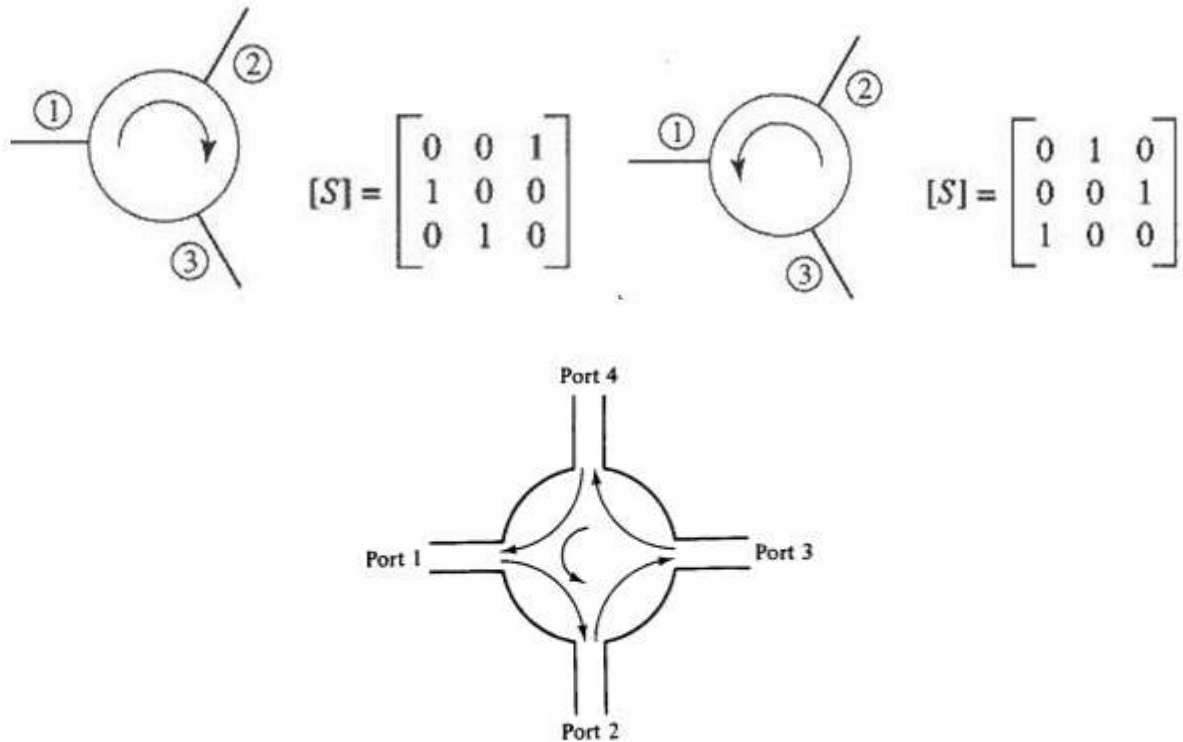
CIRCULATORS AND ISOLATORS:

- Both microwave circulators and microwave isolators are non-reciprocal transmission devices that use Faraday rotation in the ferrite material.

CIRCULATOR:

- A microwave circulator is a multiport waveguide junction in which the wave can flow only in one direction i.e. from the n th port to the $(n+1)$ th port.

- It has no restriction on the number of ports
- 4-port microwave circulator is most common.
- One of its types is a combination of two 3-dB side hole directional couplers and a rectangular waveguide with two non reciprocal phase shifters.



- Each of the two 3db couplers introduce phase shift of 90 degrees
- Each of the two phase shifters produce a fixed phase change in a certain direction.
- Wave incident to port-1 splits into 2 components by coupler-1.
- The wave in primary guide arrives at port-2 with 180 degrees phase shift.
- The second wave propagates through two couplers and secondary guide and arrives at port-2 with a relative phase shift of 180 degrees.
- But at port-4 the wave travelling through primary guide phase shifter and coupler-2 arrives with 270 degrees phase change.
- Wave from coupler-1 and secondary guide arrives at port-4 with phase shift of 90 degrees.
- Power transmission from port-1 to port-4 = 0 as the two waves reaching at port-4 are out of phase by 180 degrees.

$$w_1 - w_3 = (2m+1) \pi \text{ rad/s}$$

$$w_2 - w_4 = 2n\pi \text{ rad/s}$$

Power flow sequence: 1-> 2 -> 3 -> 4-> 1

A perfectly matched, lossless, and nonreciprocal four-port circulator has an S matrix of the form

$$S = \begin{bmatrix} 0 & S_{12} & S_{13} & S_{14} \\ S_{21} & 0 & S_{23} & S_{24} \\ S_{31} & S_{32} & 0 & S_{34} \\ S_{41} & S_{42} & S_{43} & 0 \end{bmatrix}$$

Using the properties of S parameters as described previously, the S matrix can be simplified to

$$S = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

MICROWAVE ISOLATOR:

- A non reciprocal transmission device used to isolate one component from reflections of other components in the transmission line.
- Ideally complete absorption of power takes place in one direction and lossless transmission is provided in the opposite direction
- Also called UNILINE, it is used to improve the frequency stability of microwave generators like klystrons and magnetrons in which reflections from the load affects the generated frequency.
- It can be made by terminating ports 3 and 4 of a 4-port circulator with matched loads.
- Additionally it can be made by inserting a ferrite rod along the axis of a rectangular waveguide.

Gyrator:

An important canonical nonreciprocal component is the *gyrator*, which is a two-port device having a 180° differential phase shift. The schematic symbol for a gyrator is shown in Figure 9.21, and the scattering matrix for an ideal gyrator is

$$[S] = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad 9.90$$

which shows that it is lossless, matched, and nonreciprocal. Using the gyrator as a basic nonreciprocal building block in combination with reciprocal dividers and couplers can lead to useful equivalent circuits for nonreciprocal components such as isolators and circulators. Figure 9.22, for example, shows an equivalent circuit for an isolator using a gyrator and two quadrature hybrids.

The gyrator can be implemented as a phase shifter with a 180° differential phase shift; bias can be provided with a permanent magnet, making the gyrator a passive device.

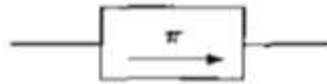


FIGURE 9.21 Symbol for a gyrator, which has a differential phase shift of 180° .

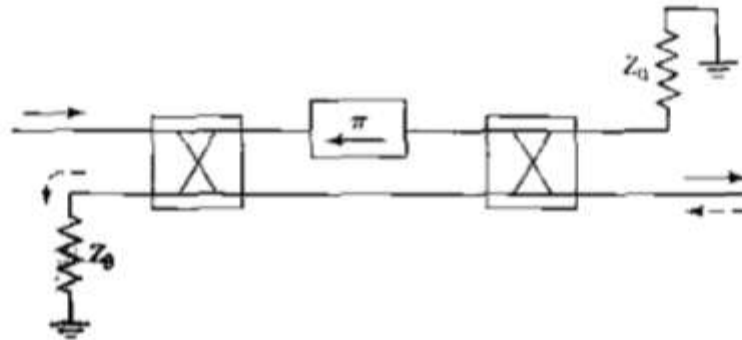


FIGURE 9.22 An isolator constructed with a gyrator and two quadrature hybrids. The forward wave (→) is passed, while the reverse wave (←) is absorbed in the matched load of the first hybrid.

Introduction to Radar Systems

- Radar is an electromagnetic system for the detection and location of objects.
- Radar is used to extend the capability of one's senses for observing the environment, especially the sense of vision. In addition, radar able to measure the distance or range to the object which is probably its most important attribute.
- An elementary form of radar consists of a transmitting antenna emitting electromagnetic radiation generated by an oscillator of some sort, a receiving antenna, and an energy-detecting device or receiver. A portion of the transmitted signal is intercepted by a reflecting object (target) and is reradiated in all directions. It is the energy reradiated in the back direction that is of prime interest to the radar.
- The receiving antenna collects the returned energy and delivers it to a receiver, where it is processed to detect the presence of the target and to extract its location and relative velocity. The distance to the target is determined by measuring the time taken for the radar signal to travel to the target and back.
- The direction, or angular position, of the target may be determined from the direction of arrival of the reflected wave front. The usual method of measuring the direction of arrival is with narrow antenna beams. If relative motion exists between target and radar, the shift in the carrier frequency of the reflected wave (doppler effect) is a measure of the target's relative (radial) velocity and may be used to distinguish moving targets from stationary objects. In radars which continuously track the movement of a target, a continuous indication of the rate of change of target position is also available.
- The most common radar waveform is a train of narrow, rectangular-shape pulses modulating a sine wave carrier. The distance, or range, to the target is determined by measuring the time T_R taken by the pulse to travel to the target and return. Since electromagnetic energy propagates at the speed of light $c = 3 \times 10^8$ m/s, the range R is

$$R = \frac{c T_R}{2}$$
- The factor 2 appears in the denominator because of the two-way propagation of radar. With the range in kilometers or nautical miles, and T_R in microseconds, then

$$R(\text{Km}) = 0.15 T_R (\mu\text{s}) \text{ or } R(\text{nmi}) = 0.081 T_R (\mu\text{s})$$

Each microsecond of round-trip travel time corresponds to a distance of 0.081 nautical mile, 0.093 statute mile, 150 meters, 164 yards, or 492. feet.

- Once the transmitted pulse is emitted by the radar, a sufficient length of time must elapse to allow any echo signals to return and be detected before the next pulse may be transmitted. Therefore the rate at which the pulses may be transmitted is determined by the longest range at which targets are expected. If the pulse repetition frequency is too high, echo signals from some targets might arrive after the transmission of the next pulse, and ambiguities in measuring range might result.
- Echoes that arrive after the transmission of the next pulse are called *second-time-around* (or multiple-time-around) echoes. Such an echo would appear to be at a much shorter range than the actual and could be misleading if it were not known to be a second-time-around echo. The range beyond which targets appear as second-time-around echoes is called the maximum unambiguous range and is given by

$$R_{\text{unamb}} = \frac{c}{2f_p}$$

Where, f_p is pulse repetition frequency, in Hz.

- Although the typical radar transmits a simple pulse-modulated waveform, there are a number of other suitable modulations that might be used. The pulse carrier might be frequency- or phase-modulated to permit the echo signals to be compressed in time after reception. This achieves the benefits of high range-resolution without the need to resort to a short pulse. The technique of using a long, modulated pulse to obtain the resolution of a short pulse, but with the energy of a long pulse, is known as pulse compression.
- Continuous waveforms (CW) also can be used by taking advantage of the doppler frequency shift to separate the received echo from the transmitted signal and the echoes from stationary clutter. Un-modulated CW waveforms do not measure range, but a range measurement can be made by applying either frequency- or phase-modulation.

Simple Form of Radar Equation:

- The radar equation relates the range of a radar to the characteristics of the transmitter, receiver, antenna, target, and environment. It is useful not just as a means for determining the maximum distance from the radar to the target, but it can serve both as a tool for understanding radar operation and as a basis for radar design.

- If the power of the radar transmitter is denoted by P_t , and if an isotropic antenna is used (one which radiates uniformly in all directions), the power density (watts per unit area) at a distance R from the radar is equal to the transmitter power divided by the surface area $4\pi R^2$ of an imaginary sphere of radius R , or

$$\text{Power density from isotropic antenna} = \frac{P_t}{4\pi R^2}$$

- Radars employ directive antennas to channel, or direct the radiated power P_t into some particular direction. The gain G of an antenna is a measure of the increased power radiated in the direction of the target as compared with the power that would have been radiated from an isotropic antenna. It may be defined as the ratio of the maximum radiation intensity from the subject antenna to the radiation intensity from a lossless, isotropic antenna with the same power input. (The radiation intensity is the power radiated per unit solid angle in a given direction.) The power density at the target from an antenna with a transmitting gain G is

$$\text{Power density from directive antenna} = \frac{P_t G}{4\pi R^2}$$

- The target intercepts a portion of the incident power and reradiates it in various directions. The measure of the amount of incident power intercepted by the target and reradiated back in the direction of the radar is denoted as the radar cross section σ , and is defined by the relation

$$\text{Power density of echo signal at radar} = \frac{P_t G}{4\pi R^2} \frac{\sigma}{4\pi R^2}$$

- The radar cross section σ has units of area. It is a characteristic of the particular target and is a measure of its size as seen by the radar. The radar antenna captures a portion of the echo power. If the effective area of the receiving antenna is denoted A_e , the power P_r received by the radar is

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

- The maximum radar range $R_{\max}$ is the distance beyond which the target cannot be detected. It occurs when the received echo signal power P_r just equals the minimum detectable signal $S_{\min}$. Therefore

$$R_{\max} = \left[\frac{P_t G A_e \sigma}{(4\pi)^2 S_{\min}} \right]^{1/4}$$

- This is the fundamental form of the radar equation. Note that the important antenna parameters are the transmitting gain and the receiving effective area. Antenna theory gives the relationship between the transmitting gain and the receiving effective area of an antenna as

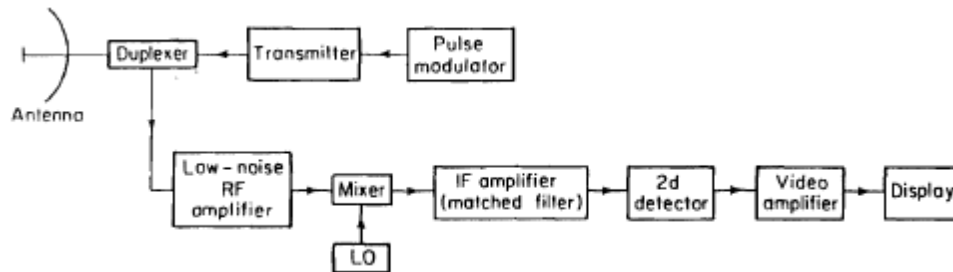
$$G = \frac{4\pi A_e}{\lambda^2}$$

- Since radars generally use the same antenna for both transmission and reception, then first for A_e , then for G , to give two other forms of the radar equation

$$R_{\max} = \left[\frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 S_{\min}} \right]^{1/4}$$

$$R_{\max} = \left[\frac{P_t A_e^2 \sigma}{4\pi \lambda^2 S_{\min}} \right]^{1/4}$$

Radar Block Diagram and Operation:



Block diagram of a pulse radar.

- A common form of radar antenna is a reflector with a parabolic shape, fed (illuminated) from a point source at its focus.
- The transmitter may be an oscillator, such as a magnetron, that is pulsed (turned on and on) by the modulator to generate a repetitive train of pulses.
- The waveform generated by the transmitter travels via a transmission line to the antenna, where it is radiated into space.
- A single antenna is generally used for both transmitting and receiving. The receiver must be protected from damage caused by the high power of the transmitter. This is the function of the duplexer.

- The duplexer also serves to channel the returned echo signals to the receiver and not to the transmitter. The duplexer might consist of two gas-discharge devices, one known as a TR (transmit-receive) and the other an ATR (anti-transmit-receive). The TR protects the receiver during transmission and the ATR directs the echo signal to the receiver during reception. Solid-state ferrite circulators and receiver protectors with gas-plasma TR devices and/or diode limiters are also employed as duplexers.
- The receiver is usually of the superheterodyne type. The first stage might be a low-noise RF amplifier, such as a parametric amplifier or a low-noise transistor. However, it is not always desirable to employ a low-noise first stage in radar.
- The mixer and local oscillator (LO) convert the RF signal to an intermediate frequency (IF).
- The IF amplifier should be designed as a matched filter i.e., its frequency-response function $H(f)$ should maximize the peak signal-to-mean-noise-power ratio at the output. This occurs when the magnitude of the frequency-response function $|H(f)|$ is equal to the magnitude of the echo signal spectrum $|S(f)|$ and the phase spectrum of the matched filter is the negative of the phase spectrum of the echo signal.
- After maximizing the signal-to-noise ratio in the IF amplifier, the pulse modulation is extracted by the second detector and amplified by the video amplifier to a level where it can be properly displayed, usually on a cathode-ray tube (CRT).

Application of Radar:

- Air Traffic Control
- Aircraft Navigation
- Ship Safety
- Space
- Remote Sensing
- Law enforcement
- Military

CW and Frequency modulated Radar

Doppler Effect:

- A radar detects the presence of objects and locates their position in space by transmitting electromagnetic energy and observing the returned echo. A pulse radar transmits a

relatively short burst of electromagnetic energy, after which the receiver is turned on to listen for the echo. The echo not only indicates that a target is present, but the time that elapses between the transmission of the pulse and the receipt of the echo is a measure of the distance to the target. Separation of the echo signal and the transmitted signal is made on the basis of differences in time.

- The radar transmitter may be operated continuously rather than pulsed if the strong transmitted signal can be separated from the weak echo. The received-echo-signal power is considerably smaller than the transmitter power
- A feasible technique for separating the received signal from the transmitted signal when there is relative motion between radar and target is based on recognizing the change in the echo-signal frequency caused by the doppler effect.
- It is well known in the fields of optics and acoustics that if either the source of oscillation or the observer of the oscillation is in motion, an apparent shift in frequency will result. This is the doppler effect and is the basis of CW radar.
- If R is the distance from the radar to target, the total number of wavelengths λ contained in the two-way path between the radar and the target is $2R/\lambda$. The distance R and the wavelength λ are assumed to be measured in the same units. Since one wavelength corresponds to an angular excursion of 2π radians, the total angular excursion ϕ made by the electromagnetic wave during its transit to and from the target is $4\pi R/\lambda$ radians. If the target is in motion, R and the phase ϕ are continually changing. A change in ϕ with respect to time is equal to a frequency. This is the doppler angular frequency ω_d and is given by

$$\omega_d = 2\pi f_d = \frac{d\phi}{dt} = \frac{4\pi}{\lambda} \frac{dR}{dt} = \frac{4\pi v_r}{\lambda}$$

where f_d is the dopplers frequency shift and V_r is the relative velocity of target with respect to radar. The dopplers frequency is

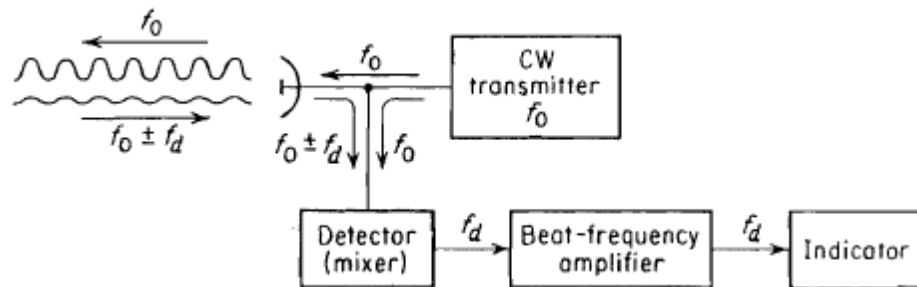
$$f_d = \frac{2v_r}{\lambda} = \frac{2v_r f_0}{c}$$

Where f_0 = transmitted frequency, c is the velocity of light, then

$$f_d = \frac{1.03v_r}{\lambda}$$

The relative velocity may be written $v_r = v \cos \theta$, where v is the target speed and θ is the angle made by the target trajectory and the line joining the radar and target. When $\theta=0$ the Doppler frequency is maximum. The Doppler is zero when the trajectory is perpendicular to the radar line of sight.

CW RADAR:



- The transmitter generates a continuous (un-modulated) oscillation of frequency f_0 which is radiated by the antenna. A portion of the radiated energy is intercepted by the target and is scattered, some of it in the direction of the radar, where it is collected by the receiving antenna.
- If the target is in motion with a velocity v_r relative to the radar, the received signal will be shifted in frequency from the transmitted frequency f_0 by an amount $\pm f_d$.
- The plus sign associated with the doppler frequency applies if the distance between target and radar is decreasing (closing target), that is, when the received signal frequency is greater than the transmitted signal frequency. The minus sign applies if the distance is increasing (receding target).
- The received echo signal at a frequency $f_0 \pm f_d$ enters the radar via the antenna and is heterodyned in the detector (mixer) with a portion of the transmitter signal f_0 to produce a doppler beat note of frequency f_d . The sign of f_d is lost in this process.
- The purpose of the doppler amplifier is to eliminate echoes from stationary targets and to amplify the doppler echo signal to a level where it can operate an indicating device.
- The low-frequency cut off must be high enough to reject the dc component caused by stationary targets, but yet it must be low enough to pass the smallest doppler frequency expected. Sometimes both conditions cannot be met simultaneously and a compromise

is necessary. The upper cut off frequency is selected to pass the highest doppler frequency expected.

- The indicator might be a pair of earphones or a frequency meter. If exact knowledge of the doppler frequency is not necessary, ear phones are especially attractive provided the Doppler frequencies lie within the audio-frequency response of the ear. Earphones are not only simple devices but the ear acts as a selective band pass filter with a pass band of the order of 50 Hz centred about the signal frequency. The narrow-band pass characteristic of the ear results in an effective increase in the signal-to-noise ratio of the echo signal. The doppler frequency can also be detected and measured by conventional frequency meters, usually one that counts cycles.

FREQUENCY-MODULATED CW RADAR

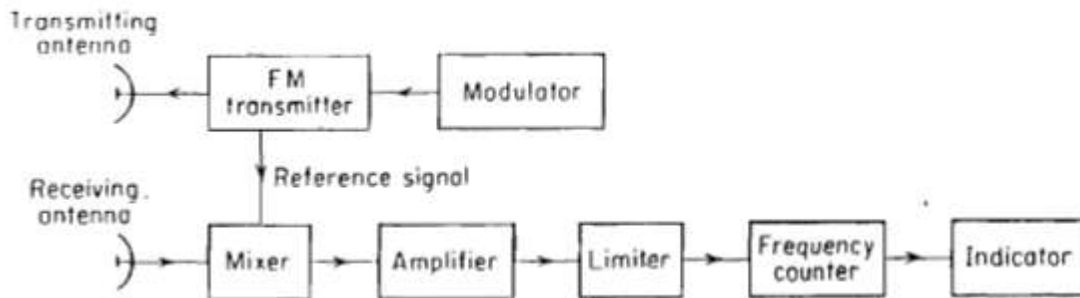
- The inability of the simple CW radar to measure range is related to the relatively narrow spectrum (bandwidth) of its transmitted waveform. Some sort of timing mark must be applied to a CW carrier if range is to be measured. The timing mark permits the time of transmission and the time of return to be recognized. The sharper or more distinct the mark, the more accurate the measurement of the transit time. But the more distinct the timing mark, the broader will be the transmitted spectrum. This follows from the properties of the Fourier transform. Therefore a finite spectrum must of necessity be transmitted if transit time or range is to be measured.
- The spectrum of a CW transmission can be broadened by the application of modulation, either amplitude, frequency, or phase. The narrower the pulse, the more accurate the measurement of range and the broader the transmitted spectrum. A widely used technique to broaden the spectrum of CW radar is to frequency-modulate the carrier. The timing mark is the changing frequency. The transit time is proportional to the difference in frequency between the echo signal and the transmitter signal. The greater the transmitter frequency deviation in a given time interval, the more accurate the measurement of the transit time and the greater will be the transmitted spectrum.
- In the frequency-modulated CW radar (abbreviated FM-CW), the transmitter frequency is changed as a function of time in a known manner. Assume that the transmitter frequency increases linearly with time.

- If there is a reflecting object at a distance R , an echo signal will return after a time $T = 2R/c$. If the echo signal is heterodyned with a portion of the transmitter signal in a nonlinear element such as a diode, a beat note f_b will be produced. If there is no doppler frequency shift, the beat note (difference frequency) is a measure of the target's range and $f_b = f_r$ where f_r is the beat frequency due only to the target's range. If the rate of change of the carrier frequency is $\dot{f}_0$ the beat frequency is

$$f_r = \dot{f}_0 T = \frac{2R}{c} \dot{f}_0$$

- In any practical CW radar, the frequency cannot be continually changed in one direction only. Periodicity in the modulation is necessary, as in the triangular frequency-modulation. The modulation need not necessarily be triangular; it can sawtooth, sinusoidal, or some other shape. The resulting beat frequency as a function of time for triangular modulation. The beat note is of constant frequency except at the turn-around region. If the frequency is modulated at a rate f_m over a range Δf the beat frequency is

$$f_r = \frac{2R}{c} 2f_m \quad \Delta f = \frac{4Rf_m}{c}$$



- A portion of the transmitter signal acts as the reference signal required to produce the beat frequency. It is introduced directly into the receiver via a cable or other direct connection.
- Ideally the isolation between transmitting and receiving antennas is made sufficiently large so as to reduce to a negligible level the transmitter leakage signal which arrives at the receiver via the coupling between antennas. The beat frequency is amplified and limited to remove any amplitude fluctuations. The frequency of the amplitude-limited beat note is measured with a cycle-counting frequency meter calibrated in distance.

- In the above, the target was assumed to be stationary. If this assumption is not applicable, a doppler frequency shift will be superimposed on the FM range beat note and an erroneous range measurement results. The doppler frequency shift causes the frequency-time plot of the echo signal to be shifted up or down. On one portion of the frequency-modulation cycle, the beat frequency is increased by the doppler shift, while on the other portion, it is decreased. If for example, the target is approaching the radar, the beat frequency f_b (up) produced during the increasing, or up, portion of the FM cycle will be the difference between the beat frequency due to the range f_r and the doppler frequency shift f_d . Similarly, on the decreasing portion, the beat frequency f_d (down) is the sum of the two

$$f_b(\text{up}) = f_r - f_d$$

$$f_b(\text{down}) = f_r + f_d$$

- The range frequency f_r may be extracted by measuring the average beat frequency; that is $\frac{1}{2}[f_b(\text{up}) + f_d(\text{down})] = f_r$.
- If $f_b(\text{up})$ and $f_d(\text{down})$ are measured separately, for example, by switching a frequency counter every half modulation cycle, one-half the difference between the frequencies will yield the doppler frequency. This assumes $f_r > f_d$. If, on the other hand $f_r < f_d$ such as might occur with a high-speed target at short range, the roles of the averaging and the difference-frequency measurements are reversed; the averaging meter will measure Doppler velocity, and the difference meter, range. If it is not known that the roles of the meters are reversed because of a change in the inequality sign between f_r and f_d an incorrect Interpretation of the measurements may result.
- When more than one target is present within the view of the radar, the mixer output will contain more than one difference frequency. If the system is linear, there will be a frequency component corresponding to each target.
- To measure the individual frequencies, they must be separated from one another. This might be accomplished with a bank of narrowband filters, or alternatively, a single frequency corresponding to a single target may be singled out and continuously observed with a narrowband tunable filter.
- But if the motion of the targets were to produce Doppler frequency shift or if the frequency-modulation waveform were nonlinear, or if the mixer were not operated in its

linear region, the problem of resolving targets and measuring the range of each becomes more complicated.

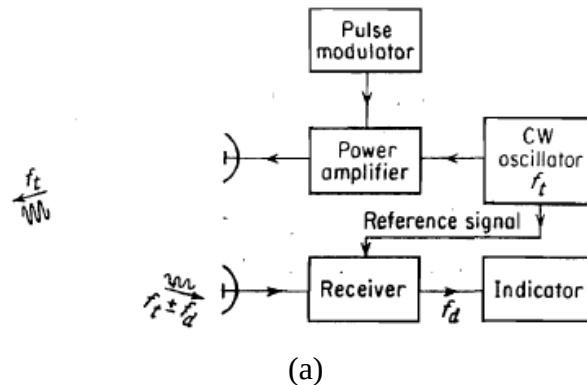
- If the FM-CW radar is used for single targets only, such as in the radio altimeter, it is not necessary to employ a linear modulation waveform. This is certainly advantageous since a sinusoidal or almost sinusoidal frequency modulation is easier to obtain with practical equipments than are linear modulations.
- The beat frequency obtained with sinusoidal modulation is not constant over the modulation cycle as it is with linear modulation. However, it may be shown that the average beat frequency measured over a modulation cycle, yields the correct value of target range.
- Any reasonable-shape modulation waveform can be used to measure the range, provided the average beat frequency is measured. If the target is in motion and the beat signal contains a component due to the doppler frequency shift, the range frequency can be extracted, as before, if the average frequency is measured. To extract the doppler frequency, the modulation waveform must have equal upswing and downswing time intervals.

MTI RADAR

- The doppler frequency shift produced by a moving target may be used in a pulse radar, just as in the CW radar to determine the relative velocity of a target or to separate desired moving targets from undesired stationary objects (clutter).
- Although there are applications of pulse radar where a determination of the target's relative velocity is made from the doppler frequency shift, the use of doppler to separate small moving targets in the presence of large clutter has probably been of far greater interest. Such a pulse radar that utilizes the doppler frequency shift as a means for discriminating moving from fixed targets is called an MTI (moving target indication) or a pulse doppler radar.
- The two are based on the same physical principle, but in practice there are generally recognizable differences between them. The MTI radar, for instance, usually operates with ambiguous doppler measurement (so-called blind speeds) but with unambiguous range measurement (no second-time-around echoes). The opposite is generally the case

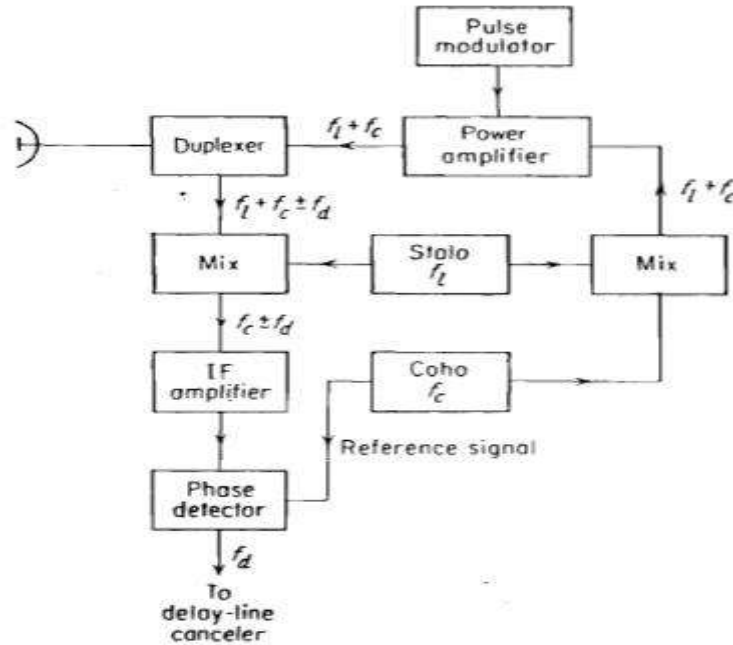
for a pulse doppler radar. Its pulse repetition frequency is usually high enough to operate with unambiguous doppler (no blind speeds) but at the expense of range ambiguities.

Block diagram and principles of operation:



The simple MTI radar

- The significant difference between this MTI configuration (b) and that of (a) is the manner in which the reference signal is generated.
- In Fig. (b), the coherent reference is supplied by a oscillator called the coho, which stands for coherent oscillator. The coho is a stable oscillator whose frequency is the same as the intermediate frequency used in the receiver. In addition to providing the reference signal, the output of coho is mixed with local oscillator frequency. The local oscillator- must also be a stable oscillator and is called stalo, for stable local oscillator.
- The RF echo signal is heterodyned with stalo signal to produce the IF signal just as in the conventional super heterodyne receiver.
- The characteristic feature of coherent MTI radar is that the transmitted signal must be coherent (in phase) with the reference signal in the receiver.
- The function of the stalo is to provide the necessary frequency translation from the IF to the transmitted (RF) frequency. Although the phase of the stalo influences the phase of the transmitted signal, any stalo phase shift is cancelled on reception because the stalo that generates the transmitted signal also acts as the local oscillator in the receiver.
- The reference signal from the coho and the IF echo signal are both fed into a mixer called the phase detector. The phase detector differs from the normal amplitude detector since its output is proportional to the phase difference between the two input signals.

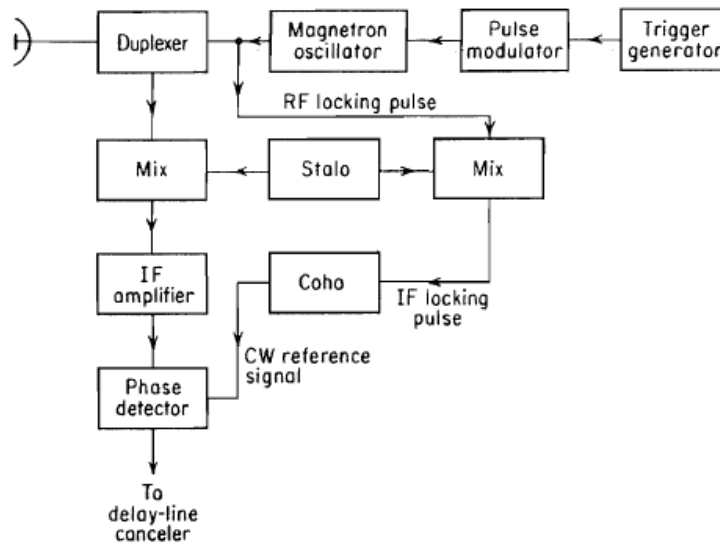


(b)

MTI radar employing a power amplifier

- Any one of a number of transmitting-tube types might be used as the power amplifier. These include the triode, tetrode, klystron, traveling-wave tube, and the crossed-field amplifier.
- A transmitter which consists of a stable low-power oscillator followed by a power amplifier is sometimes called MOPA, which stands for master-oscillator power amplifier.
- A coherent reference signal may be readily obtained with the power oscillator by readjusting the phase of the coho at the beginning of each sweep according to the phase of the transmitted pulse. The phase of the coho is locked to the phase of the transmitted pulse each time a pulse is generated.
- A block diagram of an MTI radar (with a power oscillator) is shown in Fig. (c). A portion of the transmitted signal is mixed with the stalo output to produce an IF beat signal whose phase is directly related to the phase of the transmitter.
- The phase of the coho is then related to the phase of the transmitted pulse and may be used as the reference signal for echoes received from that particular transmitted pulse.

- Upon the next transmission another IF locking pulse is generated to relock the phase of the CW coho until the next locking pulse comes along.



(c)

Block diagram of MTI radar with power-oscillator transmitter.

Delay line cancellers

- The capability of this device depends on the quality of the medium used is the delay line. The Pulse modulator delay line must introduce a time delay equal to the pulse repetition interval.
- For typical ground-based air-surveillance radars this might be several milliseconds. Delay times of this magnitude cannot be achieved with practical electromagnetic transmission lines.
- By converting the electromagnetic signal to an acoustic signal it is possible to utilize delay lines of a reasonable physical length since the velocity of propagation of acoustic waves is about 10^{-5} that of electromagnetic waves.
- After the necessary delay is introduced by the acoustic line, the signal is converted back to an electromagnetic signal for further processing.
- The use of digital delay lines requires that the output of the MTI receiver phase-detector be quantized into a sequence of digital words. The compactness and convenience of

digital processing allows the implementation of more complex delay-line cancellers with filter characteristics not practical with analog methods.

- One of the advantages of a time-domain delay-line canceller as compared to the more conventional frequency-domain filter is that a single network operates at all ranges and does not require a separate filter for each range resolution cell. Frequency-domain doppler filter banks are of interest in some forms of MTI and pulse-doppler radar.
- The delay-line canceller acts as a filter which rejects the dc component of clutter. Because of its periodic nature, the filter also rejects energy in the vicinity of the pulse repetition frequency and its harmonics.

The video signal received from a particular target at a range R_0 is

$$V_1 = k \sin (2\pi f_d t - \phi_0)$$

where ϕ_0 = phase shift and k = amplitude of video signal. The signal from the previous transmission, which is delayed by a time T = pulse repetition interval, is

$$V_2 = k \sin [2\pi f_d(t - T) - \phi_0]$$

Everything else is assumed to remain essentially constant over the interval T so that k is the same for both pulses. The output from the subtractor is

$$V = V_1 - V_2 = 2k \sin \pi f_d T \cos \left[2\pi f_d \left(t - \frac{T}{2} \right) - \phi_0 \right]$$

- It is assumed that the gain through the delay-line canceller is unity. The output from the canceller consists of a cosine wave at the doppler frequency & with an amplitude $2k \sin \pi f_d T$. Thus the amplitude of the cancelled video output is a function of the Doppler frequency shift and the pulse-repetition interval, or prf. The magnitude of the relative frequency-response of the delay-line canceller [ratio of the amplitude of the output from the delay-line canceler, $2k \sin (\pi f_d T)$, to the amplitude of the normal radar video.

Blind speeds

- The response of the single-delay-line canceller will be zero whenever the argument $\pi f_d T$ in the amplitude factor is $0, \pi, 2\pi, \dots$, etc., or when

$$f_d = \frac{n}{T} = n f_p$$

- where $n = 0, 1, 2, \dots$, and f_p = pulse repetition frequency. The delay-line canceller not only eliminates the d-c component caused by clutter ($n = 0$), but unfortunately it also rejects any moving target whose doppler frequency happens to be the same as the prf or a multiple thereof. Those relative target velocities which result in zero MTI response are called Blind speeds and are given by

$$v_n = \frac{n\lambda}{2T} = \frac{n\lambda f_p}{2} \quad n = 1, 2, 3, \dots$$

where v_n is the n th blind speed. If λ is measured in meters, f_p in Hz, and the relative velocity in knots, the blind speeds are

$$v_n = \frac{n\lambda f_p}{1.02} \approx n\lambda f_p$$

- The blind speeds are one of the limitations of pulse MTI radar which do not occur with CW radar. They are present in pulse radar because doppler is measured by discrete samples (pulses) at the prf rather than continuously.
- If the first blind speed is to be greater than the maximum radial velocity expected from the target, the product λf_p must be large. Thus the MTI radar must operate at long wavelengths (low frequencies) or with high pulse repetition frequencies, or both. Therefore blind speeds might not be easy to avoid.

Staggered PRFs

- In practice, long-range MTI radars that operate in the region of L or S band or higher and are primarily designed for the detection of aircraft must usually operate with ambiguous doppler and blind speeds if they are to operate with unambiguous range.
- The presence of blind speeds within the doppler-frequency band reduces the detection capabilities of the radar. Blind speeds can sometimes be traded for ambiguous range, so that in systems applications which require good MTI performance, the first blind speed might be placed outside the range of expected doppler frequencies if ambiguous range can be tolerated. (Pulse-doppler radars usually operate in this manner).
- The effect of blind speeds can be significantly reduced, without incurring range ambiguities, by operating with more than one pulse repetition frequency. This is called a

staggered-prf MTI. Operating at more than one RF frequency can also reduce the effect of blind speeds.

Tracking Techniques with Radar

- A tracking-radar system measures the coordinates of a target and provides data which may be used to determine the target path and to predict its future position. All or only part of the available radar data-range, elevation angle, azimuth angle, and doppler frequency shift may be used in predicting future position; that is, a radar might track in range, in angle in doppler, or with any combination.
- Almost any radar can be considered a tracking radar, provided its output information is processed properly. But, in general, it is the method by which angle tracking is accomplished that distinguishes what is normally considered a tracking radar from any other radar. It is also necessary to distinguish between a continuous tracking radar and a track-while-scan (TWS) radar. The former supplies continuous tracking data on a particular target, while the track-while-scan supplies sampled data on one or more targets. In general, the continuous tracking radar and the TWS radar employ different types of equipment.
- The antenna beam in the continuous tracking radar is positioned in angle by a servo mechanism actuated by an error signal. The various methods for generating the error signal may be classified as sequential lobing, conical scan, and simultaneous lobing or mono pulse.
- The range and doppler frequency shift can also be continuously tracked, if desired, by a servo-control loop actuated by an error signal generated in the radar receiver. The information available from a tracking radar may be presented on a cathode-ray-tube (CRT) display for action by an operator, or may be supplied to an automatic computer which determines the target path and calculate its probable future course.
- The tracking radar must first find its target before it can track. Some radars operate in a search, or acquisition, mode in order to find the target before switching to a tracking mode. Although it is possible to use a single radar for both the search and the tracking functions, such a procedure usually results in certain operational limitations.

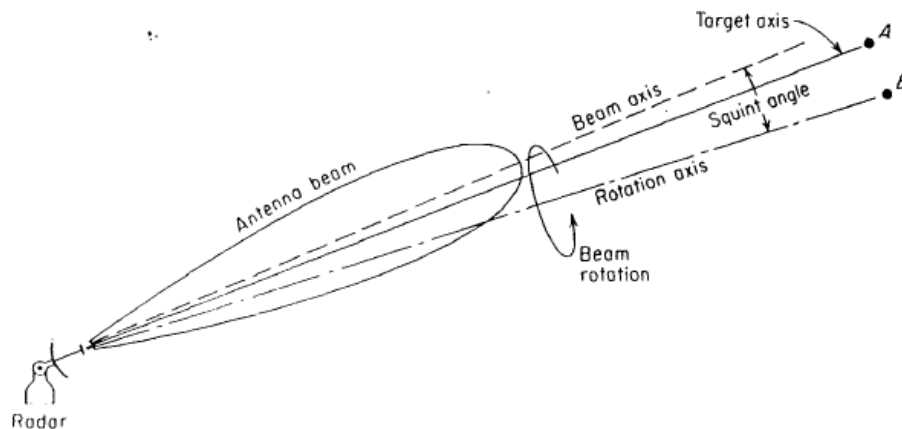
Lobe switching or Sequential lobing:

- Sequential lobing, or lobe switching, was one of the first tracking-radar techniques to be employed. Early applications were in airborne-interception radar, where it provided directional information for homing on a target, and in ground-based anti-aircraft fire-control radars. It is not used as often in modern tracking-radar applications as some of the other techniques.
- The antenna pattern commonly employed with tracking radars is the symmetrical pencil beam in which the, elevation and azimuth beam widths are approximately equal. However, a simple pencil-beam antenna is not suitable for tracking radars unless means are provided for determining the magnitude and direction of the target's angular position with respect to some reference direction, usually the axis of the antenna. The difference between the target position and the reference direction is the angular error. The tracking radar attempts to position the antenna to make the angular error zero. When the angular error is zero, the target is located along the reference direction.
- One method of obtaining the direction and the magnitude of the angular error in one coordinate is by alternately switching the antenna beam between two positions. This is called lobe switching, sequential switching or sequential lobing.
- The difference in amplitude between the voltages obtained in the two switched positions is a measure of the angular displacement of the target from the switching axis. The sign of the difference determines the direction the antenna must be moved in order to align the switching axis with the direction of the target. When the voltages in the two switched positions are equal, the target is on axis and its position may be determined from the axis direction.
- Two additional switching positions are needed to obtain the angular error in the orthogonal coordinate. Thus a two-dimensional sequentially lobing radar might consist of a cluster of four feed horns illuminating a single antenna, arranged so that the right-left, up-down sectors are covered by successive antenna positions. Both transmission and reception are accomplished at each position.
- A cluster of five feeds might also be employed, with the central feed used for transmission while the outer four feeds are used for receiving. High-power RF switches are not needed since only the receiving beams, and not the transmitting beam, are stepped in this five-feed arrangement.

- One of the limitations of a simple un-switched non-scanning pencil-beam antenna is that the angle accuracy can be no better than the size of the antenna beam width.
- An important feature of sequential lobing is that the target-position accuracy can be far better than that given by the antenna beam width. The accuracy depends on how well equality of the signals in the switched positions can be determined. The fundamental limitation to accuracy is system noise caused either by mechanical or electrical fluctuations.

Conical scan:

- A logical extension of the simultaneous lobing technique is to rotate continuously an offset antenna beam rather than discontinuously step the beam between four discrete positions. This is known as conical Scanning.
- The angle between the axis of rotation (which is usually, but not always, the axis of the antenna reflector) and the axis of antenna beam is called the squint angle.



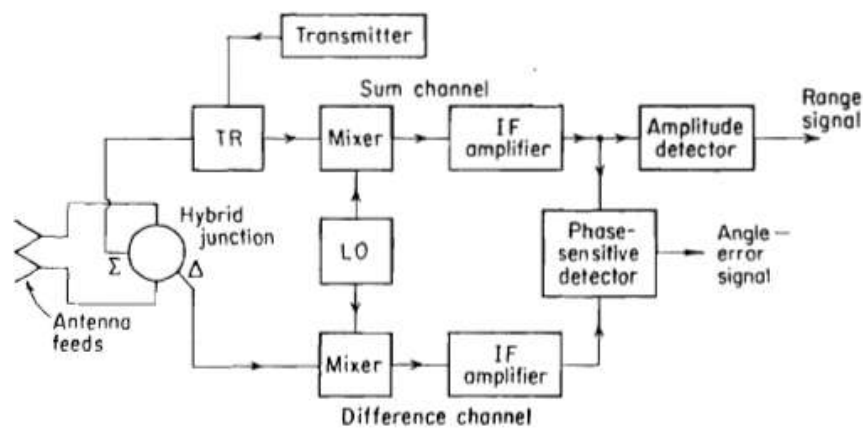
- Consider a target at position A. The echo signal will be modulated at a frequency equal to the rotation frequency of the beam.
- The amplitude of the echo signal modulation will depend upon the shape of the antenna pattern, the squint angle and the angle between the target line of sight and the rotation axis.
- The phase of the modulation depends on the angle between the target and the rotation axis.
- The conical scan modulation is extracted from the echo signal and applied to a servo-control system which continually positions the antenna on the target.
- When the antenna is on target, as in B, the line of sight to the target and the rotation axis coincide, and the conical-scan modulation is zero.

Mono Pulse Tracking Radar System

- The conical-scan and sequential-lobing tracking radars require a minimum number of pulses in order to extract the angle-error signal. In the time interval during which a measurement is made with either sequential lobing or conical scan, the train of echo pulses must contain no amplitude-modulation components other than the modulation produced by scanning. If the echo pulse-train did contain additional modulation components, caused, for example, by a fluctuating target cross section, the tracking accuracy might be degraded, especially if the frequency components of the fluctuations were at or near the conical-scan frequency or the sequential-lobing rate. The effect of the fluctuating echo can be sufficiently serious in some applications to severely limit the accuracy of those tracking radars which require many pulses to be processed in extracting the error signal.
- Pulse-to-pulse amplitude fluctuations of the echo signal have no effect on tracking accuracy if the angular measurement is made on the basis of one pulse rather than many. There are several methods by which angle-error information might be obtained with only a single pulse. More than one antenna beam is used simultaneously in these methods, in contrast to the conical-scan or lobe-switching tracker, which utilizes one antenna beam on a time-shared basis. The angle of arrival of the echo signal may be determined in a single-pulse system by measuring the relative phase or the relative amplitude of the echo pulse received in each beam. The names simultaneous lobing and mono pulse are used to describe those tracking techniques which derive angle-error information on the basis of a single pulse.
- A simultaneous-lobing technique is amplitude-comparison mono pulse or more simply, mono pulse. In this technique the RF signals received from two offset antenna beams are combined so that both the sum and the difference signals are obtained simultaneously. The sum and difference signals are multiplied in a phase-sensitive detector to obtain both the magnitude and the direction of the error signal. All the information necessary to determine the angular error is obtained on the basis of a single pulse; hence the name mono pulse is quite appropriate.

Amplitude-Comparison Mono Pulse

- The amplitude-comparison mono pulse employs two overlapping antenna patterns to obtain the angular error in one coordinate.
- The two overlapping antenna beams may be generated with a single reflector or with a lens antenna illuminated by two adjacent feeds. (A cluster of four feeds may be used if both elevation- and azimuth-error signals are wanted.)
- The sum and difference of the two antenna patterns are made.
- The sum patterns is used for transmission, while both the sum pattern and the difference pattern are used on reception. The signal received with the difference pattern provides the magnitude of the angle error.
- The sum signal provides the range measurement and is also used as a reference to extract the sign of the error signal. Signals received from the sum and the difference patterns are amplified separately and combined in a phase-sensitive detector to produce the error-signal characteristic.



Block diagram of amplitude-comparison mono pulse radar.

- A block diagram of the amplitude-comparison-monopulse tracking radar for a single angular coordinate is shown in above Figure.
- The two adjacent antenna feeds are connected to the two arms of a hybrid junction such as a "magic T," a "rat race," or a short-slot coupler. The sum and difference signals appear at the two other arms of the hybrid.
- On reception, the outputs of the sum arm and the difference arm are each heterodyned to an intermediate frequency and amplified as, in any super heterodyne receiver. The transmitter is connected to the sum arm. Range information is also extracted from the sum channel. A duplexer is included in the sum arm for the protection of the receiver.

The output of the phase-sensitive detector is an error signal whose magnitude is proportional to the angular error and whose sign is proportional to the direction.

- The output of the mono pulse radar is used to perform automatic tracking. The angular error signal actuates a servo-control system to position the antenna, and the range output from the sum channel feeds into an automatic-range-tracking unit.
- The sign of the difference signal (and the direction of the angular error) is determined by comparing the phase of the difference signal with the phase of the sum signal. If the sum signal in the IF portion of the receiver were $A_s \cos \omega_{IF}t$ the difference signal would be either $A_d \cos \omega_{IF}t$ or $-A_d \cos \omega_{IF}t$ ($A_s > 0$, $A_d > 0$) depending on which side of center is the target.
- Since, $-A_d \cos \omega_{IF}t = A_d \cos \omega_{IF}(t + \pi)$ the sign of the difference signal may be measured by determining whether the difference signal is in phase with the sum or 180° out of phase.
- Although a phase comparison is a part of the amplitude-comparison-monopulse radar, the angular-error signal is basically derived by comparing the echo amplitudes from simultaneous offset beams. The phase relationship between the signals in the offset beams is not used. The purpose of the phase-sensitive detector is to conveniently furnish the sign of the error signal.